The Presto4U project is supported by the European Commission under the 7th Framework Programme (FP7) — Grant Agreement 600845

OF COMMUNITIES AND PRACTICES

#DP Innovation & Research

FIAT/IFTA Conference, October 22, 2014

Erwin Verbruggen
Netherlands Institute for Sound and Vision
KNOWLEDGE FINDING & COMMUNITY BUILDING

- Surveys to establish knowledge base
- Regular blog & newsletters
- Webinars, Workshops & Edit-a-thon
• Media houses and broadcasters
  • i.e. private companies or public institutions working with a broad range of media services
  • for example electronic publishing, advertising and communication, entertainment etc.

• Private post-production houses
  • companies and production working with post-production and video material aimed at publications for education or infotainment / entertainment (all sizes – freelancers, permanent staff, consultants)

• Advertisers and marketing bureaus
  • i.e. commercial and public communication and advertising companies.
• Big diversity between stakeholders (livelihood, raison d’être).
  Many objectives at stake which confuses ends and means.

• Large diversity between the size and professional background of employers:
  • Private companies (freelancers, professionals, ROE, B2B and bottom lines are the main drivers)
  • Public institutions (academics, policy makers, public service and publishing are main drivers)

• Many different business and work models at stake
  • long term / short term goals
  • rights management
  • exchange of assets
  • in-house post-production
  • external post-production.
HIGH LEVEL OF HOMOGENEITY
Standards – preservation-tools – metadata- standards – technology tools
Methodologies for exchange and standardized roadmaps for handling of rights and payment Medium degree of flexibility
Business drivers are almost non-existing

MEDIUM LEVEL OF HOMOGENEITY
Mix of use of standards – preservation-tools – metadata standards – technology tools
Proprietary use of methodologies for exchange and roadmaps for handling of rights and payment
Low degree of flexibility
Business drivers exist within narrow communities

LOW LEVEL OF HOMOGENEITY
Multiple tools available but no consistency
Low level of metadata tools and standards Methodologies for exchange and roadmaps for use of rights and payment are almost not existing
High degree of flexibility and dynamical thinking and behavior
Business drivers RULE!
ISSUES – WHAT IS MISSING?

• Lack of access-facilities.
  • Making it easy to access archive material for a wide variety of uses, including the monetisation of content, faster access for clip search and a better ability to see what was there and the rights associated with content.

• Lack of standardisation.
  • A desire for consistent management of metadata + for as few codecs as possible, but more generally a demand for common standards for audio, video and graphics.

• Lack of Automation.
  • Solution based on de-humanising the workflow around archive, i.e. automated metadata generation and intelligent search – the prospect of intelligent software tools searching on content with metadata generated by intelligent software tools.

• Lack of strong B2B models which contains multiple sorts of end-users
• Lack of knowledge about solutions and opportunities
• Lack of confidence – CLOUD technologies
Sharing of preservation systems with defined content processing methods.
  - Need for making the tedious stuff automatic, and remove the need to input elementary / common preservation data.

Strong demand for tools or methods that would make it easy, simple and flexible to create versions for different destinations.
  - Intelligent management of versions, compiling, wrappers etc.

De-humanising the workflow around archive.
  - The prospect of intelligent software tools searching on content with metadata generated by intelligent software tools

Generic search, tracking and management-tools.
  - Tracking of content. More visibility of content as it moves through the workflow – e.g. by using QR-codes (like a DHL-order)

Outsourcing of distribution and delivery methods
  - external as well as internal / in-house

Sharing of storage solutions in flexible frameworks (SaaS).
  - Ability to manage workflows through centralised tooling, such as Cloud-based ones, for preservation, storage, and processing.

Flexible shopping and distribution facilities
WHAT FOOTAGE SALES ARCHIVES DO

- They exploit commercially their AV holdings by selling clips and footage for use in media productions.
- Today they generate revenues for more than $400 million per year
- They sell footage, but at the same time they act as an archive, being obliged to practice some kind of preservation activities.

THE COMMUNITY

- Memory Institutions (e.g. Istituto Luce, INA, B&G, Imperial War Museum, British Pathé):
  - usually public co-funded
  - footage sales are not their core business
  - they have at least a “moral” duty of dealing with long term preservation

- Private entities (e.g. Getty Images, ITN Source, Reuters, AP, Huntley Archive):
  - no public function
  - footage sales are their core business
  - market driven with a very pragmatic approach to digital preservation
Two main archives for digitised films:

- **2k digitised film material**
  - DPX files only
  - Priority given to unique prints, in danger reels
  - Any digitisation request that involves reels should pass first from here

- **HD digitised material**
  - ProRes 422
  - It’s mainly the delivery archives for our clients
  - It’s generated from the 2k content …or from telecine

Despite this digitisation policy:

- 10% of our subjects digitised in 2k
- 25% of our subjects digitised in HD
Most of the technical challenges in digital preservation are similar to the ones of Film archives and Broadcasters.

- Special attention is given to descriptive metadata and rights metadata.

...but the big challenge is:

- the tension between customer-driven activities/policies and digital preservation plans.
BROADCAST COMMUNITY OF PRACTICE EXPERT GROUP
broadcast archives have massive amounts of materials. Some say: of low quality. Others say: it’s the main source for 20th Century history.
WIDE RANGE OF MEDIA FORMATS

- Spectrum of formats starts expanding from 1920s onwards
- Latest technologies include ever-changing code bases
- At the vanguard of / overlapping with other media industries
Domain is highly diverse

- Many thousands of broadcast organisations (6379 broadcast companies in 40 European countries alone!)
- Deposit laws are different in each and every country
  - National AV Institutions take up tasks of public broadcasters’ archives in some countries
  - Commercial broadcasters are left alone
- Digitisation progress of legacy formats varies widely
- Key moment going on: move from tape to data recording
- Meanwhile access formats are on the move
TAPE ARCHIVES
This image is two years old. It is how many broadcasters still work.
The archive has never had such an impact on production before.
Archives are rebuilding workflows for metadata gathering, from the first steps of the production chain.
Producers need to be ‘taught’ how to work with an all-digital archive workflow.
BUT IS IT ALL DIGITAL?

And how many of these still function?

Legacy machinery is drying up – fast.
AND THEN, IT’S DIGITAL

▪ How do you keep track of the preservation process?
▪ Reading bits & pieces
▪ Maintaining quality throughout years of cataloguing

BORN-DIGITAL: NEW PARADIGMS?

▪ On the fringes of the creation world: the web as storytelling platform
▪ Web has been text-based, becoming increasingly audiovisual
▪ Technological push supported by radio organisations, arte.
▪ What chances, what futures for broadcast organisations?
Digital Maturity

Long-Term Sustainability

Certainty
Wisdom
Enlightenment
Awakening
Uncertainty

DIGITISATION STORAGE ACCESS
Enlightenment
Wisdom
Certainty
Awakening
Uncertainty
Possibility of digitising
Budgeting for digital
Knowing what's in your archive
Digital Maturity
Long-Term Sustainability
105,000 HOURS OF VIDEO

18,000 HOURS OF FILM

19,000 HOURS OF AUDIO

1,200,000 PHOTOGRAPHS

DIGITISED THANKS TO IMAGES FOR THE FUTURE
Digital Maturity

Long-Term Sustainability

- Uncertainty
- Awakening
- Enlightenment
- Wisdom
- Certainty

Connecting production & archive
## STORAGE PROVISION

<table>
<thead>
<tr>
<th>Step</th>
<th>Concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1 - INTAKE</strong></td>
<td>- Intake form</td>
</tr>
<tr>
<td></td>
<td>- Metadata template</td>
</tr>
<tr>
<td><strong>2 - PLANNING</strong></td>
<td>- “Flat storage”</td>
</tr>
<tr>
<td></td>
<td>- Available through iMMix</td>
</tr>
<tr>
<td><strong>3 - CHECKING MD</strong></td>
<td>- 4 Required fields</td>
</tr>
<tr>
<td><strong>4 - CHECKING FILES</strong></td>
<td>- DIVArchive: For MXF/WAV files</td>
</tr>
<tr>
<td></td>
<td>- DIVAdirector: FTP for other files</td>
</tr>
<tr>
<td></td>
<td>- Transcode</td>
</tr>
<tr>
<td><strong>5 - IMPORT</strong></td>
<td>- Budget, Contract, SLA</td>
</tr>
</tbody>
</table>
Enlightenment
Wisdom
Certainty
Uncertainty
Awakening
Control over QC process
Digital Maturity
Long-Term Sustainability
PRESENTATION FORMATS FOR CULTURE INFORMATION/E-ARCHIVES
<table>
<thead>
<tr>
<th>Project Coordinator + Memory Institution</th>
<th>Riksarkivet</th>
<th>Sweden</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technical + Communication Coordinator</td>
<td>Promoter Srl</td>
<td>Italy</td>
</tr>
<tr>
<td>Technical Partner</td>
<td>Packed Expertisecentrum Digitaal Erfgoed Vzw</td>
<td>Belgium</td>
</tr>
<tr>
<td>Technical Partner</td>
<td>Fraunhofer Gesellschaft Zur Foerderung Der Angewandten</td>
<td>Germany</td>
</tr>
<tr>
<td>Technical Partner</td>
<td>Hogskolan I Skovde</td>
<td>Sweden</td>
</tr>
<tr>
<td>Technical Partner</td>
<td>I University Of Skovde</td>
<td>Sweden</td>
</tr>
<tr>
<td>Technical Partner</td>
<td>Universita Degli Studi di Padova</td>
<td>Italy</td>
</tr>
<tr>
<td>Memory Institution</td>
<td>Stichting Nederlands instituut voor beeld en geluid</td>
<td>Netherlands</td>
</tr>
<tr>
<td>Memory Institution</td>
<td>Koninklijk Instituut voor het kunstpatrimonium</td>
<td>Belgium</td>
</tr>
<tr>
<td>Memory Institution</td>
<td>Greek Film Centre Ae</td>
<td>Greece</td>
</tr>
<tr>
<td>Memory Institution</td>
<td>Local Government Management Agency</td>
<td>Ireland</td>
</tr>
<tr>
<td>Memory Institution</td>
<td>An Ghiomhaimheacht Rainistiochta Rialtais Aitiuil</td>
<td>Ireland</td>
</tr>
<tr>
<td>Memory Institution</td>
<td>Stiftung Preussischer Kulturbesitz</td>
<td>Germany</td>
</tr>
<tr>
<td>Memory Institution</td>
<td>Ayuntamiento De Girona</td>
<td>Spain</td>
</tr>
<tr>
<td>Memory Institution</td>
<td>Eesti Vabariigi Kultuuriministeerium</td>
<td>Estonia</td>
</tr>
<tr>
<td>Memory Institution</td>
<td>Kungliga Biblioteket</td>
<td>Sweden</td>
</tr>
</tbody>
</table>
Empower memory institutions to gain full control over the technical properties of digital content intended for long-term preservation.
Competition-like procurement method
- Enables public sector bodies to engage with innovative businesses in development projects
- Innovative solutions that address specific public sector challenges and needs.
- Created through a phased procurement of development contracts (to reduce risk)
- More and more common within the public sectors of the European Union
The aim: to address the challenge of implementing various good quality standardised file formats for preserving data content in the long term.

The main objective: to give memory institutions full control of the process of conformity tests of files to be ingested into archives.

The main objective of the PCP launched by PREFORMA: to develop and deploy an open source software licensed reference implementation for various file format standards, aimed for any memory institution (or other organisation with a preservation task) that wish to check conformance with a specific standard.
Develop an open source conformance checker that:
- checks if a file complies with standard specifications
- checks if a file complies with the acceptance criteria of the memory institution
- reports back to human and software agents
- performs simple fixes

Establish an ecosystem around an open source reference implementation that:
- generates useful feedback for those who control software
- advances improvement of the standard specification
- advances development of new business cases for managing preservation files
Develop an open source conformance checker that:
- checks if a file complies with standard specifications
- checks if a file complies with the acceptance criteria of the memory institution
- reports back to human and software agents
- perform simple fixes

Establish an ecosystem around an open source reference implementation that:
- generates useful feedback for those who control software
- advances improvement of the standard specification
- advances development of new business cases for managing preservation files
<table>
<thead>
<tr>
<th>Content type</th>
<th>Standard specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEXT</td>
<td>ISO 32000-1:2008 (PDF 1.7)</td>
</tr>
<tr>
<td></td>
<td>ISO 19005-1:2005 (PDF/A-1)</td>
</tr>
<tr>
<td></td>
<td>ISO 19005-2:2011 (PDF/A-2)</td>
</tr>
<tr>
<td></td>
<td>ISO 19005-3:2012 (PDF/A-3)</td>
</tr>
<tr>
<td></td>
<td>ISO 12369:2004 (TIFF/IT)</td>
</tr>
<tr>
<td>MOVING IMAGE</td>
<td>Open standard:</td>
</tr>
<tr>
<td></td>
<td>• maintained by a not-for-profit organization,</td>
</tr>
<tr>
<td></td>
<td>• available either freely or at a nominal charge.</td>
</tr>
<tr>
<td></td>
<td>• IPR made irrevocably available on a royalty-free basis.</td>
</tr>
<tr>
<td></td>
<td>• no constraints on re-use</td>
</tr>
<tr>
<td></td>
<td>Capture AV-files that comply with the set of minimum technical parameters</td>
</tr>
</tbody>
</table>
Has the format been adopted by digital preservationists?
Has the format been adopted by PREFORMA stakeholders?
Has the format been adopted by service providers?
Does the license of the format allow for developing open source software?
Is the standard specification document available either freely or at a nominal charge?
<table>
<thead>
<tr>
<th>Open standard?</th>
<th>AUDIOVISUAL</th>
<th>TEXT</th>
<th>IMAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PREFORMA stakeholders</strong></td>
<td><strong>broadcast</strong></td>
<td><strong>film</strong></td>
<td><strong>Consumer</strong></td>
</tr>
<tr>
<td></td>
<td>MPEG-IMX (MXF/MPEG2)</td>
<td>DPX</td>
<td>MOV/MPEG2</td>
</tr>
<tr>
<td></td>
<td>XDCAM HD422 (MXF/MPEG4)</td>
<td>DCP (MXF/JPEG2000)</td>
<td>AVI/MPEG2</td>
</tr>
<tr>
<td></td>
<td>DPX</td>
<td>DCP (MXF/JPEG2000)</td>
<td>MPEG/ MPEG2</td>
</tr>
<tr>
<td></td>
<td>DCP (MXF/JPEG2000)</td>
<td>IMF (MXF/MPEG4)</td>
<td>MPEG/ MPEG4-AVC</td>
</tr>
<tr>
<td></td>
<td>IMF (MXF/MPEG4)</td>
<td>IMF (MXF/MPEG4)</td>
<td>MPEG/ MPEG4-AVC</td>
</tr>
<tr>
<td><strong>Industry standards</strong></td>
<td>ASI07 (MXF/MPEG2)</td>
<td>DCDM (TIFF 6.0)</td>
<td>MPEG-AF</td>
</tr>
<tr>
<td></td>
<td>(MXF/JPEG2000)</td>
<td>DCP (MXF/ JPEG2000)</td>
<td>PDF</td>
</tr>
<tr>
<td></td>
<td>FIMS (MXF/MPEG2)</td>
<td>IMF (MXF/MPEG4)</td>
<td>JPEG2000</td>
</tr>
<tr>
<td><strong>Open standards</strong></td>
<td>MKV/FFV1</td>
<td>PNG</td>
<td>WebM/VP8</td>
</tr>
<tr>
<td></td>
<td>OGG/Dirac</td>
<td></td>
<td>OGG/Theora</td>
</tr>
<tr>
<td></td>
<td>MKV (?)</td>
<td>OGG</td>
<td>PDF/A1</td>
</tr>
<tr>
<td></td>
<td>JPEG2000 (?)</td>
<td>FFV1</td>
<td>PDF/A3</td>
</tr>
<tr>
<td></td>
<td>Dirac</td>
<td>JPEG2000</td>
<td>PDF/A3</td>
</tr>
<tr>
<td></td>
<td>LPCM (?)</td>
<td>TIFF 6.0</td>
<td>JPEG2000</td>
</tr>
</tbody>
</table>

Slide with thanks to Bert Lemmens
1. A Training event for Open Source companies
Stockholm, December 2015
Includes first Prototype Demonstration

2. An Experience Workshop
Berlin, December 2016
Includes second Prototype Demonstration
PREFORMA partners will share with memory institutions their experiences of working with suppliers under R&D service agreements.

3. A final conference
Stockholm, December 2017
Results of the project.
Enlightenment

Wisdom

Certainty

Uncertainty

Awakening

Enlightenment

Data seal of Approval

Preservation plans in place

Long-Term Sustainability

Digital Maturity
TRUSTING OUR DIGITAL REPOSITORY

ingest

Virus check → Fixity check → Formaat check → Metadata extractie → Quality analysis

metadata → metadata → metadata → metadata
KNOWLEDGE FINDING & COMMUNITY BUILDING

- Surveys to establish knowledge base
- Regular blog & newsletters
- Webinars, Workshops & Edit-a-thon
Webinar: Choosing your File Format

PrestoCentre and Presto4U are organising a webinar series on diverse topics related to AV digitisation and digital preservation. Each webinar will be focused on a specific topic and hosted by experts within the field. The webinars are a unique opportunity to interact with professionals and learn from their knowledge, expertise, and experience. Get a discussion going with them and other participants during and after the webinar. For all upcoming webinars see the PrestoCentre Events Calendar.

Register now

More Information

Date: 27 October, 2014
Time: 3:00pm – 4:00pm GMT/UTC (11:00am – 12:00pm EST, 4:00pm – 5:00pm CET, 7:00am – 8:00am PST)

Summary

In celebration of World Day for Audiovisual Heritage, Presto4U organises its 12th webinar. It focuses on standardisation efforts for broadcast file preservation formats and offers a real-life example of an audiovisual archive selecting its preservation format. In the webinar Choosing your File Format, broadcast archive specialists will identify the topic of file selection. For many archives in the digital domain a recurring question is what file formats to approve of and what file formats to standardise to. Mr. Carl Fleischauer from the Library of Congress and Mr. Jörg Haupert from Cube-Tec will give an introduction to the ongoing standardisation work in the field of the MXF AS-11 standard as an ongoing archival standard for the MXF wrapper. Mr. Emanuel Joly from PACKED and Brecht Declercq from VLAIA will share their experiences and decision process of selecting file formats for the
Thank you

everbruggen@beeldengeluid.nl
@erwinverb