ICT-2007-3-231161

Deliverable D7.1.5

Audiovisual Digital Preservation

Status Report 3

Richard Wright, BBC

26/12/2011

Document administrative table

Document Identifier	PP_WP7_D7.1.5_Annual_AV_Status_3	Release	0
Filename	PP_WP7_D7.1.5_Annual_AV_Status_R0_v1.00.pdf		
Workpackage and Task(s)	WP7 Dissemination and training T1 – Dissemination and publication of results		
Authors (company)	Richard Wright BBC		
Contributors (company)	Matthew Addis, ITI		
Internal Reviewers (company)	Jacqui Gupta, BBC		
Date	26/12/2011		
Status	Release		
Type	Deliverable		
Deliverable Nature	R = Report		
Dissemination Level	PU = Public		
Planned Deliv. Date	m36 - 31/12/2011		
Actual Deliv. Date	m36 – 26/12/2011		
Abstract	The current status of audiovisual preservation as of December 2011 is described. The 2009 report introduced the new problem of digital preservation (arising from the results of digitisation) and summarise the access issues for file-based audiovisual content and contributions of PrestoPRIME. The 2010 report concentrated on practicalities of audiovisual digital preservation: century costs, a digital preservation primer, and a summary of PrestoPRIME technology. This 2011 report continues the series of technological 'brief encounters', lists European developments in 2011 and has a major section on the significance of the fact that audio and video content is, at heart, a signal – and hence fundamentally different from text and general digital data.		

DOCUMENT HISTORY

Release	Date	Reason of change	Status	Distribution
v0.01	16/11/2011	Extended Outline	incomplete	Confidential
v0.01	28/11/2011	First Draft	BBC draft - incomplete	Confidential
v0.02	01/12/2011	Second Draft	Circulated for partner contributions	Confidential
v0.03	24/12/2011	Fourth Draft	Completed following internal review	Confidential
v1.00	26/12/2011	Delivered	For public release	Public

Table of Contents

Scope....4
Executive summary....5
1 A summary of PrestoPRIME public activity in 2011....7
1.1 Public Deliverables....7
1.2 Presentations....8
1.3 Training and Workshops....11
1.4 Standardisation....11
2 PrestoPRIME Technology....13
2.1 Introduction: the tools in context....13
2.2 General descriptions....14
2.3 The comprehensive list....21
3 PrestoCentre Developments in 2011....24
4 Technology – Brief Encounters....25
4.1 Authenticity....25
4.2 Quality vs compression management....27
4.3 FRBR revisted - a preservation vocabulary....28
4.4 The end of broadcasting....29
4.5 A new approach to analogue video digitisation....29
4.6 Correcting temporal variation in audio....30
4.7 Repurposing archive audio....30
4.8 Embedded metadata....31
4.9 Containers....33
5 UK, European and International Developments....36
5.1 Europe....36
5.2 General sources of information....38
5.3 UK....39
5.4 International Developments....40
6 Preservation of Signals....42
6.1 Four stages in the life of digital audiovisual content....43
6.2 The significant properties of a signal....44
6.3 How to preserve signals....47
7 Glossary....48
8 References....53

Scope

PrestoPRIME is the European publicly-supported project that addresses **preservation of digital audiovisual content**, and **access to audiovisual content in digital libraries,** using **Europeana** as our demonstration platform.

This document is the seventh in a series of annual reviews of the status of audiovisual preservation in Europe. The first four reviews were produced by PrestoSpace. Each has had a specific focus, plus providing a general summary of annual progress toward saving Europe's audiovisual heritage.

The fifth was the first produced under the PrestoPRIME project, and so covered the general area of digital preservation, as well as introducing the project.

The sixth introduced a format that has been followed in this report:

1) PrestoPRIME activity;
2) PrestoPRIME technology;
3) PrestoCentre report;
4) Brief Encounters: short reviews of general technology affecting audiovisual preservation;
5) UK, European and international project and policy developments affecting audiovisual preservation; and
6) Specific topics: technical information on audiovisual archiving, with more depth than in the Brief Encounters.

There is one special topic in this seventh report: digital preservation of a signal, which is an exposition of the engineer's view of digital preservation of audiovisual content, and particularly of how such content differs from most other digital library content such as text.

While this year there is only one special topic, the Brief Encounters section has grown from three to nine. There is also rather more material (than last year) on UK, European and international developments.

Executive summary

This document is a product of the EU-sponsored PrestoPRIME[1] project. PrestoPRIME is the major project on digital preservation in the audiovisual sector[2]. The current status of audiovisual preservation as of December 2011 is described, as an update to the series of annual reports on audiovisual preservation previously given in January 2005 to 2008[3] as products of the EU-sponsored PrestoSpace project, and in January 2010[4] and 2011 as PrestoPRIME reports. The PrestoSpace reports concentrated on digitisation, which remains a significant issue. The PrestoPRIME reports focus on digital preservation of the content coming from digitisation as well as content which is born-digital.

This December 2011 report (on the year 2011) has the following sections:

PrestoPRIME public activity: a week-long summer school, a FIAT-IFTA workshop, and a range of other conference, presentations and papers.

PrestoPRIME Tools: the tools produced by PrestoPRIME in 2010 and 2011, demonstrated at the summer school, and publicly tested on an integrated platform (P4 = the PrestoPRIME Preservation Platform) in November in Turin. Integration with the commercial Rosetta Digital Preservation system was also demonstrated. The tools include:

- storage modelling and management, and service agreements for outsourcing;
- control of user-generated metadata and annotations;
- fingerprint technology for identify content and tracking its use;
- automatic detection of video impairments (for automated quality control);
- an ontology and software for automating rights management for archives;
- emulation technology using the Java Virtual Machine approach;
- metadata extraction (from professional audiovisual formats such as MXF that standard extractors don't process), verification and mapping; and
- the P4 integrated platform and Rosetta integration.

Technology Updates: digital preservation technology doesn't stand still. We briefly reported on three areas a year ago, and now have a follow-up on one (containers) and eight new ones, ranging from a formalism for specifying and tracking the authenticity of archive content to new technologies for recovery of audio and video from tape recordings.

UK, European and International Developments: with the growth of Europeana, the European Digital Library, there is much more to report than a year ago. As well as covering the new European-level support for actually funding digitisation, a radical departure from previous European-level policy, we review the major relevant policy announcements from Europe. But there is a wider-world, and audiovisual preservation developments from outside

[1] http://www.prestoprime.org/

[2] PrestoPRIME is the only Integrated Project on audiovisual digital preservation running under the Seventh Framework of the EC-operated IST programme: http://cordis.europa.eu/fp7/ict/telearn-digicult/digicult-projects-prestoprime_en.html

[3] All four are online PDF files, available free from PrestoSpace. Three are listed here: http://digitalpreservation.ssl.co.uk/general/#White%20Paper , and the fourth is here: http://www.prestospace.org/project/deliverables/D22-9_Preservation_Status_2008

[4] The PrestoPRIME reports from January 2010 and 2011 are also free and online: http://www.prestoprime.org/project/public.en.html

Europe are also mentioned. While PrestoPRIME has no funding (or staffing) for formally studying the global situation, the two main professional associations for audiovisual archives, FIAT-IFTA and IASA, do have a global membership. Information from their publications and annual meetings gives solid information and a genuinely global perspective.

This report also includes consideration of developments in the UK. No doubt just as much could be said about many other countries, so the UK information (which the author has *to hand,* from being based in the UK) should be seen as just one country's indication of the general level of activity around audiovisual archives.

1 A summary of PrestoPRIME public activity in 2011

This document is a status report on audiovisual preservation, concentrating on the technical needs and problems of those responsible for audiovisual content. It is produced by the PrestoPRIME project which is concerned with meeting those technical needs, hence the inclusion of a review of PrestoPRIME work.

The year 2011 was PrestoPRIME's third. The project moved from the specification stage (of 2009) through development of systems and tools for audiovisual preservation (2010) and in 2011 has produced an integrated audiovisual preservation platform – and integration with an existing commercial platform: Rosetta. Public information about these PrestoPRIME developments are given in the next three sections.

The PrestoCentre competence centre launched with a major event in March 2011. There is a separate section on Presto Centre activity: Section 3, below.

1.1 Public Deliverables

The work of PrestoPRIME was formally reviewed in March 2010 and again in March 2011. Subsequent to the first review, a set of documents was made public on the PrestoPRIME website http://www.prestoprime.eu/project/public.en.html.

A year ago (December 2010) 14 public deliverables, covering

- preservation requirements (D5.1.1 Definition of Scenarios),
- rights glossary,
- preservation status,
- preservation strategies,
- preservation process modelling,
- outsourced storage (D2.3.1 Service-Oriented Models for Audiovisual Content Storage),
- Europeana (D6.2.2 European Digital Library implementation guidelines for audiovisual archives),
- preservation toolkit
- four related major studies on preservation systems ID3.1.1 and threats ID3.2.1; use of emulation (Multivalent ID3.3.1) and services and service level agreement ID3.4.1.
- preservation technical architecture
- preservation metadata for audiovisual content

Now (December 2011) there are 19 public deliverables on http://www.prestoprime.eu/project/public.en.html. The new ones added since the last review (March 2011) are:

D2.1.2 Tools for modelling and simulating migration-based preservation
D2.2.3 Strategy for Use of Preservation Metadata within a Digital Library with examples of use in audiovisual preservation
D6.3.1 Financial Models and Calculation Mechanisms
D2.1.3 AV Data Model : Final Specification
D7.1.4 Audiovisual Digital Preservation Status Report 2

1.2 Presentations

Issues within the general area of audiovisual digital preservation have been presented by PrestoPRIME partners at major conferences during 2011.

ICT Information Day Luxembourg 8 March
Richard Wright gave a presentation (prepared by Richard and Marius Snyders): *PrestoCentre – a European Competence Centre*

PrestoCentre: Screening the Future Hilversum 14-15 March
On the 14th & 15th of March 2011, more than two hundred audiovisual archivists, librarians, museum professionals, film makers, advocates and policy managers working in cultural heritage gathered at the Netherlands Institute for Sound and Vision for a two-day meeting focused on common challenges and solutions in the audiovisual domain. Twenty-two invited experts spoke on the topics of digitisation, workflows, digital preservation and sustainability, funding, preservation strategies and much more.

In focusing on the questions, *what are we preserving*, *how can we fund our future*, *where do AV archives meet IT*,*how can we valorise our archives*, and *how will we keep our archives in good shape*, Screening the Future presented a number of benchmarks, issues and challenges, and delivered a range of guidelines, workflows and standards for the preservation of our audiovisual heritage.

All the conference presentations are online:
http://www.prestocentre.eu/screening-future-2011-documentation

PrestoCentre Launch Hilversum 14 March
Screening the Future 2011 was also marked by the inauguration of PrestoCentre.

Preservation Archiving Special Interest Group London 4-5 April
Richard Wright gave a presentation: *Preserving Media and Entertainment Content*

Digital Preservation Coalition London 8 April
Richard Wright gave a presentation: *Presto Centre: Tools for Audiovisual Preservation*

FIAT-IFTA Media Management Seminar Toronto 26-27 May
Digital preservation registries and how PrestoPRIME is facing the issues related to the digital preservation of audiovisual contents presented by Walter Allasia, EURIX..
http://www.fiatifta.org/index.php/commissions/media-management-seminar-in-toronto
Jacqui Gupta, BBC, was a seminar organiser.

Transistor 2011 Heraklion, Crete 22-25 June
Preservation Techniques and Methodologies for Digital Audiovisual Works
Daniel Teruggi and Richard Wright were invited speakers; two sessions each
Daniel Teruggi, INA: PrestoSPACE, Digi-BIC and Europeana;
Trends for Digital Contents in INA
Richard Wright, BBC: Digital Preservation – a practical workshop;
BBC online archive and Digital Public Space
http://transistor.ciant.cz/2011/module-i-preservation-techniques-and-methodologies-for-digital-audiovisual-works/

6th International Conference on Knowledge Capture Banff, Canada June 25-29
On the role of user-generated metadata in audio visual collections, Riste Gligorov, Michiel Hildebrand, Jacco van Ossenbruggen, Guus Schreiber, Lora Aroyo

IASA – International Association of Sound Archive Frankfurt, 3-8 September
Richard Wright, BBC: Storage Strategy Tools
Marius Snyders, B&G: Long term access to audiovisual material creating a platform for experimentation and sharing the results

IBC International Broadcast Conference Amsterdam, 11 September
Matthew Addis, ITI: Calculating the 'cost of compromise' for long-term AV retention and access - in a session on archive matters chaired by Jacqui Gupta, BBC

PrestoCentre Summer School: Bry-sur-Marne, 12-16 September
(see next section for details)

Theory and Practice of Digital Libraries Berlin, 25-29 September
Interactive vocabulary alignment, Jacco Van Ossenbruggen, Michiel Hildebrand and Victor De Boer

FIAT-IFTA International Federation of Television Archives Turin, Italy,
28 September: half-day PrestoPRIME workshop (see next section for details)
30 September:

- Workshop 6: Beyond the File: Containers for Archiving Complexity. Moderated; Chaired by Richard Wright, BBC & PrestoPRIME: the new world of “packages”: containers for complexity. A review of options, from the Library of Congress BagIt to the new technology of putting filing systems on data tape (LTFS): Seiji Myamoto NHK, Rod Allen ERA Ltd, Thomas Heritage BBC, Brian Campanotti (Front Porch).

1 October:

- Plenary Session: Preservation and Migration Session dedicated to: How can we keep all those digital contents forever? Presentation of prototypes and services developed in the PrestoPRIME project by Daniel Teruggi, INA & FIAT Preservation & Migration Commission
- Plenary: *Preservation Clinic* run by Daniel Teruggi and Richard Wright, FIAT Preservation and Migration Commission and PrestoPRIME

Virtual Goods 28-30 Sept Barcelona; PrestoPRIME supported the conference.

- *Broadcasting the Archive*, Richard Wright, David Jordan:
- *Digital preservation of audiovisual files within PrestoPRIME,* Francesco Gallo, Laurent Boch, Walter Allasia, Matthew Addis, Stephen Phillips, Werner Bailer and Peter Schallauer.

Europeana Tech 4-5 Oct 2011 Vienna, Austria

- Open data for the cultural masses - Mapping and the Europeana Semantic Layer, Guus Schreiber
- User-generated video annotations on the Web of Data, Michiel Hildebrand
- End-User Media Annotation with YUMA, Rainer Simon

The presentations are online:
http://www.europeanaconnect.eu/europeanatech/index.php?section=programme&page=programme

DFT Archive & Restoration Symposium, BBC, London 13 Oct 2011 -- PrestoCentre was a sponsor of this symposium

- Richard Wright (BBC) gave a paper *Century Store*, about estimating costs for maintaining a digital archive over a century. PrestoPRIME tools (from ITI) were showcased.
- Erwin Verbruggen (B&G) gave a presentation *Cine XPRES: An Archive in Search of New Film Scanning Methods*. A presentation on **Arkivum** was given my Jim Cook

Goportis Digital Preservation Summit 2011 19-20 Oct, Hamburg -- Richard Wright invited speaker; *Ingest & Metadata for Audio Materials*; also a 90-minute workshop: *Preservation and access for audio and video*

IMPACT Final Workshop, 24-25 Oct London – No formal PrestoPRIME participation, but there are three points of contact:

- IMPACT does use the PrestoSpace motto: Better, Faster, Cheaper
- PrestoPRIME partner: The Functional Extension Parser: A Document Understanding Platform by Günter Mühlberger (University of Innsbruck)
- IMPACT is launching a Competence Centre

AMIA 16-19 Nov Austin, Texas USA -- Jacqui Gupta, BBC, attended. Marius Snyders, B&G, attended and moderated a session. Martin Hall-May, ITI, gave a presentation: *Long Term Access to AV Material: Estimating the Costs*

1.3 Training and Workshops

- **Summer School** Presentations are online – but NOT in a public folder:
  PrestoCentre Summer School: Bry-sur-Marne, 12-16 September
  A one-week course on PrestoPRIME tools. The full description and course presentations are online: http://prestocentre.eu/book-page/prestocentre-training-course-2011-documentation. [If the URL redirects to the home page, search for "Bry" using the "Search our site" window, and select "PrestoCentre Training Course 2011 – Documentation".] This one-week course was oversubscribed, and included visits to INA's preservation activity and to the Bibliothèque nationale de France, two days of PrestoPRIME technology, a day of presentations and demonstrations by industry – and closed with a one-day, high-level seminar on the nature and interpretation of archive content.

- **FIAT/IFTA Workshop**
  28 September: half-day PrestoPRIME workshop
  A) Introduction to PrestoCentre: how it works, what it does
  B) PrestoPRIME tools: covering nearly a dozen specific tools, including: planning and cost models for mass storage; use of user-generated metadata; automation of rights management; metadata mapping; video quality control; ingest of audiovisual files into a formal digital repository; fingerprint and content tracking technology applied to video.

- **Integrated Testing**
  In 2010 PrestoPRIME had a public workshop, with demonstrations of seven tools and systems developed by the project; the presentations are online: https://prestoprimews.ina.fr/public/presentations/london_26_11_2010 / .

  In 2011 this technology was integrated with the PrestoPrime Preservation Platform and (where appropriate) with the PrestoPRIME extensions to the ExLibris Rosetta digital preservation system. The integrated technology was tested at a public event. The technology is described in detail in Section 2.

1.4 Standardisation

- The MPEG standards body, which is best known for defining the encoding standards used by all of digital broadcasting (MPEG 2 for standard definition television broadcasting, MPEG 4 for high definition) as well as the ubiquitous MP3 audio, met in March 2011 and formed an Ad Hoc Group on Multimedia Preservation. Walter Allasia of Eurix and PrestoPRIME is co-chair. Their initial work has been a questionnaire to determine user practice and user requirements. The questionnaire has been prepared and circulated with support from FIAT-IFTA, with Jacqui Gupta of the BBC (and PrestoPRIME) coordinating for FIAT-IFTA.

- Media Annotation Working Group. The author has great expectations for the eventual impact of a standard for time-based citation and annotation of audio and video content. This capability, if widely implemented (on a browser near you) will allow audiovisual content to be referenced with the same ease and certainty that we expect for access to text. At last, audiovisual content can be cited, a basic requirement for

use of such content in scholarly discourse. The ergonomics of practical access will be streamlined when content can start playing from the precise point of interest rather than from the beginning. Finally, the social media aspect of shared use of content will be enabled by a standard for annotation, which then makes it possible to share annotations.

The World Wide Web Consortium has a Media Annotation Working Group (MAWG); Werner Bailer, Joanneum and PrestoPRIME is a delegate, and has an online demo of their work, here http://mawg.joanneum.at/web/demo3.html.

Recent MAWG news (http://www.w3.org/News/2011#entry-9267, November 2011) announces release of their standard, and 'invites implementation':

The Media Annotations Working Group invites implementation of the Candidate Recommendation of "API for Media Resources 1.0."
This specification defines an API to access metadata information related to media resources on the Web. The overall purpose is to provide developers with a convenient access to metadata information stored in different metadata formats. The Working Group is developing a test suite during this Candidate Recommendation phase.

http://www.w3.org/2008/WebVideo/Annotations/
http://www.w3.org/TR/2011/CR-mediaont-api-1.0-20111122/
http://www.w3.org/2008/WebVideo/Annotations/drafts/API/testsuite
http://www.w3.org/2008/WebVideo/

2 PrestoPRIME Technology

In 2010 a public workshop was held in November where a range of audiovisual preservation tools was demonstrated, but the tools were working in isolation. In 2011 these tools were integrated to work with P4, the PrestoPRIME Preservation Platform – and where appropriate to also work with Rosetta. The integrated platform was publicly evaluated during a week-long session in early November 2011, held at the RAI Research premises in Turin.

The tools were presented in detail at the PrestoCentre Summer School in September, and at the PrestoPRIME workshop held on the half-day preceeding the FIAT-IFTA conference in Turin, at the very end of September.

The following sections provide information on the PrestoPRIME technology, at three levels of detail: introductory, general, and comprehensive.

2.1 Introduction: the tools in context

PrestoPRIME has publicised its developments wherever possible, including this PrestoCentre blog by Matthew Addis about the problem of estimating storage costs, and our tools addressing that problem: http://www.prestocentre.org/blog/2011/11/24/whats-it-worf).

At the FIAT-IFTA workshop, the tools were placed in a preservation context. Three aspects of digital preservation were considered:

1. Making analogue content into digital content
2. Working with digital content (lots of files)
3. Preserving the digital content

While PrestoPRIME was initially aimed at point 3 (long-term digital preservation), most people are more concerned with point 2 (getting on with their work in the new digital environment. PrestoPRIME has tools for both of these areas, as summarised in these figures adapted from the FIAT-IFTA presentations

1- Making analogue content into digital content
Planning – Budget
Workflow – Standards
Rights – Result: lots of files

PrestoSpace information online:
//wiki.prestospace.org/
//digitalpreservation.ssl.co.uk/
Now: revised for PrestoCentre
http://www.prestocentre.eu
Joanneum Quality Analysis Tool

Digitisation was the topic of PrestoSpace, the predecessor of PrestoPRIME. The focus was to define and promote efficient digitisation workflows, under the motto “Better – Faster – Cheaper”. The key idea is that all three can be achieved, using an industrial process and key ideas such as division of labour and automation that go back to Adam Smith and the 18th Century. Use of automation requires methods for automating quality control, and this is one PrestoPRIME area that applies equally to digital preservation and to digitisation. PrestoPRIME partner Joanneum Research have a Quality Analysis Tool.

With digitisation and digital technology, collections of all sorts are having to move from shelf-based processes to new workflows using files.

Digital Asset Management (DAM) or Media Asset Management (MAM) systems are commonly introduced to provide tools and controls.

2- Working with digital content (lots of files)

DAM/MAM and Trusted Repositories – what do they do, what don't they do – PrestoPRIME White Papers
Storage – ITI online free tools
Metadata – Joanneum mapping and validation ; "tag gardening" Univ Amsterdam; fingerprinting INA
Digital library technology RAI, BBC support for MXF files
Access – Joanneum Time-based navigation, annotation
Rights – RAI ontology, Eurix implementation

Long-term security is addressed by Trusted Digital Repositories. Most general systems were developed for documents, not for audiovisual files.

The figure lists areas where PrestoPRIME provides technology that supports working with audiovisual digital content.

The third area is true digital preservation. While people and institutions have to focus on maintaining their daily activity (workflows and processes) – they have to also address preservation, or risk losing their content. The PrestoPRIME Preservation Platform (P4) and the commercial Rosetta system are the integration platforms for testing and demonstrating PrestoPRIME technology for preservation, including emulation, long-term strategy and management of storage as a service rather than as physical devices or objects.

3- Preserving the digital content
Preservation Platform: P4=PrestoPRIME Preservation Platfom, Eurix; Rosetta, Ex Libris
Standards: OAIS; formal control; formal preservation actions eg migration; P4
Emulation – Multivalent, Univ of Liverpool
Formats, carriers, storage: Planning and strategy: PrestoPRIME white papers
Managing and maintaining storage into the future –SLA's for outsourced service; white papers, software for real-time SLA monitoring; modelling and simulation tools

2.2 General descriptions

The PrestoCentre Summer School provided general descriptions of each component of the PrestoPRIME technology. All these presentations are online: http://prestocentre.eu/book-page/prestocentre-training-course-2011-documentation[5].

[5] . If the URL redirects to the home page, search for "Bry" using the "Search our site" window, and select "PrestoCentre Training Course 2011 – Documentation", or use http://www.prestocentre.org/search/apachesolr_search/PrestoCentre+training+2011?filters=type%3Abook.

The descriptions covered:
DAY 1 - Monday, 12 September Preservation Planning and Management

09.30 Introduction to PrestoCentre and PrestoPRIME. *Daniel Teruggi (Ina)*

10.30 Overview of Digital Preservation *Richard Wright (BBC)*

Four stages in the life of audiovisual content:
1) A signal; 2) A recording of a signal 3) A digital/digitised signal;
4) Preservation of the digital signal

Three digital workflows:
1) Digitisation/migration of content into files as a preservation action;
2) Working with a file-based archive; 3) True digital preservation: keeping file-based content usable (and affordable)

Finally, OAIS and conservation, for analogue and digital content:
1) Handling, packaging and storing; 2) Environmental conditions;
3) Protecting the masters; 4) Condition monitoring

11:00 Overview of PrestoPRIME Technology *Laurent Boch (RAI)*

11.45 Activity: Preservation Planning - Mapping Out and Assessing Stages, Issues and Priorities *Richard Wright (BBC)*

The Presto ‘preserve the collection’ approach will be reviewed: mapping and prioritising, collection strategy, preservation strategy, preservation plan. Participants will compare their own organisation's situation: where they are, issues being dealt with, issues the course should cover, what they want to achieve at their organisation, and what they want to take away from the course. This will include a personal assessment of 6 Key issues and 6 Critical Success Factors their organisations are facing, which participants will re-examine at the course’s end.

13.30 Case Study 1: Ina and its 15 year-long Plan for Preservation and Digitisation *Daniel Teruggi (Ina)*

In 1999 Ina launched an ambitious plan to digitise its whole collections. This implied the development of an industrial environment and approach to digitisation in order to digitise the 1 million 350.000 hours of Radio and Television. This presentation will reflect on the organisation of the process and its perspectives in terms of use and access, and is a precursor for the subsequent visit to Ina’s technical premises.

15.00 - 17.00 Tour through the Institut National de l'Audiovisuel de France
Digitisation Centre, Storage Centre, Documentation Process, Legal Deposit Capture

DAY 2 - Tuesday, 13 September All About Metadata and Integration

09:30 Introduction to Metadata *Richard Wright (BBC)*

The basic types of audiovisual metadata and the critical distinction between the preserving *of* metadata vs. the use of metadata *in* the preservation of everything (essence and metadata). The need to interpret map metadata if it is to be made useful – rather than just being preserved as 'dark metadata' or 'just bits'. The relation between preservation metadata and the data model of a repository (or digital library or OAIS or even an asset-management system).

10.00 Metadata Tools: Contents Tracking *Jean-Hugues Chenot (Ina)*

Presenting a sample application of how video fingerprints can be used to help give a structure to linear TV broadcast contents, this presentation will discuss an experimentation run at Ina. Navigating within a growing database of 18 months of broadcasts on 10 different channels, the experiment allowed access to the results of unsupervised detection of repeated TV contents within this large store of material. In particular, the demonstration will highlight how results obtained without operator involvement reveal structure.

10.30 Emulation: Multivalent browse/transcode *Paul Watry (UL)*

The talk will describe the Multivalent tool which can be used to view audio-visual content and a tool to transcode MXF/D10 files for consumption by another service. The tools are written such that they are self-contained and operating system independent. The talk will also describe a policy-driven system for managing the audio-visual data which insulates the user from changes to the underlying operating system and allows policies to be defined and enacted on the data.

11.30 Case Study 2: Memnon: Use of Semantic Metadata and Linked Data
Frédéric Beaugendre (Memnon)

13.30 How Tools Integrate with Existing Standards, Technology and Workflows

MPEG Issues for Digital Preservation *Walter Allasia (Eurix)*

We live in a landscape of quickly growing audiovisual formats, and contribute to the very short lifetime of some of them. In order to be able to set up a strong global preservation action we have to agree and define formats and protocols as well as establish distributed registries for automatic obsolescence evaluation and risk assessment. During this session, a practical approach is presented involving attendees in the standardization process of multimedia preservation description information.

14.00 Rights Services *Laurent Boch (RAI)*

The goal is to provide organisations of the audiovisual domain the means to improve the management of audiovisual rights information, in order to more easily identify the possibility to re-use archived material, to optimise the fruition of owned rights, and to support rights purchase and sale activities. The final users will have greater possibilities to access more content which would otherwise be used irregularly or not at all. From the analysis of a number of real contracts a glossary of rights terms has been obtained. We then defined an extension of MVCO (Mpeg Media Value Chain Ontology) which is both a model and a language for unambiguously representing the rights situation of AV intellectual property

entities. The RAI RightsDraw services allow their users to create, view and edit the rights information related to audiovisual assets, or specific queries. A prototype installation is demonstrated.

14.30 P4, the PrestoPRIME Preservation Platform *Francesco Gallo (Eurix)*

An Open Source integration framework for the technologies developed within PrestoPRIME which support the process of running a preservation service. As an implementation of the OAIS model, it's going to offer the basic services and interfaces for the Ingestion and Access of/to AV items, which will be accepted from the Producers in the forms of Submission Information Packages (SIP) and delivered to Consumers as Dissemination Information Packages. Software tools develped by the partners or third party tools can be integrated according to the defined Reference Architecture. P4 software is developed in Java language. The human users will typically access to the services through their normal browser. During the session the P4 architecture will be presented, and information will be given about the current progress of work. Usage of the Ingest and Access interfaces will be shown from the prototype version.

15.00 Rosetta *Nir Sherwinter (ExLibris)*
Ex Libris' Rosetta, a digital preservation solution, addresses libraries' and archives' need to collect, manage, and preserve a wide variety of digital objects in different formats and structures. The system was developed in partnership with the National Library of New Zealand and reviewed by a peer review group of world-renowned preservation experts. As part of the PrestoPRIME project Rosetta is used as the commercial solution integrating with the different tools developed by the project partners. This talk includes presentation of the system openness, integration points and specifically the unique plug-in framework. In addition a live demo of Rosetta will demonstrate integrating with the MXF-Tech-MD-Extractor tool (a PrestoPRIME tool developed by RAI).

DAY 3 - Wednesday, 14 September
Predicting and Planning the Future

9:30 Introduction to Predicting and Planning the Future *Richard Wright (BBC)*

Why decision making begins with understanding risks, why a risk is really a prediction, and why preservation planning needs to have models in order to understand the probable consequences of decisions -- and thus make good and cost-effective decisions. The Presto 'cost of risk of loss' approach will be presented, to introduce the PrestoSpace and PrestoPRIME 'what-if' modelling and simulation tools.

10.00 Cost and Risk Modelling *Matthew Addis (IT Innovation)*

Long-term retention and access to AV assets as part of a preservation strategy inevitably involves compromise in order to achieve acceptable costs, especially for the huge volumes of archive material often involved. These compromises include quality control and throughput in transfer chains, data safety and accessibility in digital mass storage systems, and metadata quality and completeness when supporting access to archive materials. But how can these compromises be objectively and quantitatively assessed? How can an archive be assured that when they design a preservation and access approach that it will function as expected and can cope with the inevitable yet unpredictable variations that will happen in operation?

This session presented modelling tools which allow a range of preservation and access activities to be planned and analysed including cost projections, availability of operators and other resources, and simulation of 'disaster scenarios'.

The tools from ITI are all free and online here: http://prestoprime.it-innovation.soton.ac.uk/ which includes documentation, example videos, user guide and FAQs as well as access to the source code, bug tracking and feature requests. "

Further information on these tools: Preservation planning tools from PrestoPRIME include: iModel, which is a simulation tool for storage, transcoding and file-format migration of digital audio visual assets; and iWorkflow, which is a simulation tool for digitisation/migration workflows of discrete assets (e.g. digital video tapes) and has been developed for a specific scenario at the BBC for their D3 project. These tools currently work in a 'stand-alone' mode, but work is already underway to more tightly integrate then with preservation systems in PrestoPRIME.

The storage and format migration tool (iModel) is intended to allow a wide range of questions to be considered when planning, selecting or operating a storage and access system. The tool focuses on the storage and access to digital content in files using IT systems. The tool does not include metadata management, rights management and other issues that are of course important to consider. However, the tool does allow the following questions to be investigated:

* When storing content, how many copies should be made, what technologies should be used, how much will it cost, what are the long-term risks of losing files?
* What impact does the choice of codec (e.g. compressed or uncompressed video) have on costs and risks?
* What are the pros and cons of just in time generation of access copies compared to creating and storing a full set of proxies in advance?
* When storing data, how often should it be checked to make sure integrity is intact, and when does this become counter-productive (e.g. act of checking causes more damage than it might repair)?

* When should media migration take place (e.g. between LTO generations): regularly or at the point of obsolescence?
* What is the impact of ingest and access on shared resources for storage and data safety: what level of resources is needed to support both?

10.30 Service Management *Stephen Phillips (IT Innovation)*

Audiovisual archives are increasingly file-based and becoming an active element of the production, post-production and distribution process. Often archive systems are in-house, but increasingly parts are out-sourced and even off-site. How can archive IT systems be managed to ensure they meet key performance indicators for ingest and access whilst not compromising the safety of the assets? PrestoPRIME addresses this problem through policy-based automation applied to outwardly facing archive services and internal preservation processes alike, both defined through service level agreements (SLAs) and actively measured and controlled against metrics for performance, data integrity and availability.

We present an automated service management tool and demonstrate how it can be used to monitor and manage the services required for audiovisual preservation, define SLAs for the different users of the system and maintain the quality of service defined in those SLAs. We also show how the modelling from the previous session can provide a decision support system for service management.

Further information: IT Innovation has developed two tools for service monitoring and management: MServe is a RESTFul Web Framework for Service Providers to use when they want to deploy and manage a set of services for handling AV media (e.g. storage, metadata extraction, checksums, proxy generation). Ting is a service monitoring and management framework that can be used by a service provider to offer services to customers, and to monitor and manage use of these services. It can also be used by customers to monitor and manage their use of resources. MServe and Ting work together to provide a solution for providing, monitoring and managing a wide range of services.

MServe's purpose is to provide human and machine usable interfaces to control the ingest, access, processing and manipulation of digital AV content using computer resources (e.g. IT storage systems, compute clusters). There are 3 main interfaces in MServe: HTML (web browser) interface for human manipulation of the content; HTTP/REST interface for machine workflows and automated systems; and WebDAV interface to provide file system access to the content. In PrestoPRIME the WebDAV interface is used by P4 as archive storage location via davfs2. Ting include a GUI that allows a user to operate Ting from the point of view of each one of the stakeholders of a preservation/archiving system (following the OAID model). This GUI highlights the specific information that is relevant to the consumers of data, to the providers of data, and to the archivists of that data. In addition, Ting also provides a REST API that allows other clients (graphical or otherwise) to make use of the Ting functionality.

11.00 Video Quality Assessment *Peter Schallauer (Joanneum Research)*

Manual quality assessment of audiovisual media is a very time-consuming, and therefore expensive activity within the archive related processes of ingest, migration and access. For automatic quality control of digital audiovisual media only some tools have been developed. Mainly syntactical properties of the digital files can be checked, e.g. stream encoding or file wrapper standards compliance. Tools for checking the quality of the video or film images itself in an automatic manner are rare today.

We present our recent results for automatically detecting various video and film impairments including video breakup's, noise and grain levels, and more. We also show the interactive 'AV-Inspector Summary' tool which allows for efficient human verification of automatically detected visual impairments.

11.30 Metadata Mapping and Validation *Werner Bailer (Joanneum Research)*

There is a large heterogeneity in terms of metadata formats between different audiovisual collections, e.g. broadcast archives, film archives, libraries holding a/v assets. Making contents accessible to wider communities in both B2B and B2C scenarios, e.g. on portals such as Europeana, increase the number of metadata formats involved. We will briefly discuss some recent relevant developments, such as the MPEG-7 AVDP, the Europeana data model and the W3C Ontology for Media Resources on the Web. The variety of producers and consumers of metadata creates the need for mappings or crosswalks between the different metadata formats. We present novel approaches for automating metadata mapping and semantic validation of metadata documents, that aim at overcoming the limitations of 1:1 mappings between each pair of formats.

The Joanneum metadata tools are available online: http://prestoprime.joanneum.at/

12.00 Management of User-Generated Metadata *Michiel Hildebrand (VUA)*

In the audiovisual domain tagging games are explored as a method to collect user-generated metadata. The Netherlands Institute for Sound and Vision deployed the video labelling game Waisda? to collect time-based tags for videos from their collection. We demonstrate Waisda? and discuss the results of a study into the role of such user tags for audiovisual collections. In addition, we describe a workflow that allows the moderation of user tags by domain experts. We discuss a first prototype to semi-automatically link tags to concepts from the Linked Open Data cloud.

The remainder of the week was devoted to other issues beyond PrestoPRIME tools: a visit to the French National Library, a day of presentations and demonstrations by service providers, and a day-long general seminar on the interpretation, value and uses of audiovisual collections.

2.3 The comprehensive list

The full list of PrestoPRIME tools (and integration of tools) is detailed hereafter.

Tool Name Partner WP/Task	Tool Description Interface Mechanism Development Status	P4 Support Integration Status	Rosetta Support Integration Status
TechMD Extractor (MXF Tools) **RAI** WP3T1	Technical Metadata Extractor for MXF files. **Java-based**: provided as Jar library, including Java interface for the tool, source code available (GPLv3) Development completed. New release expected supporting extraction of Material Package ID	Fully supported (ingest module), integration completed, tested, results stored in SIP	Fully supported in Rosetta, integration completed
D10 Sum Checker (MXF Tools) **RAI** WP3T1	Sum Checker for frames of D10 files. **Script-based**: provided as Linux binary written in C, source code available (GPLv3) Development completed.	Fully supported, integration partially completed (ingest module), successfully tested on D10 samples, missing representation of results (DNX?) in SIP	Fully supported, TBC
Fingerprint **INA** WP4TX	Fingerprint computation and comparison. **Web server, non blocking call**(requires long time to compute fingerprint and show results). Getting results is always blocking. Development of web server interface in progress.	Possible support. Integration in progress.	Possible support. Integration in progress.
Multivalent **UniLiv** WP3T3	Multivalent player **Java-based**: tool is written in Java and is provided as Jar library, also available JNLP for Java We Start (JavaWS), source code is available (GPLv3?) Development in progress. Multivalent player available. Evaluate development of Multivalent decoder.	Fully supported (access module). Integration in progress.	Fully supported (as Rosetta player). Integration in progress.

Tool Name Partner WP/Task	Tool Description Interface Mechanism Development Status	P4 Support Integration Status	Rosetta Support Integration Status
Service Management Framework **ITInnov** WP4	Framework for monitoring services **External framework**, interacts with services to be monitored and managed using specific RESTful interface Development in progress	In the first prototype implementation which was demonstrated at the Turin test-bed event, a simple integration of MServe and P4 was used for sharing AV material through the WebDAV interface. MServe was used as a storage solution and the final location of AV material deposited to MServe was stored in an AIP. A better integration of MServe and P4, including some advanced MServe features, such as the workflow management, is planned for 2012.	Not supported. Rosetta has its own monitoring module
RightsDraw **RAI**,**EURIX** WP4T4	Editor for contracts mapping based on rights ontology **Web Application**: cgi-bin, source code is available (GPLv3?) Development almost completed, missing query-by-sample implementation.	Fully supported (ingest and access modules)	Partially supported: only ingest new SIP with rights information, SIP update to be done, ACL for accessing rights not supported (low priority), query-by-sample not supported
OAI-PMH **EURIX** WP5	Web service for content search using OpenArchives Protocol for publication. Important for Europeana. **Web server, blocking call**: REST web server implementing OAI-PMH specification for query format and results. Available as core service in P4. Development in progress: service interface completed, missing test on PrestoPRIME data samples	Fully supported (part of P4 core services)	Rosetta already has the same functionality
Annotation tools, Tag Gardening **VUA** WP4T2	**Web service**: interface defined in ID4.0.2c using HTTP, results provided in RDF/XML format. Development almost completed: second pilot with additional features available, further improvements in progress	Possible support. Integration in progress.	N/A

Tool Name Partner WP/Task	Tool Description Interface Mechanism Development Status	P4 Support Integration Status	Rosetta Support Integration Status
Metadata tools **JRS** WP4T1	**Web server, non blocking call**: development in progress, two services (mapping and validation) available by June. Published by JRS as REST-ful services. Available also a client application (command line tool) for validation.	Fully supported (ingest module)	Partially supported (validation tool)
Preservation Planning Tools **ITInnov** WP2T1	Simulation and modelling tool **External tool (interactive)**. Development ongoing, with several public releases made.	Possible support. Tool can be run in 'batch mode' on the command line using a configuration file. Results can be recorded and used for helping to take decisions.	Possible support. Tool can be run in 'batch mode' on the command line using a configuration file. Results can be recorded and used for helping to take decisions.
Storage Services **RAI**, **ITInnov**,**B&G** WP5	Broadcaster storage devices. **DRACMA** (RAI): deployed as REST service, development in progress **MServe storage service(ItInnov)** **LTO storage** (RAI, B&G): LTO5 to be checked by BenG, RAI will also investigate for LTO-based storage to be used in test-beds	Possible support of DRACMA. MServe storage service from ITInnovation provides a WebDAV interface (can also be accessed as a filesystem using davfs2 client). The service includes replication and integrity checking.	Required further details for DRACMA. MServe storage service from ITInnovation provides a WebDAV interface (can also be accessed as a filesystem using davfs2 client). The service includes replication and integrity checking.
QA and summarization tools **JRS** WP3T2	**Web service, SOAP based**: tools based on PrestoSpace interface (SOAP), output is XML MPEG-7, currently DAVP, AVDP in the future. Development in progress, interface likely needs to be changed.	Possible support. Integration in progress. Support analysis tool also internally, summarisation to be used from outside	Possible support. Integration in progress. Support for tool used before submission.

3 PrestoCentre Developments in 2011

PrestoCentre had many activities in 2011

- Its launch in March at the Screening the Future Conference in Hilversum. The launch was documented in PrestoPRIME deliverable D6.1.4.
- The week-long PrestoCentre workshop in September, which has been described in Sections 1.3 and 2.2 , above.
- The website has been undergoing continuous development.
- Finally, PrestoCentre has been conducting a survey of archive preservation cost and funding issues.

Highlights from the launch:

Announced March 14th at the Screening the Future 2011 conference held in the Sound and Vision Institute -- one of the largest audiovisual archives in Europe -- PrestoCentre is the outcome of more than ten years' work by several major projects funded by the European Union, conducted by over a dozen organisations.

Screening the Future Conference brought together more than 250 AV media companies, archivists, film makers, advocates and politicians for a two day meeting to listen to 22 keynote speakers on the topics of digitisation, digital preservation and access of AV content.

PrestoCentre's founding members - British Broadcasting Corporation (BBC, UK), Institut National de l'Audiovisuel (INA, France), Institute for Sound and Vision (Beeld en Geluid, Netherlands); Österreichischer Rundfunk (ORF, Austria), and Radiotelevisione Italiana (RAI, Italy) - publicly expressed their intention to work together to share their expertise with each other and with other audiovisual content providers around the world.

PrestoCentre was officially launched on the evening event, on 15/03/2011, where the representative of the five founding members - Jan Müller (B&G), Daniel Teruggi (INA), Sarah Hayes (BBC), Roberto Rossetto (RAI), Herbert Hayduck (ORF) - publicly signed a letter of intention for founding the PrestoCentre.

The Survey

PrestoCentre is gathering data about planned and completed film, audio and video digitisation projects. Information and experiences about digitisation, storage, metadata and access will be of enormous help to other cultural heritage institutions and their funders. We're looking for ways of thinking about budgets, particularly long-term, and for insights into the hardest challenges facing your organisation. This information is being gathered to map out an encompassing, yet detailed, understanding of the challenges facing AV preservation initiatives. The compiled results will be useful for the AV archiving community, helping us all learn about the wider context of experiences, issues and practices in which we operate.
http://www.prestocentre.org/surveys/costs

4 Technology – Brief Encounters

This section gives short summaries of recent developments in areas where audiovisual preservation technology and practice is rapidly developing.

4.1 Authenticity

A conference was arranged in November 2010 by the Dutch audiovisual network AVA_Net, coordinated by Annemieke de Jong of the Netherlands Institute of Sound and Vision (Beeld & Geluid). The general subject was authenticity:

- what does authenticity mean for audiovisual collections, given the many difficulties around the definition of 'the original' or even 'the master copy'?
- how is authenticity determined?
- how is authenticity managed?

The full proceedings were published in October 2011. There is much useful material in the proceedings, but one proposal that can be briefly summarised is a formal hierarchy for authenticity, given in one of the keynote papers (Wright, 2011). The following is a condensed extract from that paper:

Levels of Authenticity

Where do we start, to have some secure sense of use of such terms as 'authentic recording', authentic copy, authentic master? The following is a structure for charting the distance between an item that could be found in an audiovisual archive, and the origin of the content that the item carries.

1. **The original event**
   Throughout, we are discussing sound and image (particularly moving image) *recordings*, so the starting point is: what was being recorded? An engineer's view of authenticity includes the idea that we know what things look like, and sound like, and so we can judge (and to some extent measure) a good recording as contrasted with a bad recording. An archive can hold an authenticated original recording (meaning there is a provenance record establishing that the recording is indeed an original and was made at the time and place indicated) but it can still be a *bad* recording.

   A more interesting situation is where various copies or partial copies exist, and an archive wants to select "the best" for preservation. There are social and historical dimensions to such a selection, but there is an important engineering dimension: which version has good images and good sound?

   In general, the realities of the *original event* can only be imagined. We have images or sound, or both, but we can't prove how much was missed or poorly captured. Although audio and image recording techniques were wonders when first invented, it is worth remembering how much of an original event is never captured. The sound field is a three-dimensional acoustical space that is only sampled by placing one or more microphones inside that space. The pattern of light is similarly three-dimensional, and usually is only captured by a single electronic eye at one position, or by editing the outputs of several such monocular viewpoints in the case of multi-camera film and TV productions. Stereo images have been made for over a century, but little archive

content is in stereo. We can imagine a future with holographic video (and audio) recordings that will allow a much more complete recreation of the original sound field and light field, but archives holdings are closer to a view through a pin-hole camera than to the effect of actually being at *the original event*.

2. **The original recording**

Frances Densmore records Mountain Chief

As has already been stated, true originals are hard to find. In the illustration, a wax cylinder is being cut directly from the acoustical energy produced by Mountain Chief, captured through the horn and fed to a cutter that carves into the blank cylinder. Such 'direct recording' cylinders exist, but most collections of wax cylinders contain commercial recordings that are copies generated from an original recording – and of course would probably by now have been migrated to audiotape or a digital file. Wax cylinders are susceptible to mould; they are fragile to begin with, and become increasingly fragile as they age and dry out; and they can be damaged simply by being played (the groove is altered by the vey needle used to play the cylinder).

However when originals have been preserved, they certainly have a special status – they are as close as we will ever get to *the original event*, and so have their own place in this hierarchy.

3. **The preservation copy**
   Logically, after an original recording come the various kinds of copies, and copies of copies. But there is a kind of copy that deserves its own place in this authenticity hierarchy: the *preservation copy* or *archive master copy*. Just because actual originals are so rare, archives mainly deal in copies. One such copy (the only one the archive has, in many cases) carries the best available version of the content that we want to preserve, and so is designated as the *preservation copy*.

4. **The copy**
   Various access copies can be generated from the preservation copy, such as prints from film negatives, audio CDs and MP3 files from master sound recordings, VHS or DVD or highly-compressed file formats made from video masters. Such proxies have their own level. Regarding authenticity, if the provenance can be traced with certainty, then all *complete* copies have the same provenance in terms of being copies of the same original event, and all derived from the same original.

If this structure for authenticity proves sensible and useful, it can be promoted by the PrestoCentre as a guideline.

4.2 Quality vs compression management

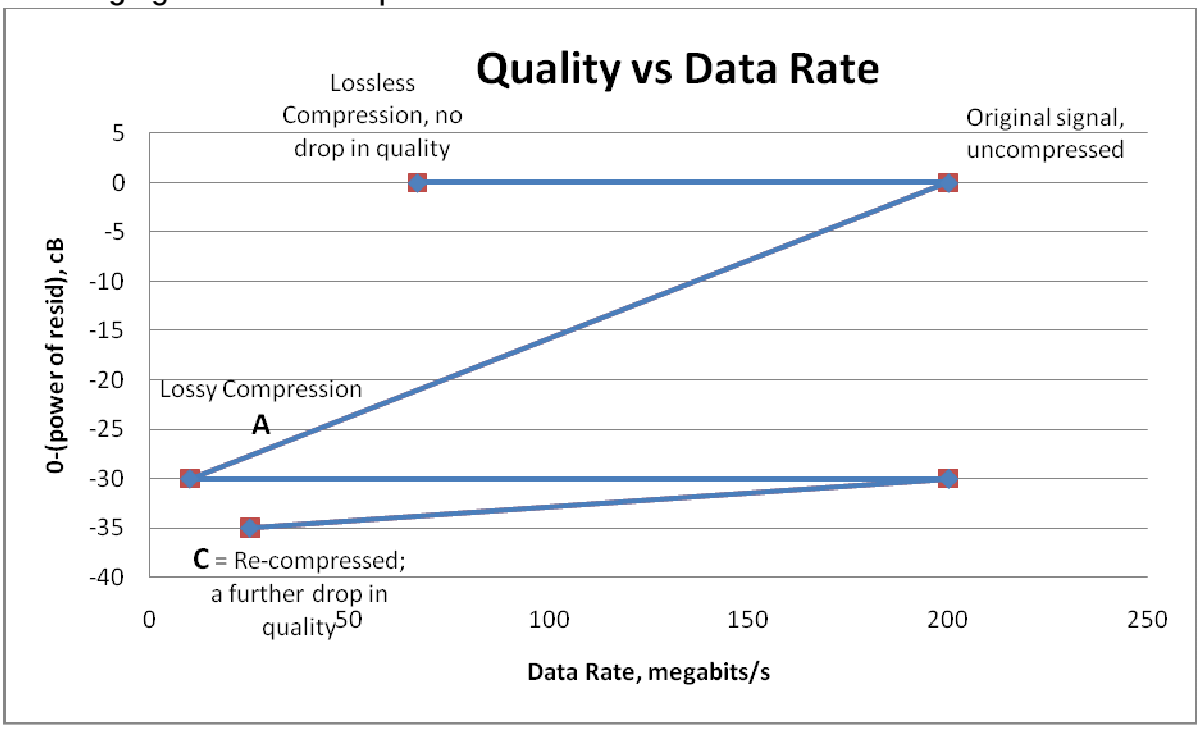
Another BBC R&D publication (at present and internal Technical Note) explores the idea of using a two-dimensional diagramme to show the results of use of data compression on images and moving images.

Lossless compression reduces the datarate without reducing quality, though there are other dimensions to consider when deciding whether or not to use any form of data reduction (compression). There is the issue of the time and equipment involved (a cost issue) and the issue of resistance to errors, as compression reduces the resilience of any type of file (a risk issue).

However the major effects can be captured by simply plotting quality vs datarate. The real utility of the plot is for tracking the life-history of content, as it passes through multiple stages of encoding and decoding.

The basic issue regarding quality is that can be kept as is, or it can be reduced. Quality can never be increased, and once reduced it can never be increased above the reduced level. So the life-history of a file on such a diagram is a zig-zag of datarates, but always going downward in quality, or at best staying level.

The following figure is an example:



Cascading compression - leading to more and more loss of quality

In the figure, the lossy-compressed signal at A can be decompressed back to a full data rate, but it cannot recover any lost quality, and so 'moves' horizontally to point B. If a relatively modest compression is now applied, the quality can only drop. So the signal at C has a higher data rate but a lower quality than the signal at A.

4.3 FRBR revisted - a preservation vocabulary

FRBR is an odd acronym for a systematic approach to describing content, and complex relationships between different items in an archive. It stands for Functional Requirements for Bibliographic Records, and is the result of a great deal of work organised by IFLA, the International Federation of Libraries and Archives (IFLA, 2011).

IFLA proposed, with FRBR, a four-level hierarchy for describing content, beginning at the highest (and most abstract) level with the work (such as a Shakespeare play, a Wagner opera, or a Hollywood brand such as Pirates of the Caribbean. The other levels are then:

- an expression of the work, such as a performance;
- a manifestation of the expression, such as an audio production of an LP (remember those?) or CD of an opera, or a DVD version of a film; and
- an item: an instance of a manifestation: a physical CD or DVD or other physical object – or a file if the manifestation is 'born digital'. The item is the object that is held in a collection such as an audiovisual archive.

The 'IFLA hierarchy' is well-known, and has been applied to many collections. The Netherlands Institute of Sound and Vision has used the FRBR approach for their overall catalogue, IMMIX. What is less well-known is that FRBR defines a whole set of relations, making it an early example of the ontology approach to content description.

Existing FRBR relationships include:

- contains
- is a retelling of
- is a (non-series) sequel to

For audiovisual content, with its peculiar problems of preservation and the need for periodic migrations or creation of 'new masters', another set of relations is needed. These would define the relationships (and by extension, the history and provenance) between the different versions of content held by an archive.

Such a set of relationships could start with:

- is a first generation copy of
- is an nth generation copy of
- is a subset (edited and shortened version) of
- is a different version of (editorial differences, eg new or alternative footage)
- is a remastering of
- is a reduced-quality version of
- is a new preservation copy of
- is a restoration of

If this proposal can be taken seriously, documentation of audiovisual items (especially in those archives that are already serious about following IFLA standards and using FRBR) could include really specific information about Authenticity Type (see section 4.1) and relationships. A cornerstone of any approach to authenticity has to be the ability to simply know what we're talking about. Without specific Type and relationship data (or some equivalent) it is far too easy to be vague and uncertain about audiovisual holdings.

It is daunting to propose any more metadata, or standards for metadata or approaches to metadata – but audiovisual catalogues could be more specific, and if the above approach is useful then it could be promoted by the PrestoCentre.

4.4 The end of broadcasting

In a PrestoPRIME paper for the Virtual Goods conference in Barcelona at the end of September, 2011 (*Broadcasting the Archive*, Wright and Jordan, 2011), the following table was introduced as a general description of media:

Media	**one-to-one**	**one-to-many**
real time	Telephony	broadcasting
non real time	messaging: post, email, SMS	publishing

There is a full discussion in the paper, but the conclusion is that the direction of broadcasting, and the direction of audiovisual archives, is toward web access – which is *non real time* and therefore a form of publishing, and so has more in common (in terms of perception, behaviour, ergonomics and economics) with book publishing than with the historical definition of broadcasting. The basic broadcast concepts were of a schedule, and of a group of simultaneous listeners. These concepts simply disappear when “broadcast” content goes into the archive and comes out, on-demand, on a website.

The paper sees this as a dramatic and positive change:

> **Viewers who are not looking at content in real time can be looking at content over all time.** The totality of broadcast content – the accumulation of everything that has been preserved from the beginning of broadcasting – can become available to the user. Whether “the schedule” remains or not, a broadcaster can point (somehow) to all its own past output, the sum of all past schedules. Exactly as the importance of the schedule diminishes, the importance of the *accumulation of content* increases: as we cease to watch what has been arranged, at the times arranged – so we develop the capacity to follow a trajectory of our own making. We will do this, because the technology makes it possible. (Wright and Jordan, 2011, pp3-4)

4.5 A new approach to analogue video digitisation

At the IASA and FIAT-IFTA conferences in Frankfurt in September, 2011, the company Cube-Tec showed the result of their work with Fraunhofer on a new method of processing signals from videotape: a conventional videotape recorder is used for playback, but virtually ALL the playback electronics are bypassed, and replaced by direct capture and digitisation of the signal directly from the playback head.

The physical playback is the same, but the head-signal capture essentially replaces the 1980s-style signal processing with an entirely new signal processing system, a ‘software videotape machine’ using digital signal processing tools on the RF signal. Fraunhofer and Cube-Tec are developing this technology (Houpert, 2011);

This development changes the game rules for digitisation of video. Where conventional playback succeeds, digitisation can proceed as usual. But where conventional playback fails, there is now a new technique that should be tried because it has potential to succeed.

The PrestoCentre recommendation should be:

- do not wait (do not hold up a digitisation project) until such technology is further advanced and more widely available; but
- for any playback problem that cannot be resolved, the original video format should be kept so that such advanced technology can be applied in the future.

4.6 Correcting temporal variation in audio

As with the new possibility for video digitisation, there is a related possibility for audio tape: increasing the complexity of the digitisation to also capture of bias frequency on audio tape. That information can be used to correct for any errors or distortions in the time domain, such as the wow associated with a cyclical variation in transport speed regulation, a common problem in analogue audio technology. The Clarity service from Plangent Processes (http://www.plangentprocesses.com/) can correct such problems, but only on an original recording (because it relies on the bias signal of the original recording). There is alternative technology from Cedar and Cube-Tec (developed partly as part of the PrestoSpace project) that can (in some cases) correct for temporal distortions without use of the bias signal.

A sensible way forward would be to use conventional digitisation, but where an original audiotape recording can be identified, that original should be saved just in case someone in the future deems it worthwhile to capture the bias for timing correction. As a general issue of best practice, old analogue media should be saved wherever there are doubts about the quality of digitisation.

4.7 Repurposing archive audio

Much is made of image restoration, and just because film is so easily damaged, there is extensive technology for image (and video) restoration. In PrestoSpace, we made a clear distiction between digitisation and restoration. The first was a necessity to bring analogue content into the digital world. The second was separate from digitisation – and was not so much a preservation issue as an access issue.

Restoration has thorny ethical problems, because material can be ‘restored’ in ways that are more about modern expectations than about the original characteristics of the content. The aspect ratio can be altered to suit, for instance, TV broadcasting. The colour can be ‘modernised’, and camera shake can be removed, though it was inherent in the original material.

Audio has been less contentious, but now there is technology to do things to audio that definitely were not part of the original recordings. Source-separation technology has been developed in various research labs such as the Centre for Digital Music at Queen Mary University (London) and has been exploited commercially by, for instance, Audionamix. The technology allows:

- stereo or even mono recordings to be converted to an approximation to the surround sound (five channels plus bass boost, known as 5.1 channels) used in cinemas, home cinemas and HD broadcasting;

- individual elements in a sound track can be identified and either pulled out for emphasis, or eliminated.

Some examples are:

- background music that has attached copyright issues can be seperated from dialogue, supporting re-issue of the content with different background music;
- unwanted sounds (like the infamous vuvuzela plastic trumpets from the 2010 South African football world cup) can be suppressed from the broadcast signal (producing recordings of the same football match that differ markedly in their vuvuzela content – how will future generations know just how loud they were?)
- Particular instruments in ensemble music recordings can be highlighted, to emphasise a particular performer, or for 5.1 channel purposes, or to increase a stereo effect. The trumpet player can be, effectively, moved closer to centre stage or can be moved right to the back. How will anyone determine and document 'the real sound'?

The good news is the success Audionamix has made of ventures for repurposing archive content and increasing its access and commercial value. Their website (http://www.audionamix.com/credits/index.php#2) lists the following projects:

- Music
  - Music Dissociation and Vocal/instrumental Isolation on *West Side Story* along with Chace for the Leonard Berstein Foundation
  - Vocal/Instrumental Extraction for *Sucker Punch* Björk's "Army of Me"
  - Vocal/Instrumental Extraction for *Time to Move*
  - Instrumental Creation for *Sleight Bells*
  - Instrumental creation for *Song Music Published Compilation*
  - Instrumental creation on "Potion" by Morphine and "Dock of the Bay" by Otis Redding.

- Cinema
  - Vocal/Instrumental Extraction on *Inception*
  - Stem Creation for *Gerry Boulet*
  - Stem Creation and Vocal/Instrumental Extraction on *Psycho*
  - Vocal/Instrumental Creation on *The Blues Brothers*.

The not-so-good news is the same as with image restoration: the technology creates new content, which may be further and further away from the original sound as recorded on original carriers.

4.8 Embedded metadata

Metadata about a digital item can be held in essentially just two places: in the item, or 'somewhere else'. As asset management and digital repository and library systems become sophisticated enough to hold objects that are composed of multiple files, and multiple kinds of files (in containers, as discussed in the next section) the problems of embedded metadata should get simpler. At present, the usual object is a file, and so if metadata is to be kept with the object, it has to be embedded in the file. Containers support keeping files together, so metadata can be in the sort of file that is best suited to metadata, such as an XML file, and audio or video can be kept in files best suited to media. But for new, the metadata and the media are mixed (wrapped) in one file, and this mixing of different kinds of information in a single file leads to problems.

The basic problem is that ANY application that works on the file has to keep the metadata intact and usable by subsequent applications. As varieties of metadata proliferate, and use of embedded metadata increases, the instances of lost or scrambled metadata are becoming apparent.

The Association of Recorded Sound Collections sponsored a study of what happens to embedded metadata passing through common applications, which was released in October 2011 (ARSC, 2011). The study was authored by Chris Lacinak of AVPS (Audio Video Preservation Solutions). The study concentrated on audio – where one would expect the situation to be better than for video, for the following reasons:

- only one wrapper format was studied, the Broadcast Wave File format;
- audio applications have been processing the BWF format for over a decade, so one would expect mature applications, wrinkle-free.

Instead, the ARSC Technical Committee concluded that “persistence and integrity issues are prevalent across the audio software applications studied”.

From their summary:

> *Results from this study demonstrate that few of the standard metadata chunks are supported in their entirety by any software application. Rather, applications tend to display and provide access to selected fields of their choosing from each chunk standard. In general, the Broadcast Wave Format bext chunk is the most widely supported, followed by selected fields within the LIST-INFO chunk. Least supported were the XML-‐based chunks—there was some support for selected fields within iXML, but no support for axml, while XMP was only supported by its creator, Adobe.* (ARSC, 2011, p 4)

The next problem arises when trying to use the metadata. Any collection (worthy of the name) needs to have a list of the contents of the collection: an inventory or catalogue. For content arriving with embedded metadata, the most efficient process is to extract the embedded metadata and use it to build the catalogue. At this point either all the files coming into the collection share a single metadata standard, or immediately there is a problem of correct interpretation of metadata.

As an example, the aspect ratio of an image might be called ‘aspect ratio’ in one kind of file, and might be called ‘format’ in another. Even if they agree on the name, the actual way to indicate aspect ratio could vary. For instance, 4:3, 4x3, 4 by 3, 16:12, 1.33 and “1 1/3” are all ways to indicate the standard definition aspect ratio for television content.

For some kinds of metadata these differences don’t matter, as all the collection has to do is hold the metadata, not interpret it (the process for so-called dark metadata). But if a user of a collection is to search for material with a certain aspect ratio, then these differences do matter and need to be solved.

For still images the problems should be solvable. An initiative to agree on IPTC image data is the Metadata Manifesto (http://metadatamanifesto.blogspot.com/). But even for images data, supported by all major camera companies and interpreted by all the main applications dealing with digital photography, there are severe problems. A survey of social media sites (http://www.controlledvocabulary.com/socialmedia/) “*make it look like the people behind*

these services have little awareness of the importance of preserving embedded photo metadata".

The alternative to embedded metadata is to have the metadata within the catalogue or inventory of the asset management or library/repository system. Assuming the metadata can be collected in one application, it should then follow a single metadata model, solving the aspect ratio vs format problem. If there are cataloguing standards and rules, the various ways of expressing an aspect ratio will also be reduced to a single preferred (imposed) method.

However there will still be compatibility problems with non-embedded metadata, of two sorts:

- combining the metadata with the content (the audio and video) to make a file for exchange purposes, or anytime the file leaves the managed environment;
- obsolescence: what to do with the metadata if the catalogue application itself needs replacement. This is a standard library system migration issue, and the bane of all librarians.

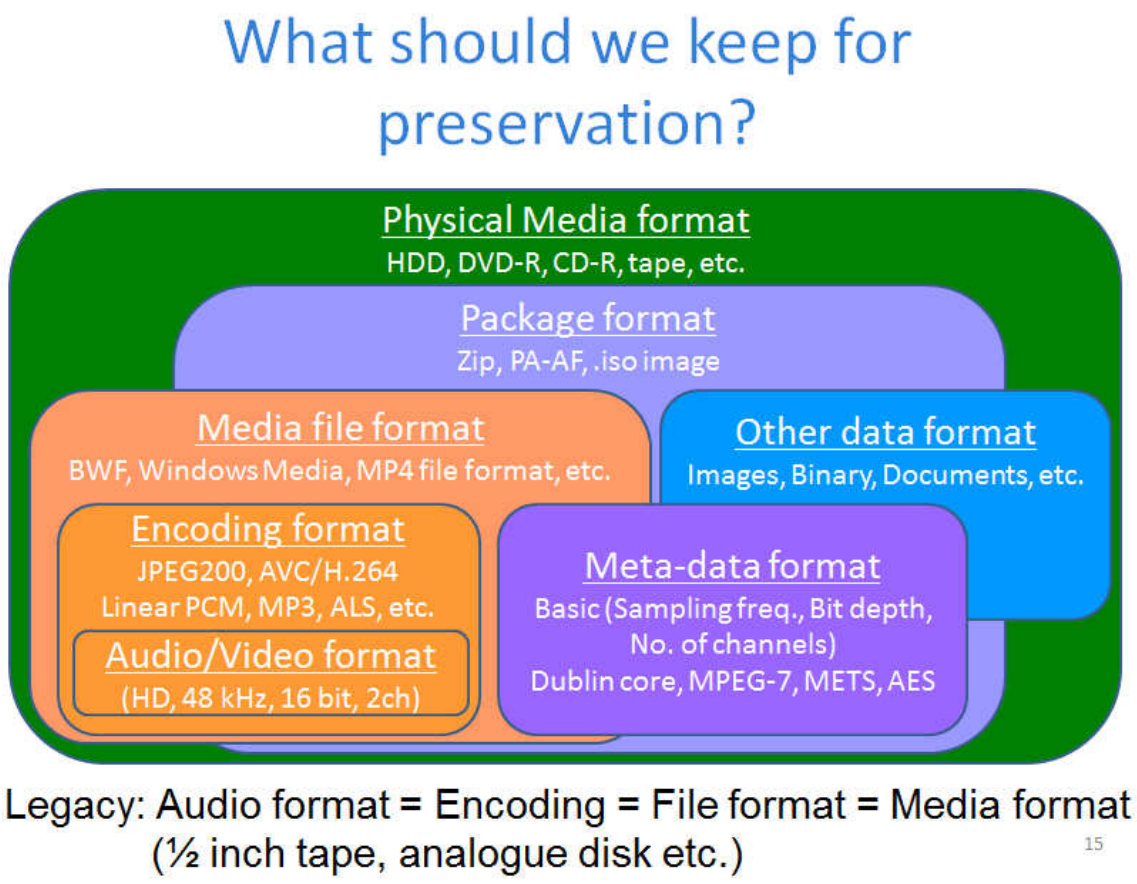
A solution to the obsolescence issue could arise from use of open standards for all the metadata AND for all the linking of all the kinds of metadata and the media. Any catalogue or asset management system uses a database to hold these links -- they are the heart of the system.

There is no open standard for the complete content and relationships within asset management systems, though Front Porch is proposing AXF as a general standard http://www.fpdigital.com/Company/Press.aspx?c=205. AXF, the Archive eXchange Format, is actually a container – a structure for holding content and metadata. Whether or not to support AXF requires more information on the range of possibilities and problems associated with containers (also called packages), which is presented in the next section.

4.9 Containers

BBC R&D is reviewing the possibilities for organising audiovisual content at a unit above the wrapper/file; a public version of that review should be available in early 2012 (Heritage, 2012).

Here is the basic problem:



As indicated, when content came on dedicated carriers there was little ambiguity. A DigiBeta videotape defined the encoding and the physical format. Now we are in the files, and audiovisual content has developed advanced and complicated file formats called wrappers, which hold various kinds of media, in any of a wide range of encodings – plus holding various kinds of metadata. An advanced wrapper such as MXF can hold everything needed for broadcast purposes, but that very complexity leads to problems of incompatible versions of MXF files, and to applications that only support subsets of the MXF format.

The worst part about the last point is that users typically find out which subset they are using, and which different subset is supported by an application, only at point of failure.

As we address the need to put more and more information into audiovisual archives, there is a clear choice:

- wrappers can be made even more complex;
- files can be simplified, and the ”wrapping it all together” function can be carried out by a higher level of information organisation: the container (or package)

There is another path: keep the file as the wrapper, but tie down the specification with an Application Specification – Archive Profile (AS-AP), which is basically a clear label for a subset of a specification – a subset that has no ambiguities, and a subset that works perfectly on all applications that support that profile. The AS-AP route is being promoted by AMWA (Advanced Media Workflow Association, and industry association supporting development of MXF). http://www.amwa.tv/projects/application_specifications.shtml

Candidates for containers include:

- **folder-based**: a folder / directory is used to contain the items making up the object. Of course, this is only possible if the device on which the package is to be stored has a file system that offers folders. This is not possible on LTO tape without using the Linear Tape File System (LTFS; for an example LTFS implementation, see http://www-03.ibm.com/systems/storage/tape/ltfs/);
- **index-file only**: the package consists purely of an 'index' or metadata file. This describes the package and its contents with links to where the content files can be located; and
- **composite:** the 'package' consists of a single file which contains all the package details, metadata and content files, similar to a ZIP or TAR file.

Examples of folder-based approaches are:

- BagIt, a very simple approach developed by the Library of Congress. Because of its simplicity, it is widely used. The Stanford University Digital Library project has switched from METS to bagit as its container (Cramer and Kott, 2010).
- NHK/IBM DPP (digital program package), a more sophisticated approach that has also been used by FOR-A in their video archiving product. This approach has been formalised by the Japanese Association of Region Industries and Businesses (ARIB) as TR-B31 http://www.arib.or.jp/english/html/overview/rb_ej.html and is a candidate AXF standard.

The main examples of the index-file approach is METS, used by Ex Libris in their Fedora digital preservation system, and also used by PrestoPRIME for their P4 platform. Though an accepted standard for digital libraries and repositories, METS is probably unknown elsewhere in broadcasting. In coordinating audiovisual standards across all branches of the US government, the Library of Congress found no use of METS for audiovisual content (FADGI, 2010).

Examples of the composite approach are:

- MPEG-A PA-AF, a detailed proposal, but with no information on take-up;
- Front Porch Digital AXF, a proprietary format that the company (FPD) is offering as an explicit version of the more general (and more vague) AXF standard.

It is too soon for PrestoPRIME to promote one of the above approaches against all others, but this is an area that PrestoPRIME is watching closely. The PrestoCentre should be able to promote best practice in this area at some point during 2012.

5 UK, European and International Developments

This section reviews significant activities by governments, public bodies and major projects and institutions that affect preservation and use of archived audiovisual content. There are four sections:

1. Developments in Europe;
2. Sources of further information, focused on Europe;
3. At the national level: developments in the UK. The report cannot begin to describe all activity in Europe at the national level, so this report on the UK is as an example; and
4. Developments outside Europe

5.1 Europe

5.1.1 push for digital preservation

The Future of the Past – Shaping new visions for EU-research in digital preservation Workshop, 4-5 May 2011, Luxembourg
http://cordis.europa.eu/fp7/ict/telearn-digicult/future-of-the-past_en.pdf

Vilnius University Library is carrying out a survey of training opportunities in digital curation and long-term preservation within Europe and internationally as part of the DigCurV (**Digital Curator Vocational Education Europe**) project.
http://www.digcur-education.org/ or contact us on info@digcur-education.org

5.1.2 push for digitisation

This is an area of major change: EU support for actual digitization. For the first time, the European level is committing funding to more than just the Delos-style coordination of national activity: Europe is actually funding digitization.

The European Commission Encourages The Digitisation Of Culture
http://www.edri.org/edrigram/number9.21/culture-digitisation-european-union

Digital Agenda: encouraging digitisation of EU culture to help boost growth (28.10.2011) press release
http://europa.eu/rapid/pressReleasesAction.do?reference=IP/11/1292&...
website http://ec.europa.eu/information_society/digital-agenda/index_en.htm

5.1.3 developments in copyright

Green Paper on the online distribution of audiovisual works in the European Union
http://ec.europa.eu/internal_market/consultations/2011/audiovisual_en.htm
JISC response: http://www.jisc.ac.uk/news/Onlinenewsroom/Statements.aspx

European Commission summary **Copyright in the Information Society**
http://ec.europa.eu/internal_market/copyright/copyright-infso/copyright-infso_en.htm

Consultation on the Commission Report on the enforcement of intellectual property rights
http://ec.europa.eu/internal_market/consultations/2011/intellectual_property_rights_en.htm

Orphan works: http://ec.europa.eu/internal_market/copyright/orphan_works_en.htm

WIPO Broadcasters' Treaty
http://ec.europa.eu/internal_market/copyright/broadcaster/index_en.htm

5.1.4 Europeana update

Strategic Plan: Accessibility of European cultural Heritage: 2011-2015
http://www.europeanaconnect.eu/news.php?area=News&pag=63 The Strategic Plan reflects Europeana's commitment to improve the accessibility of European cultural heritage. The three challanges that Europeana and its stakeholders will be grappling with are: intellectual propoerty barriers, acceleration of digitisation in Europe, long-term funding. The answer to these challenges will include a series of actions, namely, content aggregation, facilitation of knowledge transfer, content distribution and new ways of engagement wth cultural resources for users.

Libraries Project: launched in February 2011, the project brings together 19 leading research and university libraries to offer on-line access to about 5 million digital objects from their valuable collections. The project will enable access to Google Books, dissertations, open-access journals and many other useful resources.
http://www.europeanaconnect.eu/news.php?area=News&pag=65

Linked Open Data Pilot July 2011; the Europeana Linked Open Data Pilot contains metadata on 3.5 million texts, images, videos and sounds gathered by Europeana. These objects belong to 10 direct Europeana content providers, including about 300 cultural institutions from 17 countries, who have reacted early and positively to Europeana's initiative of better promoting open data and new data exchange agreements.

Data Model publications:

> **The EDM Mapping Guidelines**: This document describes subset of classes and properties from the model that will be used in the first implementation of EDM. It has an example of original data and that same data converted to EDM showing the distribution of the properties amongst the classes. It also outlines the process for migrating from ESE to EDM.
>
> The **EDM Primer**: this is a revised version of the original EDM Primer giving the "story" of the EDM and explaining how the classes and properties are used together to model data and support Europeana functionality. It has been revised to give more prominence to the aspects of EDM of greatest interest to providers and to reflect the latest modifications to EDM itself.
>
> In addition, we also provide a two-page EDM Factsheet that sums up the rationale and expected benefits of EDM. All documents are available at:
> http://version1.europeana.eu/web/europeana-project/technicaldocuments/

Europeana: Commission sets precise targets to states (28.10.2011)
http://www.europolitics.info/social/europeana-commission-sets-precise-targets-to-states-art317036-22.html

5.1.5 Centres of Excellence

IMPACT Centre of Excellence: launched Oct 2011: http://www.slideshare.net/bne/centre-of-competence-in-digitisation
project: http://www.impact-project.eu (not working 21 Nov)

IMPACT promotes better technology and workflow for document scanning. This launch means there are now two Competence Centres hosted at the British Library: Open Planets Foundation and Impact.

Information on all the digitisation and digital preservation competence centre activity is presented here:
http://cordis.europa.eu/fp7/ict/telearn-digicult/digicult-digitisation-capacity_en.html

5.1.6 Digital Preservation Next Steps:

Presentation by *Javier Hernández-Ros*
European Commission, Information Society and MediaUnit "Cultural Heritage and Technology Enhanced Learning" ***APA 2011 Conference, 8-9 Nov., London***

Summary:

- Now: preservation of **files / data**
- Near Future: preservation of **complex material:** heterogeneous and evolving content (web), applications, processes…

Moving beyond files means moving into the issue of 'containers for complexity' and how OAIS can cope. PrestoPRIME is active in this area, as audiovisual archiving moves toward archiving of objects with their full context (scripts, contracts, publicity materials, production paperwork, reviews), not just their technical core (video, audio, technical metadata such as subtitles and timecode). See Packages, Section **Erreur ! Source du renvoi introuvable.**

5.2 General sources of information

5.2.1 UK Digital Preservation Coalition 'what's new' –

http://www.dpconline.org/newsroom/whats-new
William Kilbride: Editorial: Too Big to Fail: economics of digital preservation
http://www.dpconline.org/newsroom/whats-new/765-whats-new-issue-39-november-2011#whatswhat39 which identifies economic and usability dimenstions that underly the motivation for people and institutions to undertake digital preservation activity.

Examples:

- digital preservation as a *derived demand*
- digital preservation is about turning off legacy systems (confidently)
- Imagine if instead of being faced with a dialogue box that says 'I can't open your file', the computer said 'I've not seen a file like this since 1994 - would you like to choose migration or emulation to access it?'

5.2.2 Girl in the Archive –

A very active blog from the BBC, concentrating on asset management, librarianship, digitisation and digital preservation – as they pertain to audiovisual content.
http://girlinthearchive.wordpress.com/

5.2.3 Information Society: Digital Libraries

A summary of year's activity:
http://ec.europa.eu/information_society/activities/digital_libraries/index_en.htm

Digital Libraries Initiative: Jan 2011 **Comité des Sages Report The New Renaissance**
http://ec.europa.eu/information_society/activities/digital_libraries/comite_des_sages/index_en.htm

5.2.4 Commissioner Neelie Kroes:

http://ec.europa.eu/commission_2010-2014/kroes/

5.3 UK

5.3.1 Audiopedia

Mark Thompson, Director general of BBC, has revealed the BBC plan to share distribution of its content through an online audio archive, Audiopedia.
http://www.telegraph.co.uk/technology/news/8865357/BBC-to-open-vast-radio-archive-online.html

5.3.2 Hargreaves review of copyright

There has been a general review of copyright law in the UK. The academic body JISC, which is responsible for infrastructure (including university-level digital libraries and repositories) has responded. The full response is online:
http://www.jisc.ac.uk/news/Onlinenewsroom/Statements.aspx

Summary: JISC: Letter to Daily Telegraph, 13 July 2011
By recognising that the entertainment industry's intellectual property (IP) incentives are different from those needed for education and research, the Hargreaves review can strengthen UK R&D, and support growth (report, July 4).

We support the introduction of modern digital "exceptions" in copyright law to support archiving, teaching and non-commercial research. The argument for an exception for text and data mining to speed up scientific and medical discovery is compelling. None of these exceptions in copyright law should be able to be subsequently over-ridden by private contract.

We must find a solution for "orphan" works, support better licensing mechanisms to clear rights *en masse* through collecting societies (that adhere to minimum standards), as well as ensure that in the future IP policy is evidence based and less susceptible to "lobbynomics."

5.3.3 UK Digital Public Space

A group of national cultural heritage organisations in the UK (BBC, British Library, British Film Institute, National Archive and more) are working together on how to make their content publicly available over the Internet, without getting caught up in issues of copyright or unfair competition with the commercial sector, This **digital public space** is a major development, as it envisages a common data model and linked metadata structure for all the public space content, going far beyond the 'common catalogue' model of Europeana.
Keynote: The finest digital drawing room in all Europe: how can we build a digital public space? by Bill Thompson.

5.3.4 Access Projects in the UK

Screen Heritage UK project: SHUK is a large (£22.8 million) and complex project (involving 12 regional film archives in addition to the British Film Institute) that had its funding announced late in 2007. The project was complicated by changes in the structure and funding of the BFI, as well as a change of government and a raft of other issues. Nevertheless the project has produced major achievements:

- conservation: construction of a £6 million vault for film conservation;
- digitisation: film scanning and digital storage for the regional film archives
- access: online catalogues of regional film archive content, available to the general public (for example http://sasesearch.brighton.ac.uk/yourfilmarchives/ and http://www.eafa.org.uk)

SHUK launched 5 September 2011 with a BBC-BFI co-production, The Reel History of Britain. More information is here http://www.bfi.org.uk/about/policy/screen-heritage.html?q=screenheritage

Access to Video Assets project: this is an access and re-use project of the Open University. The focus is to digitise (where necessary) audiovisual assets previously created by the OU, and place them in an asset management system so that current OU teaching and other activity can find and re-use these assets. Preservation is a by-product of the project rather than an end in itself. This project provides an important example of combining preservation of content with use of content – something of value to the institution in order to obtain a budget and deliver a benefit.

The project was presented (Allcock and Alexander, 2011) at the DPC Briefing Preserving Digital Sound and Vision (DPC, 2011). The project digitised 1200 videotapes and films, and placed the results in a Fedora digital repository. Also, 145,000 pages of documentation were digitised, providing the overall educational framework around the 1200 items, to give them context and enhance their ability to be re-used. The user interface provides granularity and time-based navigation. Overall this project is an outstanding example of best practice.

5.4 International Developments

5.4.1 new audiovisual faults database:

http://preservation.bavc.org/artifactatlas/

5.4.2 a US digital library:

Europeana and the newly-founded Digital Public Library of America [DPLA] have announced that they will work together: http://europeana.eu/portal/aboutus_press.html

Digital Public Library of America Update by Maura Marx.

5.4.3 Internet Archive: 911 broadcast news

Internet Archive has launched web access to one month of output from 20 broadcasters – but not to the ensuing ten years! At the PrestoCentre launch in Hilversum in March 2011, Brewster Kahle the originator of Internet Archive announced that he has been collecting broadcast content for the whole decade since the 911 events, and would launch it on the 10th anniversary. So there is more to come from Internet Archive, as only one month of content has so far been made available. http://www.archive.org/details/911

5.4.4 Open Video conference:

http://www.intelligenttelevision.com/video/project/open-video-conference/

5.4.5 American Archive Content Inventory Project

http://americanarchiveinventory.org/project/

Courtney Michael, Project Manager

This project has some similarity in aims to Europeana, but in a narrower context: public broadcasting from one country. It has interesting methods of reporting activity using maps: http://americanarchiveinventory.org/dashboard/map

Europeana has related search and retrieval technology under experimental development http://eurohack.pictura-dp.nl/ and http://www.version1.europeana.eu/web/api/hackathon-prototypes#geog but not a geographical display of all content (which may not make sense if done literally for tens of millions of items, so Europeana would need to think of something like density of items, with the unit of area being in some sensible proportion to the area being displayed, and the representation of density also being something ergonomically simple and useful – perhaps colour).

5.4.6 U.S. Copyright Office: “Priorities and Special Projects”

http://girlinthearchive.wordpress.com/2011/10/31/u-s-copyright-office-outlines-priorities-and-special-projects/ : “Orphan works, preservation for libraries, mass digitization, and fighting digital piracy are among the priorities set by the Register of Copyrights Maria A. Pallante this week in a paper outlining the U.S. Copyyright Office’s “priorities and special projects” for the next two years.”
http://www.copyright.gov/docs/priorities.pdf

6 Preservation of Signals

The following is an engineer's view of digital preservation and audiovisual content, concentrating on the fact that *the bits* in an audiovisual file represent a signal.

Formally, audio and video can be considered one-dimensional signals: a voltage (or other parameter) that varies as a function of time. For audio, the voltage variation out of a microphone is the electronic *analogue* of the variation in air pressure at the surface of the microphone, and hence a recording of that voltage is an analogue recording of the sound picked up by the microphone. In this sense, audiovisual files are very different from text: audiovisual files hold data, much like any other measurement data in scientific experiments – and unlike word-processing files or PDFs.

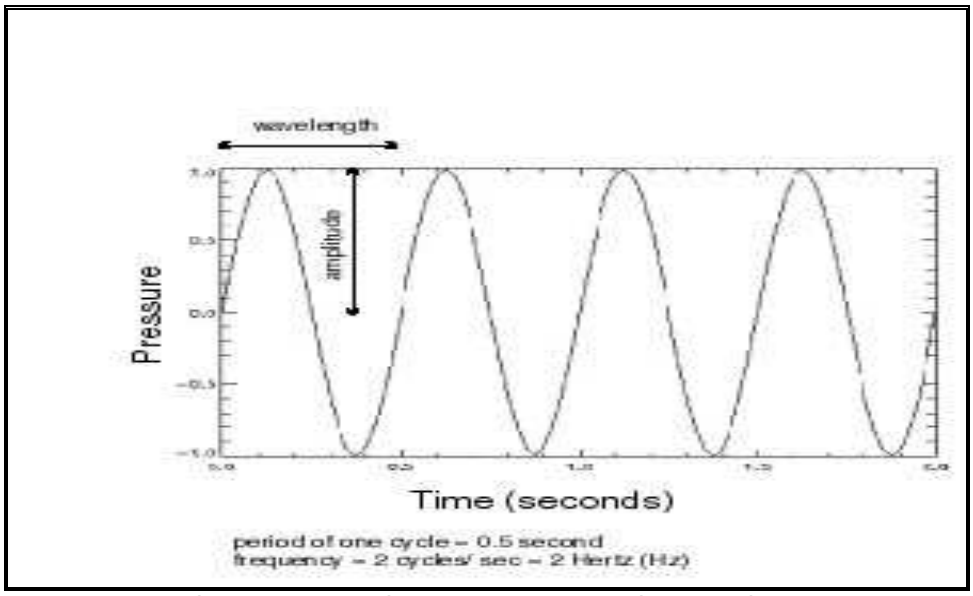


Figure 1- A sound wave: pressure varying over time.
The signal out of a microphone would look the same

A video signal as produced by a television camera is also one dimension vs time, despite the visual field being two dimensional. Part of the electronics of a camera is a *raster scanning* mechanism (see ***Figure 2***), similar to the scanning in a photocopier or facsimile (fax) machine – a mechanism that divides an area into a number of lines, allowing one measurement (response to *brightness* in a monochrome camera) to represent the field as, again, voltage vs time. Colour video is more complex, and can be thought of as three simultaneous scans, and three parallel signals. Component video is indeed three signals, while for transmission down one wire (or for broadcasting as one signal), there is the alternative of composite video where colour and brightness information are combined into what looks like one signal.

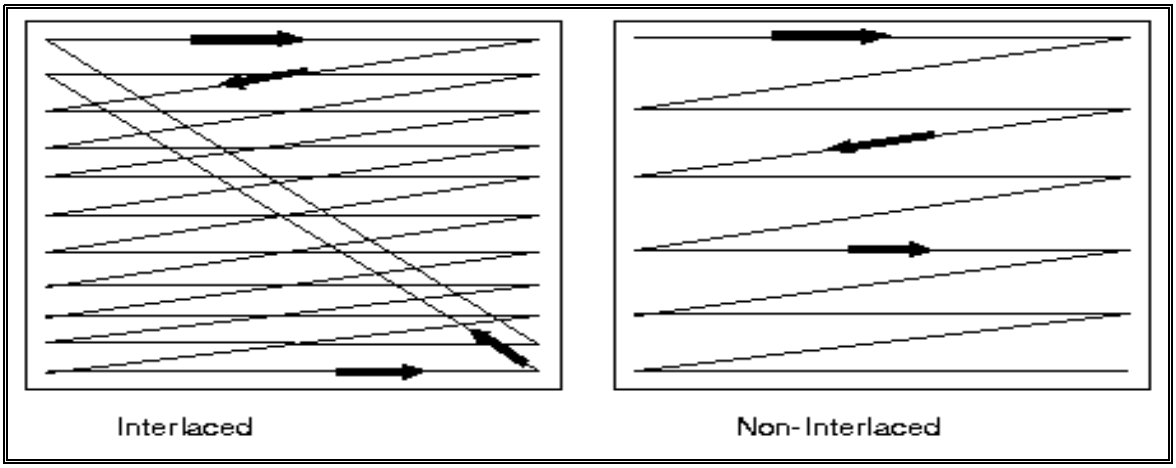


Figure 2- two types of raster scan, used to convert an image into a signal. Interlaced scanning allows twice as high a refresh rate, reducing flicker.

6.1 Four stages in the life of digital audiovisual content

For sound and moving image preservation, the following stages in the creation and digital preservation of sound and moving images need to be kept clear:

1. **the original signal**: the audio from a microphone, the video signal coming out of a video camera. These signals have physical properties (bandwidth; dynamic range) which can be defined and measured. The quality of a recording and the success or failure of any process of copying, digitisation or preservation can be reduced (in large part) to how well that process maintains these two physical properties of the original signal.

2. **capture of the original signal:** recording of a signal onto a carrier (also called support, physical medium or recording format). For a century, the methods of capturing a signal were tied to the carrier of the signal: a wax cylinder or film reel or videotape. Digital technology produces recordings that are independent of carriers. **Carrier independence is liberation**: discs, tapes and films deteriorate or get damaged. Born-digital recordings are liberated from these carrier-based problems, leading to a desire to liberate analogue recordings – by digitisation.

3. **digitisation of the recording** (of the original signal)**,** which has three cases**:**
   - **digitisation of analogue media:** analogue recordings (vinyl discs, magnetic audio tape, VHS and U-matic video tape – and film) are all analogue technology. The recording medium – the carrier – is an integral part of every analogue recording system. Analogue recordings can be played back and re-recorded onto a new carriers, or digitised and so released from carrier dependence. Digitisation has to ensure that the digital version has the same bandwidth and dynamic range as in the original, to capture the original quality;
   - **ripping of digital carriers:** digital recordings on dedicated physical carriers (audio CDs, minidiscs; video DVDs, DV tape) are already... digital. These are recordings that use numbers to represent the sound or moving image, but as with analogue recording there is a specific medium or carrier (and its associated read/write technology) that is an integral part of the recording method. These recordings are NOT files. Specific technology and workflow is needed to move such recording into files. For CDs and

DVDs either the individual sounds or video clips are *ripped* from the original carrier, or a special file is created that is meant to be an exact image (clone) of the data on the CD or DVD. The process of moving content from these digital carriers and into files is commonly referred to as digitisation, but strictly the process is not digitisation because the starting point is already digital. **Ripping** is a clearer term, though it is more used for audio than video; and

- **files:** digital recordings that exist as files on digital storage. Modern equipment can record sound and moving image signals directly into the memory in the device, at which point the recordings are files and have an existence that is independent of any particular storage media. Files are the final stage (so far!) in the technology for audio and video. Files come from the three cases just set out: analogue content is digitised; digital carriers are ripped; or content can be recorded directly into files at moment of capture.

4. **preservation** (of a file-based digital recording of the original signal): the final stage is the life-cycle of files, the problem of **preserving the ability to use** the digital representation of a signal, meaning preserving the numbers, but also preserving the technology needed to decode (render) the numbers.

Audio and moving image content has a particular preservation problem: the coding of the signal can be a compromised: not actually capturing the full signal, but instead losing some of it (lossy encoding) to get a more compact representation (reducing storage and transmission costs). Unfortunately coders/decoders (codecs) go out of use and are replaced by newer technology. The file format holding the coded signal, the wrapper, is also subject to obsolescence. The failure and obsolescence of storage technology and the obsolescence of encode/decode methods and wrapper formats are major digital preservation problems for audiovisual content.

6.2 The significant properties of a signal

The technical characteristics of an audio signal can be pictured as a rectangle: it has height and width. The height is how loud (for audio) or how bright (for video) the signal is compared to the background noise; this is called the dynamic range, or signal-to-noise ratio, and is measured in decibels, dB. The width is the range of frequencies, called bandwidth, and is measured in Hertz, Hz.

How tall is the rectangle for audio? A sound, an acoustical signal, is a variation in pressure. This variation is continuous: a waveform. Digitisation means assigning numbers to this waveform. A signal with a wide dynamic range needs an equivalently wide range of assigned numbers. It is conventional practice to specify this range in terms of how many binary digits are allowed. Eight-bit sampling uses the numbers from 0 to 255 (because 255 is one below the 8^{th} power of 2); 16 bit sampling allows numbers up to 65535. Eight-bit sampling has a maximum dynamic range of 48 dB (= 20 log 256, to a very close approximation). Sixteen-bit sampling has twice the dynamic range: 96 dB. The human ear has, at its best, a dynamic range of 130 dB. However the background noise present in even the quietest situations means that the human ear would almost never hear a sound more than 100 dB above the background noise. For sound reproduction, a 40 dB range is low quality, like a pocket radio or cheap cassette recorder. Studio quality recordings on tape achieve a 70 dB range, and recordings on the audio-CD format have a 96 dB maximum (because a CD uses 16-bit sampling).

Audio digitisation at more than 16 bits has been possible for many years, and the limiting factor for professional digitisation is now the dynamic range of the original signal. Systems that digitise with more than 16 bits are commonly referred to as 24-bit systems, but some of those 24 bits will be unused, as the state of the art for practical installations is reached at something closer to 20 bits.

How wide is the rectangle for audio? The range of human hearing varies significantly with age, and begins to deteriorate at puberty. Young adults (age 18 to 20) can be expected to hear sounds up to 18 kHz. The highest note on a piano has a fundamental frequency (inverse of the periodicity) of just over 4kHz, though to hear the sound properly requires hearing its harmonics (multiples), and the fourth harmonic of that highest piano note would be above 16 kHz and near the extremes of normal human hearing. The fifth harmonic would probably be inaudible to most everyone.

When digitising, samples must be taken at twice the highest audio frequency in the signal, at a minumum. Usually a factor of 2.3 or 2.4 is used. This means that an 18 kHz sound requires (theoretically) at least 36 kHz sampling, and in practice requires more like 43 kHz. One reason that the audio-CD standard uses 44.1 kHz sampling is to have a width (literally, a bandwidth) of sound up to 18 kHz.

When digitising, it is vital that the incoming sound is band-limited to the bandwidth of the digitising process. If not, high frequency sounds will be mirrored (aliased) as though they were within band. If a high frequency sound is present, such as 20 kHz, and the sound is sampled too low, say at 30 kHz, the signal at 20 will alias around half of 30 (namely 15), and appear at 15-5 kHz = 10 kHz, instead of where it belongs at 15+5 kHz. An inaudible sound at 20 kHz can be turned by aliasing into a very audible sound. One way to minimise aliasing is to sample at a rate much higher than the standard 2.4 times the maximum frequency (oversampling). Such oversampling allows much simpler anti-aliasing filters to be used, which in turn means that bulky hardware associated with digitisation can be eliminated. This consideration is one of the reasons for recommending that archived audio be sampled at 96 kHz rather than the 48 kHz used in broadcasting, and the 44.1 kHz used for audio CDs.

How big is the rectangle for video? Audio and video are very different. Audio engineering can quite easily capture sound so that when reproduced, it can be mistaken for the original. The famous Victor image of the dog listening to “his master’s voice” coming from the trumpet on a wind-up gramophone is an illustration of the aspirations – and achievements – of audio technology since its inception.

Figure 3 - His Master's Voice, a registered trademark in 1900

For video, nobody seriously expects the image on a TV screen to be mistaken for the original event being recorded. The 'little people on the screen' are the wrong size, for a start. For audio we considered the range of human hearing, and how to match that. For digital video, the question is the characteristics of the signal coming out of a television camera: its bandwidth and signal-to-noise ratio.

Fortunately, there is a long-established standard for digital video, the ITU Recommendation 601[6], and particularly the 10-bit version standardised as SMPTE 259M. Why 10 bits, when for audio we expect a minimum of 16 and commonly encounter 20-bit samples? Here is where human perception is the basic issue: the dynamic range of hearing (of the power of a sound) is simply far wider than the perceived range of brightness from very dark to very bright. However the full dynamic range of vision is not captured by any camera, and not reproduced by any projection or display mechanism. The preservation issue thus reduces to preserving the dynamic range and bandwidth of the signal coming out of the camera – and we have Rec 601 as an agreement for how that should be done, if done digitally.

The problem with video – is that so much video in use and in archives does not follow the Rec 601 standard. The principal reason is the high datarate (the product of the bandwidth and dynamic range). Full Rec. 601 has a datarate of 270 megabits per second (Mb/s), which corresponds to roughly 125 gigabytes per hour (GB/hr). That's nearly 200 CD-ROMs full of data for every hour of full Rec 601 video – or about 30 standard DVD-ROMs.

For audio, it is now accepted standard practice to archive as uncompressed data. For video, compression is used at every stage: in the camera, in the original recording mechanism, in production and in distribution. One of the effects of lossy compression which can be measured is the reduction in the dynamic range – but it can only be measured if the compressed and uncompressed signals are both available for comparison. Typically only the compressed version is available, as in the data on a DV (8x compression) or DigiBeta (3x compression) videotape.

[6] http://en.wikipedia.org/wiki/Rec._601

The further problem for video is that so long as it is represented using compression, rather than as straight Rec 601 uncompressed data, it is subject to cycles of decoding and recoding whenever accessed, processed or transmitted. These cycles of cascaded coding-decoding tend to proceed relatively invisibly until there is sudden and catastrophic breakdown, and major image impairment.

The insidious problem with video is that standard engineering cannot predict the point of breakdown. The significant properties of bandwidth and dynamic range may be only slightly reduced by a coding system, and still when decoded and recoded in another way there can be coding breakdown and very large errors. This process is one of the reasons that audio and video engineers may prefer analogue signals and systems to digital ones: they degrade in a continuous rather than discontinuous fashion, and the significant properties of dynamic range and bandwidth remain good indicators of the quality of the signal.

The mechanism used to reduce the datarate of digital signals (compression, coding) is to reduce the redundancy. By definition, this reduces the resilience of the signal. As an analogy, a processed signal may be a rectangle of nearly the same height (dynamic range) and width (bandwidth) after compression (data reduction), but it is as though the walls of the rectangle were made thinner and thinner, until eventually the box collapses despite not showing reduction in size as a warning.

6.3 How to preserve signals

From the above discussion, there are two conclusions:

1. The bandwidth and dynamic range of originals should be known, so that they can be maintained through subsequent processing. These two significant properties can be estimated by experience listeners, and their maximum values are already well-known for specific carriers. All digital processing has to respect, to *capture*, these properties for digitisation to be adequate, and for any subsequent digital processing to be acceptable.
2. Use of lossy compression makes it impossible to maintain quality simply by maintaining the significant properties of bandwidth and dynamic range. Lossy encoding makes a signal effectively brittle and easier to corrupt – but there is no direct way to measure or even estimate the *degree of brittleness*. For this reason, among many more, the only secure way to archive digital content is by representing it without lossy compression, and even better without any form of compression. Even lossless compression adds complexity and removes redundancy, increasing risk.

7 Glossary

Term	Definition
4:2:2	The allocation of digital samples to the luminance (black and white dimension) and chrominance (colour) dimensions in the digital representation of video, as used in the SDI standard
AMIA	Association of Moving Image Archives
Viewing proxies	Access copies of digital content, usually in reduced quality but smaller files, faster to transmit over networks or Internet
Analogue recordings	The recording method is an analogue of the original signal: a groove in a vinyl disc is the analogue of the sound pressure into the microphone that eventually produced the groove
Asset management system	Software that organises a collection of files, usually including separate metadata and search and edit tools
AVI	A wrapper format used by Microsoft
Bagit	A simple container for digital library contributions, developed by the US Library of Congress http://www.digitalpreservation.gov/tools/index.html#b
Bandwidth	The frequency range of a signal, from lowest to highest frequencies
Betacam, BetacamSP, BetaSP	Varieties of the last Sony analogue videotape format
BFI	British Film Institute
Bias signal	A frequency above the range of human hearing, used in recording on analogue audio tape
Bit (b), byte (B)	These are general terms, defined elsewhere; in audiovisual media, data rate for real-time playback is measured in bits per second, while file size is measured in bytes; to convert: storage in bytes per hour = ½ of data rate in bits/sec, times 1000. Example: CD audio at 1.4 Mb/s takes 0.7 GB per hour for storage; this conversion is 90 per cent accurate
Born digital	Files that did not come from digitising an analogue source
Broadcast WAV files	The EBU standard for a WAV file, with extra metadata
Carrier	Something physical that can hold the content
CD, CD-ROM	Optical media for audio and general data; stores about 0.7 GB
Clone	An exact copy of a digital object
Codec	The abbreviation for coder/decoder
Coder	The method (software) for assigning numbers to a signal
Coding	The process of assigning numbers to a signal using a coder

Term	Definition
Colour video	A combination of a black and white signal and a separate signal giving colour information
Component	Any method of handling video that uses separate signal, either as black and white separate from colour, or as three colour signals (red, green, blue)
Composite	A video signal that mixes the black and white and colour information into a single signal
Concealment	Replacing a missing line or group of lines or even an entire frame of information by using a previous line, lines or frame
Conservation	Keeping what you have without changing it
Container	As used in this document, a container is a way to keep a group of files together, ideally with metadata describing the composition of the container. A container is a unit of information management that is above the level of the individual file. Examples are bagit and METS (cf).
Data tape	Magnetic tape that holds numbers
dB, decibel	A unit of measurement of the relative amplitude of signals. When one signal is twice as large as another (equivalently, has one more bit) it is approximately 6 dB greater; human hearing has a range of about 120 dB, though our noisy world leaves us with a usable range of about 70 dB; the best analogue recording equipment has roughly a 70 dB dynamic range, while an audio CD with 16-bit word length has a 96 dB range (6x16)
Decoder	The reverse of the method used by a coder
Decoding	Using a decoder to get an unencoded (plain) signal
DigiBeta	The Sony professional digital videotape format
Digital intermediate	A digital representation of film frames, used in restoration, animation, computer graphics and computer generated images
Digital recordings	Stored representations of signals, using numbers
Digitisation	The process of assigning numbers to analogue signals
DVD, DVD-ROM	Optical media for video and general data; a basic DVD stores about 4.7 GB, though larger amounts are now possible
Dynamic range	The range between maximum size of a signal and either the minimum size or the noise level (whichever is greater); the word length (number of bits) of a sample limits the dynamic range (to approximately 6 dB per bit)
Embedded metadata	Metadata carried within a file
Encoder	The same as coder
Encoding	The same as coding

Term	Definition
FIAT, FIAT-IFTA	International Federation of Television Archives
File	A unit of digital storage; it is surprisingly hard to know what exactly a file is; everything we do with files relies upon multiple levels of technology to create what a file is and does; a file is a performance
Fixity check	A method for ensuring the integrity of a file http://id.loc.gov/vocabulary/preservationEvents/fixityCheck.html Further details about implementation: http://digitalpreservation.ncdcr.gov/newtodp.html and http://archivematica.org/wiki/index.php?title=Overview
Flash memory card	Storage with no mechanical or moving parts, as in a memory stick
Hard drive	Also disc or spinning disc; magnetic storage using spinning discs
high definition, HD	The new standards for television, usually 720 or 1080 vertical lines instead of 576 (onscreen) lines for standard UK TV
Hz	Hertz; cycles per second, the unit of frequency measurement
IASA	International Association of Sound and Audiovisual Archives
Interlaced	Television reduces visual flicker by sending images in two halves: all the odd-numbered lines and then all the even-numbered ones; this allows the rate of screen refreshment to be doubled, without increasing the overall data rate
JISC	Joint Information Systems Committee, the infrastructure body supporting UK higher education
JPEG	A (compressed) coding and file type for images, developed by the Joint Photographic Experts Group http://www.jpeg.org/ cf MPEG
JPEG2000	An updated JPEG codec, which includes lossless coding and supports video as well as still images
k, M, G, T, P	Kilo, mega, giga, tera, peta are the prefixes used to describe large numbers, increasing by a factor of 1000 for each step in the series – so 10 GB is 1000 times as much as 10 MB, and a million times as much as 10 kB; unfortunately when dealing with computer memory the factor is 1024, not 1000 – and the capitalisation (or not) of k is used inconsistently
Lossy encoding	Representing a signal with fewer numbers, with the result that the original signal is only approximate and cannot be exactly reproduced; usually there is a measureable difference between the original and the lossy version, though the difference may (or may not) be perceptible
Master	The reference version of an object; the version with the highest quality
Memory stick	A storage device using solid-state memory

Term	Definition
METS	Metadata Encoding and Transmission Standard: a comprehensive and fully structured approach to a unified description of complex object, designed for digital library uses. METS can be used to describe a container (of multiple files). http://www.loc.gov/standards/mets/
Mezzanine	In managing various versions of an object, a computationally efficient way to produces needed proxies may be to use a version that is not the master version, but instead is encoded in a way that supports production of proxies
Minidisc	A digital audio recording and storage method from Sony
MOV	A wrapper format used by Apple and associated companies
MPEG	A very large set of standards for audio and video, covering encoding, wrapping, metadata and rights, developed by the Moving Picture Expert Group http://mpeg.chiariglione.org/
MPDI	Multimedia Preservation Description Information, an MPEG initiative on audiovisual preservation started in 2011; one recent update: http://multimediacommunication.blogspot.com/
MXF	A SMPTE standard wrapper used in broadcasting, digital cinema and other professional contexts
OAIS	Open Archive Information System, the general standard for digital preservation
Preservation	'Preservation is the totality of the steps necessary to ensure the permanent accessibility – forever - of an audiovisual document with the maximum integrity' (Edmondson, 2004)
Preservation action	An intervention in the life of content, taken in order to keep the content usable; digitisation is one such action
Preservation factory	A concept promoted by the Presto series of projects, involving use of division of labour and other industrial techniques in order to make digitisation 'better, faster, cheaper'
Quality (of a recording)	An estimate of how closely a recording matches an original signal, as assessed by physical measurements (mainly bandwidth and dynamic range) and perceptual judgements
ITU Rec 601	The digital video standard from the International Telecommunications Union, a standards body
Recording (of a signal)	A way to make a permanent version of signal
Restoration	Changing a recording to correct defects
Sampling rate	How often number are assigned to a signal, usually per second; sampling rate controls the upper limit to the bandwidth of the digitised signal

Term	Definition
Scrubbing	Checking files on a storage system for errors
SDI	Serial Digital Interface, the wiring system for ITU Rec 601 digital video; sometimes used interchangeably with Rec 601
Signal	For our purposes, a variation against time. Sound is a variation in air pressure (over time); the variation can be captured by a microphone to produce a time-varying voltage, a signal that is the analogue of the original sound; video is also a signal, and film definitely is not a signal
SMPTE	Society of Motion Picture and Television Engineers, a professional and standards body
Standard definition, SD	The lines per frame for video that has been used in the UK from the 1960's, with 625 lines total, 576 being visible
Sticky shed	A problem with aging magnetic tape in which the oxide starts to come off the tape (shed) and causes the playback device to momentarily stick (again and again), and eventually jam
TBC	Time-base corrector; a specialist device that restores the timing of a video signal played back from videotape
Temporary Archiving	A concept promoted by PrestoSpace (2006); a recommended roadmap for format change and migration; the principle is to capture the bandwidth and dynamic range of an analogue original, and then move in steps to an uncompressed format
TIFF	A format for image files; TIFF is a compressed format, either lossless or lossy
Time code generator	A specialist device that makes a new time code, possibly in response to codes recovered from an old recording, in order to put time code without defects into a new recording
TDR	trusted digital repository
Vinegar syndrome	http://www.filmpreservation.org/preservation-basics/vinegar-syndrome
WAV file	The standard wrapper for audio
Word length	The number of bits in the digital representation of the amplitude of a signal
Wow	A defect in recorded audio resulting in a slow, cyclic variation in speed or amplitude, or both
wrapper	A method of packaging a coded signal and all related signals and information into a file: commonly video with multiple tracks of audio plus time code, subtitles and metadata; see MXF, MOV, AVI

8 References

Allcock, S. and Alexander, J. (2011) *AVA: Access to Video Assets Project*. Presentation at DPC Briefing *Preserving Digital Sound and Vision*, 08.04.2011. Available at http://www.dpconline.org/component/docman/doc_download/579-soundavision-ou-allcock-alexander

ARSC Technical Committee (2011) *Study of Embedded Metadata Support in Audio Recording Software: Summary of Findings and Conclusions* http://www.arsc-audio.org/pdf/ARSC_TC_MD_Study.pdf

Cramer, T. and Kott, K. (2010) *Designing and Implementing Second Generation Digital Preservation Services: A Scalable Model for the Stanford Digital Repository,* D-Lib Magazine, September/October 2010, Volume 16, Number 9/10. http://www.dlib.org/dlib/september10/cramer/09cramer.html

Digital Preservation Coalition (2011) DPC Briefing *Preserving Digital Sound and Vision,* London, 08.04.2011. Available at http://www.dpconline.org/events/details/27-SoundAndVision?xref=26

FADGI (2010) Federal Agencies Digitization Guidelines Initiative; report from meeting held in association with the October 2010 meeting of AMIA, the Association of Moving Image Archives, Philadelphia, PA.

Heritage, Thomas (2012) *Notes on Packages for Archives.* BBC R&D; public version in preparation.

Houpert, J. (2011) Direct RF Capture for Video. Information provided by Cube-Tec at their exhibition stands at the IASA and FIAT conferences in September 2011.

IFLA (2011) the FRBR Review Group website: http://www.ifla.org/publications/functional-requirements-for-bibliographic-records

PrestoCentre http://www.prestocentre.eu

PrestoPRIME http://www.prestoprime.eu/

PrestoSpace project website: http://www.prestospace.eu . The full integrated results of the project are here: http://digitalpreservation.ssl.co.uk/

Wright, R. (2011) *The Real McCoy: What Audiovisual Collections Preserve.* In de Jong, A. (Ed) Zorgen Voor Onzichtbare Assets (Proceedings of the AVA_Net Conference on Authenticity, Hilversum, 17.11.2010). A publication of the project Kennisdisseminatie, National Institute of Sound and Vision, Hilversum, the Netherlands. Also published as a White Paper by BBC R&D: http://www.bbc.co.uk/rd/publications/whitepaper211.shtml

Wright, R. and Jordan, D. (2011) *Broadcasting the Archive* In: Virtual Goods conference, Barcelona September 2011 http://www.virtualgoods.org/2011/program.html