The Arts and Humanities Data Service (AHDS) is a national service funded by the Joint Information Systems Committee (JISC) of the Higher and Further Education Funding Councils and the Arts and Humanities Research Council to collect, preserve and promote the electronic resources which result from research and teaching in the arts and humanities.

DIGITAL MOVING IMAGES AND SOUND ARCHIVING STUDY

Authors:	Andrew Wilson, AHDS Richard Wright, BBC Malcolm Polfreman, AHDS Sheila Anderson, AHDS Simon Tanner, KCL Digital Consultancy Services Emma Beer, AHDS
Date:	August 2006
Version:	Release Final
Document Name	Digital Moving Images and Sound Archiving Study
Notes:	Limited Circulation

Acknowledgements
Many thanks to those who contributed to this Archiving Study by participation in interviews and through responding to the online survey promoted through AHDS Executive and AHDS Performing Arts. Thanks also to those who submitted comments on the release draft of this report, and in particular Carl Fleischhauer of the Library of Congress for his generous advice and comments.

Table of Contents

Table of Contents 2

Chapter 1: Summary and Recommendations 4
1.1 Introduction 4
1.2 Curator survey 5
1.3 Current research and practice 6
1.4 Characteristics of moving images and sound 7
1.5 Preservation of moving images and sound 9
1.6 Metadata 10
1.7 Social, economic and legal aspects 11

Chapter 2: Introduction and Scope of Study 13
2.1 Outline of study 13
2.2 Outline of scope and methodology 14
2.3 Project outcomes 14

Chapter 3: Overview of current research 16
3.1 Background 16
3.2 Hidden Treasures: The UK Audiovisual Archive Strategic Framework 16
3.3 Overview of current research 17
3.4 Some key audiovisual preservation initiatives 19

Chapter 4: Survey on Preservation of Digital Moving Images and Sound 21
4.1 Introduction 21
Analysis of results 21
The most important issues 21
4.2 The key barriers 23
4.3 Conclusion 25

Chapter 5: Properties of Digital Moving Images and Sound 27
5.1 Format overview 27
5.2 Moving Images 29
5.3 Digital Sound 47

Chapter 6: Digital Preservation Methods: Moving Images and Sound 55
6.1 Overview and assessment of preservation methods 55
6.2 Assessment of methods for preservation of moving image and sound 60
6.3 Recommendations for preservation of moving images and sound 62

Chapter 7: Moving Images and Sound Metadata Review and Requirements 66
7.1 Introduction 66
7.2 Technical metadata 66
7.3 Management metadata 68
7.4 Discovery metadata 69
7.5 Review of existing metadata standards 70
7.6 Metadata extraction and storage 79
7.7 Conclusion 80

Chapter 8: Life Cycle and Organisational Models 90
8.1 Functional requirements: OAIS 90
8.2 OAIS functional entities (simplified) 90
8.3 Trusted digital repositories (TDR) 92
8.4 Digital moving image and sound life-cycle model: OAIS and TDR 93
8.5 Organisational models 98
8.6 Disaggregated models 99

Chapter 9: Assessing Costs for Preserving Digital Moving Images and Sound 101
9.1 Introduction 101
Reducing costs – automation and storage? 102
9.2 Summary 105

Chapter 10 References 106

Appendix: Collections and Support Bodies 110

Chapter 1: Summary and Recommendations

1.1 Introduction

Management and preservation requirements for digital materials are fundamentally different from analogue materials. Digital materials can be created using a wide range of technologies and formats, whether born digital or digital surrogates of existing analogue materials. They can be described and documented in a variety of ways – or not at all. They are subject to both physical deterioration and technical obsolescence. More than one copy can be easily and simply created. Access may be provided through more than one point, and may be distributed. All these factors will impinge upon the approach taken to their management and long-term preservation.

These differences present the curators of digital materials with some fundamental challenges. The way in which materials are created, particularly the technologies used, will determine how conducive to long-term preservation the materials are, and will present varied challenges to curators charged with the subsequent management and preservation of the materials. Curators will need adequate metadata about the resource if they are to successfully manage, preserve and make the materials accessible. Multiple copies may also imply multiple versions – the digital resource curator must somehow ensure the integrity and authenticity of the resource. They must be aware of changing technologies and fragility of media and take these into consideration from an early stage in the ingest process.

Jones and Beagrie define digital preservation as: "…the series of managed activities necessary to ensure continued access to digital materials for as long as necessary. Digital preservation….refers to all the actions required to maintain access to digital materials beyond the limits of media failure or technological change."[1]

All this suggests that digital curation and preservation requires a more pro-active approach beginning at an earlier stage in the material's lifecycle than would traditionally be the case with analogue materials. Within the digital preservation community, the concept of the life-cycle management of digital resources has emerged to describe and document the active management processes that need to take place, and the key decision making and intervention points along the continuum. The life-cycle concept has been incorporated into OAIS Reference Model, now adopted as an ISO standard for digital preservation. The OAIS model is proving a strong foundation for the development of digital archiving projects and services, and is increasingly being implemented by digital libraries and archives.

However, despite these developments, the difficulty for those undertaking digital preservation or with responsibility for providing access in the long term to digital resources is the lack of practical advice, and of robust tools and mature techniques. A number of digital preservation strategies have been proposed, but there is no definitive approach to the problem of maintaining digital content across multiple generations of technology, nor is there likely to be one. The problem is particularly acute in the area of moving images and sound, where development of technical solutions and agreed metadata sets lags behind that of other resource types, primarily due to the complex nature of digital moving image and sound materials.

This complexity makes preservation challenging if one wishes to put a positive spin on it, or downright difficult, if one chooses to use plain language. Currently we are in a state of transition between the analogue and the digital, and where not all digital audiovisual materials are stored as computer files. For example, radio and television signals are broadcast as real-time bit streams that do not resemble regular computer files. Decisions

[1] Jones, M and Beagrie, N, 'Preservation Management of Digital Materials: a Handbook', British Library, London, 2002

about how these are created and managed are driven, quite rightly, by how they are to be delivered to the user, rather than by the requirements of long-term preservation. Over time, we might expect to see a move towards digital audiovisual materials that more closely resemble the type of files with which we are more used to dealing with. Until then, the preservation community will need to understand how to get non-file-format audiovisual materials into and out of file formats (that is, into and out of computers), and to evaluate and decide upon the appropriate formats and codecs that are best used for preservation purposes.

At the present time (2006) the state of moving image preservation is still unsettled and uncertain. Little practical research into moving image reformatting has been carried out, and until this situation has been remedied much of the advice given in this report is necessarily interim and speculative. Recommendations offered are the best that can be provided given our current state of knowledge but it must be acknowledged that they are based largely in abstract ideas and not in practical experience.

A first draft of this report was released for comment in July 2006 and a number of comments were received. One feature of the comments that has ramifications for the topic of this report is the lack of agreement among the expert commentators on preservation of moving images. There were differences of opinion about strategies, and about formats, and it became apparent that, while there is general agreement about preservation of digital sound, there is, as yet, no emerging consensus as to the best approach for the long-term preservation of digital moving image resources. In compiling this Report, the AHDS has done the best it can within the constraints of the resources (both financial and human) available to it, but it is clear that more investigative work needs to be done in the area of digital moving images.

We recommend, therefore, that JISC institute as a priority, a technology watch to keep up-to-date information on developments in preservation of moving image and video content, and to ensure that the digital preservation community is kept abreast of developments in the area.

We further recommend that JISC undertake further investigative and analytical studies of the whole area of moving images, in particular that JISC carry out individual studies of specific moving image technologies, including digital TV, digital broadcast technologies, digital cinema, and digital video.

In the meantime, it will help to understand more fully the source of audiovisual materials for use in education, and thus what the preservation challenges are. For example, who is producing them and for what purpose; how are they currently managed; what is their educational use, and how are they best presented for educational use; what are the key sources for further and higher education; who has responsibility for these sources? This report offers a summary of some of the major collections and curators of those collections in Appendix A. However, there is much investigative work that needs to be done to ascertain the preservation and use issues associated with the place of digital audio-visual resources in education.

Therefore, we recommend that a more detailed piece of research investigating these issues be undertaken to inform decision making for preservation, and for assessing the likely future costs of preserving much of this material.

1.2 Curator survey

A key part of this project was to find out what creators, curators, and others with responsibility for managing and preserving digital audiovisual materials were doing, what their challenges were, and what they perceived as the greatest needs when pursuing the preservation of this kind of material. To that end, the Project undertook an on-line survey, followed up by a more in-depth questionnaire to a selected group of respondents.

The project received responses from some 92 individuals and organisations working to preserve audiovisual materials. The clear message was that most were at an early stage of development and whilst preservation was at the forefront of their plans, much of the work and effort was going into the more pragmatic area of establishing systems and workflows for capture, access and dissemination, deciding upon metadata, and clearing rights issues.

The majority of respondents were bringing their content into a managed environment with the expectation that, given time, the preservation issues could be resolved and that best practice would emerge from community effort. Many felt that the more immediate concerns of dealing with resourcing for this activity, dealing with copyright and so on, were difficult enough without the added burden of developing detailed preservation plans and strategies, especially in the absence of agreed standards and protocols. However, the number and fullness of the responses would indicate that a high level of concern exists regarding preservation, and that more could be done given an adequate level of funding. Much of the work going on in the area was project funded and therefore wholly inadequate in terms of developing long-term plans.

Those that were considering preservation were struggling with the lack of agreed practice and standards and particular issues arose with the many and varied file formats and codecs, and which of these may eventually emerge as a de facto standard for preservation. Similarly, a set of significant properties that the community could agree upon seemed lacking and little attempt seems to have been made to draw together efforts in this area with many respondents feeling they were making it up as they went along. Interestingly, several respondents mentioned the prevalence of commercial players in this field who promised much. Respondents felt they didn't have the necessary expertise to assess the validity of these claims, and to separate the reality of what they offered from the marketing hyperbole.

The key requirements that emerged from the survey were:

- An agreed set of standard formats for the preservation of moving image and sound
- An agreed set of significant properties for preservation purposes
- An agreed, flexible, metadata set to cover all aspects from search and retrieve through to preservation metadata
- Adequate storage for large volumes of content - this is particularly acute for uncompressed content
- Adequate funding!

The first three must be seen as three of the key recommendations from this Report, and the fourth should be addressed by exploring how a distributed environment for the shared preservation of content might be developed. The Report recommends that the JISC fund a study to investigate the use of the Grid and Grid technologies to support the distributed preservation of large-scale audio-visual content.

1.3 Current research and practice

A major study of the audiovisual landscape in the UK was the Museums, Libraries and Archives Council funded study: *Hidden Treasures: The UK Audiovisual Archive Strategic Framework.* The report provided a blueprint for the future development of audiovisual archives across the UK. It argues that the lack of strategic focus for national and regional planning in the audiovisual archives sector has led to insufficient levels of funding and investment for core activities, and a very real concern that a significant amount of the UK's moving image and sound archive heritage could be lost. The document proposed that the national strategic and funding bodies should work with the audiovisual archives to develop a national framework of institutional vision in which national, regional and local responsibilities are respectively understood and well resourced, with the aim of ensuring comprehensive coverage for audiovisual archive activity across the UK. The document, rather than proposing a strategy, proposed a strategic framework from which (it hoped) a comprehensive national strategy would arise. The report's key recommendations are, nevertheless, worth following up and acting upon. **We recommend that JISC convene a meeting of the relevant bodies named in the *Hidden Treasures* report to progress adoption of the recommendations of that report.**

Current and recent projects address a variety of issues relating to digital preservation. For instance, the 'Presto' and 'Prestospace' projects look at preservation and access for audio, video and film.[2] 'Archiving the Avant Garde' looks at strategies for describing and preserving for performance, installation, conceptual and digital art.[3] Two other projects work to provide the means for 'systematic preservation'. The 'International Database of Virtual Art' provides individual documentation – from video documentation, technical data, interfaces and displays.[4] 'Permanence Through Change: The Variable Media Approach' has created a network for tools, methods and standards development for their unconventional new preservation strategy for conceptual, minimalist and video art.[5]

There is also a significant body of work developing across a number of organisations and projects, including the British Library Sound Archive, the BBC, BUFVC, and the Library of Congress NDIPP project. Projects focusing on the preservation of moving pictures and sound are investigating the use of a range of methodologies to tackle the challenges thrown up by new technologies under the digital sound and moving pictures umbrella.

Many of these projects have been established in response to gaps in the field, and this report summarises and lists many of these in Chapter 3 and in the Appendix. However, there is still more work to be done and the development of a 'roadmap' that could identify areas of work and interest would be a vital step forward in bringing together many of the important recommendations and investigations into preservation methodology, data loss, strategy, understanding of preservation methods and so on.

We therefore recommend that JISC consider funding for the development of just such a roadmap, or perhaps an on-line information source (perhaps through the DCC) that would act as a continually updated source of information for practitioners engaged in the preservation of audiovisual materials.

1.4 Characteristics of moving images and sound

The main body of the Report seeks to identify and discuss the complexity of digital audiovisual materials, and to tease out the properties that relate to both. Moving image and sound resources are time-based, and have significantly more complex characteristics than most other digital resources. Some terminological distinctions must be made when discussing moving image and sound formats. It is most important to understand the concepts of:

- **coding/encoding**: how bitstreams are structured digitally.
- **file format**: structures for formatting bitstreams.

[2] http://www.prestospace.org/
[3] http://www.bampfa.berkeley.edu/about_bampfa/avantgarde.html
[4] http://www2.hu-berlin.de/grau/database.htm
[5] http://variablemedia.net/e/welcome.html

- **wrapper format**: structures for packaging together various types of data.
- **codec**: software (and some hardware) methods for actually carrying out encoding, usually involving a form of compression.

Moving images

Moving images, whether digital or analogue, consist of a succession of still images. Playing back these images at particular speeds (frame rate) conveys the illusion of movement. The nominal frame rate for video is 30 frames per second (fps), for film it is 24 fps in North America and 25 fps in Europe. However, there are significant differences in the technologies of film and video/television that need to be understood. Film is in some ways simpler to deal with than television, because the technologies involved are less complex. A film is a discrete sequence of still images that when played back on a projector (whatever its form) create a continuous moving image over time.

Television images are produced by a process that scans the original and then broadcasts that two-dimensional scanned image as a one-dimensional structured image (similar in some ways to broadcasting audio). These signals can be interlaced or non-progressive, although the advent of LCD displays means that interlacing, introduced to overcome the problem of flicker on displays that use scanning processes to reproduce images, is not necessary for reproducing video on such monitors.

The important characteristics to be considered when storing, transmitting, and preserving film and video are:

- **resolution**: the number of lines per scan in television (video) signals, or the pixel depth of film images
- **size**: the actual dimensions of the still images which together form the moving image file.
- **aspect ratio**: the shape of the image
- **frame rate and fields**: frame rate refers to the speed at which the images are shown in succession
- **bit rate**: represents the amount of information stored per unit of time
- **bit depth**: how many bits are used to represent the colour of each pixel in an image.
- **compression method (codec)**: the manner in which the digital video file has been compressed for transmission and storage. Video compression results in loss of picture quality and the compressed files are more prone to bit corruption.

Sound

Sound/audio files are in some ways less complex than moving image formats, since they are essentially a moving waveform that can be represented more easily than moving images. Digital sound represents the analogue waveforms using sampling methods that convert the shape of the wave over time into a series of numbers. The most common method for quantifying sound (as this process is called) is linear Pulse Code Modulation (LPCM), a technique that has existed since 1937.

The significant properties of sound files are:

- **bit depth**: the number of bits that are used by each sample to represent the audio signal
- **sampling rate**: the number of times per second the audio signal is measured.
- **compression method (codec)**: the manner in which the digital audio file has been compressed for transmission and storage.
- **number of channels**: this property directly influences the size of the resulting file (as do sampling rate and bit depth).

1.5 Preservation of moving images and sound

Currently, it seems that some form of migration offers the best options for preservation of digital moving images and sound, although when emulation concepts and technologies mature, an approach adopting both migration and emulation may prove to be the best overall strategy. Preservation activities today are dealing with a range of difficult issues and the territory remains complex and unsettled. Advice in this report concerning preservation formats and encoding methods is at best based on abstract ideas rather than practical experience, and is to some extent still speculative. Indeed, one of the most useful pieces of guidance for the community would be an in-depth case study of an actual moving image preservation (reformatting) implementation. This would be of great benefit to practitioners attempting to manage such difficult issues as: the rationale for reformatting decisions; obtaining playback equipment for obsolete formats/media; how are missing technology components compensated for; and how and when is metadata captured and created. Readers must be aware that the answers these and other questions are still being developed and refined, and that the work necessary to develop answers is in a relatively early stage. Only modest amounts of advice can be given until all the issues have been addressed and new practices, based in real life activities, are established.

In general the following requirements are recommended for moving images:

- Larger picture size preferred over smaller picture size. Picture size is expressed as horizontal lines and number of samples per line, or as horizontal and vertical pixel counts.
- Content from high definition sources preferred over content from standard definition, assuming picture size is equal or greater.
- Encodings that maintain frame integrity preferred over formats that use temporal compression (measuring across multiple frames to remove additional redundancy).
- Uncompressed or lossless compressed preferred over lossy compression.
- Higher bit rate preferred over lower for same lossy compression scheme (because the higher the bit rate, the smaller the loss).
- Extended dynamic range (scene brightness) preferred over "normal" dynamic range for such items as Digital Cinema or scanned motion picture film.
- Because film is inherently superior to video for dynamic range, conversion from analogue film originals should be to digital file formats rather than to digital video tape or proprietary video file formats, as this allows the full dynamic range of the original to be captured.
- Surround sound encoding only necessary if essential to creator's intent. In other cases, stereo or monaural sound is preferred.

This means that actual preservation decisions may differ according to the original quality of the digital content, its significance as a resource, and the end-use possibilities. Generally speaking, high-quality versions are preferred for preservation. In practice this will often mean that the best option is to retain moving image content as uncompressed or losslessly compressed JPEG2000 image frames inside a wrapper such as MXF. For commercial movies the best recommendation that can be made at the present time is to use the DCDM format. Readers are advised that these recommendations are interim and the best that can be made given the current state of knowledge and research in dealing with issues of moving image preservation.

For sound the general requirements are:

- Sampling rates higher than 48kHz are to be preferred over lower sampling rates
- 24-bit sample word-length preferred over shorter
- Linear PCM (uncompressed) preferred over compressed (lossy or lossless)
- Higher data rate (e.g. 128 kilobits per second) preferred over lower data rate for same compression scheme and sampling rate.
- If compression is unavoidable then AAC compression preferred over MPEG-layer 2 (MP3)
- Surround sound (5.1 or 7.1) encoding only necessary if essential to creator's intent. In other cases, uncompressed encoding in stereo is preferred

In practice this means the use of either uncompressed broadcast wave format (BWF) or AIFF, using Linear PCM as the encoding method, for audio files, is widely accepted as the best archival practice.

Issues in moving image and sound preservation
There are a number of complex issues in digital moving images and sound that currently make it difficult to both store and preserve these types of resources. Further research into a number of significant issues would benefit the preservation of such resources. The outcomes of research projects dealing with these issues would go along way to allowing the development of properly practice-based recommendations about moving image preservation formats and methods.

Areas for further action:

- **As a matter of priority JISC needs to undertake further investigation into the use of digital television/broadcasting technologies (perhaps in the form of a technology watch) and an assessment of the adoption of digital formats for broadcast television.**
- **Undertake research into the issue of the effect of gamut (the range of colors a device can reproduce) on preservation and access strategies.**
- **Assessment of JPEG 2000 as a moving image preservation format.**
- **Analysis of compression techniques to attempt to determine the feasibility of preservation using uncompressed or lossless formats for both moving images and sound.**
- **Investigation of 'wrapper formats' to establish the essential characteristics of each format and its suitability for moving image and sound file preservation.**
- **Investigation of digital video formats used in professional and consumer digital video devices, as well as an analysis of the issues around transferring digital video from such devices to computers.**
- **Development and assessment of 'migration on request' techniques for preservation of different moving image and sound formats.**
- **Need for a common file format for digitised and digital film. There are moves toward the adoption of the DPX format as a common standard, but its suitability will need to be monitored and assessed.**

1.6 Metadata

This report discusses three types of metadata, all of which must be considered in the preservation of digital moving images and sound. These are technical metadata, management (administrative) metadata, and discovery/use metadata. Technical metadata is necessary to describe the physical attributes of digital objects and is particularly important for preservation and rendering. Management metadata is essential to ensure authenticity, rights, ownership and provenance are properly addressed. Discovery and use metadata is essential to ensure future use of digital objects – being able to locate, access and use digital content in the long-term is arguably the raison d'être of preservation.

The complexity of digital audiovisual materials is reflected in the metadata requirements for his type of material. The report outlines the essential technical, management and discovery/use metadata required for video and sound. The technical metadata in particular is complex and includes the file format and encoding methods; fixity information; horizontal and vertical dimensions and number of pixels; display ratio; whether interlaced or progressive; bits per sample; luminance and chrominance; the frame rate; bit rate (measured in bits per second) and file size (measured in bytes); if sound is included this must be indicated and associated with an instance of audio metadata; sampling frequency; number of bits per sample and so on and so on. This report provides a minimum element set which can act as a starting point for curators of this kind of material, but it is done so with the expectation that curators will add to it their own requirements.

The report also discusses in some detail the existing standards - particularly pertinent as this was identified as one of the key issues facing the archival moving image and sound community.

Sophisticated metadata standards (e.g. MPEG-7, TV-Anytime, SMIL) and container schemas (MPEG-21 and METS) are now available for the archiving of moving image and sound resources. However, manual metadata creation is expensive and, in reality, little file level metadata is likely to exist separately from the content other than that which can be auto-generated. Few metadata extraction tools currently exist for this purpose and further research and development work is needed in this area.

The following actions are recommended:

- **Further research should be undertaken to compare the suitability of MPEG-21 and METS as containers for the metadata required for moving images and sound resources. The output of the research should include case studies and exemplars.**
- **Efforts should be made to create or modify metadata extraction tools (e.g. JHOVE, NLNZ Metadata Extractor Tool) so that they can automatically generate standardized and compatible metadata for audio and video formats.**
    - **Such tools should adopt elements from established namespaces (e.g. PREMIS and DC) and avoid the use of application-specific schemas.**
    - **Where possible such tools should be capable of batch processing a large number of files within a single digital object and generating appropriately structured XML metadata for each.**
    - **Where appropriate, the XML output should be into a directory structure that mirrors the original in order to maintain the validity of the original resource.**
- **Work should be undertaken to assess the feasibility of automatic extraction of subject keywords from audio and video formats using pattern recognition software.**
- **Scenarios for integrating the manual and automatic production of metadata for moving image and sound resources should be investigated further and possible workflow issues examined.**

1.7 Social, economic and legal aspects

A strategic approach to managing the life cycle of digital collections has been broadly advocated and a significant amount of work has been undertaken over the last 10 years or so to identify and map the life cycle of digital objects. The concept of the life-cycle is a useful tool that allows us to identify the key events that take place in the life of a digital object, to identify the actions that we need to take at these key events, and the supporting policies and documentation that are needed to underpin these events and actions. It allows us to plan for the long-term preservation of digital objects in a constructive and time-based fashion, reflecting the more active, interventionist methods of preservation that are required for digital objects, and it allows us to construct organisational models that can support these activities.

The Report summarises the lifecycle processes, and places them in the context of the OAIS model and the concept of a Trusted Digital Repository. The workflow outlined is then enriched with actions that need to take place at the different stages of the workflow process.

The lifecycle approach enables us to think about the type of organisation models that might facilitate the preservation of digital audiovisual materials. It is clear that we need to understand how the materials were created and to be able to identify the significant properties and characteristics of the files. We also need to know if the content is born digital or created from analogue, we need to be able to assess the size and storage requirements, and to place some assessment of the likely costs. Accurate assessment of preservation cost is challenging but increasingly the community is assisted in this rather difficult process by projects developing appropriate cost models such as the JISC-funded LIFE project.

The complexity and the potential significant large size of uncompressed moving image and sound digital materials suggest that we may need to look for creative solutions to cater for their preservation. Whilst some institutions may have the resources to curate and preserve large amounts of audiovisual material, many will not, and so some form of distributed model is likely to offer the best solution. Of particular interest in developing such a distributed system are the possibilities offered by grid technologies such as Storage Resource Broker (SRB). SRB manages large amounts of digital content across multiple distributed servers over and integrated repository infrastructure. It may offer one solution to the problems posed by audiovisual materials and may also allow sharing of responsibilities according to skill set and subject expertise.

This report makes a number of recommendations in regard to issues associated with lifecycle models, organisation, costs and copyright. These recommendations are:

- **JISC investigate the use of grid technologies and assess their value for the distributed curation and preservation of audiovisual materials.**
- **JISC should consider if it has a role in defining benchmarks against which digital moving imaging and sound projects could measure their compliance with standards and best practice; and and in developing tools to capture project processes.**
- **JISC commission research into when re-creation might be an appropriate strategy. This could include an assessment of the costs of continuing to maintain an inadequate audio-visual resource, against the cost of re-creating it.**
- **JISC investigate the feasibility and form of disaggregated models for the preservation of digital audio-visual materials of all kinds.**
- **In particular, JISC is urged to undertake in-depth investigations of the technology, encoding/file formats, and positioning of digital television broadcasting.**

A final recommendation made in this report concerns the management of technological obsolescence, perhaps the greatest threat to the long-term preservation of audiovisual materials. **This report recommends that JISC consider collaborative work with TNA to enhance the PRONOM database and make it as comprehensive as possible, so that it functions as a widely useful and useable tool for the digital preservation community.**

Chapter 2: Introduction and Scope of Study

2.1 Outline of study

The last decade has witnessed a rapid growth in the creation of a variety of digital sound and moving pictures formats. The prominence of the World Wide Web and growth in digital technologies has enabled users to create, download, and manipulate these born-digital files. Digital art forms are emerging from these multimedia technologies, and large scale use of digital film and music files has become a prominent feature of our cultural and intellectual life.

Another focus over the last few years has been the conversion of analogue sound and film resources to digital format, to the extent that much of the knowledge base and intellectual assets of learning institutions are in digital form. Digital library technology now enables large, rare film and audio collections to be digitised. The Joint Information Systems Committee (JISC) for example, is currently undertaking two major digitisation programmes to make available key digital moving images and sounds resources for learning, teaching and research. The British Library archival sound recordings project[6] will deliver up to 12,000 items totaling 3,900 hours of segmented recordings from oral history, field and location recordings of traditional and improvised music, rare or deleted classical and popular music recordings, soundscape and educational material. The NewsFilm online project aims to to digitise the broadest possible range of material to ensure relevance across a whole range of academic disciplines. When completed, the project will deliver up to 60,000 segmented encodings, totaling 3,000 hours and associated materials from the archives of ITN and Reuters television.[7]

Despite substantial research into digital preservation in the UK over the last few years, it is recognised that relatively little attention has been paid to the long-term durability of multimedia files. While much of the preservation knowledge base centres around static digital files, it is essential that the HE/FE community is aware of the issues of preservation for a whole range of file types, especially complex multi-object files types, such as moving pictures files.

JISC recognises the increasing importance of preserving digital resources for a wide range of activities and materials. JISC's continuing commitment to developing the UK digital preservation agenda was set out in the JISC *Continuing Access and Digital Preservation Strategy 2002-5* (Beagrie, 2002). In this Strategy JISC envisaged responsibility for digital preservation activities spread between national services, individual institutions and, potentially, institutional consortia. Digital preservation represents a complex set of challenges. Some of these are exceptionally difficult for institutions to address individually. National action in this field is therefore appropriate to the community and remit and mission of the JISC. JISC for example funded the Digital Curation Centre (DCC) to act as a conduit for sharing expertise and developing best practice in digital preservation. The DCC itself does not hold digital resources, but provides a set of central services, standards and tools for digital repositories.

Over the past few years, JISC has funded a series of feasibility studies which aim to assess the preservation risk and retention criteria for digital content, and to help inform and prioritise the development of future services and calls in digital preservation. The Moving Images and Sound Archiving Study forms part of the feasibility study programme and has been funded to

[6] British Library archival sound recordings project: http://www.jisc.ac.uk/whatwedo/programmes/programme_digitisation/project_bl_sound_archive

[7] Newsfilm online: http://www.jisc.ac.uk/whatwedo/programmes/programme_digitisation/project_bufvc

understand more fully the preservation challenges of digital moving image and sound files, to scope the preservation requirements and to determine archiving methodologies and future research directions. The authors expect the report to impact in two key areas:

- First, as a source of knowledge and information available to the creators and curators of both moving image and sound digital content (also known in this report as audiovisual materials) within the UK HE and FE community.
- As a source of information and knowledge for JISC as it develops its priorities for preservation activities over the coming years.

2.2 Outline of scope and methodology

The study was conducted from September 2005 to May 2006 by a team from the Arts and Humanities Data Service (AHDS), the BBC Archives in the person of Richard Wright, and King's College Digital Consultancy Services in the person of Simon Tanner. It was conducted in response to the JISC Invitation to Tender for a Digital Moving Image and Sound Archiving Study.

This report is concerned with time-based resources, ie. digital moving images and digital sound. The properties and characteristics of moving images and sound are complex and significantly different from other digital resources. Although standards and formats differ between moving images and sound, the issues relating to the capture, storage and preservation of both types of resources are similar enough that they can be considered together. Moving image resources include streaming video (eg. digital television broadcasts), the outputs of moving image capture devices such as consumer and professional video cameras, and digitised versions of analogue video formats. Sound resources include digitally recorded audio and digitised versions of analogue sound files.

2.3 Project outcomes

This report presents the findings of the study, arranged into five main areas:

- Curator experiences and requirements
- Significant Characteristics
- Preservation methodology, and metadata
- Life-cycle and organisational models
- Costs and copyright

The recommendations made in this report will primarily affect repository managers, but in some cases are also highly relevant to both creators and users. It is the belief of the authors of this report that all those with an interest in using digital content, whether in the short-term or the long-term should be concerned about ensuring continued access to digital content – and by default, its preservation.

Even when recommendations are targeted at those who fund, plan and manage repositories, the importance of ensuring appropriate involvement from creators and users of digital images should not be forgotten. Ultimately, the value of digital audiovisual materials is in their value to users, both now and in the future and repository managers must meet the needs of long-term preservation in a way that does not conflict with the requirements of the user. The logical conclusion to draw from this approach is that any preservation strategy and actions must be flexible. Preservation is not a one-time activity – it is an ongoing and changing process where policies and practices will need to be adapted and change.

Comments on this report may be directed to:

Sheila Anderson
Director, Arts and Humanities Data Service
King's College London
26-29 Drury Lane
London, WC2B 5RL
Email: sheila.anderson@ahds.ac.uk
URL: http://ahds.ac.uk

Chapter 3: Overview of current research

3.1 Background

Librarians and archivists have considered digital preservation issues in varying detail for the last twenty years, but no consensus has been reached regarding a shared digital preservation strategy.[8] A number of digital preservation strategies have been proposed, but there is no definitive approach to the problem of maintaining digital content across multiple generations of technology, nor is there likely to be one.

Management and preservation requirements for digital materials are fundamentally different from analogue materials. Digital materials can be created using a wide range of technologies and formats, whether born digital or digital surrogates of existing analogue materials. They can be described and documented in a variety of ways – or not at all. They are subject to both physical deterioration and technical obsolescence. More than one copy can be easily and simply created. Access may be provided through more than one point, in more than one form, and may be done in a distributed manner. All these factors will impinge upon the approach taken to the management and long-term preservation of digital materials, none more so than digital audiovisual resources.

The differences outlined above present the curators of digital resources with some fundamental challenges. The way in which audiovisual materials are created, particularly the technologies used, will determine how conducive to long-term preservation the materials are, and will present varied challenges to curators charged with the subsequent management and preservation of such materials. Curators will need adequate metadata about the resource if they are to successfully manage, preserve and make the materials accessible. Multiple copies may also imply multiple versions which must be controlled and managed, and the digital resource curator must also somehow ensure the integrity and authenticity of the resources. The curator must be aware of changing technologies and the fragility of media and take these into consideration from an early stage in the management and preservation process.

3.2 Hidden Treasures: The UK Audiovisual Archive Strategic Framework

The scale of the issues facing HE and an examination of future infrastructure needs was provided in the Museums, Libraries and Archives Council funded study: *Hidden Treasures: The UK Audiovisual Archive Strategic Framework*.[9] The report was the result of a joint study involving the British Library National Sound Archive, the Film Archive Forum, and the Museums, Libraries and Archives Council. The study aimed to:

- produce a document for audiovisual archival development across the UK;
- provide information on the strengths and weaknesses of audiovisual archival provision across the UK;
- recognise the important strategic benefits to be gained from the audiovisual sector working closely with the wider museums, libraries and archives sector;
- do the above within the broader context of changing national and regional policy making and structures;
- inform public and private funding bodies on the priorities for capital and revenue investment in audiovisual archives sector.

[8] Seamus Ross, 'Changing Trains at Wigan: Digital Preservation and the Future of Scholarship', NPO Preservation Guidance Occasional Papers', (London, National Preservation Office, 2000) p. 13.
[9] British Universities Film & Video Council, Hidden Treasures: The UK Audiovisual Archive Strategic Framework, 2004. Available at: http://www.bufvc.ac.uk/faf/publications.htm.

The completed document was launched in London in June 2003. The report provides a blueprint for the future development of audiovisual archives across the UK. It argues that the lack of strategic focus for national and regional planning in the audiovisual archives sector has led to insufficient levels of funding and investment for core activities, and a very real concern that a significant amount of the UK's moving image and sound archive heritage could be lost. The document proposes that the national strategic and funding bodies should work with the audiovisual archives to develop a national framework of institutional vision in which national, regional and local responsibilities are respectively understood and well resourced, with the aim of ensuring comprehensive coverage for audiovisual archive activity across the UK.

The strategy document sought first and foremost to provide a structural framework for the development of common voice for audiovisual organizations in the UK. This concept is clearly worthwhile and necessary and the development of such a framework is an essential first step to determine the future direction of audiovisual archives across the UK. It seems, however, that the Strategy's recommendations have not yet been taken up. Perhaps, in part this was because the document itself seems not to have been directed at the correct audience – funders and policy makers – but at the already converted. Also, the document, rather than proposing a strategy, proposed a strategic framework from which (it hoped) a comprehensive national strategy would arise. The report's key recommendations are, nevertheless, worth following up and acting upon. **We recommend that JISC convene a meeting of the relevant bodies named in the *Hidden Treasures* report to progress adoption of the recommendations of that report.**

3.3 Overview of current research

A 2004 PrestoSpace report noted that 'projects and policies under the label 'digitisation' tend to be all about access, and not at all about preservation'.[10] The authors of the document *Digital Curation and Preservation: Defining the research agenda for the next decade,* similarly noted that 'There is in general a low level of awareness of the need for digital preservation and a lack of researchers' participation'.[11] The situation has changed a little in the intervening period but the multiplicity of collections and research projects across the UK suggest that more than ever a national strategy is urgently needed in order that audiovisual resources may, as *Hidden Treasures* puts it, "make a full contribution to the UK's cultural and educational life…".[12]

Current and recent projects address a wide variety of issues relating to digital preservation. The EU funded 'Presto' and 'Prestospace' projects look at preservation and access for audio, video and film.[13] Presto, which concluded its work in 2002, identified the major problems of access to audiovisual material, and concentrated on solving the technical issue of reducing the cost of preservation transfers. The newer PrestoSpace project, begun in 2004, intends to push the limits of the current technology in order to provide products and services for bringing effective automated preservation and access solutions to Europe's audiovisual collections.

In the area of non-traditional audiovisual forms, 'Archiving the Avant Garde', based at the University of California at Berkeley, is a collaborative project investigating strategies for describing and preserving non-traditional, intermedia, and variable media art forms, such as

[10] Richard Wright, 'Annual Report on Preservation Issues for European Audiovisual Collections' January 2005, p. 17.
[11] Digital Curation and Preservation: Defining the Research Agenda for the Next Decade, Report of the Warwick Workshop, 7 & 8 November 2005, p. 13.
[12] Hidden Treasures, p. 4.
[13] Presto: http://presto.joanneum.ac.at/index.asp; PrestoSpace: http://www.prestospace.org/.

performance, installation, conceptual, and digital art.[14] The partners in this project, which include Berkeley Art Museum and Pacific Film Archive, Solomon R. Guggenheim Museum, Rhizome.org and Franklin Furnace Archive and Cleveland Performance Art Festival and Archive aim to develop ways to catalogue and preserve existing and future collections of these less traditional audiovisual forms.

Two other projects are working to provide the means for 'systematic preservation' of less traditional audiovisual materials. The 'International Database of Virtual Art' provides individual documentation of virtual art works, including video documentation, technical data, interfaces and displays.[15] The aim of the Database is "to make transparent developments in the field of virtual art and its subgenres (virtual reality, genetic and telematic art) as well as to present the rapidly growing oeuvre of the artists that create it". The 'Variable Art Network' is attempting to implement a new strategy for impermanent visual art forms. The network says it is directed at "artists working in ephemeral formats who want posterity to experience their work more directly than through second-hand documentation or anecdote" and it "encourages artists to define their work independently from medium so that the work can be translated once its current medium is obsolete". The network offers tools, methods and standards development for their preservation strategy for conceptual, minimalist and video art.[16] The Variable Media network strategies require artists to envision acceptable forms their work might take in new mediums, and to pass on guidelines for recasting their work in a new form once the original has expired.

Innovative projects that test digital formats and standards through practical application, rather than nationally adopted strategies, are leading the way in the development of standards for audiovisual resources. However, there is still much to be done and although "best practices have been and are being developed for the initial digitization process, they do not exist in many areas of the preservation chain".[17] Some projects have worked to fill this gap, with exciting developments in the field of moving images and audio. For instance, the 'Digital Video Preservation Reformatting Project' carried out in 2004, for the Dance Heritage Coalition, demonstrated the superior retention of video image quality when lossless compression was employed, as compared to a variety of degradations resulting from lossy compression.[18] The Dance Heritage Coalition was a collaborative project which aimed to establish standards for digital preservation of analogue audiovisual formats. The project was funded by an Andrew W. Mellon grant and was completed in 2004, recommending the use of JPEG2000 and the Materials Exchange Format (MXF) as the format standards. These recommendations have been adopted by the Library of Congress. Another project, based at Harvard University, 'Sound Directions' aims to create and test emerging standards and existing practices for digital preservation[19] and is also likely to produce influential work in the area of standards. 'Sound Directions' is a collaborative research and development project between Indiana University Archives of Traditional Music (ATM) and the Archive of World Music (AWM) to create best practices and test both emerging standards and existing practices for digital preservation. This project aims to contribute to the creation of interoperable digital audio preservation packages, containing audio essence and metadata, following the OAIS model. The project is expected to be completed in late 2006.

Two other projects related to the Library of Congress bear mention. One is the first offering from a group of audio-visual prototyping projects. This project ran from 1998-2003 and had as its focus the reformatting of sound recordings. It helped move the institution to adopt file-

---

[14] http://www.bampfa.berkeley.edu/about_bampfa/avantgarde.html.
[15] http://www2.hu-berlin.de/grau/database.htm.
[16] http://variablemedia.net/e/welcome.html.
[17] 'Sound Directions: Digital Preservation and Access for Global Audio Heritage', http://www.dlib.indiana.edu/projects/sounddirections, p. 2.
[18] See: http://www.danceheritage.org/preservation/Digital_Video_Preservation_Report.doc
[19] See: http://dlib.indiana.edu/projects/sounddirections/

based approaches to the reformatting of audio and to discontinue the practice of making analogue tape preservation copies.[20] The second project is the Preserving Digital Public Television project, being carried out by WNET/Thirteen in New York, WGBH in Boston, PBS in Alexandria, Virginia, and New York University.[21] This project has received matching funds from the Library's National Digital Information Infrastructure and Preservation Program (NDIIPP).

A European project aimed at digital audio is the MARTlab project.[22] The MARTlab research and production centre for musical technologies, was founded as a joint project of the Conservatory of Music of Florence "*Luigi Cherubini*" and the Institute ISTI of National Research Council of Italy, with the external contribution of the Tuscany branch of the National Broadcasting Company RAI.

The main mission of MARTlab is the preservation and restoration of analogue audio files by transferring them to digital domain. A second but important component of the project is looking at media issues, taking into account the limited life span of the present day available media for recording digital sound.

Recently the plight of broadcast archives was raised by FIAT/IFTA (International Federation of Television Archive) at the World Electronic Media Forum (WEMF). Sue Malden delivered a speech to 300 delegates, including the Secretary General of the United Nations, Kofi Annan. The main conclusions were:

- The world's key broadcast audiovisual archives (including internet output) must be preserved for future generations
- Audiovisual records must become more accessible to all
- An adhoc group should be established with representatives from UNESCO, WBU and FIAT and other relevant bodies and financial institutions
- This group to propose and implement an action plan for the preservation of endangered archives – embracing training, digital storage, access and specific help for developing countries[23]

3.4 Some key audiovisual preservation initiatives

Other audiovisual preservation initiatives provide useful advice and guidance for curators of audiovisual resources. In the UK, the Arts and Humanities Research Council (AHRC) established the Centre for British Film and Television Studies in 2000.[24] The objective's of the Centre are to:

- develop of new areas of knowledge through academic access to archival materials.
- introduce scholars from other disciplines and creative arts practitioners to film and television archive material and promote of interdisciplinary collaboration.
- provide an authoritative contribution to the formation of public policy in Britain and Europe in the areas of film and broadcast media

Among the Centre's publications is the report 'Moving History: A Guide to UK film and television archives in the public' which provides information on twelve public sector moving image archives.[25]

---

[20] Web site: http://www.loc.gov/rr/mopic/avprot/avprhome.html.
[21] See the Web site (http://www.ptvdigitalarchive.org/index.htm) for more information.
[22] http://www.martlab.it/.
[23] *Archive Zones,* Spring 2006, p. 8.
[24] http://www.bftv.ac.uk/home.htm.
[25] See: http://www.movinghistory.ac.uk/index.html.

In the United States a number of initiatives are offering advice on a range of audiovisual topics. The Library of Congress has established the National Recording Preservation Board, an advisory group bringing together a number of professional organizations and expert individuals concerned with the preservation of recorded sound. The Board is one of three components established by the legislation to form a comprehensive national program to ensure the survival, conservation, and increased public availability of America's sound recording heritage.[26]

The National Film Preservation Foundation (NFPF) is a non-profit organization created by the U.S. Congress to help save America's film heritage. The NFPF supports activities that preserve American films and improve film access for study, education, and exhibition. The NFPF started operations in November 1997, through the support of the Academy of Motion Picture Arts and Sciences and The Film Foundation. Although the focus of the NFPF has been analogue film formats, the NFPF preservation strategy includes digitisation of analogue originals for long-term preservation in digital formats.

A final invaluable initiative of the Library of Congress must be mentioned. This is the Library's 'Sustainability of Digital Formats' website. This website contains a wide range of information about digital audiovisual formats, sustainability factors, content categories, and detailed format descriptions. For anyone working in the audiovisual arena this website is a mandatory resource. The diversity, clarity, and thoughtfulness of the information it provides is currently unsurpassed.

The Appendix to this report provides an overview of moving image and sound collections in the UK, describes some JISC funded collections, and lists the audiovisual collections held by the Arts & Humanities Data Service. Also provided in the Appendix is a list of relevant advisory bodies.

[26] http://www.loc.gov/rr/record/nrpb/nrpb-about.html.

Chapter 4: Survey on Preservation of Digital Moving Images and Sound

4.1 Introduction

An online survey[27] was created to elucidate from the community information in the types of projects being run and their feelings about what constitute the main issues and barriers to this activity. The questions allowed a free form answer but were constructed to allow us to gather numeric information as well.

Analysis of results

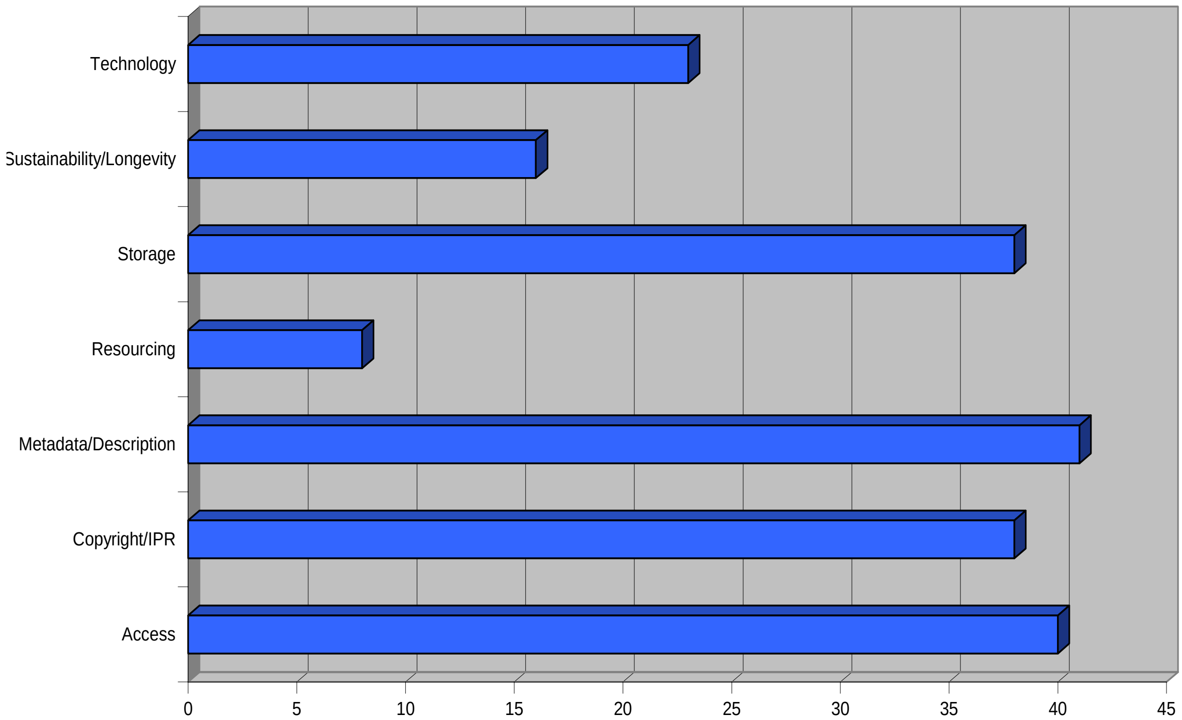
There were a total of 92 responses from the international community. Of these responses 35% were working on just digital moving images, 25% on digital audio, and 40% working with both media types. This suggests that the results will tend toward a moving image bias but will still be statistically significant.

The responses also tended to fall naturally into several clear categories for both key issues and barriers. These are:

- Access: relating to end-user access and the ability to maintain access and usability.
- Copyright/IPR: the rights to store, preserve and subsequently use content.
- Metadata/Description: the means of finding or describing the content.
- Resourcing: issues of cost, staffing and infrastructure.
- Storage: the means of storing digital content.
- Sustainability/Longevity: the ability to sustain the digital objects and services over time.
- Technology: the mechanisms and tools used for preserving digital content.

Respondents were free to include as many issues or barriers as they wished.

The most important issues



[27] http://www.ahds.ac.uk/moving-images-survey.htm

Metadata, access, copyright and storage are clearly of greater importance to the respondents than any of the other issues found. This suggests that these are difficult issues which almost every project has to take account of.

Indicative comments from respondents are listed below in descending order of ranking:

Metadata (41 responses):

- Much metadata does not get to the heart of the information content, but concentrates only on technical metadata and somewhat banal and uninformative titles. The trouble is that providing really helpful metadata enabling resource discovery is very expensive.
- Metadata - technical and content details. Clarity about what is required, what is machine-readable, how the audio file and the metadata should be connected.
- The metadata standards are still being developed. METS / MODS / MIX / PBCore, etc. seem to be the most mature.

Access (40 responses):

- Access is crucial I think, else what is the point... and good access is dependant upon decent search methods.
- Effective access mechanisms; digital rights management; appraisal, i.e. understanding which content to preserve to what level of quality
- The most important access issue is cataloguing of content - rather than metadata in its entirety - in the form of synopses, followed by copyright and the method of access.

Copyright/IPR (38 responses):

- The key problem is copyright for materials which are strictly educational. The difficulty is finding a balance between the reasonable commercial interests of organizations and the long term educational and intellectual use of important materials. The issue is not one of commercial copyright, but of the recognition by commercial organizations and some museums, libraries and universities that inappropriate restrictions make work on, and use of, media materials very difficult. Yet it is increasingly evident these are crucial to any critical understanding of the twenty-first century
- Copyright is a big issue for 'orphaned works' - e.g. cassette recordings published by small no-longer-contactable publishers. We would like to be able to preserve access to this sort of material but it is at high risk of falling off the list because of the administrative cost of pursuing copyright clearance.
- Copyright is a big issue for performance work, as is consent. We have to ensure that all members of groups are aware of not just recording but of the storage of their image - what if someone became famous and someone sold the tape to 'before they were famous'?
- Unpublished field recordings involve performers and donors who sometimes feel strongly about widespread dissemination when the performers are not compensated, or in the case of possible misuse of music that was originally performed in a certain cultural context.

Storage (38 responses):

- Digital storage media will only have a short life compared to film.
- Generally I would have said digital storage in terms of a safe repository plus the backups, simply because the places who do this in Australia are mostly there by Federal government good will, even the ones put in place by legislation.

- Storage in multiple forms and formats, lack of standardization, and knowledge that whatever is produced is subject to repeated technology transfers (lack of a stable electronic medium).

Technology (23 responses):

- The whole zoo-full of file formats codecs standards etc, many of them often proprietary.
- Formats (would be good if somebody could standardise like TIFF for images). What is the standard format for Video? MPEG, QUICKTIME, WINDOWS MEDIA?
- Constantly changing technologies, that is broadcast industry-led.
- 1. preservation of the physical media so that it can be read
  2. preservation of equipment to read that media
  3. preservation of the knowledge of how to use and maintain that technology
  4. ability to decode or decompress often proprietary digital bitstreams into contemporary formats with fidelity

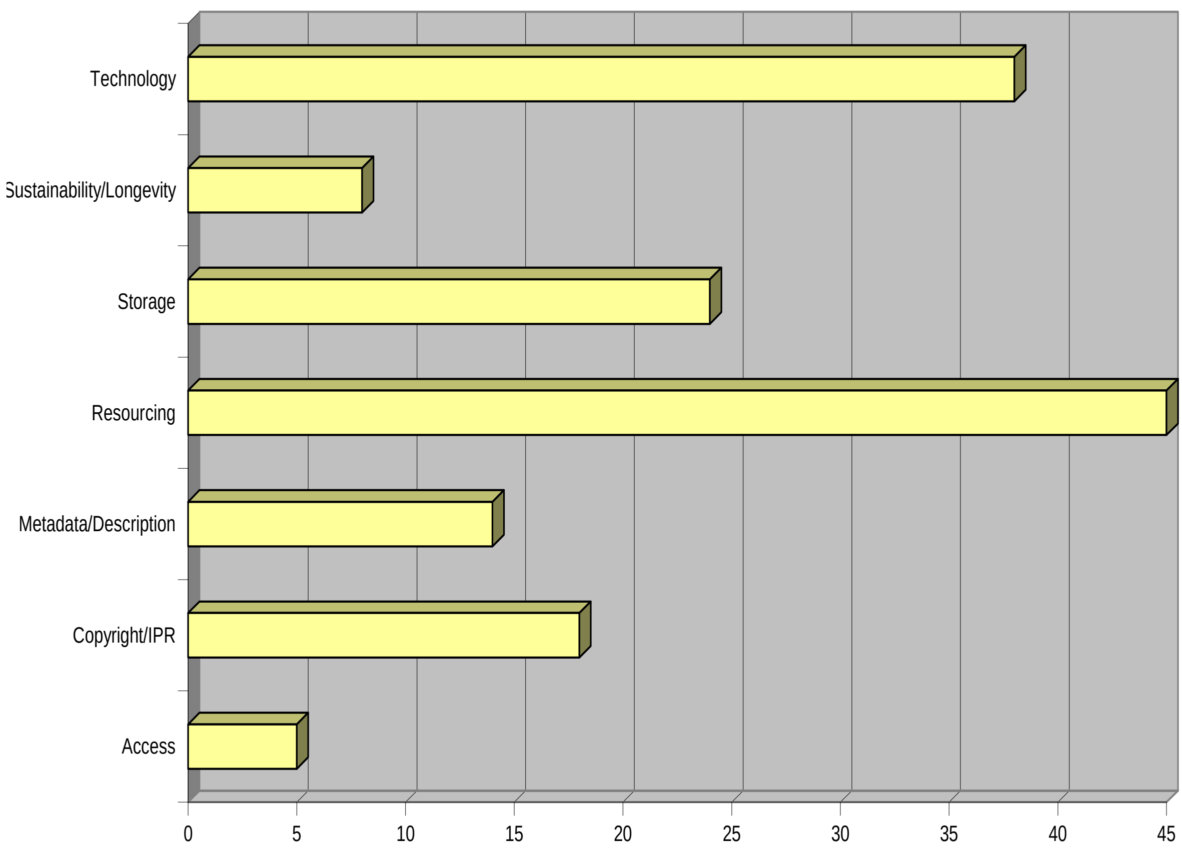
Sustainability/longevity (16 responses):

- What happens if the internet ceases to be or changes beyond recognition understanding future forms of digital formats might help to flag up problems now.
- The most important issue may be compatibility. We will require agreed standards which will have industry support in the long term.

Resourcing (8 responses):

- staff time for curatorial/documentation work

4.2 The key barriers



Resourcing and technology far outstrip any other responses, whilst storage, copyright and IPR make up the rearguard. This about face from the issues responses strongly suggests that these are barriers which are very common across most of the projects but are not deemed to be central to the activity itself. In other words resourcing is essential to enable the activity but is not seem as an integral issue of digital preservation is the way that metadata was perceived.

Indicative comments from respondents on key barriers are listed below in descending order of ranking:

Resourcing (45 responses):

- Knowing what will last, with least staff intervention!
- The unknown cost and the lack of clear open standards which conform to an overall system. Technical skills in this area are evolving and so future skills required are both unknown and difficult to plan for.
- Technical know-how + cost of setting up and maintaining the equipment. Capture of preservation-worthy video is a very complex and time/money/staff intensive activity and requires so much specialized equipment & knowledge I can imagine it will be prohibitively complex/expensive for all but a few digital library programs. While outsourcing is a possibility, we've found that even the major vendors working in this area (transfer of magnetic media) don't work in a way that we'd consider appropriate for creating preservation masters.
- There is about fifteen years worth of digitizing time, at the current work flow rate.
- Convincing financial and political stakeholders that it is worth preserving.

Technology (38 responses):

- The key barriers all derive from structure of the marketplace for digital media technology. Cultural Institutions must continually appropriate AV hardware and formats designed for consumer or professional use which naturally have a short product in order to meet their preservation goals. No manufacturer produces digital AV equipment with preservation in mind, and the complexity of contemporary digital technologies makes it impossible for an individual institution to design and build that equipment.
- The rapid change of formats and the cost of migration.

Storage (24 responses):

- The cost of storage - access, metadata and copyright can be negotiated on, storage costs are the big one.
- Storage is still an issue for many institutions, obtaining uncompressed material, the need to provide alternate versions for access (web, downloadable in various formats) requires additional storage and processing.

Copyright/IPR (18 responses):

- The key problem is copyright for materials which are strictly educational. The difficulty is finding a balance between the reasonable commercial interests of organizations and the long term educational and intellectual use of important materials.
- Clearing of copyright is the main barrier - this is time-consuming and intricate.

Metadata (14 responses):

- The lack of metadata recorded by equipment, different levels of metadata being recorded by various equipment (same issue as digital cameras),obtaining appropriate descriptive metadata.

Sustainability/longevity (8 responses):

- Digitisation is carried out constantly by many different staff in this institution - at the point of capture, decision about preservation and long-term usefulness are the lowest priority.
- As always the key barrier is a lack of sustainable resources.
- Uncertainty about future sustainability of digital repositories
- This material is intangible. We are essentially attempting to contain the horribly long list of ones and zeros in the right order in a format available to be read over a period of time which is untested. Seems a bit daunting.

Access (5 responses):

- Archivists can find solutions to all of the issues noted in question 2 (except perhaps Copyright which requires government intervention), but the extent to which archives will be able to preserve and provide access to this type of material will be determined by the level of secure funding in the long term.

4.3 Conclusion

The number and fullness of the responses to a short online survey suggest a high level of interest and concern regarding digital moving image and sound preservation. The responder's experience was varied. Only a few have completed implementations of digital preservation for moving images and sound. Most were in the midst of setting up or planning projects or trying to work out how to address digital preservation. Whatever the levels of experience, the survey responder's comments carried a common sense that digital archiving of moving images and sound remains an emerging field with many questions yet to be answered. Clearly, the maturing of experience and project implementations has not happened on a wide scale.

The responses seem to suggest that almost all projects have to address issues such as metadata, access, copyright and storage but that it is the levels of resource and the nature of the technology itself which are the major barriers to be overcome in achieving success.

In assessing cost models, we found that there were few case studies or exemplars for costing digital preservation of moving image and sound. This is reflected in at least half the survey responses. Comments were focussed upon finding the money to carry out the project in the first place and many commented on the difficulty of sustaining the technology and staff over the medium to long term. In some cases, the problem was how to plan effectively because of the difficulty in costing all the factors needed to achieve success. This lack of cost information and methodology may act as an insuperable barrier for many who wish to make projects happen.

Whilst this gap remains between understanding the stakeholders' needs and justifying the expenditure to preserve content, then projects will continue to be under developed and lack a sound foundation for long term sustainability. The research agenda for this sector is often focussed upon the technical barriers; but further research into the costs and economic factors behind these activities is required to give the community tools, methods and exemplars to enable this financial barrier to be overcome.

The technology itself is a barrier, particularly to smaller institutions without the necessary infrastructure or experience in house. Moving images and sound files tend to be of very large file sizes and can be in many formats.This was both a key issue and barrier for many projects – how to deal with storing all this content and then maintaining the usability when there are so many formats and standards to consider.

The survey showed metadata is a common theme of concern for the majority of projects. This report has identified the complexities and layers of information required for storing metadata to control the intellectual property and rights in the moving image and sound content. As one of the survey respondents commented, “much metadata does not get to the heart of the information content, but concentrates only on technical metadata and somewhat banal and uninformative titles”. The description of content to enable future access, at a time when the context of the recording may have been lost or become opaque, remains a challenge both for resourcing metadata creation and defining the most appropriate method for capture and storage of metadata. At the very least, clearly identified minimal requirements to enable future retrieval and current management of these digital objects would provide useful guidance to those attempting to implement digital archiving projects. There was much comment about the proliferation of standards (such as METS / MODS / MIX / PBCore) and yet many of these standards are still developing. Some confusion or contradictory opinions were expressed regarding which standards to implement and what metadata could be gathered automatically or had to be created by humans.

Chapter 5: Properties of Digital Moving Images and Sound

This report is concerned with time-based resources, ie. digital moving images and digital sound. The properties and characteristics of moving images and sound are complex and significantly different from other digital resources. Although standards and formats differ between moving images and sound, the issues relating to the capture, storage and preservation of both types of resources are similar enough that they can be considered together. Moving image resources include streaming video (e.g. digital television broadcasts), the outputs of moving image capture devices such as consumer and professional video cameras, and digitised versions of analogue video formats. Sound resources include digitally recorded audio and digitised versions of analogue sound files. This chapter identifies the defining characteristics of both moving image and sound file types, analyses and assesses the principle file formats used in their creation, and provides information as to their suitability for long term preservation.

5.1 Format overview

We are in a transition period where not all digital audiovisual materials are stored as standard computer files. Digital videotape and audio CDs are digital media which store streams of bit sequences that are often not held in the usual form of computer files.[28] Digital radio and television signals are broadcast and received as real-time bit streams (suitably encoded), in a manner that does not resemble regular computer files. An additional issue is how to get non-file-format audiovisual materials (ie data streams) into and out of file formats and computers).

For long-term preservation of audiovisual materials, we have to foresee the ‘death of videotape’ and possibly the end of physical media such as CD and DVD, and possibly even the replacement of broadcasting itself by file transfers across a super Internet. Nevertheless, currently all digital resources, whether contiguous bit streams or distinct files, have a format (or encoding) which is what enables them to be used by equipment such as computers, CD and DVD players, and digicams.[29] This chapter will discuss file formats and encodings (codecs) for both digital audio and video.

5.1.1 Encoding, file formats and wrappers

When discussing digital video and audio, it is important to clearly define and understand several related concepts:

1. coding/encoding: the particular way numbers are used to represent sound or images. The digital video/audio community uses the term ‘encoding’ to refer to both the ‘format’ of video/audio streams and to the file type of digital video or audio files. For example, both .mp3 and .dv are referred to as ‘encodings’, although the former is an audio file format and the latter is an encoding format for digital video streams. Secondly, a number of the bitstream encodings are media-dependent formats, such as audio CD and digital video tape formats like DigiBeta (by and large these are all recording formats). Thirdly, a number of the encoding formats that will be discussed here are ‘wrapper formats’ (see under wrapper

---

[28] For the purposes of this chapter a computer file is defined as “a stream (sequence) of bits stored as a single unit, typically in a file system on disk or magnetic tape. Usually, the length of a file can be any positive integral number of whole bytes (groups of 8 bits each), from zero up to a file system-dependent upper limit. Files are usually associated with and identified by file names,”. Source: Wikipedia: http://en.wikipedia.org/wiki/File_(computing) [last checked 10/01/2006].

[29] For the purposes of this chapter, format (or encoding) is defined as: “*packages of information that can be stored as data files or sent via network as data streams (aka bitstreams, byte streams)*”. Source: Library of Congress Digital formats website: http://www.digitalpreservation.gov/formats/intro/format_eval_rel.shtml [Last checked 10/01/2006]. As a corollary, a ‘file format’ is a way of encoding data for storage in a computer file. Encoding refers to the manner in which the byte sequences are structured for use.

formats below). When a bitstream is captured and stored on computer storage media, a computer file is created. All video and audio found on the Internet and in computer systems is encoded in some format or other. There are proprietary encoding methods such as Real and Windows Media, and more open standards such as MPEG. In this document the word encoding will generally only be used in reference to the structure of a bitstream.

2. file format: The usual name for the structure of a discrete chunk of digital data is a file. A file is really only a series of encoded bits that has been captured and stored somewhere. Because the structure of a file is tied to the content (word processing, spreadsheet, image and dozens more), and is often proprietary (the files of one brand of word processor may not be compatible with any other brand), and varies from computer system to computer system (mainframes of various sorts, Apple, PC), there is a huge number of file types. This whole area of the variety of kinds of computer file is a major challenge for preservation, with sound and moving images only occupying a small corner of the battlefield (or minefield).

Often the file format is specifically for one type (or closely-related family) of encoding, leading people to think of, for instance, 'mp3' files as being just an encoding type rather than a file format. As far as the computer system is concerned, it is a file format because it is a way of capturing an encoded data stream as a logical unit. The computer's file management system (which stores and retrieves files) needs to be able to move such files about, but doesn't need to understand them. It is up to the decoder (or codec software) to interpret the file. The reason one .wav or .avi or .mpg file may play and another may not is due to the presence or absence of the appropriate piece of software (or codec – see below) for decoding the encoded bitstream. By itself a three-letter extension on a file name isn't much information, and really only refers to a family of related file structures. For instance, although most .wav files are uncompressed, there are actually around 100 bitstream encodings defined as content for the .wav file type. So, the .wav and .avi extensions on file names do not specify the method by which the encoded bitstream can be understood – they are only indications of the file format' used.

3. wrapper (or packaging) format: in this report 'wrapper' is used to indicate a file format for encapsulating different constituent bitstreams (or, in fact, files) into a single file. So, in an audio 'wrapper format' like WAVE (.wav), audio streams or files and other chunks of (meta)data are individually encapsulated in a single file. Similarly, AVI, the standard windows media wrapper, and MOV, the standard QuickTime container, encapsulate audio, video, metadata, and other control information together as a single package. It is only when the package is opened that the detailed requirements – the most important being the method of encoding of each constituent bitstream – are brought to light. If the software codec (see below) is specified and available, the file will play. If not, there is a serious problem that will probably require investigative software to diagnose – and a friendly local bank of historical and contemporary codecs to solve.

4. codec: short hand for (en)**co**der/**dec**oder or **co**mpression/**dec**ompression. A codec is the method by which an audio or video bitstream is encoded for transmission (interchange) and decoded for use. Data reduction is important for getting signals across restricted information channels, such as a slow internet connection, so many codecs compress audio and moving image data in order to reduce the size of the data streams. Often data reduction is accompanied by data loss, so many compression methods are called 'lossy' for this reason. Sound can be represented just as a sequence of numbers, or various mathematical procedures can be applied to these numbers to reduce the amount of data. For video, where it takes millions of numbers to represent one image, data reduction is essential in many applications. Lossy compression methods for audio and video are an archivist's nightmare, as data loss implies loss of quality or robustness of the signal. For archival purposes use of lossless compression or uncompressed encodings is the best choice since it retains the highest data quality.

However, it is possible to perform 'lossless compression' of audio and video, in which case reductions in size (data rate) are usually around a factor of 2 to 3 (for many people this small reduction is hardly enough to maybe justify the extra effort). Lossy compression, encoding of video to formats such as MPEG-2, can reduce the data by factors of over 100. Digital broadcasting in the UK uses MPEG-2 to reduce the data in the video signal from 270 megabits per second down to about 4 Mb/s. This allows 50 channels in a space otherwise occupied by only one, and makes digital broadcasting viable. Codecs have been used for decades for long-distance telephones (again, an application where bandwidth was at a premium), and were essential to the success of mobile telephones. Without an appropriate codec an encoded signal (bitstream) cannot be played. Today most codecs are software, although for real-time operations broadcasting has an immense investment in MPEG-2 encoding hardware, with the decoders also in hardware in the viewer's set-top box.

The usual form of uncompressed audio encoding (the straightforward sequence of numbers) is called linear PCM (pulse code modulation, a name carried over from pre-computer days). The recent development of a method of lossless encoding of video, JPEG2000, is gaining interest amongst archivists. For archiving and re-use of audio or video, one absolute essential is knowledge of how the signal was encoded, if at all. The second essential is to have a working decoder (a codec).

5.2 Moving Images

5.2.1 Moving images overview and significant properties

The history of using a succession of fixed images to produce the impression of motion dates back at least to the zoetrope of 1832[30] (an animation is available online[31]). The early devices used hand-drawn images – the early 19th C precursors of today's computer animation and special effects. Because photography was also being developed in the early 19th Century, it was inevitable that eventually it would be possible to expose film sufficiently rapidly (many times per second) to allow real images to be viewed in the manner of the Zoetrope. It then was one more stop for such images to be projected to a general audience, and the cinema was born.

Film is impressive. How can a few square centimetres of light-sensitive material on celluloid produce a visual field, much less motion pictures? The answer requires a brief reminder of basic properties of human vision.

The sea of light: Life on earth has developed sensitivity to the electromagnetic energy coming from the sun, in particular to that part of the electromagnetic spectrum referred to as light. Light travels in straight lines, and bounces off solid surfaces. We move about immersed in a 'sea' of bouncing electromagnetic energy – and we have the ability through our eyes to analyse this mass of bouncing energy in order to detect the shapes of objects in our environment.

The Visual field: Humans are sensitive to light – but then so is a sunflower. Creatures with eyes do much more than respond to light – we focus the light through a lens onto a set of receptor cells. Certain cells (the majority) respond just to the presence or absence of light,

---

[30] en.wikipedia.org/wiki/Zoetrope Invented by George Horner. "It consists of a cylinder with slits cut vertically in the sides. Beneath the slits, on the inner surface of the cylinder, is a band which has either individual frames from a video/film or images from a set of sequenced drawings or photographs. As the cylinder spins, the user looks through the slits at the pictures on the opposite side of the cylinder's interior. The scanning of the slits keeps the pictures from simply blurring together, so that the user sees a rapid succession of images producing the illusion of motion, the equivalent of a motion picture."

[31] www.centres.ex.ac.uk/bill.douglas/Schools/animation/animation4.htm

allowing us to see in black and white. Two further types of cells respond selectively to frequencies within that band, giving us colour vision.

As we move through the sea of light, we are bombarded from all directions. Most of this light does not enter our eyes, because like a camera we point our visual receptors (eyes) in one direction. Also human vision is non-uniform. Although we have peripheral vision out to nearly 180° (a hemisphere), we can only actually focus and build up a detailed image using the cells of the centre of the retina (the fovea). The angle for this “high-resolution imaging” is quite small: about 5° (a very narrow cone)[32]. It is this narrow cone of detailed vision that allows cinema (and its small-screen relatives television and computer monitors) to succeed in producing images that are at least to a degree believable. We can tolerate what we see on a small screen because human vision has a built-in ‘small screen’.

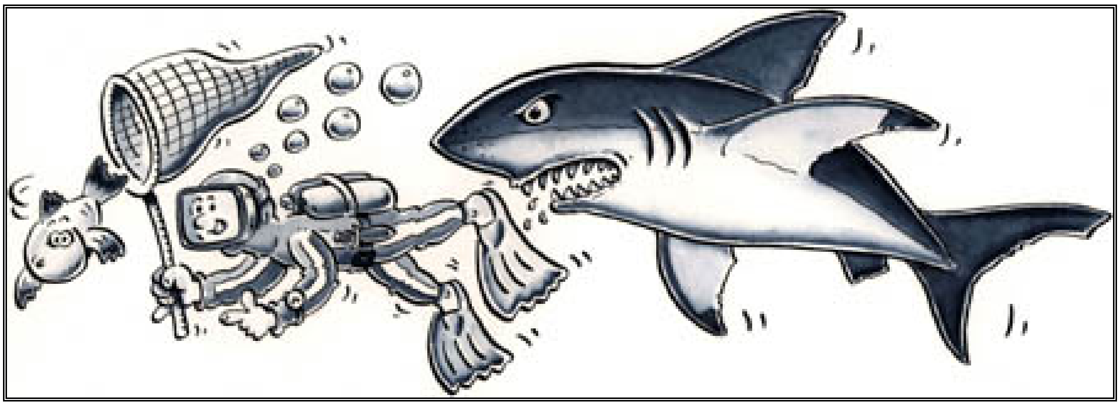


Figure 1- We don’t see what’s behind us – and we can’t focus on much of what’s in front of us either!

Colour space: There are two reasons for going into some basics of colour vision:
Moving images can be in colour, or they can be in black and white (monochrome).
The ways film, video and digitisation deal with colour should make more sense if put into the context of how human perception deals with colour.

The basic facts about colour vision are these:
There are separate receptors in the eye for colour, and there are proportionally fewer colour receptors than general “light/dark” receptors. The consequence is that colour is perceived with lower detail.

Colour is perceived in a two-dimensional space, familiar to us as the “colour wheel”. In combination with brightness information (which is an independent dimension), there are in total three dimensions to vision – and there are an equivalent three dimensions to colour film, colour TV cameras (and TV sets and colour scanners), colour printers, transmission of colour television signals, and representation of colour on computer systems.

[32] www.du.edu/~jcalvert/optics/colour.htm

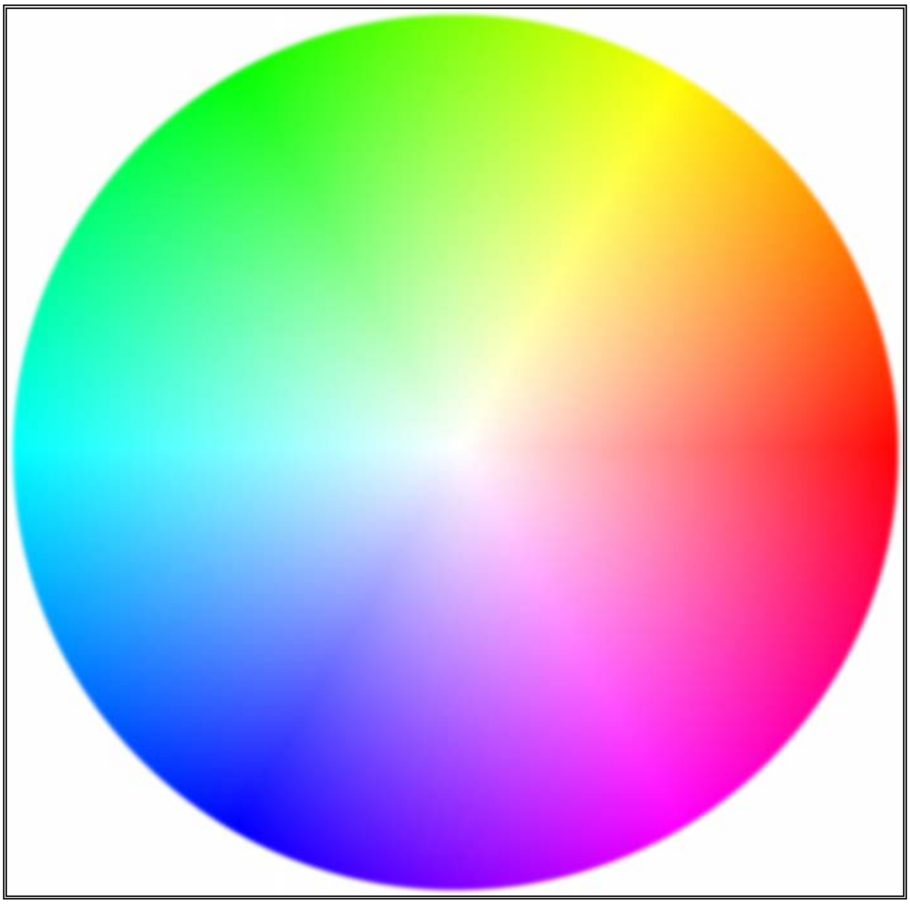


Figure 2- A colour wheel, showing colour arranged in two dimensions.

There is much more that could be said, and TASI (Technical Advisory Service for Images) has a general paper on colour theory[33].

Flicker Fusion: The final fact about vision for moving images is that the eye integrates visual information over a time period on the order of 20 milliseconds. This means that if an image flickers on and off, the flickering disappears from human perception if the rate is above something like 50 Hz (50 on-and-off cycles per second). So all the inventors of cinema had to do was find a way to take photographs 50 times per second, and the rest was in comparison relatively easy. It took many seconds to make a single exposure of film in the early days of photography (around 1850), and took many minutes to actual take a photograph (remove the wet plate, re-insert a new plate, re-charge the flash powder, re-focus) – so the development of cinema was a considerable challenge.

The earliest forms of cinema operated at considerably less than 50 exposures per second, and 14 frames (exposures) per second was a common speed until the 1920's – and so the perception of flicker is a common effect in early film. In addition to raising the speed to 25 (24 in North America) frames per second, film projectors also flashed twice on each frame, so the flicker rate was 48 flashes per second (or 50) and flicker disappeared for the average viewer.

Film: A succession of still images, and an appropriate projector, are all that is needed for cinema. Each image is *continuous*: a record across the two dimensions of the film frame (across the area of the exposure, and generally corresponding to height and width of the visual field) of the pattern of light reaching the film. In this sense film is continuous: the

[33] *Colour Theory: Understanding and Modelling Colour*: www.tasi.ac.uk/advice/creating/colour.html

images are continuous in space. There is no quantisation or division of the analogue image into discrete elements or pixels. Under a magnifying glass, a film frame does not break up into dots or little squares – although there is an underlying physical quantum: the grain size of the crystals of photosensitive material used to produce the film. Grains are not like pixels in that grains are not of uniform size and shape.

However film is discrete in time, as just discussed: 24 or 25 discrete events per second. Because film is both continuous and discrete, it does not fit well into the division of phenomena into analogue vs digital. This complexity is inherent in film; television signals are even more complicated.

Moving images preceded the dominance of the computer and the digital age by a century for film, and by a half-century for video. Film provided the only storage medium for these moving images (film and television) up until the advent of videotape in the late 1950's. Film was a widely-used production and storage medium in television right up until the 1980's, and as a consequence a large portion of television archives remains on film (40%, totalling nearly 300,000 hours, in the case of the BBC archive).

Television: A key problem in trying to combine cinema technology with the long-distance properties of radio, was the question of how to turn an image (an essentially continuous phenomenon over two dimensions) into something more like sound – which has one dimension (the amplitude of the air pressure at a point, for sound in air picked up by a microphone) which varies over time**.**

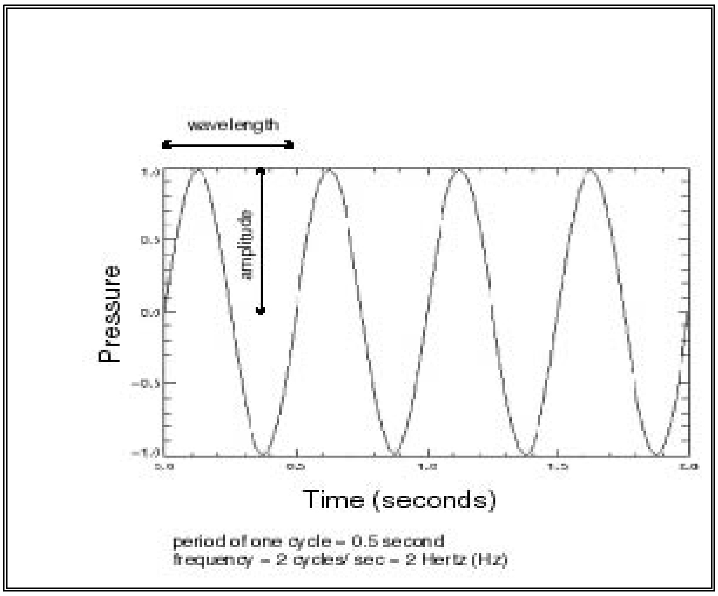


Figure 3- A sound wave: pressure varying over time.

Sound, when picked up by a microphone, is a rather simple *signal*: a voltage that varies over time. An image isn't a signal – even without colour to worry about, a monochrome image various continuously in two dimensions, and doesn't appear to be anything like a signal. The problem of a one-dimensional representation of moving images occupied many people, including a young American farm boy named Philo Farnsworth who discovered an answer while ploughing: just as a plot of land could be covered by the lines of the plough, an image could be converted to a signal by covering the image with lines[34].

[34] www.time.com/time/time100/scientist/profile/farnsworth.html

In principle, it just takes a very tiny photocell, and a very accurate method for moving that cell around an image, in order to scan a frame of film (or a photographical negative) and produce a signal that could then be transmitted by the methods already developed for radio – and indeed Farnsworth transmitted a photo by radio waves in 1927[35]. Many other people were simultaneously developing television, but the common feature was conversion of a two-dimensional, continuous image into a one-dimensional and highly structured signal, using some form of scanning.

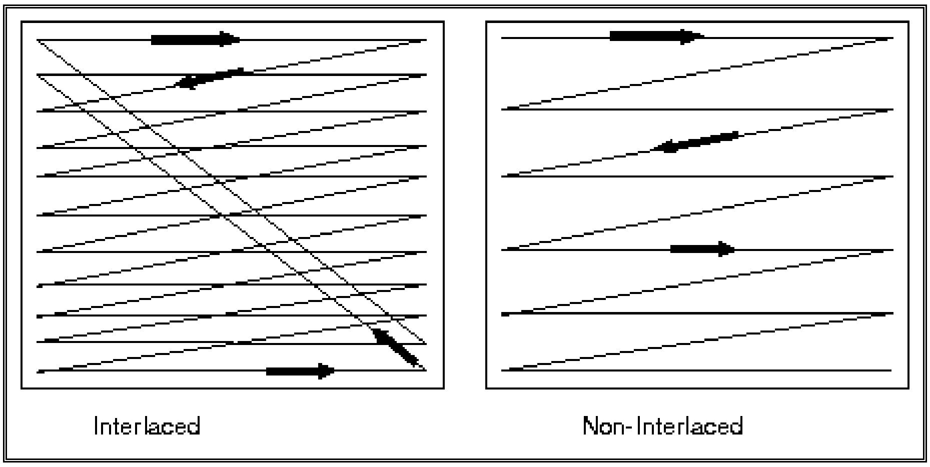


Figure 4- two types of raster scan, used to convert an image into a signal. Interlaced scanning allows twice as high a refresh rate, reducing flicker.

As mentioned, film consists of continuous images in space (except for film grain) that are discontinuous in time. Television relies on a signal that is discontinuous in both time and space, although during the course of a single line (which in modern television is only a duration of 50 microseconds) the brightness of the image can be continuously tracked. So television is discrete in time and in the vertical dimension, and continuous in the horizontal one. Standard broadcast television signals are analogue signals (because they have not been digitised), but certainly not one continuous signal like sound from a microphone, which isn't quantised in any way.

5.2.2 Digital moving images

Audio is a single continuous signal, which can be represented for recording and reproduction purposes as an *analogue* of the original sound pressure (at one point, the location of the microphone). It can also be *digitised*: divided up into samples, and each sample can be represented as a number. So, as an alternative to analogue recording, sound for use by computers and the whole range digital processing technologies (ranging from CD recording media to digital radio, with computers somewhere in the middle) can be converted to a *digital* format, and represented as a sequence of numbers.

Film and video can also be in digital form but the situation is more complex for two reasons. Firstly, film and video are not continuous data streams; both are discrete in time, and video is also formed of discrete lines which scan an image. Secondly, this structure (time sampling and scanning) is information (metadata) that has to be included with a digital signal in order for it to be correctly displayed.

A monochrome scanned image is a single signal, but a colour image requires adding two further dimensions to represent location in the colour wheel. Hence coloured images require three signals, which when digitised convert to three sequences of numbers. These three

[35] www.museum.tv/archives/etv/F/htmlF/farnsworthp/farnsworthp.htm

sequences can be handled in various ways, and it is part of the definition of a video signal to specify how the colour information is combined with the brightness (light vs dark; monochrome) information.

Because of this complexity, it took video engineers about a decade to agree on a standard for the digital representation of video, the European Broadcasting Union and CCIR Recommendation 601[36]; this standard was agreed in 1980. By contrast, the standard method of digitising sound (linear PCM) has been around since 1937[37].

Finally, as with all technology, digital video (and film digitised as video) is in a transitional state. The standard television display technology for 80 years has been the cathode ray tube (CRT), a scanning technology that fits exactly with the scanning technology used in television cameras. Now most domestic cameras and some professional cameras use technology which does not scan, but instead ‘grabs’ an entire visual field. At the viewer’s end, CRT displays are being replaced by LCD and plasma technology which also do not need scanning to produce an image. Therefore some of the complexity of video signals is no longer necessary. For instance, in a world of LCD displays there is no need for *interlacing* the scanning in order to reduce flicker – because there is no scanning and there is no flicker! However issues of compatibility (with viewing on a TV playout, and with storage technology that relies on TV-based viewing, eg DVD) mean that many forms of digital moving images still use video standards with all their complexity.

There are many characteristics of digital moving image which are necessary to provide a moving image as it was intended:

- Resolution (or definition): the number of lines per image, and the rate at which each line is sampled;
- Aspect ratio: an image is a rectangle of a definite shape;
- Frame rate and fields: a frame is a full image, and the frame rate is the number of full images per second. To reduce flicker, video divides a frame into two interlaced halves (called fields);
- Monochrome, or component or composite colour video: the colour signals can be kept as separate components (three signals), or combined with the monochrome signal into one composite.
- Gamut: the range of colour values which can be captured in devices like cameras and sensors or which can be displayed by a rendering device such as an LCD monitor or a CRT television.

The resulting moving image can be held on digital videotape or DVD (or even certain formats on CD), in a format that plays in real time and produces an image on a television receiver. All digital videotape formats are both proprietary and incompatible with competing formats. DVD has a standard format, although higher-definition (and again incompatible) formats (HDVD, BluRay) are just coming to the market.

Resolution: television started in the UK with scanning at 30 lines per image. This was not very detailed, and so the BBC launched its first ‘high definition’ service in 1937, with 405 lines per image. This was gradually replaced after 1964 with the current 625 line system. Ironically, the television industry is again in the throes of upgrading to another ‘high definition’ standard, with a resolution of 720 or 1080 lines (both are already in use). In North America and Japan television launched with 525 lines.

---

[36] *Digital Video – 25th Anniversary*, three articles in the EBU Technical review of October 2005: www.ebu.ch/en/technical/trev/trev_304-contents.html

[37] “PCM was invented by the British engineer Alec Reeves in 1937 while working for the International Telephone and Telegraph in France.” From: en.wikipedia.org/wiki/Pulse-code_modulation

Aspect ratio: a digital image is almost always a rectangle, but the shape of the rectangle may be unspecified. Video was a 12:9 image until the 1990s, but 16:9 is the new 'widescreen' standard. The problem is, there is nothing in the sequence of numbers that constitutes an image that automatically guarantees the correct aspect ratio. There will be a certain number of samples per line (horizontal resolution), and lines per frame (vertical resolution) – but these two resolutions do not have to be equal. The only real solution is adequate metadata to define the aspect ratio.

Video encoded for web access and computer display (and not for display on a conventional television) is usually described in terms of the number of pixels (picture elements) in the horizontal and vertical dimension. Again, this does not unambiguously define the aspect ratio, as there is no inherent necessity for a display to have a square pixel. A typical computer monitor does have a square pixel, but encoded video is always played on a "virtual monitor", a software player that opens in a window on the display. Many such players can be manipulated by the user (sized) to have virtually any aspect ratio, and consequent image distortion. A television engineer yearns in vain for a standard test pattern at the start of each video file. Again, the best that can be done is to ensure that the intended aspect ratio is in the metadata.

Frame rate and fields: television uses the mains power signal as a reference frequency, so in Europe we have 50/2 = 25 frames per second, and in North America and some other places they have 60/2 = 30 frames. A frame is a full image.

When television was being developed, the limits of the technology of the time meant that it was just too difficult, and used too much bandwidth, to send a full frame 50 times per second. A full frame could be sent at half that rate, but that would be a refresh rate of only 25 Hz, well below the 'flicker-fusion threshold' and hence unacceptable. The solution was to divide the video signal into two halves, which can be thought of as the even-numbered and odd-numbered lines (starting with 1). The odd-numbered lines are sent in one *field*, followed by the even-numbered lines. One field is sent for each cycle of the mains frequency, and as a field is distributed across the whole TV screen and updates at 50Hz, the flicker is acceptable. This method of dividing a frame into fields is *interlaced* video, which will eventually be made obsolete by modern display technology.

The alternative to *interlaced* lines is simply to scan, transmit and display lines in one sequence: 1, 2, 3… up to 625 or 720 or 1080. This method is called *progressive.* The conventional notation for specifying resolution is to use a 'p' or 'i' after the number of lines per frame, to show whether the format is progressive or interlaced. 1080i is a common high-definition format, as is 720p.

Video that is encoded for web-access and computer display is no longer tied to mains frequency, so many frame rates are now in common use, including those used by web-cam displays that refresh only once every few seconds (although whether such an image is really "moving" is questionable).

Monochrome, or component or composite colour video: Most new moving images are now in colour, but much historical material is in black and white. When such material is digitised, it will probably be digitised 'in colour' simply because modern video digitisation and display standards assume colour, regardless of the colour of the analogue source.

However the real complications arise with the various methods for mixing colour and monochrome signals. Computer file formats can handle three simultaneous sequences without difficulty, but historically a three-dimensional signal was a challenge, and so the colour information was mixed in with the monochrome information in one *composite* signal. The alternative was to keep the signals separate, and use three wires to make every video connection. This approach is labelled *component*, as it keeps the components apart. Unfortunately, all the analogue methods of broadcasting video required composite signals, so composite video is still very much with us, even in the digital world.

Some digital video formats are composite (eg D3 videotape) and others are component (Digibeta). The information on a DVD (which is a digital media format) is encoded in a way that is much more like a computer file format, and so will be component, but the domestic DVD player expects to connect to a standard television receiver, and so will transmit an analogue, composite signal.

There are two main forms of component signal:

- RGB: using the three primary colours red, green and blue. These are associated with the three elements in a colour camera and the three phosphors in a colour monitor, so RGB is a straightforward way to deal with colour – but all three are needed to make white (or any shade of grey down to black) so it is a complicated way to deal with monochrome should the need arise;
- YUV: using one signal (Y) for the monochrome information, and two others (U, V) strictly for colour.

To make matters worse, there are three standard forms of composite video, dating back to the competition to develop colour television in the 1950's and 1960's:

- PAL (phase alternate lines) used in the UK and widely in Europe and elsewhere;
- NTSC (National Television System(s) Committee) used in North America, Japan and Korea, Central America and parts of South America;
- SÉCAM (Séquentiel couleur avec mémoire) used in France, Russia and parts of Eastern Europe, and parts of francophone Africa.

Gamut: the whole area of digital moving images is made more complex by the issue of gamut, and the ability of current technologies to represent this. Gamut refers to the range of colours that can be captured or reproduced by a device. Analogue colour film has a much larger gamut than can be reproduced by most display or reproduction technologies. For this reason reformatting strategies for preserving film originals and video originals are different. Digital camera sensors capture a different gamut from analogue film. It can be crucial to record information about the gamut (often called 'colour space') about digital images in order to ensure that they are reproduced or preserved correctly.

It is recommended that JISC undertake research into the issue of the effect of gamut on preservation and access strategies.

5.2.3 Digital moving image formats

TFile formats and encodings are complex and fraught with terminological distinctions. Many of the recognised formats include specifications for encodings, file formats, and wrappers under the same generic name (e.g. JPEG 2000) and it is sometimes not clear where the distinction between a file format and an encoding lies, or between a file format and a wrapper format. Rather than trying to categorise formats as encodings, wrappers, or file formats this section describes the most common formats and identifies within each format the different format types. Much of the information here is based on the excellent Library of Congress web site on formats at: http://www.digitalpreservation.gov/formats/. Readers desiring more complete and detailed description of formats should visit that site.

Common moving image formats

Formats that are specifically for audio are dealt with below in section 5.3. However, some wrapper formats which can be used to wrap audio streams are listed in this section because they also wrap moving images. The information on those formats will not be repeated in the section on audio formats.

Advanced Authoring Format (AAF)

A format that wraps metadata together with video, audio, and other bitstreams, designed for the interchange of content across platforms and between applications. AAF is intended for video post-production use and not for use as a playback format. The Material Exchange Format (MXF) is a subtype of AAF which may therefore be classified as a meta-wrapper format. Not yet widely adopted but interest in the format is growing.

Advanced Systems Format (ASF)

A proprietary wrapper format developed by Microsoft that can wrap various content bitstreams, including audio, video, script command, JPEG-compressed still images, binary, and other streams defined by content developers. Its primary use is as an end-user delivery format. ASF is the supertype for both the WMA (Windows media audio) and WMV (Windows media video) formats and may therefore be described as a meta-wrapper format. It can also be used to wrap non-Windows based media streams. ASF is primarily used as an end-user delivery format.

Audio/Video Interleaved (AVI)

A proprietary Microsoft family of specifications based on the RIFF meta-format specification. The AVI family includes a wrapper format *(.avi)* usually used to wrap video streams with other data chunks, such as audio, and a number of file format specifications for wrapping specifically encoded bitstreams, such as AVI Indeo, AVI Cinepak, and AVI DivX, with other data such as audio. AVI files may contain multiple streams of different types of data (e.g. DV, MJPEG, J2K). The format has now been largely replaced by MPEG and the Windows media formats. AVI files are usually large files of very good quality, but they must be encoded/decoded properly for the quality to be apparent. Widely used for digital video production and filmmaking.

Cinepak (cvid)

One of the most widely used older video codecs for encoding video bitstreams. It was the primary video codec used in early versions of Apple's QuickTime and Microsoft Video for Windows, but was later superseded by Sorenson Video, Intel Indeo, and most recently MPEG-4. Cinepak encoded bitstreams can be wrapped by a number of file formats including AVI and Quicktime.

Digital Cinema Initiative Distribution Master (DCDM)

The DCDM is a relatively recent specification for a distribution package for digital cinema. DCDM is not really a wrapper format, and may perhaps be better categorised as a 'packaging' format. A DCDM package consists of a set of uncompressed and unencrypted files that represent moving image content optimized for the electronic playback in cinemas. The specification includes not only feature films but is also intended for content like trailers and advertisements. The Digital Cinema Initiative Package (DCP) format is based on DCDM.

DivX (.divx)

A proprietary bitstream encoding format for video, created by DivX Inc. and based on MPEG-4. DivX is widely used as an end-user format for distributing video over the Web primarily because of its ability to achieve high compression ratios. An open source version is available (Open DivX). Version 6.0 of DivX is also a wrapper format but is usually considered as being just a rebadged AVI format. Can be used by a number of file formats, including AVI and ASF.

Digital Video formats (DV, DVCAM, DVCPRO)
A related family of open bitstream encoding formats for recording digital video on physical media (tapes, hard disks) through digital video devices (digicams, camcorders). Now, with the MiniDV tape format, the standard for consumer and semi-professional video production. Digital video streams are encoded using discrete cosine transforms (DCT). Digital video streams can be stored as files in raw form (.dv or .dif extension) or wrapped in file formats like AVI, QuickTime, and MXF.

Videotape gradually replaced film as a working medium in television, and digital videotape appeared in 1987. There are about 15 digital videotape formats, from the original D1 (Sony, 1987) to the consumer-oriented DV formats (miniDV) that started in 1996. Variants on the DV format include DVCAM and DVCPRO, and all three have versions for high-definition recording: DV-HD, DVCAM-HD and DVCPRO-HD. All of these consumer or 'pro-sumer' formats are likely to be in use for some time, along with professional digital formats such as digibeta, D5 and IMX. There are variants in tape sizes and other physical characteristics (e.g. speed) of camcorders (miniDV, D-9, Digital8), and it is becoming more usual to find DV recorded on hard disk drives installed in DV equipment. Table 1 below, provides an indication of the formats available, the differences in tape capacity, cross-compatibility, image and audio quality.

Table 1: Digital video formats

Format	**DV**	**DVCAM**	**DVCPRO**	**Digital8**
Supplier(s)	Consortium of 60 manufacturers including Sony, Panasonic, JVC, Canon, Sharp.	Sony, Ikegami	Panasonic; also Philips, Ikegami, Hitachi.	Sony, Hitachi
Type of Cassette	MiniDV	MiniDV	Small	Video8, Hi8 standard NTSC 120 minute tape: 60 min; standard PAL 90 min tape: 60 min.
Maximum tape capacity	80/120 min. (SP/LP)	184 minutes	63 minutes (AJ-D400/610/700/810); 123 min. (AJ-D200/210) 184 min. (AJ-D410)**	60 minutes
Resolution & Sampling	720x480, 4:1:1 (NTSC) 720x576, 4:2:0 (PAL)	720x480, 4:1:1 (NTSC) 720x576, 4:2:0 (PAL)	720x480, 4:1:1 (NTSC) 720x576, 4:1:1 (PAL)	720x480, 4:1:1 (NTSC) 720x576, 4:2:0 (PAL)
Audio Quality	2 ch @ 48 kHz, 16 bits; 4 ch @ 32 kHz,	2 ch @ 48 kHz, 16 bits 4 ch @ 32 kHz,	2 ch @ 48 kHz	2 ch @ 48 kHz, 16 bits 4 ch @ 32 kHz,

	12 bits;	12 bits		12 bits

Indeo (iv30-50)
A bitstream encoding for video, originally developed by Intel. Both Windows implementations and Mac versions exist. Indeo is generally used as an end-user delivery format. Indeo version 3 is said to use vector quantization, while the later version, Indeo 4 and 5, use wavelet transforms. Indeo is used by AVI (Audio Video Interleaved) and QuickTime wrapper formats. The encoding was widely used in the 1990s but has now been largely superseded by the MPEG family of encodings.

(ITU) Recommendation BT 601 (Rec. 601)
A standard for encoding interlaced analogue video signals in digital form, developed by the SMPTE with input from the European Broadcasting Union and issued by the International Telecommunications Union (Radiocommunication Sector) as a standard for digital TV in 1982[38]. ITU-R BT 601 specifies the image format, acquisition semantic, and parts of the coding for digital "standard" (ie. PAL, NTSC, SECAM) television signals. The latest version is ITU-R BT 601-2, issued in 1990. With 'born' digital TV broadcasts the situation is more complicated with a number of different standards vying for adoption: DVB-C, DVB-S, DVB-T, J83.B, ISDB-T, DMB-T, SDI and HDTV.

JPEG 2000
A large family of format specifications for digital images. JPEG 2000 is a wavelet based format developed from Apple's QuickTime by the Joint Photographic Experts Group (JPEG), and specified formally in ISO 15444. The JPEG2000 specification encompasses both lossy and lossless compression methods. The JPEG 2000 family includes 3 file formats (.jp2, .jpf, .jpm), 10 bitstream encodings (including lossy and lossless compression encodings - .j2c) , and 8 formats referred to by the Library of Congress as *combo-packs*, ie. file formats with specific bitstream encodings, e.g. .jp2 with lossy compression, .jp2 with lossless compression, etc. For more information on the JPEG 2000 formats see the Library of Congress file formats web site at:
http://www.digitalpreservation.gov/formats/fdd/browse_list.shtml#j.

Since JP2 is a relatively recent standard it is not yet widely used, although the Library of Congress now specifies JPEG 2000 as its format of choice for preserving moving image files. The Library of Congress currently proposes to use MXF as its wrapper format for JP2 files, although other wrapper formats, such as MJPEG 2000 (defined in the JPEG 2000 specification), and AAF can be used to wrap JPEG 2000 files.

Motion Picture Experts Group (MPEG) standards
A large family of specifications for encodings, file formats and wrappers developed by the Motion Picture Experts Group. Each of the significant MPEG standards is treated separately below.

MPEG-1 (1993)
MPEG-1 is a family of video and audio bitstream encodings, designed for storage of audio and video on digital media such as CDs. MPEG-1 was later used as the standard for Video CD, and also includes the very popular MP3 compression standard for digital audio.

MPEG-2 (1994)
MPEG-2 is a family of encoding formats for compressed video and audio data multiplexed with signalling information in a serial bitstream. MPEG-2 was initially developed for the

[38] A full discussion of the development of Rec. 601 can be found at: http://www.ebu.ch/en/technical/trev/trev_304-rec601_wood.pdf?display=EN.

transmission of television programs via broadcast, cable, and satellite, and was subsequently adopted as the audio and video standard for DVD production and for some online delivery systems. MPEG-2 is widely adopted, and is required by the ATSC digital television services specification for use in digital television broadcasting in North America. The MPEG-2 family also includes the popular MP3 audio encoding format.

MPEG 4 (1998)
MPEG-4 is a newer family of encoding formats using lossy compression that incorporates many of the features of MPEG-1 and MPEG-2 and is based on Apple's QuickTime format. MPEG-4 includes 2 compression standards taken from the European ITU-T specifications H.263 (MPEG-4 part 2) and H.264 (AVC, ie. MPEG-4 part 10).The primary uses for the MPEG-4 standard are web (streaming media) and CD distribution, conversation (e.g. videophones), and broadcast television. MPEG-4 adds new features such as (extended) VRML support for 3D rendering, object-oriented composite files (including audio, video and VRML objects), support for externally-specified digital rights management and various types of interactivity. The MPEG-4 AVC (advanced video coding) encoding format is compatible with Apple's video iPod introduced in 2005. Satellite and mobile device distribution of moving image content will use the very efficient H.264 specification. MPEG-4 also includes a wrapper format (part 14 of the family of MPEG-4 formats) that allows for encapsulating video and audio objects, 3D content, with support for Digital Rights Management. Widely used for streaming multimedia files over the Internet.

Motion JPEG (mj2, mjp2)
Strictly speaking, Motion JPEG2000 is a wrapper for JPEG2000 frame images, audio, and related elements. Its specification is part 3 of the multi-part ISO 15444 specification for JPEG2000. Each frame of video is compressed into a JPEG2000 still image and the overall format can be seen as a successor to Motion JPEG, which never had a specified file format of its own. Taken in sequence, the series of JPEG2000 images represent the source video. It is worth noting that a series of JPEG2000 frame images can also be wrapped in such formats as AVI. Plans at the Library of Congress call for the investigation of JPEG2000 frame images in an MXF wrapper. All of the JPEG2000 frame-based approaches are likely to be for mastering rather than for distribution.

Material Exchange Format (MXF)
An open wrapper format developed by the Society of Motion Picture and Television Engineers (SMPTE) which supports a number of different bitstreams encoded with any of a variety of codecs, together with a metadata wrapper which describes the material contained within the MXF file. The Library of Congress notes that "most commentators say "MXF should be seen as the 'digital equivalent of videotape,'" an allusion to tape's simple, linear structure.".[39]MXF is related to the AAF format (see above). It was specifically developed for optimized interchange and archiving of multimedia content. Although currently (2006) too new to be widely used it is emerging as a standard. The Library of Congress intends to use MXF as its wrapper format of choice for JPEG 2000 image frames.

Ogg format (.ogg)
An open wrapper file format, that can encapsulate audio, video and other bitstreams. Ogg can encapsulate the following encoding formats: Speex, Vorbis and FLAC for audio streams; Theora and Tarkin for video streams; and WRIT, a text encoding format designed for embedding captions and sub-titles. Most often used for encapsulating Ogg Vorbis (see below) encoded audio streams.

Open Media Framework (omf)

[39] Library of Congress file formats website. See: http://www.digitalpreservation.gov/formats/fdd/fdd000013.shtml

A video format developed by and used primarily by Avid editing systems, but has been adopted by other professional video applications and has become a high-end standard for platform-independent transfer of digital media between different software applications. Usually these files are not playable from the desktop but rather only inside the applications that use them so they are application dependent.

QuickTime (.mov, .moov, qt)
A multimedia wrapper format developed by Apple to encapsulate audio, video and other bitstreams. Often used with Cinepak or Sorenson codecs for encoding video bitstreams. Apart from the wrapper format, the QuickTime specification includes a number of what the Library of Congress refer to as combo-packs, ie. specifications for QuickTime files containing bitstreams encoded with specific codecs, e.g. QuickTime video (QTV) with Sorenson, QuickTime Video (QTV) with MP4, and QuickTime audio (QTA) with AAC. QuickTime is not platform specific and is used on Macintosh, Windows and Unix/Linux based computers. It has been widely adopted.

Real Media (RM)
A family of wrapper and encoding formats developed by the Real Networks company. multimedia wrapper format created by RealNetworks. The Real Media family includes the Real Media wrapper format that is typically used in conjunction with Real Video and Real Audio codecs, and is popular for streaming content over the Internet. The RealVideo format is a proprietary video codec developed as part of the real Media family of formats. Usually used for streaming video across the Internet, and now starting to appear in video sent to mobile (cellular) telephones.

Sorenson (svq)
A digital video codec for encoding bitstreams devised by the company Sorenson Media, and used by Apple's QuickTime, and in the newer versions of Macromedia Flash. Generally used as an end-user delivery format. Not widely used in the Windows community but widespread in the Apple community. The Sorenson codec can be used by a number of file and wrapper formats, apart from QuickTime.

Windows Media Video format (.wmv)
A family of video formats by Microsoft. Windows Media Video is a wrapper format within the specification which can wrap encoded audio, and video bitstreams with other data streams. The bitstream encodings in the specification are compressed formats that include the WMV9 codec and the WMV9 PRO codec. WMV files are often encapsulated within an AVI or ASF wrapper file. Widely used on the Internet for streaming video, and now also used to distribute high definition video on standard DVDs in a format Microsoft has branded as WMV HD.

5.2.4 Issues in archiving and preserving digital moving images

It is certainly not enough to say “a file is just a file, so once everything is digital it can all be handled in the same way”. Documents, data, still images, sound and moving images all have distinctive characteristics, vital for the use and preservation of the content of the files.

a) Analogue vs digital moving images

While there is a simple relationship between a sequence of numbers and a sound (as measured by a microphone), for moving images the relationship is complicated. Film is a discrete sequence of analogue images, like a sequence of photographs (at 24 per second). Video is a discrete sequence of scanned images, where each image is scanned as a fixed number of lines (525 or 625 or the new high definition formats of 720 or 1080) – as although each frame of film had been sliced into hundreds of horizontal strips. The only continuous portion of a so-called analogue video signal is the brightness and associated colour information running along each strip.

While each film frame is still projected as a continuous image, and while monochrome television receivers have a light-emitting phosphor spread uniformly across the screen, colour television and colour monitors are actually constructed of thousands of individual phosphor dots (of red, green and blue). So even the continuous video signal hitting a monitor excites discrete dots of light-emitting phosphor (or discrete LED elements), leaving nothing about the signal that is not already discrete – even though none of it is digital.

A video signal is heavily structured (divided into discrete units: frames, fields, lines), so any subsequent digitisation (assigning of numbers to the discrete units) has to match the structure – or information will be lost or scrambled. The digital files made from video signals need to match the structure of the original video, which means that the complexities of traditional video will live on in the computer and file-based age.

The DVD format is an important form of digital storage of moving images. All DVDs currently contain video that uses lossy MPEG-2 compression. There is currently (2006) interest in allowing the use of other encoding methods in DVDs, and in any case encoding methods will necessarily change with the adoption of HD-DVD and BluRay formats for DVD. It is worth noting that DVD disks often contain interactive programmes, and preservation methods will need to ensure that the interactive menus and other features are retained..

b) Multiple formats

Digital moving image files can arise in a number of ways. Firstly, analogue formats can be digitised using a video capture card. In this case, the analogue format will be played into the capture card either as a composite or component signal. If composite, the colour if present will be coded as SECAM, PAL or NTSC. The video will generally have either 50 or 60 fields per second and either 525 or 625 lines. All these parameters must be known, and communicated to the video capture software, and supported by the video capture hardware, in order for the video to be correctly digitised.

Direct transfer of digital data into the computer from consumer digital videotape (DV and its variants) or DVD is the other main method of creating digital moving image files. This is a much simpler process. DV camcorders use the firewire interface (recent digital video camcorders may use USB-2) and standard software transfers the data unchanged from the camera to the computer. The complication is that cameras use encoding in various formats, so the resultant file will be compressed. This type of direct transfer is described as "acquisition of the digital camcorder data in native format" and is the best that can be done (in terms of archiving an object exactly as found), but the resultant file is compressed and is not in a standard format.

Professional digital videotape players also have digital output. Reformatting will usually start with a tape player appropriate for the tape format being reformatted (a non-trivial exercise for obsolete formats). The output from the tape player is either in the form of a "601" or SDI (Serial Digital Interface format[40]) signal, or a data stream that needs to be processed by specialised equipment (time base correctors, comb filters, etc) before it can be output as 601/SDI formatted digital video. Note that SDI is a real-time, digital, professional device interconnection format, not a file format. However some professional video capture cards accept an SDI input, which may be further processed to produce, for example, a MPEG-2 encoded data stream or a series of JPEG2000 frames. The data stream can then be file-formatted and/or wrapped into the digital object that is to be preserved.

[40] en.wikipedia.org/wiki/Serial_Digital_Interface

The files can be created digitally (graphics, special effects). In the latter case, they can still be created as *digital video* (a sequence of numbers matching the frame/field/scan line structure of television) or they can be arbitrary arrays of pixels that can only form moving images on a computer and could never be transferred to a television receiver. Again, all such files are typically in proprietary formats.

It is possible to scan film using technology similar to document scanners, and some document scanners have adapters for film. Because this technology is both slow and of limited quality, it is not likely that there will be significant amounts of film scanned by low-cost document scanners. High-quality digitisation of film is a very professional task, using high-definition scanners (*telecine* and *datacine* equipment) that cost about 1000 times as much as a desktop frame-by-frame film scanners (which cost a few hundred pounds).

There is also no single accepted file format for digitised film. The ITU-R BT601 digital TV standard is not adequate, because film needs to be scanned at higher than television 'standard definition' resolution or information (quality) is lost. It is possible that 16mm film will be adequately captured at high definition (1080 vertical samples, 1440 horizontal samples for a 4:3 aspect ration, which is standard for 16mm film, especially in broadcasting). "Digital film" as used in cinema special effects and for digital restoration is often stored as discrete images, one file per frame – and it is left up to the system administrator to ensure that all the files are kept together, with a naming convention to ensure they are used in the correct order. This approach to digital film storage is the one adopted by (Kodak) Cineon[41] or the closely related SMPTE/ANSI DPX[42] file format, although the latter is really a 'folder format' and each file is simply metadata plus a sequence of numbers representing an array of pixels forming one frame.

c) Standard format for digital video

The DV format, created by a consortium of leading electronic manufacturers, can be considered an international standard for digital video. DV uses 1/4 inch (6.35mm) metal evaporate tape to record very high quality digital video. The video is sampled at 720 pixels per scan line although the colour information is sampled at half the D-1 rate 4:1:1 in 525-Line (NTSC) and 4:2:0 in 625-line (PAL) formats. Other digital video formats - DVCAM, DVCPRO and Digital8 - are similar to DV, but have a niche market and are not considered a standard (see Table 1 above).

In broadcasting, the technology is gradually moving to the MXF[43] file format, but this is a more generic 'wrapper' format that can take many types of encoded video (or audio). If there is a 'canonical form' for broadcast video, it is ITU-R BT601 (see above) which is uncompressed and takes 270 Mb/s (megabits per second) for a 625-line colour signal.

It is possible to migrate digital videotape by copying from an old tape to a new tape. In the case of DV, a copy can be made on computer storage media - about 20 minutes of DV native format (an .avi file at 25 Mb/s) can be stored on a standard DVD-ROM. A digibeta to digibeta transfer will not use native format, as the digital connection between the two machines uses the SDI standard (270 Mb/s, uncompressed from the 80 Mb/s native format on the tape). Videotape players have built-in dropout correction, and will 'fix' a line (that fails to read) by replacing it with a neighbouring line (concealment). This feature, designed to make a videotape play despite read errors, has the unfortunate effect of interfering with the creation of an exact copy (cloning). Concealment will also affect the SDI output into a capture card – so there is no way to actually get exactly "what's on the tape" without very special equipment.

---

[41] www.cineon.com/ff_draft.php
[42] en.wikipedia.org/wiki/DPX
[43] www.pro-mpeg.org/publicdocs/mxf.html

d) Encoding and compression

Moving images simulate motion as a rapid succession of fixed images. It has long been known that there is redundancy in most images: pixels look like their neighbours, to a degree. Digital image compression began in the late 1960's with satellite imaging, and the need to send high-resolution images over low-bandwidth communication channels. Image compression is now a well-established technology[44], and all of this technology is applicable to moving images. A very common family of image compression techniques is JPEG, named for the Joint Photographic Expert Group[45] responsible for the standards.

A sequence of fixed images succeeds in fooling the eye only by virtue of each image being rather similar to the preceding and following images. So moving images have redundancy between images, which provides additional compression possibilities. Theremoval of interframe redundancy is also called temporal redundancy. There are many forms of compression, but a large group of related standards (covering images plus associated metadata and rights) has been developed by MPEG, the Moving Pictures Expert Group[46].

MPEG and other forms of data reduction have had remarkable success. While uncompressed standard definition video has a data rate of 270 Mb/s (megabits per second), the original MPEG method (MPEG-1) produces images of approximately the same quality as from a domestic VHS videotape recording, at a data rate of 1.2 Mb/s. That's a data reduction by a factor of 225. The latest versions of MPEG (MPEG-4) can produce VHS quality at 0.5 Mb/s = 500 Kb/s (kilo bits/sec). Most domestic internet connections of the 'broadband' type offer at least 1 Mb/s of bandwidth, making it possible to view VHS quality images directly from the internet. Without data reduction, a one-minute video sent uncompressed would take roughly 270 minutes to download (on a 1 Mb/s broadband link) – and most people would not wish to spend 4½ hours to get a one-minute video.

The following table gives common compression types, their resultant approximate image quality and common uses. Most types of compression algorithm can operate over a wide range, with a equivalent wide range of output data rates. So the 'Data rate' column just gives typical values.

Table 2- Compression rates and qualities

Compression Type	Data rate, Mb/s	Quality	Comment
No compression	270	studio	ITU-R BT601, standard definition TV
Lossless JPEG2000	90	studio	ITU-R BT601, standard definition TV
Lossy JPEG2000	90	studio	Used in frame-based devices such as moving image cameras.
MPEG-1	1.2	VHS	Wide internet use
MPEG-2	5	DVD	Used on DVD and digital TV broadcasting (DVB)
MPEG-4	0.5	VHS	Will replace earlier MPEGs
MPEG-4 AVC	8	HDTV	Will be used on HD DVDs, and possibly on HD TV
DVX	0.5	Near VHS	Wide internet use

[44] For instance, this review from TASI: www.tasi.ac.uk/advice/creating/fformat.html
[45] www.jpeg.org/
[46] www.chiariglione.org/mpeg/

Digibeta	80	studio	Nearly full quality
DV, DVCAM	25	Pro-sumer	Pictures near digibeta quality, quality suffers on repeated decode-encode
DVC-PRO 50	50	Near studio	Pictures near digibeta quality, quality suffers on repeated decode-encode

Compression generally looks attractive because of the reduction in file sizes achieved, and is essential for Internet access and most end-user delivery situations. However lossy compression causes problems for archiving and preservation:

- **Not full quality:** the pictures may look all right, but if full quality material is available, then a (lossy) compressed version is by definition a reduced form and should be considered a proxy, not a true substitute for the original.
- **Reliance on decoder:** clever coding requires clever decoding. Proprietary decoders come and go. Encoding adds a risk to managing the data.
- **Problems with manipulation:** typically tools for manipulation moving images work with the decoded data, although some video editors support some 'native formats'. In general, doing an edit requires decoding, and for web or DVD access the data must then be again encoded. Each cycle of decode-encode produces a loss in signal quality. Many forms of highly compressed coding produce visible flaws after two or three such cycles, whereas digibeta at 80 Mb/s is usually good for ten to twenty such cycles. Lossless coding supports an infinite number of edit cycles!
- **Problems with format migration:** when a compression format becomes obsolete, conversion to a new format will almost certainly require decompression, and possible require re-encoding in the new target format (unless a lossless format is chosen). This decode-encode is the same as an edit cycle. Hence use of lossy compression can severely limit the future life of the content!

e) Playback and Replication

There are three playback problems, which vary according to the original source of the moving images**:**

- **Analogue media:** analogue videotapes (VHS, BetaSP, U-Matic, various domestic formats and so forth) each have their required videotape player. If you don't have the machine, you can't play the tape. Given a player, the output will be analogue (and probably composite except for professional players). Digital preservation will require capture by an appropriate video capture card.
- **Digital media:** digital videotapes (such as digibeta or DV) and DVDs also need their own players, although video DVDs will also play (given the right software) on computer DVD drives. The digital signal from a digibeta videotape player will need a special capture card, whereas DV camcorders are able to connect directly to a computer via firewire or usb-2 for capture and storage. A DVD can be 'cloned' to a blank DVD (unless digital rights management intervenes!), or the files can be read and saved to mass storage. DVD files use MPEG-2 compression and need DVD-player software, which could cause problems once the current DVD format is obsolete. Of course, a standard DVD player has an analogue output (for connection to a TV), but it would be very bad practice to digitise that output for preservation purposes (not capturing native format; going digital to analogue to digital again; decode and encode cycle).
- **Computer files:** the main playback problem for files is format and software obsolescence. It's the same issue as for word-processor and other document

formats: they don't last. Encoded (compressed) images can be kept as they are and simply migrated to new mass storage for preservation, but the files will not play without the software, which may in turn depend upon particular operating systems or web browsers. Even computer filing systems change and become obsolete, but that's a standard digital preservation issue, and not specific to audiovisual data. Migration to a new format (as mentioned above: **Problems with format migration)** involves loss of quality if lossy compression was used.

5.2.5 Recommendations for file formats for preservation of digital moving images

Preservation of digital moving images will require, for the foreseeable future, adoption of a migration approach. However, file format migration for moving images is difficult. It is impractical to rely on obsolete coding formats and software, so migration is necessary. But if lossy compression is involved, signal quality is affected by each migration, and may cause visible deterioration after as few as two migrations. It is for this reason that institutions which take the long view, such as the BBC, Library of Congress, and Vienna Phonogrammarchiv, are using uncompressed or lossless compression for digital preservation of video. In general, moving images should be saved in an uncompressed or lossless format. This is particularly important when the resources to be preserved are of studio quality, or original, or unique. Dissemination / distribution versions can be in lossy formats. Compression is preferred in fact because of the great reduction in bandwidth needed for transmission.

There are two separate issues in moving image preservation, those of bitstream encoding and file format. For bitstream encoding the following requirements are recommended:

- Larger picture size preferred over smaller picture size. Picture size is expressed as horizontal lines and number of samples per line, or as horizontal and vertical pixel counts.
- Content from high definition sources preferred over content from standard definition, assuming picture size is equal or greater.
- Encodings that maintain frame integrity preferred over formats that use temporal compression.
- Uncompressed or lossless compression preferred over lossy compression.
- Higher bit rate (often expressed as *megabits per second*) preferred over lower for same lossy compression scheme.
- Extended dynamic range (scene brightnesses) preferred over "normal" dynamic range for such items as Digital Cinema or scanned motion picture film.
- Surround sound encoding only necessary if essential to creator's intent. In other cases, stereo or monaural sound is preferred.

File format preservation presents its own, different, set of challenges and issues. The following process steps are recommended:

- decide on the level of reproduction quality (ie. post-preservation access and delivery) necessary for various classes of digital resources. Some classes of content will be suitable for lower quality; some content will be of sufficient significance that the highest quality reproduction will be warranted
- the resultant benefits of using highest quality reproduction might not be apparent until subsequent migrations. It is nevertheless true that highest quality reproduction will reduce the cumulative effect of visual artefacts due to lossy compression methods
- highest quality reproduction will derive from capture of full or wide colour gamut and high bit depth in a frame-based mode. The highest quality frame images may be uncompressed or losslessly compressed using encodings such as JPEG2000

- reduced quality outputs from the preceding approach can be graduated or scaled to a particular situation. For example, a frame based mode may be used with lossy compression as is done for the digital cinema package. Modest reduction in quality would result from the use of lossy JPEG2000, or "all I-frame" MPEG-2, or H.264
- some additional compression efficiency with consequent quality reduction may be achieved by using lossy compression that includes temporal elements, e.g. the typical implementations of MPEG-2, H.263, or H.264. Quality settings for these encodings are determined by a number of factors; interested readers should familiarize themselves with the various profiles that have been defined for these formats. (It may also be reasonable to consider customary professional video formats like DVCPRO 50, or DVCPRO 25.)
- for frame based reproductions, there are a variety of approaches available for file formatting and/or wrapping, e.g. MXF, Motion JPEG2000, DPX, and more.
- for compression that includes temporal elements, file or wrapper formats such as those specified for MPEG-4 or the ad hoc MPEG-2 may be selected. As the MXF format comes into wider use it may be a good wrapper choice
- at this time (2006) the tools needed to construct files like those named above are not widely available. As a practical matter curators may be forced to select an approach based upon available technology. For professional quality work, the decisions made by broadcasters and digital cinema producers a re likely to be very influential in technology development.[47]

At this time, it is not possible to make a recommendation for digital broadcast television streams. The industry is still in a state of flux and various formats are vying for widespread adoption globally.

It is recommended that JISC undertake further investigation and analysis of the state of digital broadcasting formats and technologies as a matter of urgency.

5.3 Digital Sound

5.3.1 Overview and significant properties

Sound is best described as rapid changes in air pressure (vibrations), or as longitudinal waves[48] that propagate through a medium (such as air). We hear a sound because the changes in air pressure (or the longitudinal waves) interact with our eardrums which vibrate in response, and the vibrations are interpreted by the brain as sound. Sound waves are usually measured in terms of their frequency and amplitude, although they have other characteristics which affect the sound. Frequency is a measurement, in Hertz (Hz), of the number of vibrations that occur per second. Optimally, people can hear from 20Hz to 20000Hz (20kHz) although this upper limit decreases with age. Amplitude is a measure of how many air molecules are affected by the vibration of a sound (see Figure 1.). In analogue audio equipment sound waveforms are represented by voltage variations.

[47] The authors are greatly indebted to Carl Fleischhauer from the Library of Congress for the suggestions and wording of this entire section.
[48] Longitudinal waves are waves that have vibrations along or parallel to their direction of travel.

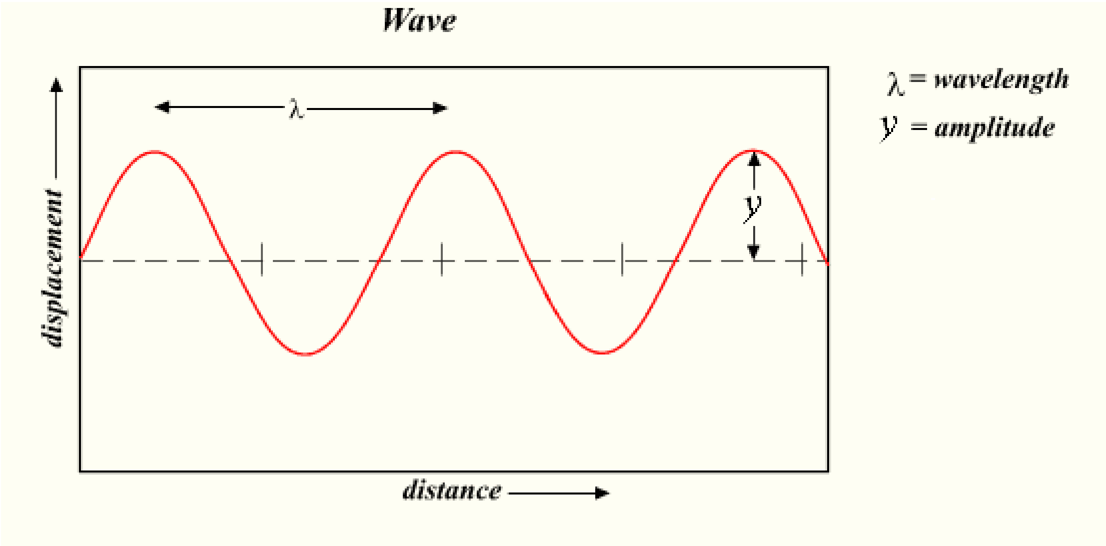


Figure 5 Amplitude and wavelength

Essentially digital sound consists of a digital reproduction of the vibrating waves that make up sound. Digital sound files can be created entirely within the digital domain, or they can be created through digitisation of existing analogue sound material. But however the sound is created, it can only be experienced by us as an analogue signal, not a digital one. Digital sound represents waveforms as a series of samples, taken at specific time intervals, whose values are given as binary numbers. The sample rate is the number of times per second that the analogue signal is measured. The sampling rate influences the quality of digital sound. The standard encoding (sampling method) for digital audio and digital video is Pulse Code Modulation (PCM).

The difference between analogue audio signals and digital sound is best shown by the following diagram:

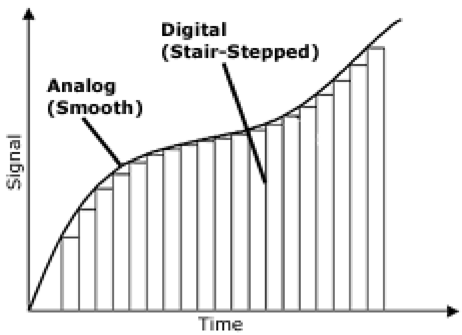


Figure 6 Quantisation of analogue sound

Clearly, the greater the sampling rate the smaller the size of each ‘stair’ and the closer the digital sample is to the original analogue waveform.

The conversion of analogue data to digital and back to analogue is accomplished by special electronic circuits nowadays fitted onto integrated chips. A chip that converts analogue to digital is called an ADC- an Analogue to Digital Converter. The ADC measures the amount of voltage at each sampling interval and converts it to a binary number. This is the process by which analogue sound is digitised. At the other end of the process is a chip called a DAC- a Digital to Analogue Converter. This chip takes the binary numbers which represent the sampled sound and converts it to an output voltage which can then be transmitted through audio speakers as analogue waveforms which our ears can hear.

Significant properties

The most significant characteristics of digital audio files are:

- bit-depth (bits per sample);
- sampling rate (kHz);
- compression method (codec);
- number of channels (mono, stereo, or higher)
- digital coding method.

Bit Depth (aka 'resolution')
The term "bit-depth" refers to how many bits (1s and 0s) are used in each sample to represent the digital signal. The bit-depth for digital audio is expressed in figures of 8-bits, 16-bits, 24-bits etc., and correlates directly to the file size and sound quality. Bit depth determines the accuracy of digitisation and the dynamic range of the audio. Lower bit-depth results in smaller file sizes but poorer sound quality.Higher bit-depth results in better quality but larger files. CD audio uses 16 bits, which provides a range of binary values for representing the size of the waveform from 0 to 65,534 (2^{16}). This gives a theoretical maximum dynamic range of 96dB (decibels). Higher quality digital audio can be achieved by using 24-bit signals which give a theoretical maximum of around 144dB, although in practice a dynamic range of 110dB is about the maximum possible.

Sampling Rate
The sampling rate is how many times per second the voltage of the analogue signal is measured. Sampling rate has a direct effect on sound quality. CD audio is sampled at a rate of 44,100 times per second (44.1 kHz). DAT (Digital Audio Tape) supports sampling rates of 32, 44.1 and 48 kHz. Sampling rates of 96kHz and 192kHz are becoming more common with increased use of 24-bit signals, although there is a lot of dispute about the value of using these higher rates. The dispute concerns whether or not analogue audio signals contain enough information to justify such high sampling rates.

Codec (encoder/decoder)
Because of the amount of data needed to represent a sound file it is most common for digital audio to be compressed in some fashion for end-user delivery. This reduces file sizes and enables the transmission of sound over the Internet, where bandwidth and download quantities for most people are limited. The issue with most compression methods is loss of quality or robustness of the signal – hence the term 'lossy' compression. The codec refers to the method (usually software) of compressing and de-compressing the source files. Without the correct codec a compressed file is not playable so codec information (if applicable) is crucial to accessing digital sound files. Not all compression methods are lossy, and use of lossless compression such as in the Free Lossless Audio Codec (FLAC) is an important consideration in digital sound preservation. For preservation purposes, uncompressed files should be used whenever possible, in which case no codec is used.

Number of channels (mono, stereo, or higher)
The number of channels greatly influences the size of the audio file. Stereo files are generally twice the size of mono, while surround sound, which contains 5 channels is much larger still. Channel information is necessary to correctly reproduce digital audio, although this is in turn affected by the playback equipment available.

5.3.2. Common audio formats

(Note: Since wrapper formats are usually methods for encapsulating both audio and video files and bitstreams they are discussed in the preceding section [5.2]. That information will not be repeated here.)

Advanced Audio Coding (.aac; .m4a)

An open standard developed by the Motion Picture Experts Group (MPEG) as a high quality encoding for audio streams, originally part of the MPEG-2 standard. AAC uses perceptual audio encoding and lossy compression methods. The AAC compression approach adopted as a part of MPEG-2 has been further refined within MPEG-4, where the format is referred to as M4A. MPEG hopes that AAC/M4A will eventually replace MP3. Widely used as the standard audio format in Apple's iTunes service. Bitstreams encoded with AAC can be wrapped in QuickTime and other file and wrapper formats.

Audio Interchange File Format (.aif; .aiff)

An audio file format developed by Apple for wrapping high quality audio streams, primarily for interchange but is flexible enough to be used as a storage format. The audio streams within AIFF files are commonly encoded with linear PCM (see below) but AIFF can wrap many other audio stream formats, from uncompressed waveforms to MIDI. Audio is of high quality although file sizes are large because the bitstreams are uncompressed. Functionally equivalent to the WAVE wrapper format.

Compact Disc Audio (CDDA)

The standard specification for audio Compact Disc media (CD), often referred to as the 'red Book' standard. The standard specifies the use of 2-channel signed 16-bit pulse code modulation (PCM) sampled at 44100 Hz for audio bitstreams. Audio in this format cannot be directly stored on computer hard disks but needs to be converted to a file format first ('ripped').

Digital Audio Compression (AC-3; Dolby Digital)

A lossy format for the encoding of 'surround sound' using lossy compression, developed by Dolby Laboratories principally for cinemas and 'home theatres' and often known as 'Dolby Digital'. AC-3 was standardised by the US Advanced Televisions Systems Committee as standard A/52. AC-3 is widely used in home cinema systems, commercial cinemas, DVDs, and has now been adopted as the audio component for High Definition Television (HDTV).

Extensible Music Format (.xmf)

An audio wrapper format and a number of *combo-packs* developed by the MIDI Manufacturers' Association. An XMF file can contain audio encoded as standard MIDI Files, DLS (downloadable sound) instrument files, WAV or other digital audio files. The purpose of the XMF format is encapsulate all the resources needed to present a musical piece, an interactive webpage soundtrack, or any other piece of media using pre-produced sound elements.

Free Lossless Audio Codec (FLAC)

A popular open format for lossless compression of PCM bitstreams. The FLAC specification includes a wrapper for FLAC bitstreams plus the lossless compression codec itself. FLAC can be used as an end-user delivery format and is also a suitable format for archival storage of audio streams.

Linear Pulse Code Modulated Audio (LPCM)

Pulse code modulation (PCM) with linear quantization. A bitstream encoding for digital audio streams. Linear PCM is essentially an uncompressed format resulting from sampling of the source audio. Widely used on audio CDs, digital audio tape, and is the default encoding format for audio streams in AIFF and WAVE wrapper formats. It is unusual to find LCPM files existing independently, ie. outside wrapper formats.

MPEG-1 Layer –3 (.mp3)

A compressed format developed as the encoding for audio streams within MPEG-1 files. File sizes vary depending on sampling and bit rate. Typical compression of 10:1. Widely used across the Internet, for non-commercial copying and recording of audio files to CD, and interchange of audio files. (see also section 5.2.2)

Music Module formats (MODS)

A family of file formats created by a variety of software applications which are known as 'trackers' (e.g. *Soundtracker*, the first such application). Module formats store musical data in a form similar to a spreadsheet. A MOD file contains a set of instrument data in the form of samples, a number of patterns indicating how and when the samples are to be played, and a list of what patterns to play in what order. The format allows for up to 32 channels of music playback and 31 instruments can be represented. There are many types of MODS formats, each with their own file extensions e.g. 669, .amf, .dmf, .far, .it, .med, .mod, etc. Like MIDI messages MOD files are not waveform based.

Ogg Vorbis

An open audio encoding format designed for efficient streaming and lossy compression of audio streams, associated with the Ogg wrapper format. The term "ogg" is often used to refer just to the audio file format Ogg Vorbis, ie. Vorbis-encoded audio in an Ogg container. Quite widely adopted as a format for sound in more recent computer games, and by the open source community more generally. Ogg Vorbis can be used within other wrapper formats. (See section 5.2.2 above for more information on Ogg formats)

Real Audio (.ra; .rm; .ram)

A family of proprietary formats developed by RealNetworks for audio encoding. The Real Audio file format provides for a number of lossy codecs, and includes one for lossless compression. One of the most widely used streaming formats in web applications, particularly by Internet Radio Stations for streaming their audio signals. Compresses sound up to 10:1. Sound quality is passable, but not high.

Standard Musical Instrument Digital Interface (MIDI) File (.smf; .mid)

A wrapper format for the MIDI bitstream encoding specification (MIDI sequence data) used to support the dynamic construction of sound through combinations of hardware and software. A MIDI "messages" is a set of instructions that tell a digital 'instrument' (e.g. a synthesiser) how to play a piece of music. MIDI is not based on waveform (ie. the shape and form of the audio signal) unlike most other encoding methods. The instructions in a MIDI message are real time digital data which provide information about such aspects of a musical performance as the type and intensity of the musical notes, and technical cues. Note that the XMF version of MIDI (DLS) may be accompanied by waveform samples. MIDI was widely adopted but requires some form of general MIDI player to reproduce the sounds. MIDI 'messages' may be wrapped in SMF (standard MIDI file format) or XMF (extensible music format) wrappers.

SUN Audio (.au, .snd)

An older format developed by SUN Microsystems and still used on Unix systems. Specifies an arbitrary sampling rate. Can contain 8, 16, 24 & 32 bit data and uncompressed or lossy compressed encodings. In comparison to other 8 bit samples it has a larger dynamic range. Slow decompression rates because of large file size, but audio is of high quality.

Wave (.wav)

A proprietary Microsoft and IBM audio wrapper format that is the standard for storing audio on Windows PCs. WAVE is a subtype of the RIFF meta-format so is similar to Apple's AIFF format. The most common encoding format used with WAVE files is PCM encoding but a WAVE file can contain bitstreams encoded with other codecs, lossy and lossless A very popular format for file sharing over the Internet but limited to files less than 4GB in size. WAVE is the basis for the European Broadcast Union's Broadcast Wave (BWF) format, a standard and open format for the exchange of audio files between different broadcast and computer systems. BWF is the audio format used by most file-based non-linear digital recorders used for motion picture and television production, and has become the de facto standard for audio preservation.

Windows Media Audio format (.wma)
A proprietary file format for wrapping audio streams, developed by Microsoft to compete with MP3. WMA is based on the meta-format ASF and uses one of the Windows Media Audio codecs for encoding audio data. WMA files are often encapsulated within an ASF wrapper file. Used widely for streaming audio and also as a format for high-quality audio archiving.

5.3.3 Issues in archiving and preserving digital sound

Advantages of Digital Audio

There have been ongoing debates for years concerning the merits of digital audio versus high-end analogue systems. Some listeners question whether digital sound quality is quite as good as analogue sound, but it can be very good indeed, and most people are unable to tell the difference in sound quality. The obvious advantages of digital audio are discussed below, briefly.

Wider dynamic range
Digital sound recorded at a bit rate of 16 bits has a theoretical maximum dynamic range of 96dB, compared to the dynamic range of most analogue systems of around 80dB. Higher bit rates provide higher theoretical maximum dynamic ranges, although the practical limit is about 110dB, reached with 24 bit sound. Actual dynamic range is limited to whatever component in the record / reproduce / digitise chain is the 'weakest link' – often this will be the source material itself.

Less equipment noise;
Analogue sound equipment introduces electromagnetic frequency interference as audio signals travel through the physical circuits. This sort of signal degradation is unlikely in digital systems.

Faithful copying/reproduction
Every time analogue sound is copied data losses occur and noise is introduced (see above). Over analogue sound generations quality decays noticeably. These copying/reproduction problems do not occur in properly used digital audio systems.

Error correction
Error correction in digital audio playback equipment is both a blessing and a curse. CD-DA and DVD players both contain error correcting mechanisms to allow uninterrupted audio (and video) playback in spite of bit level errors. Unfortunately, this makes playback of such media on such equipment an unreliable method of determining integrity of the digital signals. Indeed by the time errors can be heard the bitstream has decayed to such an extent that recovery through re-copying is impossible.

Issues with Digital Audio

Quantization error

Quantization in digital audio refers to the sampling representation of the analogue waveform by integers, the range of possible integer values being denoted by the bit level of the sampling. Thus, the 16-bit sampling of audio CDs allows for a possible 65,536 integer values (2^{16}). Quantization errors are introduced in digital audio because the waveform sampling of analogue signals needs to represent the value of each sample as a whole number (integer). The converter selects a whole number that is closest to the signal level at the instant of sampling. This produces small rounding errors in the sampling that are equivalent to a noise level. Without “dithering” – see next section – the noise may be in a narrow frequency range, making it more likely to be audible.

Dithering

Dithering is the introduction of 'erroneous' signal or random noise which is added to an audio signal for the purpose of minimizing the effects quantization error, by spreading the quantisation noise across the whole range of frequencies in the signal. Dither is routinely used in processing of both digital audio and digital video data. Many ‘audiophiles’ are horrified at the notion of introducing deliberate errors in digital audio signals and believe that this lowers the sound quality of digital audio. In practice, most human ears are unable to tell the difference between analogue and digital sound, although they could hear the loss in quality of digital sound that had not been dithered.

Compression and codecs

As discussed above in the section on moving images, a codec is a software device that is used to encode a bitstream – in this case an audio bitstream. The algorithm used in the encoding (and decoding) software is referred to as a codec—as in *cod*ing/*dec*oding (see discussion above)*.* There is often more than one codec for a particular format, and different codecs can vary widely in quality and speed, even for the same format. A few codecs do not compress the data stream (LCPM, AIFF, RIFF) but most codecs employ some form of compression, either lossless or lossy. Compression is usually necessary in digital audio for end-user delivery because digitisation of analogue sound or creation of digital sound files quickly produces files of very large size. File size is related to sampling rate, bit rate, number of channels, and time. 3 minutes of sound sampled at 44.1kHz, using 16-bits and in stereo (the most frequently used quality) will result in a file size of around 31 Megabytes (uncompressed), or approximately 10Mb per minute of sound. Encoding methods using compression are not generally suitable for preservation and should be avoided.

Lossless compression

Lossless compression for audio files is a specialised area, as the sorts of lossless techniques used for text and image files do not work with sound. The highest compression ratios for lossless compression of audio are usually about 2:1, ie. an uncompressed file of 30 Mb would be compressed using a lossless method, such as the Free Lossless Audio Codec (FLAC), to about 15Mb. A recent Presto study of lossless audio compression, using several codecs, found it was possible to achieve compression ratios of around 3:1. Other lossless formats include: Apple Lossless, Audio Lossless Coding (ALC - MPEG-4 ALS), Lossless Transform Audio Compression (LTAC), True Audio (TTA), and Windows Media Audio 9 Lossless.

Lossy compression

Lossy compression removes unnecessary or redundant information completely, so will always lead to a loss of quality, although the differences may be undetectable to the human ear. Lossy compression ratios of up to 10:1 result in generally insignificant quality loss, although many people can tell the difference between an MP3 file at 128 bits (a compression of about 10:1) and original CD audio. MP3, for example, compresses files up to about 11:1 without noticeable loss of quality, so a 30Mb file will be reduced to around 3 Mb using MP3. Newer compressed formats, such as AAC, can achieve similar or better compression ratios with much less loss of quality. There are many lossy encoding formats including: Advanced audio Coding (AAC – MPEG-2 and MPEG-4), Dolby Digital, MP 1 2 3, Vorbis, and Windows Media Audio (prior to version 9), but generally these should never be used in professional archiving situations,

Error correction
A problem, particularly in CD and DVD playback devices, because it disguises or hides the existence of missing or corrupted data until it is too late to rectify the problem (see above). Useful for end-user delivery because it restores missing bits and thereby ensures accurate playback of the signal.

See also the discussion of these issues in the section on Moving Image formats.

5.3.4 Recommendations for file formats for preservation of digital sound

In the best case, we would have raw unedited master source files or sound streams to preserve, but it is unlikely that such objects will be offered for preservation in most cases. Digital audio is the product of particular creation or digitisation processes, and the objects available for preservation are dependent to a large extent on the purpose for which the digital audio was created, the specific processes involved in its creation, and end user requirements. As with moving image preservation, there are two issues to consider: bitstream encoding and file format.

All other things being equal, the best object to preserve is the one using an encoding that retains the highest bit depth and sampling rate, and with no compression. Thus, we would preserve a sound stream in LPCM format in preference to the same data stream encoded in MP3 format. . Thus, for bitstream encoding the following requirements are recommended:

- Higher sampling rate (ie., 96kHz or higher) preferred over lower sampling rate (ie. 48kHz or lower).
- 24-bit sample word-length preferred over shorter
- Linear PCM (uncompressed) preferred over compressed (lossy or lossless)
- Higher data rate (e.g. 128 kilobits per second) preferred over lower data rate for same compression scheme and sampling rate.
- AAC compression preferred over MPEG-layer 2 (MP3) when no other options are available
- Uncompressed encoding in stereo is preferred.

It is recommended that where possible digital audio be saved either in broadcast wave format without compression, and using LCPM encoding, or as AIFF files, again using LCPM encoding.

Chapter 6: Digital Preservation Methods: Moving Images and Sound

6.1 Overview and assessment of preservation methods

The lack of practical advice, robust tools and mature technologies is a great difficulty facing anyone trying to undertake digital preservation. A number of digital preservation strategies have been proposed, but there is no definitive approach to the problem of maintaining digital content across multiple generations of technology. In addition, information on the likely costs, possible limitations and long-term sustainability of different strategies is still quite limited.

Few organisations with digital preservation responsibilities appear to have yet fully developed their policies (ERPANET, 2003). It is unwise to commit to a course of action before its consequences are clear. It is equally important that organisations which make some claim to preserve digital resources should declare to their stakeholders what they can do to achieve this goal at the present time. Decisions about preservation methods might usefully take into account the following three-tiered levels:

1. Preservation of the bit stream (basic sequences of binary digits) that ultimately represent the information stored in any digital resource;
2. Preservation of the *information content* (words, images, sounds etc.) stored as bits and defined by a logical data model, embodied in a file or media format;
3. Preservation of the *experience* (speed, layout, display device, input device characteristics etc.) of interacting with the information content.

Techniques for achieving the first of these objectives are well understood and include environmentally controlled storage, data replication, backup, and media refreshment. In the OAIS model, much of this activity falls into the archival storage function. The second and third objectives present a far greater challenge.

Binary data remains useful only for as long as it can be correctly *rendered* (displayed, played-back, interacted with) into meaningful content such as sound files and video clips. The process of rendering is performed by a complex mix of hardware and software, which is subject to rapid obsolescence. As a rule of thumb, it is reasonable to predict that current hardware and software will be able to correctly render a file for around ten years after its creation. By the end of this period, repositories need to have adopted a more active preservation strategy than simply preserving the bit stream of the file if they are to maintain access to information content held in the file. Either old data must be altered to operate in a new technical environment (migration, format standardisation) or the new environment must be modified so that it can render the old data (emulation, virtual computers). Within these two broad approaches there are many different techniques (figure 1).

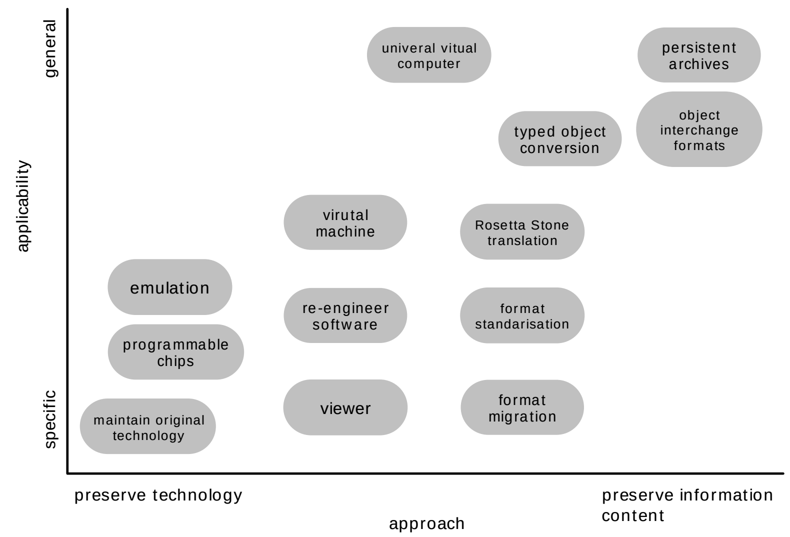


Figure 1 Digital Preservation Strategies (based on Thibodeau, 2002)[49]

This chapter reviews some of these techniques and makes recommendations for approaches to take for the preservation of both moving images and sound, and the preferred formats for preservation. It starts with a look at bitstream preservation and then spends some time looking at the preservation of content and the experience.

6.1.1 Bitstream preservation

This simply involves retaining the original data as a single sequence of binary digits (bits), ie. the original data in an uninterpreted state. Bitstreams may be preserved in 2 ways: either as the original file in the data format as received, e.g. an MP3 file, a stream of PCM encoded video or audio, etc.; or in a normalised bitstream format, e.g. as a sequence of bits contained inside XML wrappers. Bitstream preservation is widely viewed as a form of 'insurance' in that it allows for the possibility of using future techniques for making content accessible. It is an additional form of data backup and is a necessary component of any digital preservation strategy.

It is hard to see that the first method offers any real advantages over the second. In either case metadata is needed to make sense of the file. Using the first method, the metadata would need to be kept separately but associated with the data content, perhaps, for example, in a separate METS file. The second method has the advantage that metadata about the file and its format can be included within the XML wrappers surrounding the content bitstream and is thus always kept with the associated content data[50]. The second method might be preferred for the sake of human understanding of the bitstream content and for the possibilities it offers of creating and maintaining self-documenting digital objects.[51]

6.1.2 Content preservation

6.1.2.i Preservation of technology

Technology preservation is an attempt to ensure the usability of digital resources over time by preserving and maintaining software (applications and operating systems) and hardware,

[49] Thibodeau, K. 2002. 'Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years' in proceedings of *The State of Digital Preservation: An International Perspective*. Conference Proceedings. Washington. 2002.
[50] See H. James et al, *Feasibility and Requirements Study on Preservation of E-Prints*, JISC 2003, 32-33.
[51] See http://www.digitalpreservation.gov/formats/sustain/sustain.shtml#self for a brief description of self-documenting digital objects.

essentially creating an IT museum.[52] This may be a useful short-term solution but costs in terms of the maintenance and storage of equipment makes it prohibitive for any sustained strategy. As an example, AHDS Archaeology maintains a small 'computer museum' (http://ads.ahds.ac.uk/project/museum/) but 'exhibits' are intended to facilitate data recovery, not time-based access to source files, and have been used in earnest for recovery purposes. Data is moved into stable open formats and thus into a format migration strategy. At one stage the museum contained two working Amstrad computers but both of these have now failed and are not able to be repaired. A hardware preservation strategy would fail in the same way, without the expenditure of large sums of money to ensure that technologists with appropriate skills are always available. Although such an approach to preserving digital moving images and sound seems attractive because it would enable the explicit recreation of the original experience, it cannot be regarded as viable for any but the wealthiest organisations.

6.1.2.ii Migration

Migration as a preservation term can be used to describe both file format migration and media migration. Each has its place in a planned preservation strategy.

Media migration is more often known as ‘refreshing’ and is necessary to ensure that data is not lost through media degradation over time. The lifetime of media must be estimated and migration to new media undertaken before the point of no return is reached (ie. the point at which media errors make the data unrecoverable).[53] Today there is a move toward storage on hard disk with multiple disk or tape backups. Integrity between versions is maintained through the use of fixity or checksum values. Refreshing media is an essential activity in any preservation programme, but does not in itself ensure data preservation. Media migration is an essential component of a preservation strategy especially for time-based resources, such as audio and video, stored on physical media like DVD-R discs, CDs, DV tapes, etc.

File format migration is used to ensure the accessibility of a digital object when the software it depends on becomes obsolete or unusable. It can involve conversion of digital objects from one file format to another (not necessarily the same) format, for example from Word 98 to Word 2000, from Word 2000 to Adobe’s Portable Document Format (PDF), or from GIF to PNG. Some attributes of the digital object may be lost during the conversion process, so the experience may not be equivalent after migration. The level of data loss through migration depends on the number of preservation treatments applied to the record, the choice of process, the new data format, the level of human intervention, and post-migration descriptive work.

Migration as an approach has a number of variations. The traditional technique is to migrate file formats to newer versions of the same format as the earlier versions approach complete obsolescence, eg Word 95 to Word 98 to Word 2000. This method requires migration at as late a stage in the life of the digital format as possible. In the world of audiovisual resources this form of migration is useful but may not be the most effective strategy, particularly as audio and video codecs, file formats, and wrapper formats do not evolve in the same manner as other types of software applications. Instead, older codecs and formats are replaced by new codecs or formats, not by upgraded versions of the existing ones. So for example, the replacement format for MP3 is not a newer version of MP3, but the AAC/M4A codec. This common approach can be applied to moving image and sound resources, and requires organisations to purchase or otherwise acquire the full range of software applications for the formats they hold. The software would need to be maintained often past the date of its

[52] T. Hendly, “Comparison of Methods and Costs of Digital Preservation”, *British Library Research and Innovation Report 106*, 1998, 16-17.
[53] T. Hendly, 1998, 12.

commercial obsolescence, and new and updated versions of applications would have to be acquired.

A second migration technique involves migrating digital objects to limited range of standard formats at the time they are ingested into a digital repository, or earlier if possible. This involves fewer overall migrations but still requires the migration of the standard formats as existing versions of those formats become obsolete. This approach has been adopted by the Library of Congress as well as other institutions such as AHDS, and is possibly the most effective strategy for moving image and sound files at present. A variant of this approach involves the migration of all file formats to a standardised file format which is chosen for its presumed longevity as a digital format, eg XML . This migration technique is not yet widely used but has been adopted by the National Archives of Australia where it is referred to as 'normalisation'. It too may have a place in preservation strategies but depends to some extent on the availability of tools or applications that can 'normalise' source formats to the preservation format, so is not suitable for all preservation institutions. Since not many moving image and sound applications support the full range of available formats as export formats this approach does have obstacles that could prevent its successful implementation.

6.1.2.iii Migration on request

Migration on request is an interesting variation on the migration approach. Instead of migrating original objects to normalised formats on ingest, or migrating objects once formats become obsolete, this approach retains objects in their original format and develops tools to migrate the object to a current format when access to the object is required.

Originally conceived by the CEDARS project, migration on request was further developed by the JISC-funded CAMiLEON project.[54] One of its chief developers says of CEDARS: "Original objects are maintained and preserved in addition to a *migration* tool which runs on a current computing platform. This would be employed by users to convert the original bitstream of the object they want to use into a current format".[55] A separate migration tool would be needed for each data format ingested by the repository although the concept envisages the development of tools that would be able to handle multiple (similar) input formats through a modular approach. These migration tools would need to be maintained over time, ie. re-coded to cope with operating system and hardware changes. Thus, the original objects are maintained in their original formats and the migration tools are migrated as computer platforms change. Alternatively, emulators could be developed so that the initial migration tools can be operated on newer hardware and operating systems, and only the emulators would need to be migrated.

With this approach the focus of preservation moves from the object itself to the migration on request tool. Since an archive has control over the development of the tool it can be designed to last. In theory implementation of the migration on request approach over time has a lot less development redundancy than other approaches and should be cheaper to implement. The migration on request approach is a promising avenue and should certainly be considered a viable approach to preserving digital still images.[56]

6.1.2.iv Emulation

Emulation is a technique often proposed as a solution to the hardware/software obsolescence problem. Put simply, emulation involves the development of software to replicate the behaviour of obsolete processes (such as hardware configurations or operating

---

[54] See http://www.leeds.ac.uk/cedars/ and . http://www.si.umich.edu/CAMILEON/.

[55] P. Wheatley, "Migration - a CAMiLEON discussion paper", *Ariadne* vol. 29 2001, available at http://www.ariadne.ac.uk/issue29/camileon/.

[56] A number of proofs-of-concept migration tools were developed by the CAMiLEON project to demonstrate the viability if the approach. See: http://www.leeds.ac.uk/reprend/migreq/migreq.html.

systems) on current hardware. Emulation thus aims to recreate part of the original process that interprets the data to produce a modern rendering of the original performance. Much emulation work is motivated by a belief that the original 'look and feel' of a digital resource must be maintained forever. 'Look and feel' includes not only the content of the resource, be it a moving image or a sound file, but also tangible aspects of the presentation of the content, such as colour, layout, and functionality. Emulation is, at first glance, an attractive approach for preserving digital moving images in particular, since it would allow the recreation of the original experience intended by the creators of the resources. However, it has been pointed out by a number of commentators that it may simply be unfeasible to preserve digital objects unchanged over time.[57]

One of the other main concerns with emulation is that it requires the retention of at least one relevant software application for every type of file format being preserved. If the relevant software applications are able to deal with more than one file format (which is often the case) preservation of software applications will be less onerous. However, it must also be noted that this approach requires the reliable preservation of operating systems as well as software applications. As well, the whole issue of the significance of the original 'look and feel' is a discussion largely avoided by proponents of the emulation approach, although it is assumed to be of prime importance as the rationale for emulation-based approaches. Many people state that recreation of the original experience is of prime importance for digital preservation yet there is little or no quantitative evidence to back this up, and the position is neither a given nor unarguable. This is an important underpinning and justification of the emulation approach that deserves to be debated more widely. It is our view that emulation cannot be recommended as a viable preservation approach at present.

6.1.2.v Universal virtual computer

A development related to emulation as a preservation strategy is the so-called 'universal virtual computer' (UVC). This concept was proposed in 2000 by Raymond Lorie of IBM, in research paper written for IBM, and later published more widely in an article in *RLG DigiNews*.[58] In brief, the UVC is a virtual representation of a simplified computer that will run on any existing hardware platform. Its appeal appears to lie in the fact that problems of hardware and software obsolescence become irrelevant, and digital objects can be retained in their original format. It is said by its proponents to have the advantages of both the emulation approach and the format migration approach, with none of the disadvantages.

The only use of this concept to date is at the Netherlands Koninklijke Bibliotheek (KB) where a test implementation has been developed for preserving digital images (in fact PDF files, each of which is manifested as a sequence of JPEG files).[59] In this, the only real world application of the concept, it becomes clear that there is a large software development and maintenance load on implementors. In order to preserve JPEG images today the KB needs the UVC emulator, a format decoder (and will need one for each format being preserved in the future), a logical data scheme, an equivalent of a document type definition or DTD, (again it will need one for each format), and a viewer that allows viewing of the decoded file. To access the file in the future will also require development of a UVC emulator for every hardware/software configuration on which the file will be accessed throughout its useable life.

---

[57] E.g., "traditionally, preserving things meant keeping them unchanged; however…if we hold on to digital information without modifications, accessing the information will become increasingly more difficult, if not impossible." Su-Shing Chen, "The Paradox of Preservation", *Computer*, March, 2001, pp. 2-6.
[Professor Department of Computer Information Science & Engineering, University of Florida-Gainesville]

[58] R.A.. Lorie, "A Project on Preservation of Digital Data", *RLG DigiNews*, vol. 5 no. 3. At http://www.rlg.org/legacy/preserv/diginews/diginews5-3.html.

[59] See http://www.kb.nl/hrd/dd/dd_onderzoek/uvc_voor_images-en.html.

The approach may have some worth, and the KB test implementation is regarded by them as demonstrating "the only method so far that guarantees long-term accessibility of PDF files". However, this is a sweeping statement which is not based on their own experience and which, in any case, can hardly be shown to be accurate except by the passage of time (and it is only a year since the original project finished). The KB implementation itself shows the problems of the UVC approach. It was set up to as a solution for preserving PDF files, but both KB and IBM found that developing a decoder for PDF was too difficult and complex a task to be completed during the life of the project. Instead, the PDF files were migrated to JPEG - a jpeg image for every individual page in each PDF file – and the JPEG files are preserved as the preservation version of the PDF files. The KB admits that "some of the original aspects of PDF publications are lost when using this method", which would seem to compromise the aim of the project "to preserve the original object".[60]

As well, the comparisons of this approach with migration in the KB document display a shortcoming often encountered in advocates of emulation or the UVC approach: an inability to recognise that migration is not a single technique. when it comes to making comparisons between migration and emulation[61] It is our opinion that the software development burden, as well as issues around the complexity of developing decoders mitigates against the use of this approach in all except the most well-off or financially secure organisations, [62] or those with an extremely limited (and relatively simple) range of file formats being ingested.[63]

Of late, the KB has begun an interesting project testing a number of migration techniques as an approach to digital preservation, including migration. This will complement the work previously done on the UVC approach and should enable the KB to develop sounder methods for cost comparisons.[64]

6.2 Assessment of methods for preservation of moving image and sound

Retention of source bitstreams is a necessary component of any preservation approach to safeguard against migration errors and choices of preservation formats which might prove to be incorrect over time. It is also important to bear in mind that preservation formats are not necessarily the same as dissemination / delivery formats (although dissemination / delivery versions of moving images and sound resources may be in the same file format as the preservation copies).

6.2.1 Moving images

Moving images probably present the most complex preservation problem that most institutions will have to face over the next decade or so. There is a full discussion of many of the issues relating to both image and sound preservation in Chapter 5. That discussion will not be repeated hereunder.

The only currently viable approach to preserving digital moving images would seem to be an approach using a form of file migration. As discussed above, current visions of emulation do not seem viable because of the need to continue to develop emulators as well as retain

[60] Ibid.
[61] The web page cited in note 3 assumes that migration is only a technique of continual same format migration, as do other papers by KB staff, and a paper by H.M Gladney in which cost comparisons are made. See H.M. Gladney, *Trustworthy 100-Year Digital Objects: Durable Encoding for When It's Too Late to Ask*, available at: http://eprints.erpanet.org/archive/00000007/01/TDO_Durable_final_submission.pdf.
[62] The KB claim that "no periodic actions are required (unlike migration)" is demonstrably misleading, since the UVC emulator itself will need to be recoded every time the KB changes its hardware platform (or OS).
[63] A single example will suffice to demonstrate the burden such an approach could place on collecting institutions. The George Mason University (Virginia, USA) manages a 9/11 archive which consists of 57,000 digital objects (ca. 13GB). 97% of these objects are in 9 file formats, while the remaining 3% is in 100 different file formats.
[64] See http://www.kb.nl/hrd/dd/dd_projecten/projecten_migratie-en.html.

original software applications. This is only one step better (i.e. no need to retain hardware) than the clearly unviable museum of technology approach. As well, a more serious criticism of emulation has come from David Bearman who comments that “Rothenberg [the original and committed advocate of emulation] is fundamentally trying to preserve the wrong thing by preserving information systems functionality rather than records. As a consequence, the emulation solution would not preserve electronic records as evidence even if it could be made to work and is serious overkill for most electronic documents where preserving evidence is not a requirement”.[65] Likewise, Bearman’s criticism can be levelled at the UVC approach, which would have similar disadvantages for most institutions and organisations responsible for preservation. In addition, the UVC imposes its own development burden which would limit the ability of most organisations to adopt such an approach to digital preservation.

Migration on request may offer another preservation option, but it is an approach whose worth can only be tested over time. As yet no operational preservation service has implemented such an approach. Until this approach has been implemented and tested operationally (as opposed to test situations), it is hard to see wether it can be regarded as viable, despite its apparent advantages.

As recommended in Chapter 5, we believe that the use of uncompressed (motion) JPEG2000 is currently the best strategy for preservation of moving image files. Industry support for a JPEG 20000 frame-based approach is currently represented by the American digital cinema package specification and, for video, by an approach under investigation at the Library of Congress. Both propose the use of MXF as the wrapper format of choice.

Clearly, there is a need for more research into the preservation of moving image files, particularly in the area of television broadcast formats. We recommend the initiation of appropriate research projects into moving image files as a high priority for JISC.

Some principles for preservation of moving images can be specified. In the first instance the best example of a moving image resource should be preserved, when there are choices. For digital audiovisual material, ‘best’ means highest quality, often the version with the highest data rate. Given a digibeta tape and a DVD (of the same content), the digibeta has potentially higher quality. If the copying sequence is known, the copy furthest ‘up the chain’ should be preserved (as being closest to the original). However it may have deteriorated more than a later copy, so the best surviving copy should be preserved, whatever its position in a copying chain. There is a need to make a clear distinction between a preservation master which should use the highest quality content available but which may have little playback functionality, and an access copy which may be short-lived and will undoubtedly be of lower quality.

6.2.2 Sound

The discussion in the previous section regarding moving image preservation is also relevant for preserving digital audio. An approach based on a form of migration is currently the best option. Sound files that exist independently are somewhat less problematic to preserve than moving image files with associated sound. The major issue is the plethora of audio formats that will need to be migrated to the recommended preservation format, broadcast wave. Sound files that are associated with moving image files must retain their linkages to the image files and synchronisation information must be captured in metadata within the wrapper format used.

---

[65] D. Bearman, “Reality and Chimeras in the Preservation of Electronic Records”**,** *D-Lib Magazine*, vol. 5 no. 4 (April 1999). Available at: http://www.dlib.org/dlib/april99/bearman/04bearman.html. David Bearman is a seminal writer and thinker in the cultural heritage domain, specifically in the areas of archives/records and museums.

To preserve digital sound files it is recommended that the following practices be implemented by data curators:

- Use a software tool to ensure the file format, length, bit-rate and sample rate equal those provided in the documentation about the audio data;
- Locate suitable audio conversion software that can handle the source format and is capable of outputting the data as broadcast wave. There are a great many free software packages that can perform this task, such as dbPowerAmp;
- Configure the software to output at the same or higher quality bit-rate (kbps) and sample frequency (kHz) data. Some audio codecs offer higher quality than other codecs at similar bitrates (e.g. the Windows Media Audio format at 128kbps is equivalent to 160kbps for mp3);
- Ensure that associated metadata integrated into the file itself can be extracted in the software tool and stored in the preservation format. If not, the information should be manually output to an ASCII text file;
- Unusual audio formats (e.g. modules) that store audio in a non-sequential manner should be exported or resampled (by connecting the audio-in to the audio-out) and saved at a bit-rate of 1411kbps.

6.3 Recommendations for preservation of moving images and sound

It should be stressed again that the recommendations in this report are to a large extent research, and not practice, based. This qualification applies even more strongly to the information in the tables below. The recommendations are based on abstract assessment since there is little practical experience as yet on which to base more solid recommendations. These recommendations, in particular those for moving image content are still in the process of development and refinement and should be regarded as interim only. What we offer here is our best assessment of viable approaches based on our knowledge and understanding at the current time (late 2006). It is recommended that JISC pursue the development of real life case studies of audio and moving image preservation (including digitisation) projects to help inform the community of issues, processes and decision making in these areas.

6.3.1 Moving images

The following table gives suggested migration paths for standard moving image and sound media types. In all cases, DVDs can be made as access copies, but only for VHS will a DVD be an adequate ‘master copy’ replacement. Table 1 is followed by a second table recommending preservation formats for a number of the common moving image formats currently in use.

Table 1- Migration Paths for Audiovisual Media

Ingest Format	Migration format	Notes
VHS tape	DVD	Perfectly adequate for VHS playback
VHS tape	MPEG-4	Adequate for quality. Future migrations can suffer loss of quality, but VHS is already poor quality. Minimum data rates (MPEG-4): 500k b/s
U-Matic	DVD	Reduces quality; suitable only for viewing

U-Matic	MPEG-2/4	Acceptable if data rate high enough. Minimum 20k b/s (MPEG-2) for 'high-band' U-matic. MPEG-4 at 10k b/s might be adequate. Future migrations can suffer loss of quality (but only after about five of them).
BetaSP, Digibeta, other pro formats	JPEG2000	Ideally lossless compression will be used, with a resultant data rate of around 90M b/s. That is about 45 GB for one hour – meaning 10 DVD-ROMs (or part of one data tape or hard drive).
DV, DVCAM	.avi files	DV will transfer to a computer at 25 M b/s, resulting in an .avi file 'clone' of the original DV tape. These should be migrated to MPEG-2/4 or JPEG2000 for preservation.
DVCPRO50	.avi files	As for DV, but at twice the data rate. Less susceptible to loss on future migrations. Should be migrated to MPEG-2/4 or JPEG2000 for preservation.

Table 2- Migration Paths for moving images

Ingest Format (+ extension)	Notes
Motion JPEG-2000 (.mj2, .mjp2) OR JPEG2000 profiles 3 or 4 (.j2p) [for use with the Digital Cinema Initiative (DCP) specification]	The best preservation format at this time (October 2006). Ensure the use of uncompressed encoding formats for constituent bitstreams. Store unchanged.
MPEG-2 (.mpg2) MPEG-4 (.mp4'n') [where n represents any number]	Suitable for preservation of low quality digital video converted from analogue originals. Otherwise, convert MPEG-2 or MPEG-4 bitstreams to the recommended mjp2/jp2 format using an appropriate software application.
QuickTime (.mov)	Not recommended for preservation except as a wrapper format for encapsulating preferred encoding formats. Convert bitstreams encoded as .mov to the recommended jp2/mjp2 format using an appropriate software application.
Audio Video Interleave (.avi)	Not suitable for preservation, except as a wrapper format. Constituent bitstreams should be encoded using a lossless encoding format, ie. mjp2 or jp2. Convert bitstreams encoded as .avi to the recommended mjp2/jp2 format using an appropriate software application.
DivX (.divx)	Not suitable for preservation. Convert divx bitstreams to mjp2 or jp2 using an appropriate software application.
[H.264] MPEG-4 AVC (.avc)	Not suitable for preservation. Convert to recommended mjp2/jp2 using an appropriate software application.
MPEG-1 (.mpg, .mpeg)	Not suitable for preservation. Convert to the recommended mjp2/jp2 format using an appropriate software application.
Real Video (.ram, .rm)	Not suitable for preservation. Convert to the recommended mjp2/jp2 format using an appropriate software application.
Windows Media Format (.wmv)	Not suitable for preservation except as a wrapper format with constituent bitstreams encoded using jp2 or mjp2. Convert bitstreams encoded as .wmv to the recommended mjp2/jp2 format using an appropriate software application.

6.3.2 Sound

The following table contains our recommendations for the best preservation formats for digital audio files. The table does not cover every possible audio format in use but only the most common formats being used. Note that generally speaking it is not suitable to accept compressed formats for long-term preservation. Converting compressed formats to BWF will not improve the data quality but will ensure that the audio data is in a form that will remain useable for a long-period of time. In an archival situation deposit of compressed formats should be avoided unless no other options for preservation of the audio data are available.

Table 3- Migration Paths for sound

Ingest Format (+ extension)	Notes
Broadcast Wave Format (.bwf)	Suitable for preservation, using the LPCM encoding for bitstreams.
Linear Pulse Code Modulated Audio (LPCM)	Recommended encoding format for audio bitstreams. Convert or migrate audio bitstreams in other encoding formats to LPCM (may require re-sampling of originals).
Audio Interchange File Format, (.aif, .aiff)	Suitable as a preservation format. Note that AIFF files are large in size.
CD audio (CD-DA)	A PCM (Pulse Code Modulation) bitstream encoded to CD media. Suitable for preservation,.
Extensible Media Format (.xmf)	Further research needed on this format. It may be suitable as a preservation format for note-based digital music compositions.
Advanced Audio Coding (AAC MPEG-4) (.aac, .m4a)	Not suitable for preservation. Convert to bwf using an appropriate software application.
Dolby Digital Audio Compression (.ac3)	A proprietary format not suitable for preservation. Convert to bwf using an appropriate software application.
MIDI (.midi, .mid)	Unsuitable for preservation. Conversion not suitable because information is lost.
MP3 (MPEG-1 Audio Layer -III) (.mp3)	Not suitable for preservation. Convert to bwf using an appropriate software application.
Ogg vorbis (.ogg)	Not suitable for preservation. Convert to bwf using an appropriate software application.
Quicktime (.qt, .mov)	Not suitable for preservation. Convert to bwf using an appropriate software application.
Real Audio (.ra, .ram)	Not suitable for preservation. Convert to bwf using an appropriate software application.
Sun (.au)	Not suitable for preservation. Convert to bwf using an appropriate software application.
WAVE (.wav)	May be suitable for preservation if the wrapped encoding is LPCM, but it is a proprietary format generally not suitable for preservation purposes. Best practice is to convert to bwf using an appropriate software application.
Windows Media Audio (.wma)	Not suitable for preservation. Convert to bwf using an appropriate software application.

Chapter 7: Moving Images and Sound Metadata Review and Requirements

7.1 Introduction

Metadata may be defined as structured information that describes resources and events associated with those resources. The effective use of archived moving image and sound resources depends upon the suitability and adequacy of the metadata that describes them. This chapter aims to identify the metadata elements required to meet the needs of curators and end-users of archived moving images and sound. It covers the essential technical, management and intellectual aspects of the image resource[66]. It identifies and assesses the current metadata standards available for moving images and their use within METS and MPEG-21 environments. A usable, flexible and interoperable elements set for the archiving of moving images and sound is recommended and areas requiring further work are identified.

Discussion about metadata is currently bedevilled by a confusion of terminology. For the sake of convenience here, metadata will be divided into three broad categories: technical metadata, management metadata, and resource discovery metadata

7.2 Technical metadata

Technical metadata describes the physical rather than intellectual characteristics of digital objects. Some kinds of technical metadata are common to all resource types and others are specific to moving images and sound.

7.2.1 Video

The file format (eg. .wav .aiff) and encoding method, otherwise known as the codec (e.g. MPEG-1 video, CINEPAK), are fundamental details that must be identified and recorded for all digital audio and visual resources. The operating system and application software on the users' own machine will use some or all of this information to decide the application software appropriate to decode the digital files. It is important to distinguish between a file format and a codec. Even though most audio file formats, for instance, support only one audio codec, a file format may encapsulate different types of encoding, as AVI does.

Fixity (authentication) information must be recorded so that the curator of the archive can be sure that the resource has not been altered, either intentionally or unintentionally in any undocumented way. All digital objects, including moving image and sound files, are mutable and are liable, even without external intervention, to "bit-rot" –the gradual degradation of stored bits leading to partial or even complete information loss". Fixity information should include the name of the fixity method, the value of any check-sum, and the date that the fixity check was undertaken.

The number of pixels in the horizontal and vertical dimensions should also be stated explicitly within the metadata. A pixel is a standard method of measurement in still and moving digital images. A pixel appears as a square or rectangular object that contains specific colour information. The number of pixels stored horizontally and vertically indicates the resolution – the amount of information contained within the image. The pixel count is likely to be a key determinant of the suitability of a resource for some users. For example, a low-definition image is suitable for small display but will be pixelised when expanded to a larger screen.

[66] For a good introduction to this topic see Annemieke de Jong. Metadata in the audiovisual production environment. (Nederlands Instituut voor Beeld en Geluid, 2003).

The display aspect ratio should also be stated explicitly within the metadata although, in this case, the information is entirely for machine use. The display aspect ratio (also known as the frame aspect ratio) indicates the desired shape or proportions for the image output on screen in terms of the total vertical (y) frame size over total horizontal (x) frame size. Computer generated images typically have square pixels that may be changed to rectangular shapes for different types of displays. For example, 'widescreen' displays increase the horizontal size of a pixel. Moving images will be appear squashed or stretched if displayed on monitors designed for pixels of different proportion unless the software makes the necessary adjustments.

The metadata should also state whether a digital video item is scanned in an interlaced or progressive mode. This information is necessary for decoding the image. Without it, scanlines from completely different images may be blended together when resizing an interlaced video.

Bits per sample, the number of bits of information stored for each sample taken during the quantization of an analogue signal, is also essential information to record for video files. These days, eight bits are typically used to represent each colour component (red, green and blue). A 24-bit image, say, may be assisted to decay gracefully for display on a 15-bit screen if this piece of metadata is available to the software.

The sampling format expresses the luminance and chrominance (e.g. 4:2:0, 4:2:2, 2:4:4 etc.). Like bits per sample, the luminance and chrominance should be recorded for, essentially machine-processing reasons. Luminance is the component of a video signal that includes information about the amount of light to display. An incorrect luminance setting can cause visible flicker in a television display. Chrominance is the colour component of a video signal that includes information about hue (the colour of light or of a pixel, as determined by the dominant wavelength of light; hue corresponds to the radial dimension around a colour wheel) and saturation (intensity or vividness; saturation corresponds to the distance from the centre of a colour wheel). It defines the difference between a colour and a chosen reference colour of the same luminous intensity.

The frame rate of a video resource must be stated explicitly to ensure that it is played at the correct speed. If it is not, the duration of the video may differ from that of the source. The frame rate may vary according to the country of origin and the purpose of the resource (e.g. NTSC/PAL-encoded video typically operate at 25 or 30 frames per second; internet downloads typically operate at about 15 fps).

When disseminating the audio-visual object, an optimal balance between image/sound quality and download time can be achieved if the bit rate, file size and local network capabilities are known. Therefore the bit rate, measured in bits per second (bit/s) or megabits per second (Mbit/s), and the file size of the digital object, measured in bytes, should also be recorded for both the audio and visual components of multimedia resources. A variable bit rate, that uses more bits for rapidly-changing scenes than slow-motion scenes, may be used to maximise visual quality whilst minimising bit rate. If so, the maximum and minimum bit rate should be recorded in the metadata.

Finally, the technical metadata should contain an indicator as to the presence of sound in the video file. If the value is "yes", the video file needs to be associated with an instance of audio metadata.

7.2.2 Audio[67]

Many of the core technical metadata elements described above are also essential in relation to audio formats. File format and fixity (authentication), for instance, are generic features that must be recorded for all resource types. The encoding method (codec) must be stated explicitly, as for video files.

The essential element sets for audio and video resources are similar because the two media are based upon similar underlying technologies. They both divide data (in the case of audio, a sound signal) into discrete samples and quantizes them (i.e. represent each sample by a number). The two essential things to record in this respect for audio files are the rate at which the audio signal was sampled (sampling frequency), normally expressed in kHz, e.g., 22, 44.1, 48, 96, etc, and the number of bits (of information) per sample, e.g., 16, 20, 24. The sampling frequency and the number of bits determine the resolution – the amount of detail within the digital audio file. All other things being equal, the resolution determines the clarity of the music, voices and sound effects we hear. Different users will have very different ideas about what constitutes quality – according to the nature of their needs. Different types of audio also require different quality levels. A voice recording can be faithfully recorded at 22 kHz but to capture the full frequency range for music will require a higher sampling frequency. Recording the sampling frequency and bits per sample provides the software or end-user with a basis upon which to base decisions about whether an audio-visual resource is of sufficient quality to fit the intended purpose. For audio files the number of channels should also be recorded (e.g. 1, 2, 4, 5).

The data rate (expressed as kbps, e.g. 64, 128, 256, etc.), whether the data rate is fixed or variable, and the duration or elapsed time of the entire file, expressed using ISO 8601 syntax[68], are other items of metadata that should be recorded for audio as well as video formats.

Audio and video are similar in that they are both time-dependent media types. They must be played over a period of time and accurate timing is essential for both. Audio formats divide their data into logical blocks. For instance, data on a CD is typically contained in logical blocks of 2352 bytes each, with each block representing 1/75 seconds of sound. It is important to record the size of an audio block in bytes, the location of the first valid sound byte in the file, the location of the first valid sound byte in the block, and the location of the last valid sound byte in the block.

An array of other technical metadata can be recorded. For instance, the channel configuration may be indicated by mapping the audio channel to their intended aural position/loudspeakers (e.g.1=left,_front, 2=right_front, 3=center, 4=left).The aural space arrangement of the sound recording (e.g. monaural, stereo, joint stereo, surround sound DTS5.1, etc.) may also be relevant in some cases. However, the elements identified above are the essential ones. Some are solely for machine or curator use. Others may help the end-user to make conscious decisions about how to use the resource – although, with MPEG-21, the trend is towards the seamless delivery of content across networks with minimal human intervention.

7.3 Management metadata

Management metadata is information documenting the life cycle of an electronic resource, including data about ordering, acquisition, maintenance, licensing, rights, ownership, and provenance.

[67] This section is indebted particularly to: http://www.loc.gov/rr/mopic/avprot/DD_AMD.html
[68] http://www.w3.org/TR/NOTE-datetime

It is essential that the provenance (custodial history) of any digital object is recorded from, where possible, the time of its creation through all successive changes in custody or ownership. Users and curators must be provided with a sound basis for confidence that a moving image or sound resource is exactly what it is purported to be.

Likewise, all processing actions or changes that are made to the digital file(s) throughout the life cycle of the resource should be recorded and, in particular, any changes that result to the significant properties of the resource as a result. There should be a clear audit trail of all changes.

Any binding intellectual property rights attached to the digital object must be documented in the metadata. Rights are crucial in two respects. They may limit the archive's powers to undertake preservation action on the resource. They may also prohibit or restrict its dissemination to users.

A small amount of administrative metadata should also be created for each resource, by which is meant metadata about the metadata-creation process itself. (i.e. the 'who', 'when' and possibly 'how' of its creation).

As part of the 'business rules' or 'preservation planning' process of the archive, the curators will probably also wish to document their schedule for future preservation action in relation to the various classes of digital resource (e.g. next preservation action; date next preservation action due, retention periods, risk assessment, and significant properties of each class of resource). However, it is not essential - nor perhaps desirable - for such information to be within the formal metadata of individual digital objects.

Some structural metadata is also necessary to describe the internal organisation of the digital object. For example, when a boxed set of vinyl records is digitised, several dozen digital files may be required to cover the various parts - the various disc sides, the images of the labels, the box, and the pages in the accompanying booklet. The relationship of these files to one another is documented in the structural metadata. The structural metadata enables the software to display and navigate the multi-file digital object for the user. METS[69] and MPEG-21[70] are two standards that have been designed for this purpose and their suitability for audio-visual resources will be considered below.

7.4 Discovery metadata

Technical and management *must* be augmented by resource discovery metadata. The purpose of archiving moving image and sound resources is to facilitate their long-term retention *and use* in the future. Resource discovery metadata enables users to identify, locate and retrieve suitable resources – normally via the medium of a searchable index or catalogue.

Intellectual content is a key criterion upon which resource discovery decisions are usually based. The content or subject of an audio-visual resource should be described so that the user can determine what it is 'about' or contains. This is particularly important because of the inherent lack of machine-readable keywords within the content of audio-visual resources. The bare minimum is that each resource has a unique permanent identifier but, wherever possible, suitable titles and free-text descriptions should be applied. Keywords from a controlled vocabulary may be added to enhance computerised retrieval. The latter is unlikely to be an automated process despite recent developments in image and sound recognition software and is normally very labour-intensive.

[69] http://www.loc.gov/standards/mets/
[70] http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm

Particularly for time-based content such as audio and video, what is sometimes called "logging" metadata may be just as important as the classic forms of bibliographic/descriptive metadata. The simplest form of logging data is a transcription of the words heard in a soundtrack and/or a set of "keyframes". (Since a soundtrack transcription is a bit like the searchable text for a scanned book, we can see that logging data is not limited to time-based formats). For video, logging data is provided by systems like Virage, and their sales demos suggest its efficacy in discovery. It is in the logging-data arena that MPEG-7 is promised to Excel[71].

Content is not, however, the sole determinant of suitability. Suitable resources are those that are *fit for the purpose* the user has in mind. This brings us back to the importance of adequate technical and administrative metadata because suitability is related to the quality of the digital object - measured, in the case of moving images, by technical attributes such as bit depth and aspect ratio.

What is not in doubt is that, wherever possible, metadata should be expressed as XML – which extensible, supports the separation of data from the application, and is widely used, thereby encouraging data exchange.

7.5 Review of existing metadata standards

A review of metadata standards is timely because metadata provision has been identified as *the* chief issue facing the archival moving image and sound community.

7.5.1 MPEG21-didl[72]

MPEG-21 is probably the most sophisticated metadata standard yet devised by the audiovisual community to meet the requirements of its characteristic resources. They may be summarised as complex, multi-file, digital objects to which multiple rights typically pertain, that are played over a period of time, and which typically require description at various levels of granularity from the digital object as a whole down to the level of, say, individual frames or seconds. MPEG-21 is a wrapper that can bring together a complicated array of metadata elements from a variety of standards into a single machine-readable and navigable structure. A remarkably flexible structure with a consistent logic can be achieved because the same basic building blocks are repeated throughout and at all levels of granularity:

The basic MPEG-21 concept is the digital item, which is a structured digital object (e.g. a digital recording of an RSC performance of Hamlet) together with its metadata. The item may have sub-items (e.g. acts or scenes). An act or scene may itself comprise a number of separate files. These individually identifiable data streams, such as video files, images, audio clips or textual assets, that go to make up the complex digital item or its sub-items are known as resources. In MPEG-21, the actual metadata information that relates to a digital item, sub-item or resource is either held directly within descriptor statements or within sub-elements from other standards (e.g. Dublin Core, METS, MPEG-7 or Indecs) that are embedded within the descriptor statements. Components are wrapper elements that bind a resource or sub-item to its relevant descriptors but do not contain information themselves. A particular instance or expression concerning rights, duration, identifiers, keywords, and so on, can relate to the entire digital item, a specific sub-item or resource (DataStream), and the purpose of MPEG-21 component is to relate it efficiently just to the specific portion of the digital item to which it actually applies.

---

[71] This paragraph is based on notes submitted by Carl Fleischhauer
[72] http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm

In fact, resources can be further divided into fragments. A fragment designates a specific point or range within a DataStream (resource) – such as a point in time of an audio track. Anchors (broadly equivalent to components) then bind the fragments to their descriptor statement(s).

A final set of entities, containing the choice, selection, condition and assertion entities provides a means for distributing content via the most appropriate channels to various consumption platforms in a seamless fashion. Based on the selection made by a User, the conditions attached to (an entity) of a digital Item may be fulfilled and the digital item may become available. Content may be produced only once and distributed, without human intervention, in the most appropriate format for the target channel (e.g. a high resolution DataStream for a high bandwidth Internet connection or a compressed file for a dial-up connection) according to the particular conditions, user preferences and technological capabilities that are expressed about client systems within these entities. To maximise machine-to-machine processing across the widest range of networks and devices used by different communities, MPEG-21 has supplied a controlled vocabulary, the MPEG-21 Rights Expression Language (MPEG-21 REL)[73]. Partly due to its flexibility, MPEG-21 is a difficult standard to understand - particularly when considered in the abstract. However, the seamless "create once, publish anywhere" model[74] that MPEG-21 facilitates potentially opens up exciting new possibilities for data use and re-use within digital repositories[75].

MPEG21 was designed with access and distribution in mind. It was to provide a framework that enables users to "exchange, access, consume, trade and otherwise manipulate digital materials"[76] across the widest range of user communities, domains, systems and technologies. However, there appears to be no logical reason why its generic structure cannot accommodate technical preservation metadata too.

MPEG-21 is not restricted to the audio-visual domain. Having been designed with complex [77]multimedia objects in mind, it appears to accommodate other (mostly much simpler) resource types well – something of relevance to repositories with multiple resource types.

7.5.2 MPEG-7

MPEG-21 – although difficult to understand - provides a flexible wrapper for structuring audio-visual metadata. As has been shown, archives will often use the <didl:Descriptor> element as a wrapper for sub-elements from richer or more specialised metadata standards.

One of these is MPEG-7, which provides many descriptors for the searching and filtering of specific types of content. For example, for audio content there is waveform, fundamental frequency, spoken content, timbre, etc. For visual content there is colour, texture, shape, motion, etc.

With the right software and appropriate MPEG-7 descriptors, some very sophisticated searching may be possible:

- Music: It should be possible to play a few notes on a keyboard, or even sing them, and get a list of works containing the required tune, or something close to it.

---

[73] http://www.chiariglione.org/mpeg/technologies/mp21-rel/index.htm
[74] http://www2003.org/cdrom/papers/poster/p300/p300-vanassche.html
[75] Belle L. Tseng, Ching-Yung Lin, and John R. Smith. Using MPEG-7 and MPEG-21 for Personalizing Video. January/March 2004 (Vol. 11, No. 1) pp. 42-53 http://ieeexplore.ieee.org/iel5/93/28183/01261105.pdf?tp=&arnumber=1261105&isnumber=28183
[76] Annemieke de Jong. Metdata in the audiovisual production environment. (Nederlands Instituut voor Beeld en Geluid, 2003), p.36
[77] See Jeroen Bekaert, Patrick Hochstenbach and Herbert Van de Sompel. Using MPEG-21 DIDL to Represent Complex Digital Objects in the Los Alamos National Laboratory Digital Libraryhttp://www.dlib.org/dlib/november03/bekaert/11bekaert.html

- Graphics: Draw a simple sketch and have returned to you similar graphics, including logos, trade marks etc.
- Images: Describe objects, colours, textures and get in return example among which you select interesting objects to compose your image.
- Movement: On a given set of objects, describe movements and relationships between objects and get in return a list of animations fulfilling the described temporal and spatial relationships.
- Voice: "Using an excerpt of Pavarotti's voice, and getting a list of Pavarotti's records, video clips where Pavarotti is singing or video clips where Pavarotti is present." [78]

MPEG-7 is enormously flexible because the descriptors may be applied at any level of granularity. For example, they may relate to an entire production or a segment of it such as a scene or a single frame.

Various levels of abstraction are also possible - from low-level description (e.g. shape, size, texture colour, position, trajectory, mood, and tempo) to high-level semantic information (e.g., "This is a scene with a barking brown dog on the left and a blue ball that falls down on the right, with the sound of passing cars in the background"[79]). MPEG-7 allows the same material to be described in different ways according the need – although higher-level abstractions are normally least amenable to automated generation and even for low-level entities, the tools to make and search the data are not yet widely available.

MPEG-7 can, of course, be used without the broader framework provided by MPEG-21. MPEG-7 addressed the problem of content description and also, to some extent, the issues of rights and user preferences and, when it was drafted, it represented a major advance. Prior to MPEG-7, access to audio-visual content was hampered by resource discovery metadata being embedded within a variety proprietary file formats. MPEG-7 addressed the problem by standardising a set of metadata descriptors and a description definition language that is applicable across as broad a range of applications as possible. It was designed to take into account all the viewpoints of other leading standards, such as SMPTE, Dublin Core, EBU P/Meta, and TV-Anytime.[80]

However, MPEG-7 was better suited to single-file based media types than content consisting of multiple files. Nor did it standardise how applications can interact with such content in an interoperable way. MPEG-21 was developed to ensure the transfer, adaptation and dissemination of content in ways that protect the content and the rights of the rights holders. MPEG-7 remains a key metadata standard for audio-visual resource discovery but has been subsumed within the broader framework provided by the more recent MPEG-21.

7.5.3 TV-Anytime extended metadata schema[81]

Television archives may also consider using the metadata schema/data dictionary that has been designed specifically for television resources by the global TV-Anytime Forum. TV-Anytime facilitates richer description of this type of resource than more generic standards. It is harmonised with the MPEG-7 Description Definition Language (DDL), from which it has borrowed much of its structure and vocabulary, and is therefore broadly interoperable. Indeed, it could be considered an application of MPEG-7[82]. TV-Anytime elements can be used either on their own or within an MPEG-21 framework. TV-Anytime covers:

---

[78] 'A Sound Person's Guide To Video' http://www.spgv.com/columns/mpeg7.html
[79] http://www.cs.cmu.edu/~hdw/Wactlar_CLIR-final.pdf, p5.
[80] http://www.cs.cmu.edu/~hdw/Wactlar_CLIR-final.pdf
[81] http://www.tv-anytime.org/
[82] Silvia Pfeiffer and Urea Srinivasan. TV Anytime as an application scenario for MPEG-7, CSIRO Mathematical and Information Sciences. http://www.citidel.org/?op=getobj&identifier=oai:ACMDL:articles.357885

- *Content Description Metadata* – i.e. aspects of content that are independent of any particular instantiation. This includes resource discovery metadata (e.g. program title, genre, production date, language, credits, synopsis) and technical metadata (e.g. FileFormat, FileSize, NumOfChannels, (audio) MixType, HorizontalSize, AspectRatio).
- *Instance Description Metadata* – i.e. description that relates to particular instantiations of the audio-visual resource (such as broadcast and Web delivery). TV-Anytime is better suited for this than generic metadata standards because there is a dedicated section for the details of each instantiation, such as published time and duration, repeat, first/last showing, URL, title (instances of a program can have a different titles).
- *Consumer metadata* – TV-Anytime has sophisticated mechanisms for recording personal usage history from which user preferences can be generated. The reasons were, no doubt, largely commercial (e.g. to allow the sale of viewing history to advertisers). However, they represent an area of opportunity within educational contexts too – e.g. by allowing the building of personalized TV guides, the tailoring of future resource-creation to user needs, or the configuring of systems to extract content based on users' known preferences.
- *Segmentation metadata* –TV-Anytime has more flexible mechanisms for defining, accessing and manipulating temporal intervals (*i.e.*, segments) within an AV stream than more generic schemes, such as Dublin Core, which lack the precision or vocabulary to do this effectively. As with MPEG-7, resources can be described at any level from series or programme to the individual frame. Highlights, previews, shots, events, themes, key frames, and so on, can all be identified, together with their location, timing and duration. For instance, China Central TV's "Oriental Space Time" television programme has three sections, "Oriental People", "Live Broadcast" and "Stories of Grassroots"[83], which can each be treated as a segment.

TVAnytime metadata also has the advantage of being broadly interoperable because of its use of XML and the MPEG-7 Description Definition Language (DDL) for its structure and vocabulary. Rather than adopt a standard, such as TV-Anytime, in its entirety, archives may legitimately 'mix and match' elements from a variety of standards, including TV-Anytime, to create application profiles suited to their specific needs. MPEG-21 and METS provide useful wrappers for such an metadata.

7.5.4 SMIL (Synchronous Multimedia Interchange Language)[84]

Although it is perhaps more correctly described as a markup language than a genuine metadata standard, SMIL deserves a mention because it enables rich, television-style multimedia of the sort mentioned above to now be brought within the scope of educational Web pages.

Audio and visual resources have an additional dimension that is not shared by other resource types in that they are time based. SMIL is a flexible XML markup language that gives you the power to control the timing of such media objects with respect to each other. An audio clip, for instance, can be carefully timed so that it plays at the appropriate point in time in relation to a video presentation, still image or text caption. It can stipulate not only *which* audio or video file will play but also which *part* of the audio or video clip and for *how long*. Any point within a clip can be called up as its starting point. It works at a fine (frame) level of granularity.

---

[83] http://www.slais.ubc.ca/PEOPLE/faculty/tennis-p/dcpapers2004/Paper_17.pdf
[84] http://www.w3.org/AudioVideo/ for official SMIL documentation. See also Michael Morrison, SAMS Teach yourself XML in 24 hours, 2nd ed. (Indianapolis, 2002) for a good introduction to SMIL.

Usefully, SMIL can also divide the output on screen into two or more sections or regions (sometimes known as 'channels'). For instance, visual archival footage from the Second World War may be shown across the majority of the screen accompanied by an audio commentary whilst troop movements are traced on a map in a corner of the screen and a transcript below keeps time with the audio commentary.

SMIL can also assess the capabilities of local systems and disseminate media objects that are appropriate to them. For instance, high quality images and sounds may be provided to users with high-bandwidth connections and lower quality versions of the same object to users with slower connections.

SMIL achieves all this using a small range of element attributes:

begin – (start playing/displaying the digital media object)
dur – (play/display for this length of time/duration)
end – (stop playing/displaying)
par – (play/display in parallel, i.e. simultaneously)
region – (play/display in this region of the screen)
seq – (play/display in sequence, i.e. one after the other)

e.g.
<audio src="sound.au" region="background_audio" type="audio/x=au" dur="68s"/>

It must be borne in mind, though, that SMIL can stipulate 'play the video for 20 seconds, starting at the 35th second', but detailed resource discovery metadata, such as MPEG-7 or TV-Anytime, will have been necessary for the Web page developer to identify that as the relevant portion of the video in the first place.

7.5.5 Dublin Core[85]

The metadata standards reviewed so far have all permitted rich levels of resource description but such metadata is normally expensive to produce – particularly if it cannot be generated automatically. A simpler standard, such as Simple Dublin Core, may sometimes be more appropriate – particularly under circumstances when all that is required is access to the digital object or item as a whole rather than a segment of it.

Simple Dublin Core is probably the best base standard upon which to develop a *minimum* generally applicable element set for the archiving of digital images. The Dublin Core standard (formally known as the Dublin Core Metadata Element Set) has defined fifteen core elements: Title, Creator, Subject, Description, Contributor, Publisher, Date, Type, Format, Identifier, Source, Language, Relation, Coverage and Rights. The worldwide DCMI consortium has designed the standard to accommodate the fundamental resource discovery requirements of *all* data types, subjects and domains. Because of its inherently cross-sector design it is well suited for use within a repository environment where digital images may be only one amongst a variety of data types. It is unlikely that repositories will find it feasible to implement separate schemas for each resource type. Dublin Core is already widely adopted and is therefore a major source of interoperability between repositories. Simple DC metadata can be shared between all repositories, which support the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). Widespread use of Dublin Core facilitates consistent results when users are searching across holdings in multiple repositories or browsing metadata records gathered from multiple repositories.

According to DCMI, all Dublin Core elements are optional and all fifteen are relevant to at least a subset moving images and sound resources. However, only Title, Creator, Subject,

[85] http://dublincore.org/documents/usageguide/

Date, Identifier, and Rights are likely to be essential for all audio and video objects and this should be regarded as the minimum element set for this type of resource. All Dublin Core elements – including Rights - are repeatable and, to this extent, Dublin Core may be regarded as being well-suited to the description of audio-visual objects, which typically display a complicated array of associated rights,. However DC has no built-in mechanisms for making fine distinctions between different kinds of entry within each of its broad elements. Therefore there is no way of differentiating between, for instance, copyright, design rights, performance rights, literary rights, moral rights, trademark rights and conditional rights of access, which are all bracketed together under rights.

In particular, <dc:relation> is too blunt an instrument for adequately defining the correct relationships between the various components of a complex digital resource (e.g. part/whole, parent/child, hasVersion, hasFormat, relationships, etc.). <dc:relation> is prone to being overloaded and unsuitable for machine processing because of the variety of information it has to accommodate – which may include the title and location (e.g. URL) of the related source and also, in the absence of any dedicated element for this purpose, the precise nature of the relationship.

One of the richer metadata standards should generally be preferred when feasible but Simple Dublin Core may provide a summary description, suitable for metadata exchange and harvesting, of the resource at item or digital object level. That is all that some non-specialist libraries and archives may require or find feasible. However, greater precision/granularity of description may be possible if simple DC elements are wrapped within MPEG-21 components and descriptors or METS divisions and pointers that relate to particular segments of the resource.

7.5.6 DC Terms[86]

One solution is DC Terms (previously known as Qualified Dublin Core), a standard that accommodates somewhat fuller description than simple Dublin Core. DC Terms presents three additional elements - audience, provenance and rightsHolder – and a set of element refinements (also called qualifiers) that refine the semantics of the elements and afford a greater degree of descriptive precision. Provenance, in particular, is a key piece of information for archives to record and it is recommended here that the dcterms:provenance element should be included in the minimum element set for the archiving of moving images and sound resources.

For the sake of interoperability and resource sharing, the key thing, though, is to choose a metadata standard that maps well to Dublin Core. Metadata in any format can be oai harvested, although only simple Dublin Core elements approach being universally interoperable. Archives must judge for themselves (in relation to their particular circumstances) the optimal balance to strike between richness of description and maximum interoperability because they stand in opposition to one another.

7.5.7 PBCore[87]

The emerging metadata specification for American public broadcasting, the Public Broadcasting Metadata Dictionary (PBCore), is an example of a richer standard that has been built on the foundation of the Dublin Core, maps to it, and which the Dublin Core Metadata Initiative Usage Board has reviewed. PBCore presents what its creators consider to be the smallest set of elements that could adequately describe the media items produced by Public Broadcasting radio and television stations so that their media items may be shared between stations, regional and national distributors, independent producers, and even

[86] http://dublincore.org/documents/usageguide/qualifiers.shtml
[87] http://www.utah.edu/cpbmetadata/

vendors of Digital Asset Management systems. It is designed to allow traditional programs to be parsed into short segments for Web distribution or as niche content for specific community, service and institutional needs.

Thirteen of the PBCore elements describe *intellectual content*, seven elements relate to the media items as *intellectual property* (e.g. are concerned with the creation, creators and usage of the media asset or resource), and twenty-eight are concerned with *instantiation* in that they identify the nature of the media asset as it exists in some form or format in the physical world or digitally.

Encouragingly, PBCore's *instantiation* section contains many of the essential technical metadata elements identified by this study. E.g *formatDuration, formatAspectRatio, formatFrameSize*.

The most fundamental PBCore elements map to Dublin Core, e.g. *creator* (which records the person or organization primarily responsible for creating a media item), but PBCore provides additional richer description than DC. For instance, the *creator* element is accompanied by *creatorRole* – an element that provides a controlled list of terms (such as *artist, composer, director*) for defining the creator's contribution more precisely.

7.5.8 METS[88]

Archives must also decide, in relation to their specific circumstances, whether to employ METS, MPEG-21, or neither, as a container for their metadata. METS (Metadata Encoding and Transmission Standard) provides a good over all data structure for wrapping the metadata and (if desired) the data required for displaying, exchanging, and preserving audio-visual and other types of digital object within a digital repository. METS is a realistic alternative to MPEG-21 for audio-visual resources[89]. Metadata expressed using one or multiple schemas can be held within separate sections of a METS document and linked to the relevant data file or portion of it. The METS document, in turn, is an instantiation of an Information Package as stipulated in the Open Archival Information System (OAIS).[90]

A METS document consists of seven major sections[91]:

1. The METS Header <metsHdr> contains metadata describing the METS document itself, including such information as creator, editor, etc.
2. The descriptive metadata section <dmdSec> provides a container for resource discovery metadata elements drawn from Dublin Core or other resource discovery metadata standards.
3. The Administrative metadata section <amdSec> can accommodate four kinds of metadata that are crucial for archiving and preservation:
    a. <techMD>, Technical Metadata (information regarding a files' creation, format, and use characteristics). The sub-elements within this section may, for instance, be drawn from the PREMIS data dictionary and relate to a file's preparation
    b. <rightsMD>, Intellectual Property Rights Metadata (copyright and license information)
    c. <sourceMD>, Source Metadata (descriptive and administrative metadata regarding the analog source from which a digital library object derives)

[88] http://www.loc.gov/standards/mets/
[89] For a sample Library of Congress METS document describing an audio compact disc see: http://www.loc.gov/standards/mets/examples-profiles/sample7.xml
[90] E.g., see DSpace METS Document Profile for Submission Information Packages (SIP) at: http://cwspace.mit.edu/docs/xsd/METS/SIP/profilev0p9p1/metssipv0p9p1.pdf
[91] http://www.loc.gov/standards/mets/METSOverview.v2.html

    d. <digiprovMD>, Digital Provenance Metadata (information regarding source/destination relationships between files, including master/derivative relationships between files and information regarding migrations/transformations employed on files between original digitisation of an artefact and its current incarnation as a digital library object).
4. The file section <fileSec> lists all files containing content that comprise the digital object.
5. The structural map <structMap> outlines a hierarchical structure that enables multiple instances of the administrative metadata within the <amdSec> and <techMD> to be linked with the pertinent content file <fileSec>.
6. Structural Links <structLink> allows the recording of hyperlinks that exist between items within the structural map
7. The behaviour section <behavior> can be used to associate executable behaviours with content in the METS object.

METS is emerging as the metadata wrapper of choice for many digital repositories. Specialist audio-visual archives *may* be more disposed to MPEG-21, while more general (e.g. University) repositories that archive limited amounts of audio-visual data might be expected to gravitate more towards METS because of its library origins.

Crucially, METS, like MPEG-21, is sufficiently flexible for metadata to be associated with the relevant fragments or portions of files. The METS website describes an oral history project (with Mayor Abraham Beame of New York City) that includes three subsections: an opening introduction by the interviewer, some family history from Mayor Beame, and a discussion of how he came to be involved with the teachers' union in New York. Each of these subsections/divisions is linked to three files: an XML transcription, and a master and derivative audio file. A subsidiary <area> element is used in each <fptr> to indicate that this division corresponds with only a portion of the linked file, and to identify the exact portion of each linked file. In this case, the metadata enables the interviewer introduction to be found in both audio files in the segment beginning at time 00:00:00 in the file and extending through time 00:01:47[92].

There are, however, some grounds for suspecting that MPEG-21 offers the greater flexibility in relation to complex digital resources – although its flexibility can make it difficult to understand. The METS approach is to declare one or more namespaces (e.g. dc or TV-Anytime) and provide each schema with its own dedicated portion of the <dmdSec> into which its relevant elements are grouped regardless of the level of resource to which they refer. ID's then reference the relevant files from there. MPEG-21, in contrast, permits an arbitrarily large number of levels of nested description each of which can contain elements drawn from any number of the separate schemas. The container/descriptor/statement syntax permits all of the elements pertaining to an entity (at whatever level) to be grouped logically at the correct level of granularity and in direct relation to the given entity regardless of which schema they derive from. However, more research is necessary in this area before archives can make decisions on this issue with confidence. Few case studies and exemplars are currently available.

7.5.9 Collection-level metadata for image collections

Resource discovery metadata should, where possible, be created at collection, as well as item, level. Collection-level records enhance resource discovery and provide a home for contextual information about the wider environment within which the digital object was created – for instance, the name, purpose and nature of the parent project, any grant number associated with it, and the other images related. Such information gives important

[92] http://www.loc.gov/standards/mets/METSOverview.v2.html

clues for interpreting the image and both from a logical and practical point of view, is best recorded once and referenced by each item-level record in the collection.

7.5.10 Digital reproductions

Digital reproductions are usually more complicated to describe than 'born digital' resources because there is a source entity to consider. Metadata should focus unambiguously on the resource to hand – which in this case means the digital object. In the case of a sound recording, DC:date, for instance, should be the date the digital sound file was created rather than the date that the analogue source (e.g. vinyl record) was recorded/manufactured. DC:source, dc:relation, and dc:description may provide important information about the source item, but the orientation as a whole should be towards the digital object.

7.5.11 PREMIS[93]

Dublin Core is designed for purposes of resource discovery and lacks technical and management metadata elements. Therefore, the minimal Dublin Core element set should therefore probably be augmented by elements from the PREMIS data dictionary[94]. PREMIS defines a set of core management and technical metadata elements that are needed to support the preservation of all kinds of digital resources, regardless of their data type. It covers the entities *Objects, Events, Agents, and Rights.* Objects includes all of the essential technical metadata requirements of digital moving images and sound that are of a generic sort - such as fixity, format name and format version. Rights covers all of the rights and permissions required for preservation activity but, as the PREMIS working group itself acknowledges, the focus is specific so that rights associated with access and/or dissemination are out of scope - as are all resource discovery requirements on the basis that they are provided for by other standards such as Dublin Core. PREMIS is intentionally not a comprehensive standard that meets all the metadata needs of a digital repository but should provide many of the essential technical and management metadata elements necessary for the archiving of moving image and sound digital resources.

PREMIS, however, is of little use in relation to the significant characteristics of audio and visual resources that were mentioned earlier but that are *uniqu*e to this type of resource. These include:

- Indication of the audio or video codec, if it differs from the file format
- Detailed Codec information - Frame rate, recording length, sampling frequency, bit rate etc.
    - The length of the recording (in minutes and seconds)
    - Video dimensions, frames per second (FPS) and bit-rate
    - Audio bit-rate (kbps) and sampling frequency range (KHz), which should be recorded for all channels, if applicable
- Description of any (descriptive) metadata associated with the audio or video file.
- Information on the purpose/use of the digital footage

There was an expectation that the separate digital communities would produce their own PREMIS implementations or extensions for specific resource types. However, nothing comprehensive like the draft standard NISO Z39.87 for raster images yet exists for audio-visual resources.

[93] See *Preservation Metadata: Implementation Strategies (PREMIS)* international working group home page at: http://www.loc.gov/standards/premis/
[94] http://www.oclc.org/research/projects/pmwg/premis-final.pdf

7.5.12 AudioMD (Audio Technical Metadata Schema)

Some developmental work on technical metadata for moving images and sound has, nevertheless, been done elsewhere. The Library of Congress Audio-Visual Prototyping Project, 1st phase, (1999-2004)[95] has produced a draft METS extension metadata schema and data dictionary for audio resources:

Draft data dictionary: http://www.loc.gov/rr/mopic/avprot/DD_AMD.html
Draft schema: http://www.loc.gov/rr/mopic/avprot/AMD_020409.xsd

AudioMD contains 37 technical metadata elements for describing an audio object. The data dictionary isn't much more than a table of elements. It has been developed jointly by the Library of Congress and the Audio Engineering Society and has been adopted by the Library of Congress as an interim measure.

7.5.13 VideoMD (Video Technical Metadata Schema)

The Library of Congress Audio-Visual Prototyping Project also produced VideoMD, a draft METS extension metadata schema and data dictionary for video objects that contains 16 technical metadata elements.

Draft data dictionary: http://www.loc.gov/rr/mopic/avprot/DD_VMD.html
Draft schema: http://lcweb-2.loc.gov/mets/Schemas/VMD.xsd

Unlike other official METS extension element sets, no tools yet exist for automated extraction of AudioMD and VideoMD data from digital objects. Furthermore, VideoMD is an early draft and there is no evidence of any further development work having been done on it since Spring 2003.

7.5.14 Contextual documentation

If a suitable metadata schema is adopted within an OAIS compliant structure, there should be little or no need for additional documentation outside of the formal metadata. This is certainly the case for moving images and sound resources.

7.6 Metadata extraction and storage

7.6.1 Creating metadata

The extraction of technical metadata manually from moving image and sound files is largely impractical and few tools have yet been developed to do it automatically. The uptake of specific metadata standards, such as PREMIS, is likely to be influenced in part by the quality of the programmes designed for them.

7.6.2 Storing metadata

It is possible to record the metadata relating to moving image and sound resources in several ways. The metadata may be embedded in the digital object itself (in the case of file formats that support descriptive headers), be held separately as a distinct metadata record (e.g. an external catalogue), or held within a separate file but linked to the digital object within a repository structure.

In a database:

Separate storage of metadata is generally recommended. The fundamental technical information, which the computer system relies upon in order to access the file and understand it, *must* be recorded within an explicit metadata record separate from, or

[95] http://www.loc.gov/rr/mopic/avprot/avprhome.html

additional to, the digital object file itself. Operating systems recognise a file by first associating its file extension with a particular type of software, and then reading the coded technical metadata in the file header. If a future operating system lacks the ability to play, say, a .wav file, then the software player will be unable to recognise the file extension and will not be able to access the file to extract the necessary information to understand it."[96] The resource will effectively be unusable. Keeping this data separate from, but linked appropriately to, the resource is important for its preservation and access in the long term.

A second advantage of this approach is that the resource discovery metadata is normally easier to update if stored separately. This may be particularly important within the broadcast environment, where the digital materials are dynamic and subject to frequent adaptation or within the heritage sector, where descriptive metadata relating to heritage collections is prone to incompleteness. This approach allows one to update the records easily as information and resources become available. One cannot assume that embedded metadata will be adequately maintained – not least because of the difficulty of providing permanent *online* accesses to all audiovisual files and their embedded meta-information.

There is, however, a down side to this approach. Storing the content and its associated metadata within separate files works fine when they remain within their original environment, such as within an online catalogue. Problems arise when the digital object is downloaded and removed from its original setting.

Embedded in the moving image/sound file:
IASA, the International Association of Sound and Audiovisual Archives, says that "Though there are many advantages to maintaining metadata and content separately, in principle, metadata can be integrated with the audio files".[97] The challenge has been to create a standard way for it to be represented. Broadcast and film industry formats that have been standardised by SMPTE now often carry embedded metadata expressed as KLV (Key-Length-Value) triplets, in which key identifies the data, length specifies the data's length, and value is the data itself. KLV is defined in SMPTE 336M-2001 (Television-Data Encoding Protocol Using Key-Length Value). AAF and MXF are two emerging specifications that, like MPEG-21 and METS enable the wrapping of metadata at varying levels of granularity - and they both use the KLV format. There is now a standard for the embedding of metadata within non-tape-based camcorders for broadcast production. I it the European Broadcasting Union's (EBU) Broadcast Wave Format (BWF)[98], which provides for a limited number of embedded descriptive metadata elements within .wav files. IASA is confident that this format and standard will remain viable for some time. However, we would not recommend relying on embedded information alone. There remain various issues regarding the use of the common formats and their ability to be read by common application software.

Although it is desirable to encapsulate objects, storing both the metadata and the digital object a single structure, metadata required to manage digital archives needs to be accessible and maintainable. Therefore extracting and storing all the relevant metadata in a separate file is preferable to having to access the objects themselves to retrieve metadata.

7.7 Conclusion

Sophisticated metadata standards (e.g. MPEG-7, TV-Anytime, SMIL) and container schemas (MPEG-21 and METS) are now available for the archiving of moving image and

[96] IASA-TC 04 2004. Guidelines on the production and preservation of digital audio objects. August 2004, p9. http://www.iasa-web.org/tc04/index.htm
[97] IASA-TC 04 2004. Guidelines on the production and preservation of digital audio objects. August 2004, p9. http://www.iasa-web.org/tc04/index.htm
[98] http://www.ebu.ch/CMSimages/en/tec_doc_t3301_v1-2005_tcm6-40955.pdf

sound resources. However, manual metadata creation is expensive and, in reality, little file level metadata is likely to exist separately from the content other than that which can be auto-generated. Few metadata extraction tools currently exist for this purpose and further research and development work is needed in this area.

Providing a minimum essential metadata set for the archiving of audio-visual resources is a daunting challenge. Multi-file, time-dependent moving image and sound resources are much more complex and varied in terms of the demands that they place upon metadata than, say, raster images. An approach based on format-specific standards would, most likely, be unsustainable and hinder interoperability in view of the number of formats to be found. Therefore, the proposed element set uses a cascading approach whereby resource discovery elements are drawn firstly from the most generic established standard (Simple Dublin Core). These are augmented, where necessary, with generic resource management and technical preservation elements from other recognized standards (e.g. DC Terms and PREMIS). Some format-specific elements are then added but only where that is unavoidable.

The resulting element set should be regarded as a bare minimum. Most repositories should record richer metadata than it mandates by adding elements from other recognized standards to form application profiles suited to their specific needs. Moreover, the element set makes no assumption as to the way the metadata is stored or contained (e.g. MPEG-21 or METS). Nor does it stipulate the level of granularity to which each element pertains. Some elements will apply to the digital object as a whole, others to items, component, fragment, sections or other subdivisions. Curators should bear this in mind when applying them.

7.8 A recommended minimum element set for the archiving of digital moving image and sound resources.

No.	Element	Metadata standard	Definition	Comment
Essential elements for all moving image and sound resources.				
1	Title	Dublin Core	The name given to the resource.	Typically, a Title will be a name by which the resource is formally known.
2	Creator	Dublin Core	An entity primarily responsible for making the content of the resource.	Examples of a Creator include a person, an organization, or a service. Typically the name of the Creator should be used to indicate the entity.
3	Subject	Dublin Core	The topic of the content of the resource.	Typically, a Subject will be expressed as keywords or key phrases or classification codes that describe the topic of the resource. Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme.
4	Date	Dublin Core	A date associated with an event in the life cycle of the resource. format.	Typically, Date will be associated with the creation or availability of the resource. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [Date and Time Formats, W3C Note, http://www.w3.org/TR/NOTE-datetime] and follows the YYYY-MM-DD

5	Identifier	Dublin Core	An unambiguous reference to the resource within a given context.	Recommended best practice is to identify the resource by means of a string or number conforming to a formal identification system. Examples of formal identification systems include the Uniform Resource Identifier (URI) (including the Uniform Resource Locator (URL), the Digital Object Identifier (DOI)
6	Provenance	Dublin Core	A statement of any changes in ownership and custody of the resource since its creation that are significant for its authenticity, integrity and interpretation.	The statement may include a description of any changes successive custodians made to the resource.
7	Rights	Dublin Core	Information about rights held in and over the resource.	Typically a Rights element will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights. If the rights element is absent, no assumptions can be made about the status of these and other rights with respect to the resource.
8	objectIdentiferType	PREMIS	A designation of the domain within which the object identifier is unique.	Identifier values cannot be assumed to be unique across domains; the combination of objectIdentifierType and ojectIdentifierValue should ensure

				uniqueness. Value should be taken from a controlled vocabulary.
9	objectIdentiferValue	PREMIS	A designation used to uniquely identify the object within the preservation repository system in which it is stored.	The value of the objectIdentifier.
10	preservationLevel	PREMIS	A value indicating the set of preservation functions expected to be applied to the object.	Some preservation repositories will offer multiple preservation options depending on factors such as the value or uniqueness of the material, the “preservability” of the format, the amount the customer is willing to pay, etc. Value should be taken from a controlled vocabulary.
11	ObjectCategory	PREMIS	The category of object to which the metadata applies.	Preservation repositories are likely to treat different categories of objects (representations, files, and bitstreams) differently in terms of metadata and data management functions. Value should be taken from a controlled vocabulary.
12	messageDigestAlgorithm	PREMIS	The specific algorithm used to construct the message digest for the digital object.	Value should be taken from a controlled vocabulary.
13	messageDigest	PREMIS	The output of the message digest algorithm.	This must be stored so that it can be compared in future fixity checks.
14	formatName	PREMIS	A designation of the format	Value should be taken from a

			of the digital object.	controlled vocabulary.
15	formatVersion	PREMIS	The version of the format named in formatName.	Many authority lists of format names are not granular enough to indicate version, for example, MIME Media types.
16	storageMedium	PREMIS	The physical medium on which the object is stored (e.g., magnetic tape, hard disk, CD-ROM, DVD).	The repository needs to know the medium on which an object is stored in order to know how and when to do media refreshment and media migration.
17	eventIdentifierType	PREMIS	A designation of the domain within which the event identifier is unique.	For most preservation repositories, the eventIdentifierType will be their own internal numbering system. It can be implicit within the system and provided explicitly only if the data is exported.
18	eventIdentifierValue	PREMIS	The value of the eventIdentifier.	
19	eventType	PREMIS	A categorization of the nature of the event.	Categorizing events will aid the preservation repository in machine processing of event information, particularly in reporting. Value should be taken from a controlled vocabulary.
20	eventDateTime	PREMIS	The single date and time, or date and time range, at or during which the event occurred.	Any date/time convention may be used, as long as it is consistent and can be translated into ISO 8601 for export if necessary.
21	agentIdentifierType	PREMIS	A designation of the domain in which the agent identifier	Value should be taken from a controlled vocabulary.

			is unique.	
22	agentIdentifierValue	PREMIS	The value of the agentIdentifier.	May be a unique key or a controlled textual form of name.
23	MediaFormat/BitRate	MPEG-7	Indicates the nominal bit rate in bits/s or kbps of the audio or video instance.	e.g., 64, 128, 256, etc. Use the attributes *minimum* and *maximum* to record the minimum and maximum numerical value for the BitRate in cases of variable bit rate.
24	MediaInformation/MediaProfile/MediaFormat/Duration	MPEG-7	Duration of the Audio or Video content. The elapsed time of the entire file.	Use ISO 8601 syntax; see http://www.w3.org/TR/NOTE-datetime.
25	Size	PREMIS	Indicates the size, in bytes, of the file where the video instance is stored.	Optional. Not strictly needed for preservation purposes. There are better ways of finding out whether you've got the whole resource (e.g. by doing checksums). It is useful for dissemination purposes. For example, for knowing how long something will take to download. But is not essential.
Essential additional elements for moving images.				
26	VisualCoding/Frame/@aspectRatio	MPEG-7	The desired aspect ratio of the image on screen.	e.g. 4:3, etc. Some files produced for display on non-square-pixel monitors have a desired aspect ratio that differs from the ratio of horizontal to vertical pixels.
27	VisualCoding/Pixel/@bitsPer	MPEG-7	The number of bits of sample depth.	e.g. 8, 24, etc

28	VisualCoding/Frame/@rate	MPEG-7	The number of frames per second at which the video source item was digitized.	e.g. <Frame rate="25"/>
29	VisualCoding/Format/Frame/@ width	MPEG-7	The horizontal size of the video frame measured by number of pixels.	e.g. <Frame height="288" width="352"/>
30	VisualCoding/Format/Frame/@ height	MPEG-7	The vertical size of the video frame measured by number of pixels.	e.g. <Frame height="288" width="352"/>
31	VisualCoding/Pixel/@resolution	MPEG-7	Resolution of digital video source item expressed as horizontal lines.	
32	Sampling	VideoMD	The video sampling format (in terms of luminance and chrominance).	e.g. 4:2:0, 4:2:2, 2:4:4
33	Scan	TV-Anytime	An indication whether the digital video item is scanned in an interlaced or progressive mode.	Can take on the values: *interlaced* or *progressive*.
34	VisualCoding/Format/Name	MPEG-7	The encoding method of the visual component of a resource. Also described as the codec.	Defined by a controlled vocabulary, e.g. MPEG-1 video
35	Sound	VideoMD	Indication of the presence of sound in the video file.	Can take on the values: *yes* or *no*. If the value *yes* is selected, then the video file will also be associated with an instance of audio metadata. This element may or may not be required depending on the metadata wrapper (e.g. MPEG-

				21, METS) used.
Essential additional elements for sound.				
36	AudioCoding/Format/Name	MPEG-7	The encoding method of the sound component of a resource. Also described as the codec.	Defined by a controlled vocabulary.
37	AudioCoding/AudioChannels	MPEG-7	The number of channels of audio	e.g. 1, 2, 4, 5.
38	AudioCoding/bitPer	MPEG-7	Number of bits per audio sample.	e.g., 16, 20, 24, etc.
39	sampling_frequency	audioMD	Rate at which the audio was sampled.	Expressed in kHz, e.g., 22, 44.1, 48, 96, etc.
40	audio_block_size	audioMD	Size of an audio block in bytes.	Large signed number.
41	first_sample_offset	audioMD	Location of the first valid sound byte in the file.	Large signed number.
42	first_valid_byte_block	audioMD	Location of the first valid sound byte in the block.	Large signed number.
43	last_valid_byte_block	audioMD	Location of the last valid sound byte in the block.	Large signed number.

Summary of recommendations for future action

- **Further research should be undertaken to compare the suitability of MPEG-21 and METS as containers for the metadata required for moving images and sound resources.**
    - **The output of the research should include case studies and exemplars.**
- **Efforts should be made to create or modify metadata extraction tools (e.g. JHOVE, NLNZ Metadata Extractor Tool) so that they can automatically generate standardized and compatible metadata for audio and video formats.**
    - **Such tools should adopt elements from established namespaces (e.g. PREMIS and DC) and avoid the use of application-specific schemas.**
    - **Where possible such tools should be capable of batch processing a large number of files within a single digital object and generating appropriately structured XML metadata for each.**
    - **Where appropriate, the XML output should be into a directory structure that mirrors the original in order to maintain the validity of the original resource.**
- **Work should be undertaken to assess the feasibility of automatic extraction of subject keywords from audio and video formats using pattern recognition software.**
- **Scenarios for integrating the manual and automatic production of metadata for moving image and sound resources should be investigated further and possible workflow issues examined.**

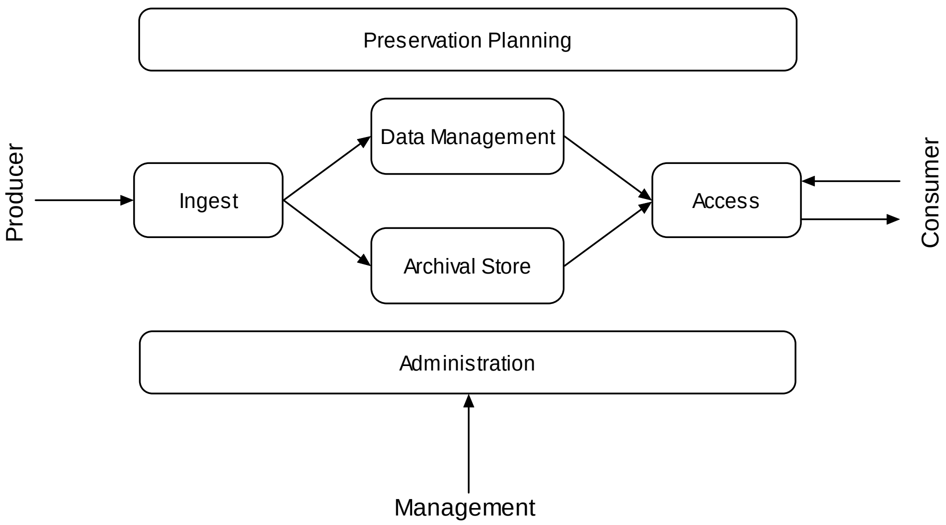
Chapter 8: Life Cycle and Organisational Models

A strategic approach to managing the life cycle of digital collections has been broadly advocated and a significant amount of work has been undertaken over the last 10 years or so to identify and map the life cycle of digital objects. The concept of the life-cycle is a useful tool that allows us to identify the key events that take place in the life of a digital object, to identify the actions that we need to take at these key events, and the supporting policies and documentation that are needed to underpin these events and actions. It allows us to plan for the long-term preservation of digital objects in a constructive and time-based fashion, reflecting the more active, interventionist methods of preservation that are required for digital objects, and it allows us to construct organisational models that can support these activities.

Before presenting the life-cycle and organisational models developed for audio-visual resources it is useful to have a more detailed understanding of the functional requirements of a preservation repository for such resources – one that is capable of preserving digital audio-visual resources in the long-term – and what attributes it would need to possess to be regarded as a secure place of deposit for long-term preservation of digital audio-visual resources to inform our thinking.

8.1 Functional requirements: OAIS

The functional requirements for the preservation of digital information have been the focus of considerable attention and the *Reference Model for an Open Archival Information System (OAIS)* (Consultative Committee for Space Data Systems [CCSDS], 2002) has become the accepted standard.



8.2 OAIS functional entities (simplified)

Source: Based on Figure 4-1 in CCSDS, 2002, p. 4-1

The OAIS functional model, shown above, identifies the main tasks that any type of repository must perform in order to secure the long-term preservation of digital material. The model defines six main functional entities that describe the activity of a digital repository as a flow of digital material, from the arrival of new material in the repository, its storage and management, and through to its delivery to a user (consumer).

Ingest

Ingest includes the physical transfer of files and the legal transfer of rights through the signing of licences or other agreements that establish the OAIS repository's right to maintain the ingested material. During ingest, descriptive information (resource discovery metadata) should be created to describe the material, and the submitted files are checked to ensure that they are consistent with the OAIS repository's data formatting and documentation standards. This may include tasks such as file format conversions or other changes to the technical representation and organisation of the submitted material.

Archival Storage

This functional entity is concerned with the bit storage of the submitted digital material including tasks such as backup, mirroring, security and disaster recovery.

Access

All the services and functions needed for users to find and access the contents of the repository.

Data Management

Data management involves the collection, management and retrieval of both resource discovery, administrative and preservation metadata.

Administration

The administration functional entity involves the entire range of administrative activities that an archival organisation should undertake. Notable tasks include managing, monitoring and developing the repository's software systems, negotiating submission agreements with producers (authors), and the establishment of policies and standards for the repository.

Preservation Planning

This functional includes four sub-entities associated with identifying preservation risks and developing plans to address them:

Monitor Designated Community – the designated community is an OAIS term that refers to the community of stakeholders who have an interest in the content of the repository. An OAIS repository needs to monitor its designated community's adoption of new technology, and other trends that may affect preservation of the community's digital output. In the case of digital images, this would refer in part to the user communities identified in Chapter Three.

Monitor Technology – The monitor technology function ensures that the OAIS repository is constantly aware of technological changes that may render its current holdings obsolete or difficult to access.

Develop Preservation Strategies and Standards – The development of strategies and standards for preservation that are informed by the current and future requirements of the producers and consumers of the OAIS repository.

Develop Packaging Designs and Migration Plans – This function accepts standards for file formats, metadata and documentation (generated as part of the administration functional entity) and creates tools or defines techniques that apply these standards to submissions.

8.3 Trusted digital repositories (TDR)

Over recent years work has been carried first by RLG and OCLC, and subsequently by RLG and NARA (further information can be found at: http://www.rlg.org/), to develop a model and checklist for the attributes of a trusted digital repository. The concept of TDR is intended to provide an environment of trust between content owners and creators and those who are be responsible for its long-term preservation, similar to that which exists between authors, publishers and libraries. A trusted digital repository is one whose mission is to provide reliable, long-term access to managed resources now and in the future. The organisational structure behind such a trusted digital repository is regarded as flexible. What is important is the ability to meet expectations and to be able to demonstrate the following attributes:

8.3.1 Organisation

i. governance and organisational viability: a repository must demonstrate an explicit, tangible, and long-term commitment to compliance with prevailing standards, policies and practices
ii. organisational structure and staffing: a repository must have designated staff with requisite skills and training and must provide ongoing development
iii. procedural accountability and policy framework: a repository must provide clear and explicit documentation of its requirements, decisions, development and action to ensure long-term access to digital content in its care
iv. financial sustainability: a TDR should be able to prove its financial sustainability over time
v. contracts, licences and liabilities: a repository has and maintains appropriate contracts and deposit agreements

8.3.2 Repository functions, processes and procedures

These will differ between repositories but the key issues are that the policies, procedures, functions and processes are properly documented and available for public scrutiny, and that the repository is following best practice as outlined in the OAIS functional model and as specified for the content types with which it is working.

i. ingest/acquisition of content
ii. archival storage and management of archived information
iii. preservation planning, migration, and other strategies
iv. data management
v. access management

8.3.3 The designated user community and the usability of information

i. documentation: the repository has a definition of its designated user community/ies and what levels of service it expects, and that this is a public document
ii. descriptive metadata: a repository must articulate minimum metadata requirements to enable the user community to discover and identify materials of interest
iii. use and usability: access and delivery options are open and transparent and fit for purpose, and that ensure that all legal aspects are complied with
iv. verifying understandability: to have mechanisms in place that ensure the repository obtains, manages, and makes available information in forms that allow digital objects to be understandable and usable over time

8.3.4 Technologies and technical infrastructure

i. system infrastructure: the repository must provide a secure and trusted infrastructure to manage its digital content
ii. appropriate technologies: a repository should use strategies and standards relevant to its designated communities and its digital technologies
iii. security: the system must be secure and protected, including m-m interaction and human – machine interaction

8.4 Digital moving image and sound life-cycle model: OAIS and TDR

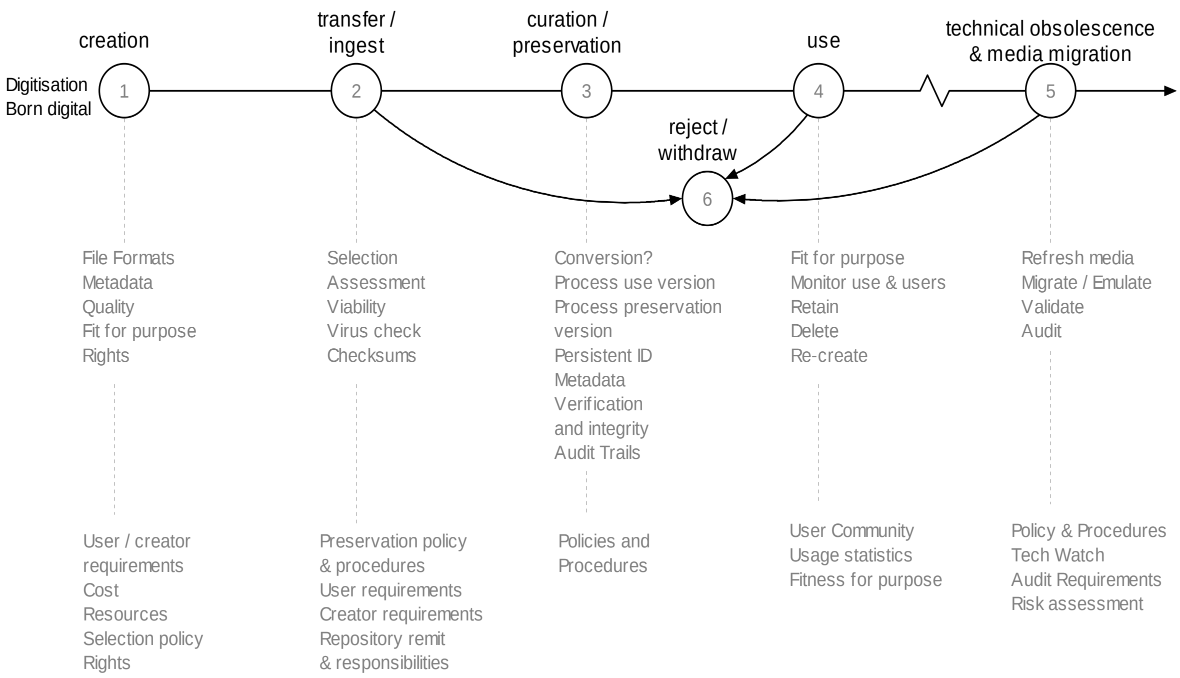
The OAIS model offers a functional model for any digital preservation system. OAIS itself does not provide a practical implementation of the framework however, and institutions will adopt their own locally appropriate workflow solutions. Similarly, the Trusted Digital Repositories framework provides a checklist of things a repository should do if it wishes to be taken seriously in its preservation efforts, but does not specify how that might relate to workflows.

The life-cycle model presented here attempts to incorporate a sense, in a somewhat simplified schematic model, of how the life-cycle meets OAIS and TDR. It identifies the key events that take place, the activities that should take place at those events, and the policies and processes which underpin them. Thus the top layer conveys the idea that this is part of continuum where key actions points are identified; the second layer outlines the actions and decisions that are likely to made at this time; and the third layer identifies the requirements, policies and processes on which these decisions are likely to be based. Readers will hopefully recognise elements from both OAIS and the TDR.

What follows is intended to bring together much of what has gone before in the previous chapters of this report by highlighting the key issues, challenges and decisions to be taken at key points in the digital image life-cycle. It is also intended to provide practical and useful advice for those responsible for the preservation and curation of digital moving image and sound collections. The model presents six key events that may occur in the full lifecycle of digital moving images and sound:

1. Creation
2. Transfer/Ingest
3. Curation/Preservation
4. Access and Use
5. Technical Obsolescence and media migration
6. Withdraw / reject

At each key event a range of actions are, or should, be taken that will affect the future of the digital audio-visual resources. Many of these actions will affect the longer term survival of the moving image and sound resources and will determine if they are merely a collection of bits, or digital objects that remain fit for purpose and usable.



Life-cycle model for moving images and sound

8.4.1 Creation

The ability to approach the curation and preservation of digital audio-visual resources with some degree of confidence starts with the decisions taken at the point of creation or capture. It is at this point that decisions will be made on such concerns as the quality of the moving images, the colour resolution, frame rates, bit-depth, number of audio channels, the metadata that is to be captured and created, the technical standards and metadata standards to be used, etc. – or not used, as the case may be. It is here where decisions about software and hardware platforms are made that may impact on how embedded the moving images and sound files are within a particular platform. It is also here that rights are cleared (or should be cleared) and that care must be taken to ensure that any licence agreement embody a right to preserve the audio-visual resources – that is, to take the necessary preservation actions to ensure future accessibility.

These decisions are likely to be driven by user and/or creator requirements, the costs involved, and the resources in terms of expertise, people and equipment that are available. Decisions may also be driven by a selection policy. It is important that all those who create and capture digital audio-visual material record information on the decision making process in order that these decisions can feed into the ongoing sustainability and preservation of the resources. To this end **JISC may wish to consider if it has a role in defining benchmarks against which digital moving imaging and sound projects measure their compliance with standards and best practice; and tools to capture their project processes.**

The type of information that should be recorded at this stage includes:

- Technical and metadata standards used
- Details of any pilot or feasibility testing
- The process and methodology by which the moving images and/or sound were captured or created, including equipment used, quality standards applied, and software used
- Quality assessment procedures
- Subsequent management or transformation actions, such as de-compression

- Naming system used (unique IDs etc.)
- Rights and legal actions undertaken

Those responsible for preservation may or may not have control over the creation or capture of digital moving image and sound collections. If they don't then the preservation repository should ascertain the standards and methods used to create the moving images and sound files, and to assess what actions might be required to bring these resources into the managed environment of the repository and, if necessary, to assess the likely costs. This assessment process will feed into decisions made at transfer and ingest, and indeed into the longer term curation of digital audio-visual collections. Depending upon the decisions made at this stage, the suitability and viability of the moving images and sound for longer term access and preservation will be determined. Wherever possible, repositories should seek to liaise with content creators through the process of creation.

8.4.2 Transfer / Ingest

Once a collection has been selected for long term preservation (and not all audio-visual material necessarily will be so selected, depending upon the decisions made at creation, and the selection policy of the repository) then a range of actions will need to be undertaken to ensure the successful transfer to the preservation environment. These will need to be underpinned by a range of policies including the selection policy of the repository (informed by its role and responsibilities), and the preservation policy which describes the requirements for content to be ingested. These policies are likely to be driven by its designated user community, or in some cases the designated creator community – repositories policies may be driven by a remit to collect *from* a community, or by a remit to collect *for* a community. Acquiring some understanding of the practices of the content creators, and of subsequent content users, as well as understanding the designated community for the repository or institution, can only help the repository do a better and more efficient job.

Preservation policies should define the collecting community and the user community, the range and type of material it seeks to preserve, and its preservation responsibilities. It should state the levels of service offered and any dependencies involved, such as requiring deposit of particular formats – these might be separated into recommended and preferred. Repositories may well wish to exclude certain types of content of formats if they do not reach the accepted standards, or to limit the preservation actions that may be undertaken. The policies should also outline the metadata required to accompany the audio-visual collection.

The point of transfer is the first, and often only, point of substantial contact between the content owner and the preservation repository. This is a crucial opportunity to provide feedback to the creator that may improve the preservation characteristics of later submissions. It is the repository's best opportunity to collect resource discovery and administrative metadata needed to manage the moving images and sound in the long-term, but more importantly, it is the repository's only real chance to establish a formal legal agreement to govern the long-term care of the audio-visual resources.

On transfer for preservation the following steps should be taken:

- Check that the audio-visual files are in an acceptable preservation format – if not they need to be migrated to an acceptable format if the producer has technology available to do so;
- Carry out a virus check to ascertain that the files are virus-free before transfer; if files are infected they should be replaced with uninfected versions;
- Create checksums for the individual moving image and sound files; if the files are to be transferred on a portable medium (e.g. CDs, DVDs) create a checksum for each instance of the medium;

- Prepare the documentation and metadata that must accompany the audio-visual files and which are necessary to ensure that they remain accessible and useable over time;
- Digital files are ingested by the preservation service initially onto a staging/processing server; this includes but is not limited to the following steps:
    - media and file readability check
    - check file counts/formats/names against any documentation provided
    - compare checksums to ensure no data has been corrupted
    - check documentation is adequate for data provided
    - copy the data to appropriate place on server
    - data validation and consistency checks
    - create receipt to send depositor

It is at this point that a collection of audio-visual resources might be rejected if they don't meet the necessary quality or other standards. It is important to bear in mind that not everything must or should be preserved.

8.4.3 Curation / Preservation

Once the initial set of actions have been undertaken at transfer then the process of curation proper may start. It is at this point that decisions must be taken on the level of format normalisation that occurs, the number of copies that will be created and preserved, the assignment of persistent identifiers, and verification and integrity checks that will be carried out. It is also important at this stage to define access rights for future use and the method of providing access.

It is recommended that this process is carried out on dedicated server separate from the primary preservation repository. Actions to be carried out include:

- Carry out any preservation actions that are necessary to ensure long-term usability of the audio-visual resources, for example any migration or normalisation actions that might be necessary;
- Validate any conversions/migrations that have been performed;
- Assign persistent identifiers
- Assign version number(s)
- Prepare dissemination version(s) of image as appropriate;
- Ensure that preservation actions are adequately documented (prepare an audit trail);
- Move preservation copies to preservation repository and check for data integrity;
- Copy original (source) bitstreams to the preservation server;
- Ensure file and directory structures are consistent;
- Carry out ongoing verification of the objects in the repository using checksum techniques.

Ideally, the preservation repository will be a stand alone system not connected to any external network. This is the best way of ensuring that data authenticity, integrity, and security is not compromised.

It is recommended that repositories responsible for long-term preservation create a set of policies and procedural documents that provide in some detail the processes they will undertake to ensure future access to the audio-visual collections. The procedures documents (what might be termed ingest manuals) will document the actions that will be undertaken for each set of moving images and sound, and will specify the audit trail to be documented during the ingest/curation process. It is important that clear and transparent procedures are followed at all stages. This will ensure that it is always possible to retrace the steps taken should that prove necessary in the future.

8.4.4 Access and use

The primary aim of preservation is to ensure continued access to digital content. The key concept here is fitness for purpose. There is little point (other than historic or to allow replication of research) in continuing to provide access to a digital video or set of audio files in a way that no longer meet the needs of the community of users. It is therefore to be recommended that repositories monitor access and use of their collections, and the behaviour of their users, especially with regard to the software they are using to manage their images for research and/or teaching and learning purposes. It may well become necessary to create new delivery versions (AIP) for users. In these cases it is recommended that in order to maintain the integrity of the delivered version of the audio-visual resources, repositories use the master version held in the preservation system to re-create a new version for access and delivery. This process should be properly recorded and verified as before.

Monitoring fitness for purpose in this way may also lead the repository to the decision to withdraw the moving image or sound file from its collection. If its quality and usefulness have deteriorated to such an extent that it no longer serves its purpose, then it may be necessary to withdraw a single audio-visual resource or set of resources, or to make a decision to re-create them (where this is possible).

It would be useful for JISC to commission research into when re-creation might be an appropriate strategy. This could include an assessment of the costs of continuing to maintain an inadequate audio-visual resource, against the cost of re-creating it.

Best practice would include the following practices:
Monitor use of audio-visual resources
Monitor user practices and use of software
Creation of new version of delivery versions (with appropriate integrity checks)
Re-create audio-visual resources if appropriate (suitable only for digital representations of an analogue original)
Assign version numbers as necessary
Fully document this process

8.4.5 Technological obsolescence

Technological obsolescence and media degradation are the central problems to overcome when planning for the long-term preservation of digital audio-visual collections. A process of monitoring the collections and the software and hardware in which they are contained for possible obsolescence and degradation is necessary, alongside a set of procedures which explain the necessary actions needed to overcome these. When action is deemed necessary then a programme of media refreshment, further migration or invoking an emulation process should be undertaken and fully documented. The same process of validation and integrity checking should be undertaken at all stages in this process.

Underpinning this must be a technology watch process to manage the risk as technology evolves and to keep up to date with new technologies such as new emulators which may be emerging. A process of risk assessment of the audio-visual content is necessary – understanding what content is held, how it is held, and what the risks are to it due to format or media obsolescence is essential. A Risk Assessment Policy such as that recommended by the ERPANET Risk Communication Tool would be a useful aid to this process.

A general source of technical information about digital file formats, which can be of use in risk management strategies, is the UK National Archives' PRONOM database, containing detailed information about digital file formats, including a range of audiovisual formats.[99] In addition, the JISC recently undertook a survey and assessment of sources of information on file formats and software documentation. The overall objective of the survey was to assess how the availability of file format information affects development of tools and strategies to enable digital preservation to be successfully performed. The final report of the survey was released in 2003.[100]

Technological obsolescence is the greatest threat to long-term useability of digital formats. Dealing with this is a technological and management issue that must be addressed by digital archivists, data curators, and digital librarians. **This report recommends that JISC consider collaborative work with TNA to enhance the PRONOM database and make it as comprehensive as possible, so that it functions as a widely useful and useable tool for the digital preservation community.**

8.4.6 Reject or withdraw audio-visual resources

The life-cycle model highlights three places where this process may happen, and these have been further explained in the text above. It is vital that any preservation strategy builds in the ability to reject or withdraw content at designated places in the life-cycle and that these are transparent and fully documented. Rejecting or withdrawing a collection is not a decision to be made lightly (unless there are concrete legal reasons for doing so of course) and should therefore be underpinned by well-developed policies and procedures.

When audio-visual resources are withdrawn they may be physically deleted from the repository, but the better practice already followed by many repositories, is to maintain the original resources, but to mark them as no longer available or as superseded by a newer version. A reference or link to the new version is recommended.

8.5 Organisational models

While galleries, libraries and museums undertake established roles in the preservation of analogue audio-visual resources, the allocation of responsibilities for the preservation of digital moving images and sound material is still evolving. Audio-visual collections are collected, stored, and delivered within a variety of organisational settings, some of which are better equipped than others to meet the functional and non-functional requirements for the long-term preservation of digital audio-visual materials.

Within the preservation community there is a growing awareness that in the digital world responsibility for preserving information will need to be distributed in new ways. The way forward envisioned by many is to disaggregate the tasks undertaken by a digital repository, so that not all repositories need undertake all tasks.

> Fundamental to implementing this disaggregated model is the logical separation of the content and service components…. This separation allows for distributed open access content repositories to be maintained independently of value-added services fulfilled discretely by multiple service providers. Crow (2002a)

Digital preservation could be seen as one of these ‘value-added’ services, and could be provided in a number of ways, as suggested in the JISC Continuing Access and Digital Preservation Strategy 2002-5 (Beagrie, 2002, p. A13). Preservation of audio-visual materials could take place within an institutional repository, or be part of a collaborative service with a group of other repositories, or could be undertaken by an external agency or a national service of some kind.

---

[99] Available at: http://www.nationalarchives.gov.uk/pronom/#.
[100] Final report available at: http://www.jisc.ac.uk/uploaded_documents/FileFormatsreport.pdf

In the JISC e-infrastructure, a number of organisational models for the provision of archival e-print repositories are in the process of being developed. These models are not mutually exclusive, and disaggregated provision of archival repository functions does not necessarily require the establishment of national services. Institutions, or consortiums of institutions could provide their own preservation services, while commercial solutions could also play a role.

8.5.1 Single institution audio-visual repository

Audio-visual repositories operated within larger institutions may be in a position to undertake the full range of activities to meet OAIS functional requirements and the attributes of a TDR. Systems such as DSpace and Fedora offer institutions an off-the-shelf open source solution that could be implemented and used. However, the skills required to meet the TDR requirements are significant and this route requires commitment of both people and money if it to be successful. This is particularly true if a repository takes on responsibility for the preservation of broadcast digital bitstreams (e.g. digital TV).

8.5.2 Audio-visual repository with specialist support

Otherwise self-contained audio-visual repositories may need, or prefer, to call upon external services with specialist expertise in digital preservation. Services such as the CCLRC, the JISC-funded AHDS, or ESDS, further supported by the DCC, could provide these services.

8.5.3 Audio-visual repository with outsourced preservation services

Following the model currently under development by the SHERPA DP project (http://www.ahds.ac.uk/about/projects/sherpa-dp/index.htm) preservation planning and activity could be outsourced to an external organisation which then works in partnership with the audio-visual repository or a group of audio-visual repositories to provide an overall OAIS compliant and TDR compliant service.

8.5.4 Outsourced audio-visual repository services

An individual academic, project, interest group or institution could make use of an external repository service. More than one supplier of image repository services may emerge, such as those on offer for the e-prints. Outsourcing audio-visual repository services could prove a cost-effective solution for the audio-visual collections of smaller institutions, projects and individual academic staff. More complex audio-visual collections such as those consisting of digital television broadcast material or outputs of ‘pro-sumer’ digital movie cameras, present larger challenges and outsourcing solutions may not have sufficient capabilities to manage and preserve such audio-visual material.

8.6 Disaggregated models

Given the complexity of preserving digital audio-visual collections, and the scarcity of a pool of skilled staff trained in digital moving image and sound preservation it is recommended that JISC consider the development of a disaggregated but networked model for the long-term preservation of digital audio-visual material of all kinds. We would also recommend that this is framed around an investigation of the use of grid technologies and tools such as Storage Resource Broker (SRB). SRB offers an opportunity to curate and preserve digital content across a network of distributed computers, and has the potential of offer a solution to the large size and complexity of audiovisual materials.

The disaggregated model recommended here also suggests separating out issues to do with the content, particularly liaison with the creators of digital moving images and audio resources, the collection and retention policy, and assessment of audio-visual collections submissions, from the technical management and delivery of the audio-visual – the infrastructure services.

The former is likely to require subject expertise and experts in order to fully appreciate and understand the requirements of both start and end users of digital audio-visual resources, and to be able to assess the quality and form of the content. Subject expertise is also required to understand metadata and associated documentation requirements and, most crucially, to assess fitness for purpose of the resources. These experts should understand the technical requirements for managing and providing access to audio-visual collections, and be able to convey these requirements to systems developers and engineers, but they need not have practical experience themselves in developing technical systems or of data management and preservation.

The Infrastructure Services, namely the long-term data management, archival storage and access functions that are required to operate the repository, and the systems development and management that supports these activities could be divided up in various ways. Most obviously, there are many commercial and non-commercial organisations capable of providing the preservation functions. In an institutional setting, it may be that computer services will take on responsibility for archival storage, but that ingest, data management and access might be controlled by library services. Alternatively, it may be that groups of subject specialists take on responsibility for content matters, working with a repository service provider who provides the technical infrastructure services that support access and delivery and long-term preservation.

At one extreme all the constituent parts may be located in the same organisation, although perhaps spread across a number of sections of that organisation. At the other extreme, they may be distributed across multiple organisations. The need to ensure that work practices are compatible, communications and management are efficient, and services are technically interoperable will place some practical limits on the disaggregation of an image repository, but there is still considerable scope for a variety of solutions to emerge.

Infrastructure and specialist subject support services may be provided by a single organisation in some situations. The OCLC Digital Archive (http://www.oclc.org/digitalpreservation/services/archiving/digital/) offers this type of unified service, while within the JISC e-infrastructure existing services such as the AHDS, the Economic and Social Data Service (ESDS), EDINA or MIMAS could provide a similar combined service.

Chapter 9: Assessing Costs for Preserving Digital Moving Images and Sound

"What is the toughest thing about making a film? Putting in the little holes. The sprocket holes are the hardest thing to make. Everything else is easy, but all night you have to sit with that little puncher and make the holes on the side of the film. You could faint from that work. The rest is easy: the script is easy, the acting is easy, the directing is a breeze...but the sprockets will tear your heart out." --- Mel Brooks, comedian, writer, film director[101].

9.1 Introduction

The preservation world understands and acknowledges that preserving digital content will have costs and that these may be substantial. What is not clear is:

- who in the chain of creation, stewardship and use will bear these costs,
- just how substantial might these costs be,
- whether any one costing model will be applicable for wide usage, or
- whether currently published means of assessing costs are applicable to moving images and sound?

Brain Lavoie of OCLC, inc. is a key commentator on the costs associated with digital preservation and he states:

> Stewards of the scholarly and cultural record recognize that their curatorial responsibilities now extend beyond familiar forms of materials… If we are not to tolerate gaps in society's memory in the future, we must intervene in the present to secure the long-term availability of culturally significant digital materials… But the best of intentions are almost always reined in by practical constraints. In the case of digital preservation, it is economic realities that intrude[102].

Aschenbrenner and Kaiser suggest that "costs of a digital repository are hard to calculate due to the lack of hands-on data from other initiatives… the lack of experience with digital preservation costs obstructs a complete picture"[103].

Whilst there is a strong growth in the general application of digital preservation technologies and techniques we have yet to see a maturing of the economic basis for this activity. Mainly this is due to the costing assumptions used by institutions differing significantly from each other. Differences include the means of estimating: the cost of labour and infrastructure, the cost of investment and expected returns, and how costs are allocated between distinct activities that share resources.

Most of the published material relates to cost models, rather than fully fledged business models. A cost model provides a framework in which all costs can be recorded and allocated or apportioned to specific activities. However, a business model shows how the service functions, the method by which it can sustain itself and involves both strategy and implementation.

---

[101] Quote taken from *Video Aids to Film Preservation*, a useful resource of guides to preservation of film, at www.folkstreams.net/vafp/

[102] Brian F. Lavoie, *The Costs of Digital Preservation*, chapter from Digital preservation and the future of culture, Marilyn Deegan and Simon Tanner (Editors), to be published in 2006 by Facet Publishing.

[103] Andreas Aschenbrenner, Max Kaiser (2005) White Paper on Digital Repositories, March 2005 www2.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf

Reducing costs – automation and storage?

Two elements often assumed in the literature as being key to solving digital preservation's perceived high cost relate to reducing storage costs and the automation of tasks to remove as much human cost from the equation as possible. However, even if storage continues to drop in price and core tasks of digital preservation, such as ingest, normalization, validation, monitoring, repair, migration, etc. were fully automated there would still be costs to be met. Sizing these costs has proved not to be easy and especially so for moving images and sound,

Automation remains an aspiration at present and there is little in the literature to suggest that the costs associated with digital preservation will be significantly reduced through full automation in the near future. For instance, as Aschenbrenner states in his white paper "the generation of adequate and sufficient metadata may be one of the most costly tasks of digital repositories. Still metadata generation requires considerable manual input, although well-designed workflows, the reuse of metadata from other sources, and automation could greatly streamline efforts and reduce costs". For moving images and sound, where the content of the digital object is not easily interpreted by a machine but most readily understood by a human, there are still high costs to be expected for metadata generation. Generating metadata to describe content in moving image or sound content, where none currently exists, will not easily be automated in the near future and thus remain variable in cost and relatively expensive.

Volume of data, validity of description and the number and nature of digital objects to be managed are more important indicators of the cost elements to address rather than levels of automation. As Steve Chapman of Harvard University states:

> managed storage costs are not fixed, but arrived at collection-by-collection by judicious decision-making. The choice of repository, the scope of service, the repository pricing model, and owner's decisions regarding formats, number of items, number of versions, and number of collections to deposit: all are potential variables, and therefore instruments, to negotiate for affordable prices for managed storage services from centralized repositories. These variables apply equally to traditional and digital repositories, and in both cases one potentially finds that some formats (content types) are more favoured than others[104].

The cost of storage is important for digital moving images and sound. They tend to make greater demands upon the storage space available per object than other popularly used object types like text and images. Chapman has elucidated this issue very clearly through his groundbreaking work looking at comparative costs for physical storage of originals versus digital storage of surrogates. He states that:

> Annual pricing per gigabyte, for example, privileges relatively small objects such as datasets and structured text (e.g., XML), but potentially creates a fiscal barrier to manage comparable numbers of sound recordings, videos, or other large objects. Repositories that strive to become safe harbours to shelter as much "eligible" content as possible face significant challenges in designing simple pricing models for digital preservation[105].

This builds on his previous work in 2003 comparing the Harvard Depository (which assesses costs for analog storage per billable square foot) and the OCLC Digital

---

[104] Stephen Chapman (2003) *Counting the Costs of Digital Preservation: Is Repository Storage Affordable?*, Journal of Digital Information, Volume 4 Issue 2, Article No. 178, 2003-05-07, http://jodi.ecs.soton.ac.uk/Articles/v04/i02/Chapman/

[105] Stephen Chapman, *It's money that matters in long-term preservation*, chapter from Digital preservation and the future of culture, Marilyn Deegan and Simon Tanner (Editors), to be published in 2006 by Facet Publishing.

Archive (which assesses rates per gigabyte for storage of digital objects). His work shows that "at OCLC's current per GB rates [per year] for other formats, the unit prices to preserve digital audio or moving images would be in tens ($30.76 for audio) or thousands of dollars per hour ($3,270 for moving images). (OCLC has not yet established prices for these formats). Thus, the financial and technical challenges associated with long-term preservation of these formats are significant".

Number of objects to preserve
A further factor that is most likely to drive up the costs of preserving digital moving images and sound are that there are likely to be many more versions of the same content to be preserved than with many text or image based formats. This is because digital moving images and sound demonstrate in their creation an inversion of the commonly assumed costing truism that creating the master item (e.g. high resolution TIFF image) will be more expensive than the cost of creating surrogates for end user interaction (e.g. JPEG for Web delivery). Indeed, many projects will only preserve the 'master' image file as the other surrogates for use may be automatically generated in batch processes.

The amount of bandwidth and infrastructure loading used to deliver moving images and sound are critical factors to ensure ease of use by a wide potential audience. This is often resolved by delivery solutions which include: just edited highlights of a complete work, streaming or by varying the quality through differing amounts of compression. In many cases this means that the effort made in creating a surrogate for access makes it worthy of preservation in its own right. There is a higher likelihood with moving image or sound for more than one version to be stored for the long term and this has attendant knock-on costs for ingest, storage, metadata and version control.

Copyright and IPR
As has been noted in an earlier part of this report the Intellectual Property Rights associated with moving images and sound can be complex and multi-layered. It should be expected that the costs of creating and maintaining IPR tracking records for moving images and sound will be more costly than for other media types.

Models and exemplars

PrestoSpace, a European collaborative project, is the premier activity worldwide looking at technical solutions and integrated systems for a complete digital preservation of all kinds of audio-visual collections. The project began in February 2004 and will last for 40 months, and has 35 partner institutions. PrestoSpace have carried out surveys[106] which have so far identified 17.5 million items of audio-visual material to be preserved in Europe, which may represent roughly 6 million hours or more of content. In light of the huge volumes they try to address cost issues as much as is possible whilst recognising the very varied nature of their members.

PrestoSpace report:

> The archives have forecast a spend of nearly 60 million Euros in 2004-2006. Assuming they're talking about the 500 000 items they plan to migrate, this comes out as 120 euros per item. Again using an estimate of 20 minutes per item, the cost is 360 euros per hour… As the PrestoSpace project continues, it will be very interesting to see the extent to which we can reduce 'the going rate' for archive transfers across Europe[107].

[106] 5.3 million audio only, 6.2 million video, 6 million film formats. Estimates assume an average of 20 minutes per item. See PrestoSpace Annual Report (January 2005): http://www.prestospace.org/project/deliverables/D22-4_Report_on_Preservation_Issues_2004.pdf

[107] Richard Wright (January 2005), *Annual Report on Preservation Issues for European Audiovisual Collections*, PrestoSpace, at www.prestospace.org/project/public.en.html

To put this into context it is important to note the spend detailed above includes the archival transfer of audio-visual material from physical to digital formats. It also includes the digital infrastructure, mass storage, network, workstations etc. as most of the members are in a fluid implementation phase rather than having mature structures in place.

PrestoSpace have also identified other problems for those working on archiving digital moving images and sound. Foremost is the lack of commercial capacity to carry out archival quality transfers to digital formats. Current capacity appears to provide less than 50% of the European capacity desired and thus will increase costs in this area. There is also a lack of established mass storage facilities for this type of media and rights clearance and documentation remains a barrier.

PrestoSpace have a cost model for digitisation projects, with a simplified version available as an online tool [[ref: http://prestospace-sam.ssl.co.uk/hosted/d13.2/newcalc.php]], though the full version is not publicly available as yet. It is intended to act as a framework for medium to large archives to build or review a business case and is referenced as Report D13.1 'Planning for Digitisation and Access'[108].

Finally, PrestoSpace identified a 60% gap between the budgets available for digital preservation of moving image and sound archives and the costs forecast for identified projects in the following two years; this represents at least a 25 million Euro gap.

ERPANET, has hosted many seminars and created a number of useful guidance documents that are relevant to costing of digital preservation. They have created a cost orientation document[109] to help users of it to think through the costing issues involved in digital preservation. It does not provide costing information and ERPANET state that there "is still a lack of sound costing information to build upon". This guide builds upon the key work of Sannett[110] but adds nothing unusual to standard business and accounting cost estimation practice. There is no specific moving image or sound orientation to this guide.

ERPANET were the host of a seminar on "Business Models related to Digital Preservation"[111] in Amsterdam (September 2004). This was one of the first times that differing cost models from across the world had been presented side by side and could be thus compared. The report on the event concludes that "the landscape of business models as it currently appears, is hugely complex and hardly any practical experiences are so far available to help clarify it"[112].

The **British Library**, has established a strong theoretical model which offers a lifecycle model with a formula for costs[113]. Whilst it has not been specifically tested in relation to moving images or sound, it is a holistic model which would be applicable to most formats.

---

[108] Matthew Addis and Ant Miller (2005), *The PrestoSpace Cost model, DCC/DPC Workshop on Cost Models for preserving digital assets*, London, 26th July 2005, at http://www.dpconline.org/graphics/events/050726workshop.html
[109] http://www.erpanet.org/guidance/docs/ERPANETCostingTool.pdf
[110] Sanett, S. (2003) *The Cost to Preserve Authentic Electronic Records in Perpetuity: Comparing Costs Across Cost Models and Cost Frameworks RLG DigiNews* Vol. 7 No. 4.
[111] ERPANET, Business Models related to Digital Preservation, Amsterdam, 2004, at http://www.erpanet.org/events/2004/amsterdam/index.php
[112] ERPANET (2004), FINAL REPORT Business Models related to Digital Preservation, Amsterdam, 20-22 September 2004, at http://www.erpanet.org/events/2004/amsterdam/Amsterdam_Report.pdf
[113] Shenton, H. (2005) Real Time, Deep Time, Life Time: Spanning Digital and Traditional Collections Life Cycles. In *Archiving 2005: Final Program and Proceedings of the IS&T Archiving Conference held on 26-29 April 2005 in Washington, DC*.

The **Digital Curation Centre** and the **Digital Preservation Coalition** have acted strongly to place costs and economic factors for digital preservation towards the centre of their activities and research. They jointly hosted a workshop on cost models for preserving digital assets[114] in 2005 and this provided an excellent overview of the current state of the art. The PrestoSpace presentation by Addis and Miller was the only one to specifically cover moving images or sound but many of the other papers offer some holistic lessons in costing digital preservation.

9.2 Summary

The cost elements for digital preservation of moving images and sound have many of the same features as any other digital preservation activity. These may be divided according to the OAIS model but in short are not much more advanced as general models of wide applicability than many straightforward business planning tools. As H. M. Gladney notes "digital repository literature makes surprisingly little allusion to employing professional accountants. Cost estimation is a standard subject in all professional accounting programs"[115]. There is a significant gap in the literature on costing for digital preservation in general, but in particular, digital moving image and sound are poorly served, with the notable exception of PrestoSpace's efforts.

What does mark out digital moving image and sound as different are the ways the multitude of potential formats and compressions complicate an already difficult preservation task. This variety also creates doubt and a lack of cohesion in how to approach cost modelling especially for metadata and ingest costs. Storage costs are usually higher for temporal content but the assessment of these costs is reasonably easy to achieve. End user access to the preserved media is very difficult to cost as many of the born digital formats are fixed to proprietary delivery solutions or codecs that are volatile.

[114] DCC/DPC Workshop on Cost Models for preserving digital assets, London, 26th July 2005, at http://www.dpconline.org/graphics/events/050726workshop.html
[115] Gladney, H.M. (2005) *About Digital Preservation Costs*, Digital Document Quarterly, Perspectives on Trustworthy Information, Volume 4, Number 4, at http://home.pacbell.net/hgladney/ddq_4_4.htm

Chapter 10 References

Chapter 3

Case Study: Broadcasting: An ERPANET report
http://www.erpanet.org/studies/docs/erpaStudy_broadcasting.pdf

This ERPANET report found that broadcasters held shared concerns about digitisation. These included speed of deterioration of formats, uncertainty on the stability of new formats, technological obsolescence, quality of images (mainly those derived from digitisation of analogue material), difficulty in the management of information and metadata (due to the large quantity of material), optimisation of access to and sharing of information, lack of experienced operators.

Survey of the State of Audio Collections is available on CLIR's Web site at http://www.clir.org/pubs/abstract/pub128abst.html.

This report contains a survey undertaken in 2003 by CLIR to study the state of audio recordings in academic libraries. One purpose of the survey was to inform decision makers in those libraries, as well as in funding agencies, about the scale and extent of barriers to preservation and access. Another purpose was to elicit information that would help the participating libraries assess their own readiness to preserve and provide access to their recorded-sound collections. We also hoped that survey findings would help library leaders and funders determine how best to allocate preservation funds and thereby help ensure access to historically important sound recordings. Finally, the survey was designed to raise awareness within the larger research and funding communities of the value of audio collections and to encourage institutions with important audio holdings to seek support for their collections.

IASA-TC 03 2001 The Safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy. Version 2. September 2001, Internet on-line http://www.iasa-web.org/iasa0013.htm

Hidden Treasures: The UK Audiovisual Archive Strategic Framework

http://www.bufvc.ac.uk/faf/HiddenTreasures.pdf *completed 2004*

The Museums, Libraries and Archives Council has published *Hidden Treasures: The UK Audiovisual Archives Strategic Framework.* The document provides a blueprint for the future development of audiovisual archives across the UK.

The Moving Image Archive Framework: Policies, Standards, Guidelines

ongoing

This document (a work in progress) brings together a selection of ‘standards’ for moving image archivists in nationally and internationally-produced documents, generated by the film/audiovisual community, the wider archive and heritage community and national and international bodies and governments. The Film Archive Forum advocates that future work on standards by the UK moving image archive sector needs to be informed by (i) this framework of documents (ii) a good understanding of sectoral practice, and (iii) the wider archive community. [Version 1, July 2005]
See http://www.bufvc.ac.uk/faf/MIAFv1.pdf

Ross, Seamus and Gow, Ann, 1999 'A JISC/NPO Study within the Electronic Libraries (eLib) Programme on the Preservation of Electronic Materials'. Library Information Technology Centre, London. http://www.ukoln.ac.uk/services/elib/papers/supporting/pdf/p2.pdf

Webb, Colin 2003 'Guidelines for the Preservation of Digital Heritage UNESCO Memory of the World', Internet on-line http://unesdoc.unesco.org/images/0013/001300/130071e.pdf

Wright, Richard and Williams, Adrian 2001 Archive Preservation and Exploitation Requirements. PRESTO-W2-BBC-001218, PRESTO – Preservation Technologies for European Broadcast Archives 2001, Internet on-line:http://presto.joanneum.ac.at/Public/D2.pdf

(Best practice in digitising audio provided (in general forms) by the International Association of Sound Archives, Library of Congress, Colorado Digitization Program, Research Libraries Group, AHDS, National Initiative for a Networked Cultural Heritage, MATRIX at Michigan State University, University of Texas, University of Kansas, University of Wisconsin, Cornell University (Macaulay Library of Natural Sounds) and the book entitled "Creating Digital Audio Resources: A Guide to Good Practice'.)
best so far:
IASA-TC04 Guidelines on the Production and Preservation of Digital Audio Objects August 2004 (from *Sound Directions Project*)

Useful Papers on Assessing Digital Preservation Costs

Ashley, K (1999) Digital Archive Costs: Facts and Fallacies. Paper at the DLM Forum 1999. http://europa.eu.int/ISPO/dlm/program/abst_ka_en.html

Bickner, R. 1983. "Concepts of Economic Cost" in *Key Papers in the Economics of Information*. White Plains: Knowledge Industry Publications, Inc.

British Academy. 2004. *'That Full Complement of Riches': The Contributions of the Arts, Humanities, and Social Sciences to the Nation's Wealth*.
Available at: http://www.britac.ac.uk/news/reports/contribution/pdf/contribution.pdf

British Library. 2003. "£363 Million a Year Knowledge Dividend to UK Economy – UK National Library Reveals Results of Ground-Breaking Research". Press Release.
Available at: http://www.bl.uk/cgi-bin/press.cgi?story=1399

Chapman, S. "Counting the Costs of Digital Preservation: Is Repository Storage Affordable?" *Journal of Digital Information* Vol. 4 No. 2.
Available at: http://jodi.ecs.soton.ac.uk/Articles/v04/i02/Chapman/chapman-final.pdf

Arturo Crespo, Hector Garcia-Molina: Cost-Driven Design for Archival Repositories. Joint Conference on Digital Libraries 2001 (JCDL'01); June 24-28, 2001; Roanoke, Virginia, USA. http://wwwdb.stanford.edu/~crespo/publications/cost.pdf

Mary Feeney: The digital culture: maximising the nation's investment - Chapter 5: Estimating the costs of digital preservation. A synthesis of JISC/NPO studies on the preservation of electronic material. National Preservation Office, 1999. http://www.ukoln.ac.uk/services/elib/papers/other/jisc-npo-dig/

Galloway, P. 2003. "Preservation of Digital Objects", in *Annual Review of Information Science and Technology* Vol. 38.

Stewart Granger, Kelly Russell, and Ellis Weinberger: Cost elements of digital preservation (version 4); October 2000. http://www.leeds.ac.uk/cedars/colman/costElementsOfDP.doc

Vanessa Griffin, Kathleen Fontaine, Gregory Hunolt, Arthur Booth, David Torrealba: Cost Estimation Tool Set for NASA's Strategic Evolution of ESE Data Systems. Paper presented at the CNES Symposium "Ensuring Long-Term Preservation and Adding Value to Scientific and Technical Data"; November 5-7, 2002. http://sads.cnes.fr:8010/pvdst/DATA/5-8_griffin.pdf

Hendley, T. 1998. "Comparison of Methods and Costs of Digital Preservation". Available at: http://www.ukoln.ac.uk/services/elib/papers/tavistock/hendley/hendley.html

Kenney, A. and D. Stam. 2002. *The State of Preservation Programs in American College and Research Libraries*. Washington, D.C.: Council on Library and Information Resources. Available at: http://www.clir.org/pubs/reports/pub111/pub111.pdf

Lavoie, B. 2003. *The Incentives to Preserve Digital Materials: Roles, Scenarios, and Economic Decision-Making.* OCLC Research White Paper.
Available at: http://www.oclc.org/research/projects/digipres/incentives-dp.pdf

Lavoie, B. and L. Dempsey. 2004. "Thirteen Ways of Looking At … Digital Preservation" *D-Lib Magazine* Volume 10 No.7/8.
Available at: http://www.dlib.org/dlib/july04/lavoie/07lavoie.html

Lazinger, S. 2001. *Digital Preservation and Metadata: History, Theory, and Practice*. Englewood: Libraries Unlimited.

Library Information Update. 2004. "Demonstrating Value at the BL" *Library Information Update* Vol. 3. No. 10 p. 16-17.

Mathis, M., A. Fawcett, and L. Konda. 2003. "Valuing Nature: A Survey of the Non-Market Valuation Literature".
Available at: http://www.harc.edu/harc/Projects/Nature/Reports/LitReview.pdf

NSF and Library of Congress. 2003. *It's About Time: Research Challenges in Digital Archiving and Long-Term Preservation*. Workshop Report.
Available at: http://www.digitalpreservation.gov/repor/NSF_LC_Final_Report.pdf

OCLC. 2004. *The 2003 OCLC Environmental Scan: Pattern Recognition*. Dublin: OCLC. Available at: http://www.oclc.org/reports/2003escan.htm

Quandt, R. 2003. "Scholarly Materials: Paper or Digital?" *Library Trends* Vol. 51 No. 3 p. 349-375.

Russell, K. and E. Weinberger. 2000. "Cost Elements of Digital Preservation".
Available at: http://www.leeds.ac.uk/cedars/documents/CIW01r.html

Sanett, S. 2002. "Toward Developing a Framework of Cost Elements for Preserving Authentic Electronic Records into Perpetuity" *College and Research Libraries* Vol. 63 No. 5 p. 388-404.

Sanett, S. 2003. "The Cost to Preserve Authentic Electronic Records in Perpetuity: Comparing Costs Across Cost Models and Cost Frameworks" *RLG DigiNews* Vol. 7 No. 4.

Task Force on Archiving of Digital Information. 1996. “Report of the Task Force on Archiving of Digital Information”.
Available at: http://www.rlg.org/legacy/ftpd/pub/archtf/final-report.pdf

Travers, T. and S. Glaister. 2004. *Valuing Museums: Impact and Innovation Among National Museums*. London: National Museum Directors’ Conference.
Available at: http://nationalmuseums.org.uk/images/publications/valuing_museums.pdf

Waters, D. 2002. “Good Archives Make Good Scholars: Reflections on Recent Steps Toward the Archiving of Digital Information”, in *The State of Digital Preservation: An International Perspective*. Washington, D.C.: Council on Library and Information Resources.
Available at: http://www.clir.org/pubs/abstract/pub107abst.html

Appendix: Collections and Support Bodies

MOVING IMAGE AND SOUND COLLECTIONS

US
Moving Image Archives (listed by US National Preservation Board)
http://www.loc.gov/film/arch.html

The Library of Congress (USA)'s National Preservation Board lists over 500 public moving image archives and research centres worldwide, including the Moving Image Collections (MIC), which documents moving image collections around the world through a catalogue of titles and directory of repositories, providing a window to the world's moving image collections for discovery, access and preservation. The MIC is sponsored by the Library of Congress and the Association of Moving Image Archivists (AMIA), and funded in large part by the National Science Foundation.

120 of the collections listed by the National Preservation Board are housed in Europe.

UK
BBC Research and Collections
http://www.bbcresearchcentral.com/

The BBC has millions of hours worth of radio and television output and sound effects. It also archives key documents, commercial and sheet music. Archive access vary according to requirements and purpose. Public and research access to BBC film, television and audio material is generally available through the British Film Institute research viewing service and the British Library Sound Archive

BBC Nation on Film
http://www.bbc.co.uk/nationonfilm

Nation On Film is a site containing online archive video clips from the 20th century. It includes old newsreels, documentaries, as well as home movies. These videos have been drawn from a variety of sources including the Northern Region Film and Television Archive and the BBC's own archive.

BBC/Open University Nation on Film web site
http://www.open2.net/nationonfilm/index.html

A companion web site to the BBC/Open University TV series 'Nation on Film'. The site explores issues raised in the series, looking at "the untold story of how film both recorded and contributed to social change in the 20th century".

BBC Scotland on Film
http://www.bbc.co.uk/scotland/history/scotlandonfilm/index.shtml

Scotland on Film is a collection of online clips from film, television and radio archives. The archive film extracts are courtesy of Scottish Screen Archive - Scotland's Moving Image Collection.

Bill Douglas Centre
http://www.billdouglas.org

The Bill Douglas Centre for the History of Cinema and Popular Culture is the University of Exeter's newest research initiative in the Arts and Humanities. The Centre contains both a public museum and an academic research centre, housing one of Britain's largest public collections of books, prints, artefacts and ephemera relating to the history and prehistory of cinema.

Broadcasters' Audience Research Board
http://www.barb.co.uk

BARB (Broadcasters' Audience Research Board) is the primary provider of television audience measurement in the UK and is responsible for providing estimates of the number of people watching television. This includes which channels and programmes are being watched, at what time, and the type of people who are watching at any one time.

British Artists' Film and Video Study Collection
http://www.bftv.ac.uk/avantgarde/

The British Artists' Film and Video Study Collection, funded by the AHRB Centre for British Film and Television Studies, houses a large collection of videotapes, documentation and visual material by British artists, which is accessible to researchers. A major project of the collection has been the establishment of an extensive online database.

British Library Sound Archive
http://www.bl.uk/collections/sound-archive/nsa.html

One of the largest sound archives in the world. The Sound Archive holds over a million discs, 185,000 tapes, and many other sound and video recordings. The collections come from all over the world and cover the entire range of recorded sound from music, drama and literature, to oral history and wildlife sounds.

British Pathé
http://www.britishpathe.com

This site allows free access to preview items online from the entire 3500 hour British Pathé Film Archive which covers news, sport, social history and entertainment from 1896 to 1970.

British Universities Newsreel Database
http://www.bufvc.ac.uk/databases/newsreels/index.html

160,000 records of British cinema newsreel production, from 1910 to 1979.

Collect Britain
http://www.collectbritain.co.uk

The British Library's largest digitisation project to date. At the site you can view and hear a staggering 100,000 images and sounds from the British Library's world-renowned collections of maps, prints and drawings, photographs and documents, and rare early sound recordings from around the globe.

Film Institute of Ireland, Irish Film Archive
http://www.fii.ie/archive/index.asp

The Irish film Archive acquires, preserves and makes permanently accessible Ireland's film and audiovisual heritage. Its holdings include an extensive collection of film and magnetic tape, film stills and posters, a paper archive, film equipment and memorabilia.

Imperial War Museum Sound and Film Archive
http://collections.iwm.org.uk/server/show/nav.00g

The Imperial War Museum Sound and Film Archive contains 120 million feet of cine film, 10,000 hours of videotape and 36,000 hours of historical sound recordings. The Imperial War Museum actively digitises their collections

Media Archive for Central England (MACE)
http://www.nottingham.ac.uk/film/mace/

MACE is based at the University of Leicester (from October 2006 – before then it was based at the University of Nottingham) and is responsible for collecting, preserving and making accessible the moving image heritage of the East and West Midlands.

Mitchell & Kenyon film collection
http://www.bfi.org.uk/collections/mk

For around seventy years, 800 rolls of early nitrate film sat in sealed barrels in the basement of a shop. Now miraculously rediscovered and undergoing restoration, this amazing visual record of everyday life in Britain at the beginning of the twentieth century is the most exciting film discovery of recent times and promises to radically transform British film history. These web pages include information about the collection being preserved and restored by the BFI, and the BBC series 'The Lost World of Mitchell and Kenyon'. The National Fairground Archive (see below), is working closely with BFI Collections, to undertake to research into the collection.

Moving Image Gateway
http://www.bufvc.ac.uk/gateway/index.html

The Moving Image Gateway (MIG) is a service that collects together web sites that relate to moving images and sound and their use in higher and further education. The sites are classified by academic discipline, and collected within the four main categories of Arts & Humanities, Bio-Medical, Social Sciences and Science & Technology.

The National Archives Learning Curve - Onfilm
http://www.learningcurve.pro.gov.uk/onfilm.htm

Onfilm is a resource for teachers and learners where you can view clips from twentieth century films ranging from life in Britain to conflict in Vietnam.

National Fairground Archive
http://www.shef.ac.uk/nfa

Housed in the Main Library at the University of Sheffield, The National Fairground Archive (NFA) is a unique collection of photographic, printed, manuscript and audiovisual material covering all aspects of the culture of travelling show people, their organisation as a community, their social history and everyday life; and the artefacts and machinery of fairgrounds. The National Fairground Archive is also working closely with the BFI, to undertake major research into the Mitchell and Kenyon film collection.

National Museum of Film, Photography and Television

http://www.nmpft.org.uk

The Museum's renowned collection includes more than three million items of historical, social and cultural value. These include three key 'firsts': the world's first negative, the earliest television footage and what is regarded as the world's first example of moving images – Louis Le Prince's 1888 film of Leeds Bridge.

National Museum of Welsh Life
http://www.nmgw.ac.uk/mwl/collections/film/

The Museum has a collection of approx. 50 hours of 16mm film footage. Much of this is material shot during the 1970s by curatorial staff who were anxious to record a way of life that was fast disappearing. Using a hand-cranked cine camera, they made silent colour films of traditional farming techniques, foods, and crafts.

National Screen and Sound Archive of Wales
http://screenandsound.llgc.org.uk/

The Archive has a large collection of films, television programmes, videos, sound recordings and music relating to Wales and the Welsh. The Archive was established in 2001 and currently holds over 5.5million feet of film, over 250,000 hours of video, over 200,000 hours of sound recordings, and thousands of tapes, records and compact discs.

North West Film Archives
http://www.nwfa.mmu.ac.uk/

The NWFA is part of Library Special Collections at Manchester Metropolitan University. It collects both amateur and professional film and video footage featuring people, places, events and activities in Greater Manchester, Lancashire and Cheshire. The Archive collects material on any subject illustrating life in the north-west region and currently holds over 31,000 items from the mid 1890s to the present day.

Northern Region Film and Television Archive
http://www.nrfta.org.uk/about.html

The Northern Region Film and Television Archive (NRFTA) was founded in 1998 as the public-sector moving image archive serving County Durham, Cumbria, Northumberland, Tees Valley and Tyne and Wear. NRFTA core funding comes from the University of Teesside, Northern Film and Media, a government agency charged with promoting and funding moving image culture in the north-east, Gateshead Council, and a number of smaller sources including some commercial revenue. The NRFTA's holdings consist of five major film and TV collections, as well as a large number of smaller deposits from organisations and private individuals.

Northern Ireland Film and Television Commission
http://www.niftc.co.uk

The Northern Ireland Film and Television Commission is an integrated agency for the development of the film industry and film culture in Northern Ireland. The Commission's Digital Film Archive (DFA) is an access resource (*not* available over the internet) for researchers, students, historians, and anyone who has an interest in moving images in Northern Ireland. The DFA contains 55 hours of moving images about Northern Ireland from 1897-2000. There are 389 items covering drama, animation, documentaries, news, newsreel, amateur and actuality film.

RGO- Researcher's Guide Online
http://www.bufvc.ac.uk/databases/rgo.html

The RGO database has entries on almost 550 film, television, radio and related documentation collections in the United Kingdom and Ireland. It features national and regional archives as well stockshot libraries and collections held by local authorities, museums, institutions of further and higher education, industrial companies and private individuals.

Screenonline
http://www.screenonline.org.uk

Aimed at schools and libraries, the new Screenonline site aims to bring to life the history of British film and television, with hundreds of hours of digitised video, alongside thousands of stills, posters, production designs, press and publicity materials, unpublished scripts, personal papers, recorded interviews with filmmakers and actors, and interactive timelines.

Sonic Arts Research Archive
http://www.sara.uea.ac.uk/

SARA is an online database of composers and artists working with new technology and sound. It offers access to a wide cross-section of the UK Sonic Art Network's collection of compositions and published texts. It will also offer access to much of UEA's archive of electro-acoustic music in digitised form. It will offer video, papers on aesthetic and technical issues relating to electro-acoustic music and areas of 'electronic arts' activity associated with it - sonic art, soundscape, sound installation, and multimedia work.

The Ulster Folk and Transport Museum
http://www.uftm.org.uk

The Ulster Folk and Transport Museum, part of the Museums and Galleries of Northern Ireland, ranks among Ireland's foremost visitor attractions, recapturing a disappearing way of life, preserving traditional skills, celebrating transport history and has a related film collection.

JISC FUNDED MOVING IMAGE AND SOUND COLLECTIONS

Film and Sound Online (previously known as Education Media OnLine)
http://www.filmandsound.ac.uk/

The Education Media OnLine service (EMOL) became Film and Sound Online on 1 September 2006. The Film and Sound Online collections contain several hundred hours of high-quality film, video and sound material. These materials include seventeen separate collections with contents ranging from archaeology to medical documentaries and classical music files. The materials are downloadable, either in full, or as segments, and can be used in learning, teaching and research. The service is available free of charge to UK Further and Higher Education Institutions.

Lifesign.ac.uk
http://www.lifesign.ac.uk/

Lifesign received funding from JISC to establish the feasibility of delivering streaming media to support learning in the Life Sciences. Lifesign has now evolved into an ongoing service for students in UK Higher Education. Although funding from JISC has ended, Lifesign will continue to be hosted by The University of Portsmouth as an ongoing service.

AHDS MOVING IMAGE AND SOUND COLLECTIONS

Collection Title:	Institution:	Funder:	Funding Dates:	File Format	Number	Risk
The Acquisition of English Intonation by Spanish Students : Tesis Doctorals Microfityades 3674	Universitat de Barcelona, 2000			wav and mp3	lll-2456-l	Wav – low Mp3 - high
Just For Now	http://www.justfornow.net	AHRC	2001	mpeg	pa-1030-1	low
Designing Shakespeare	Royal Holloway, University of London			.mov	pa-1018-1	medium
Five Centuries	University of Glasgow	SCRAN	1997-98	mpg	pa-1001-1	low
Women's Writing for Performance	Lancaster University	AHRC		vob	Pa-womens-writing-for-performance	medium
Double Happiness	University of Bristol	AHRC		Vob (DVD)	Pa-1025-1	medium
Handing on Tradition by Electronic Dissemination	RSAMP/Edinburgh University School of Scottish Studies	JISC	2001	.mov .avi .mp3	Pa-1028-1 (hotbed)	Mov - medium; avi – medium, mp3 - high
Moving History www.movinghistory.ac.uk	South East Film and Video Archive			.avi	Pa-1031-1	medium
ADOLPHE APPIA AT HELLERA	University of Warwick	AHRC	2002-2004	.avi	Pa_adilpheappia_1	medium
Techniques for the Analysis of Expressive Gesture in Music Performance	King's College London			.mp3	Preservation-only\pa\pa-1024-1	high
French Lerner Language Oral Corpora (FLLOC)	University of Newcastle	AHRC		wav	LlI-2495-1	low

ADVISORY AND SUPPORT BODIES

ADM-HEA - Art Design & Media Subject Centre
http://www.brighton.ac.uk/adm-hea/html/home/home.html

The Subject Centre for Art, Design, Media (ADM) is part of the Higher Education Academy (HEA) established by the Higher Education Funding Councils for England, Northern Ireland, Scotland & Wales to promote high quality learning and teaching in subject communities.

AHRC Centre for British Film and Television Study
http://www.bftv.ac.uk

The Centre was created in October 2000 with funding from the Arts and Humanities Research Council, aims to enhance and extend the recent growth in high level academic research on British Film and Television. The centre also contributes to the formation of public policy in Britain and Europe in the area of film and broadcast media. They have developed a research guide entitled 'Moving History: A Guide to UK film and television archives in the public', providing detailed information UK's twelve public sector moving image archives: http://www.movinghistory.ac.uk/index.html.

AMIA - Association of Moving Image Archivists
http://www.amianet.org

The Association of Moving Image Archivists is an international professional association established to advance the field of moving image archiving by fostering cooperation among individuals and organizations concerned with the acquisition, preservation, description, exhibition and use of moving image materials.

Arts and Humanities Data Service (AHDS)
http://www.ahds.ac.uk

The Arts and Humanities Data Service (AHDS) is a UK national service aiding the discovery, creation and preservation of digital resources in and for research, teaching and learning in the arts and humanities.

British University Film and Video Council (BUFVC)
http://www.bufvc.ac.uk/aboutus/index.html

The BUFVC is a representative body which promotes the production, study and use of film and related media in higher and further education and research. It was founded in 1948 as the British Universities Film Council. The Council receives core grant support from the Joint Information Systems Committee (JISC) of the Higher Education Funding Councils via the Open University.

Coordinating Council of Audiovisual Archive Associations (CCAAA)
http://www.ccaaa.org/

CCAAA provides a shared platform for seven membership based organisations wishing to co-operate on influencing the development of public policy on issues of importance to professional audiovisual archivists.

DELOS – Networking of Excellence on Digital Libraries
http://www.delos.info

This EU programme has a preservation strand, but currently research their preservation research has looked at the ‘DELOS DPC Testbed: A Framework for Documenting the Behaviour and Functionality of Digital Objects and Preservation Strategies’ and ‘Digital Preservation Automated Ingest and Appraisal Metadata’.

DIGICULT
http://www.digicult.info/pages/index.php

DigiCult monitors existing and emerging technologies that provide opportunities to optimise the development, access to, and preservation of Europe’s rich cultural and scientific heritage.

ERPANET – Electronic Resource Preservation and Access Network
http://www.erpanet.org

The European Commission and Swiss Confederation funded ERPANET Project to enhance the preservation of cultural and scientific digital objects through raising awareness, providing access to experience, sharing policies and strategies, and improving practices. ERPANET constructs authoritative information resources on state-of-the-art developments in digital preservation, promotes training, and provides advice and tools.

FACT – Foundation for ART And Creative Technology
http://www.fact.co.uk/

Based in Liverpool, FACT is a leading organisation for the development, support and exhibition of film, video and new and emerging media.

FIAF - International Federation of Film Archives
http://www.fiafnet.org

The International Federation of Film Archives brings together institutions dedicated to rescuing films both as cultural heritage and as historical documents. FIAF is a collaborative association of the world's leading film archives whose purpose has always been to ensure the proper preservation and showing of motion pictures.

FIAT/IFTA - International Federation of Television Archives/Fédération Internationale des Archives de Télévision
http://www.fiatifta.org/

The International Federation of Television Archives is an international professional association established to provide a means for co-operation amongst broadcast and national audiovisual archives and libraries concerned with the collection, preservation and exploitation of moving image and recorded sound materials and associated documentation.

Focal International

http://www.focalint.org

The Federation of Commercial Audiovisual Libraries International Ltd, is a not-for-profit professional trade association providing excellent networking and marketing opportunities for people in the footage, stills and audio archive content industry, and giving users easy access to over 300 members worldwide. It has its base in the UK.

IAMHIST - The International Association for Media and History
http://www.iamhist.org

An association of professional film and television broadcasters, scholars, and others who are passionately concerned about film, radio, television and their relations to history.

IASA - International Association of Sound and Audiovisual Archives
http://www.iasa-web.org

The International Association of Sound and Audiovisual Archives was established in 1969 to function as a medium for international co-operation between archives that preserve recorded sound and audiovisual documents. IASA supports the exchange of information and fosters international co-operation between audiovisual archives in all fields, especially in the areas of: acquisition and exchange; documentation and metadata; resource discovery and access; copyright and ethics; conservation and preservation; and, research and publication.

ICA - International Council on Archives
http://www.ica.org/

The International Council on Archives is dedicated to the advancement of archives worldwide. Archives, by providing evidence of human activities and transactions, underlie the rights of individuals and States, and are fundamental to democracy and good governance. Archives safeguard the memory of mankind by preserving records of its past. In pursuing the advancement of archives, ICA works for the protection and enhancement of the memory of the world.

IFLA - International Federation of Library Associations and Institutions
http://www.ifla.org

The International Federation of Library Associations and Institutions is the leading international body representing the interests of library and information services and their users. It is the global voice of the library and information profession. The Audiovisual and Multimedia Section of IFLA is the international forum for persons working with non book media in every kind of library and information service.

National Representatives Group and Minerva
http://www.minervaeurope.org/structure/nrg.htm

The NRG was set up by Minerva, which is a network of Member States Ministries to discuss, correlate and harmonise activities carried out in digitisation of cultural and scientific content. Audiovisual is a peripheral concern of their overall programme. Annually Minerva publishes a report entitled, ‘Coordinating digitisation in Europe’.

SEAPAVAA – Southeast Asia-Pacific Audiovisual Archive Association
http://www.geocities.com/seapavaa/

The Southeast Asia-Pacific Audiovisual Archive Association aims to provide a regional forum for addressing common issues and concerns related to the collection and preservation of, and provision of access to, the audiovisual heritage of member countries. It particularly aims to promote audiovisual archiving and to preserve and provide access to the region's rich audiovisual heritage.

Sonic Arts Network
http://www.sonicartsnetwork.org/

Sonic Arts Network is a national organisation working exclusively with sound and technology in creative, innovative and experimental ways.
Every year the Sonic Arts Network supports British artists through a series of new commissions and through providing performing and exhibiting platforms for their work; the SAN nurture creative projects with people of all ages through educational workshops; and build the profile of electronic music and sound art by bringing the best artists in the world to show their work to British audiences.

Training for Audiovisual Preservation in Europe (TAPE)
http://www.tape-online.net

TAPE is a three-year project, funded under the Culture 2000 programme of the EU, aiming to raise awareness and training by expert meetings, research, publications and workshops. TAPE focuses on audiovisual collections held outside the major national institutions with specific responsibilities for audiovisual heritage. Instead they are primarily concerned with preservation and access issues of AV materials in non-dedicated institutions (ie. collections not specializing in audiovisual that happen to have AV collections)

UNESCO - United Nations Educational, Scientific and Cultural Organization, Information Society Division
http://www.unesco.org/webworld

The United Nations Educational, Scientific and Cultural Organization (UNESCO) is mandated by its Member States inter alia to promote the free flow of ideas by word and image, to foster international co-operation in the fields of communication and information in order to narrow the gap between the "information rich" and the "information poor" in these areas and to promote access for all to ICTs.