

# DPC Student Conference 2018: What I Wish I Knew Before I Started #dpc\_wiwik

## Venue Sponsor:



## Supported by:

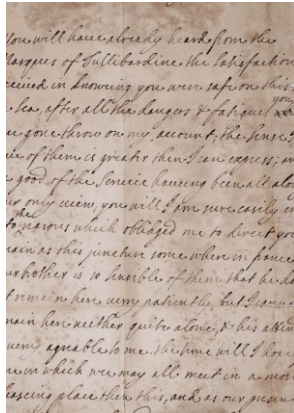


# Introduction to Digital Preservation



Digital**Preservation**Coalition

## Traditional Media



- Robust
- Tangible
- Independently understandable
- Well-developed approaches to preservation
- Experienced in assigning value

- Traditional objects are generally quite robust
- They are tangible, we can hold them in our hands
- Are generally independently understandable (if you speak the language they are written in.....)
- We are quite experienced in understanding their worth and assigning value to such objects

# Digital Information



- Ephemeral
  - Proliferation
  - Rights
  - Privacy
- Need tech to interpret
- Obsolescence
  - Media
  - Formats
  - Software/Hardware
  - Documentation
- New skills and solutions
- How to estimate value?
- But also new opportunities!



- Digital objects are ephemeral by their very nature
- They are very susceptible to obsolescence as they are entirely dependent on the media they are stored on, the accessibility of their file format and often require documentation to use and understand them
- Managing issues such as rights can also be much more difficult, from protecting copyright to ensuring personal data is protected
- They require us to gain new skills to care for them, or for us to work with new groups of colleagues with different skills groups (particularly IT specialists)
- But they do also bring a whole host of new benefits, in particular the ability to make content accessible to users.

## What's the Problem?



Digital data (images, documents etc.) have value  
and create opportunities

...but...

Access depends on software, hardware and people

...and...

Technology and people change, creating barriers to  
reuse

...therefore...

We need to actively manage data to protect and  
create opportunities

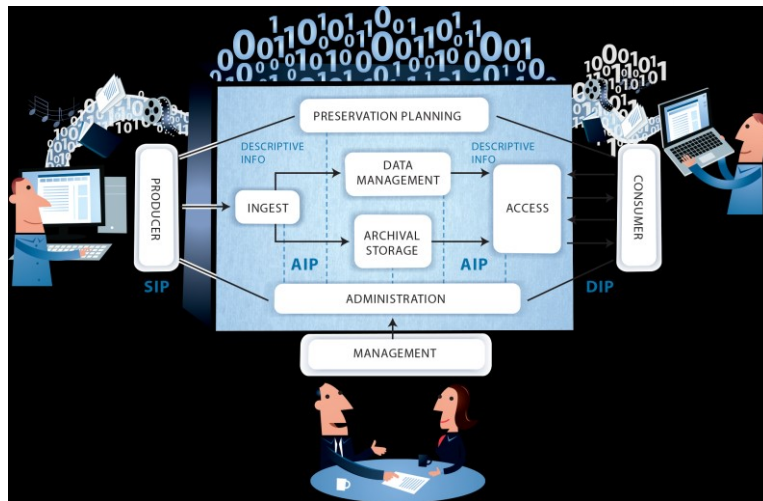
## Why We Preserve....



- Legal and Regulatory Compliance
- Increased Efficiency
- New Revenue Streams
- Improving Health
- Protecting the Environment
- Enabling Research
- Documenting Cultural Heritage
- Ensuring Transparency and Accountability



## DP Models: OAIS



While it is far from perfect the Open Archival Information System model is one of the keystones of Digital Preservation. In particular it provides much of the terminology used within the field. This diagram represents it's functional model at the highest level. As well as the key functions of an OAIS that it shows (such as Ingest, Preservation and Access) it also includes various information packages. These information packages contain the digital material to be preserved along with its accompanying metadata and within OAIS these exist in 3 different forms across the lifecycle:

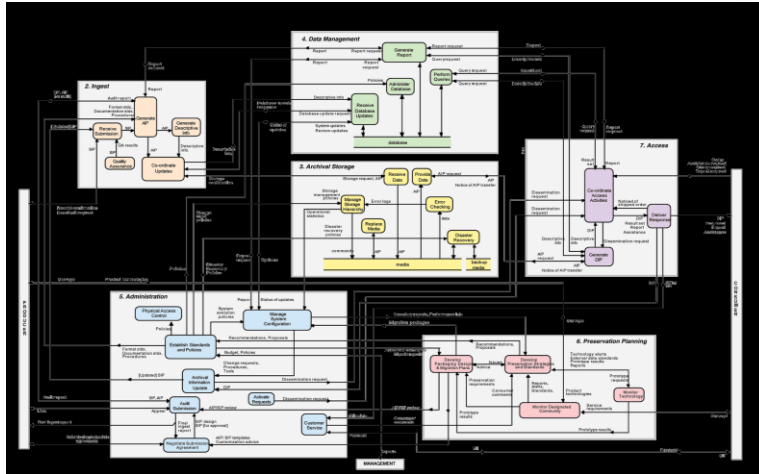
1. The Submission Information Package
2. The Archive Information Package
3. The Dissemination Information Package

To accompany this functional model the OAIS also describes an information model that lays out what types of metadata (specifically called Representation Information in OAIS) should be included in the information packages to facilitate preservation.

Full OAIS Standard: <https://public.ccsds.org/pubs/650x0m2.pdf>

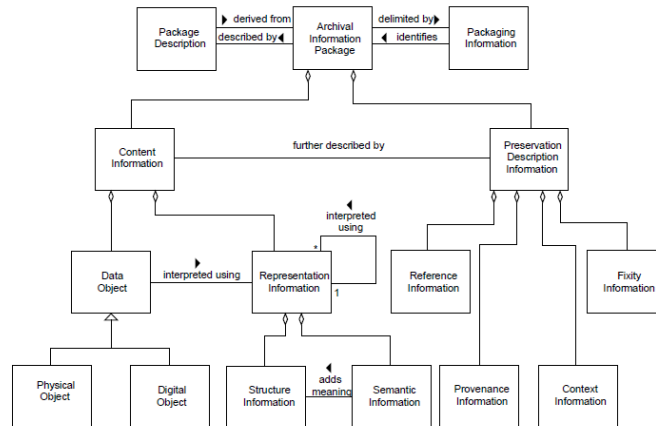
Brian Lavoie's Tech Watch Report on OAIS: <http://dx.doi.org/10.7207/twr14-02>

Where it starts getting really scary...

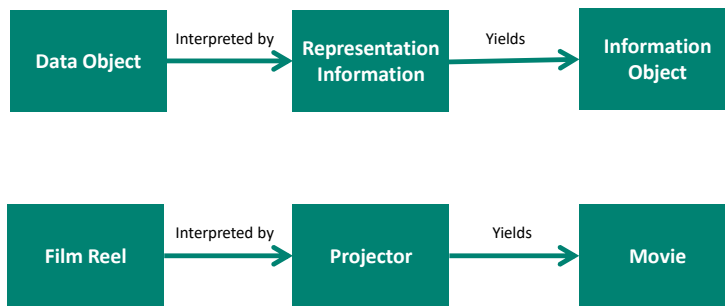




# I've Always Been More Terrified Of...

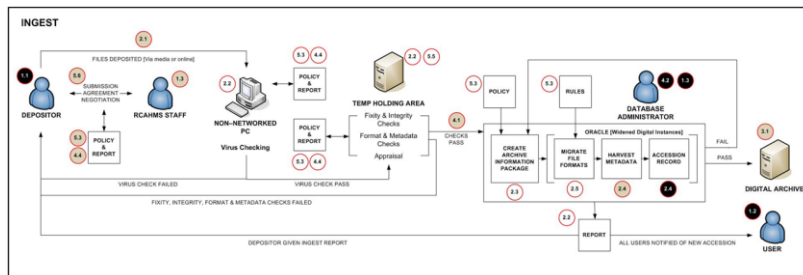


# Making Sense of the Information Model



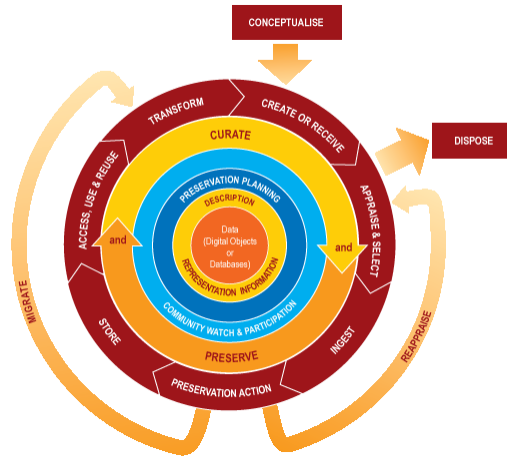
Representation information is perhaps the most difficult part of the information model to grasp. It can be useful to consider in relation an object like a reel of film. To show this onscreen a film projector is required. The projector fulfilling the role of the representation information. In the digital world representation information can be anything from the file format to detailed information about the file and the software and operating system environment in which to access it (in some cases perhaps even the software itself).

But still useful....



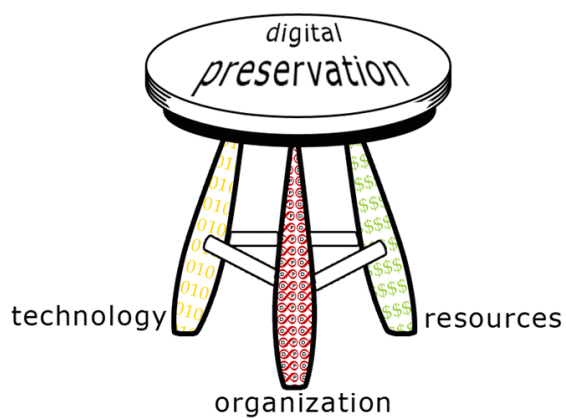
This a diagram we produced at my previous place of work, the Royal Commission on the Ancient and Historical Monuments of Scotland. The aim was to start building what a digital preservation workflow might look like then carry-out a gap analysis comparing the current systems with the OAIS functional model. The black circles are people, systems or process that were in place and meeting standards, the tan circles are those which were in place but not sufficiently developed/meeting requirements and the white circles where things that did not currently exist. It provided a very useful, clear visual aid for presenting to managers as they could clearly see the large gaps in our current systems and processes.

# DP Models: DCC Life-Cycle



<http://www.dcc.ac.uk/resources/curation-lifecycle-model>

## DP Models: Three Legged Stool



<http://dpworkshop.org/dpm-eng/conclusion.html>

## More on the Three Legged Stool



### Technology

- Storage and Back-Up
- Repository Systems
- Tools
- Security

### Organisation

- Policy & Strategy
- Planning
- Procedures
- Risks and Benefits
- Staffing

### Resources

- Business Planning
- Cost modelling
- Funding
- Sustainability
- Staff skills



Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark



## Other Standards

- Metadata – METS and PREMIS
- Information Packages – E-ARK Project
- Accreditation and Certification – Archives Accreditation, CoreTrustSeal, ISO 16363, DIN 31644...
- Data Security
- File Formats

# Approaches to Preservation



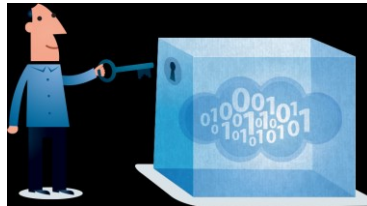
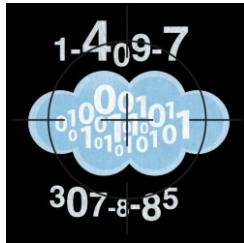
- Bit-Level
- Migration
- Emulation
- Hardware Preservation
- Digital Archaeology
- Virtualisation
- etc.....

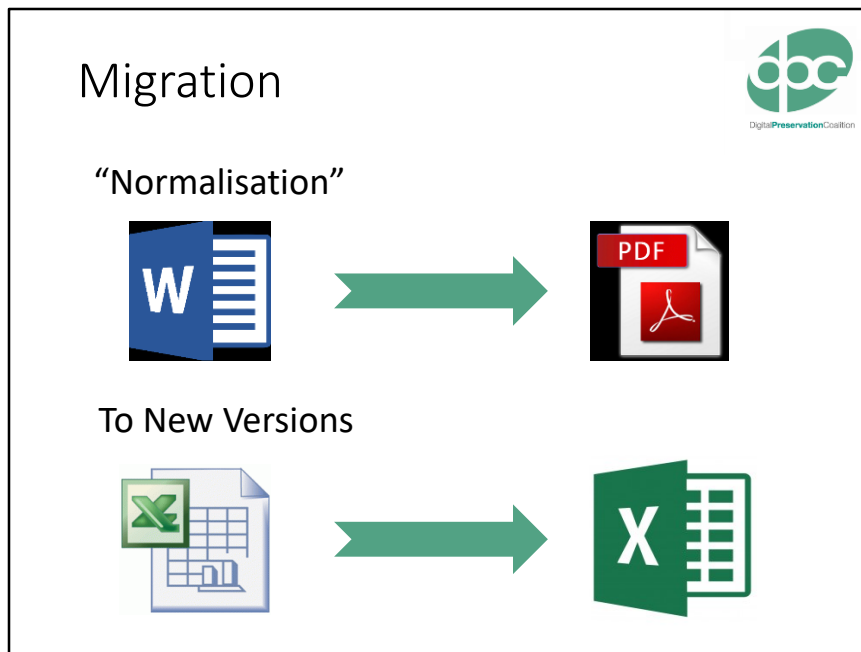


Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark



# Bit-Level Preservation





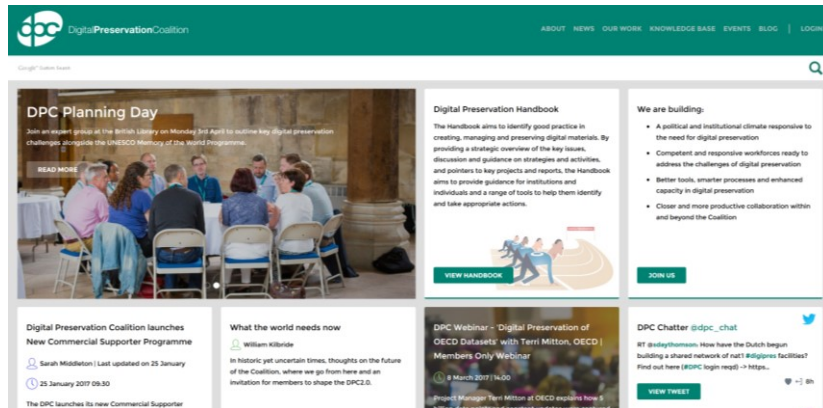
- There are two main forms of migration for digital preservation, and it is possible to use one or both.
- The first is a method often referred to as 'normalisation'. This is where all files of a particular type (for example, text documents) are 'normalised' to one file format.
- The example on the slide shows Word documents being normalised to PDF. For images this could be JPEGs and GIFs normalised to TIFFs.
- The choice of normalised files format used will depend on the needs of the organisation and its users.
- The second method involves migrating old file formats to newer versions when they are at risk of becoming obsolete.
- This could be migrating an old .xls spreadsheet to a newer .xlsx format.
- Both methods have their positives and negatives:
  - Normalisation creates homogenous, easier to manage collections and means that users need to know how to use fewer files types.
  - Migrating to new versions means that files can be accessed in current computer environments.
  - Both processes can be automated but quality control is incredibly important and careful consideration must be given to migration pathways to avoid loss of data and functionality.

# Emulation



- Emulation is the process of recreating the original environment in which a file was created and used via a layer of specially written software: the emulator.
- ☞ Emulation has been particularly successful in the world of computer games, where enthusiasts will create emulators to allow them to play older games.
- ☞ It is also an increasingly popular preservation method and several projects have produced emulators for everything from early browsers to old versions of PowerPoint. Many of these emulators are freely available, either online or as software downloads.
- Emulation perhaps seems like the ideal version of digital preservation as it allows users to access the files in their originally environment, providing a more authentic experience.
- It is however, very resource intensive and emulators will require updates (or their own emulators) as computer environments change.
- It can also be difficult to confirm the emulator truly captures the original environment unless there is still access to an original example to compare.

# Digital Preservation Coalition



<http://www.dpconline.org>

62 members

4 strategic areas

Shiny new website

- Responsive design
- Easier access to resources

# Blog and Case Notes



## Preservation Planning for Personal Digital Collections by Paul Wilson



Sara Day Thomson

🕒 Last updated on 13 January 2017

Personal Archiving

Case Note or Case Study

A condensed version of this case note also appears in the Technology Watch Report *Personal Digital Archiving* by Gabriela Redwine.

Paul Wilson's case note summarizes his attempts to find a suitable preservation planning process and associated documentation to apply to his personal digital collections. Since he could find no preservation planning process appropriate to individuals, he obtained a slide set detailing a simple preservation workflow from the Digital Preservation Coalition, and used that as a foundation on which to establish an approach to the work. This general approach and accompanying documentation was tested and refined on two of his personal digital collections (one of 800 mementos and the other of 17,000 photos). Template documents were then derived from the results.

This case note describes the solutions best suited to Wilson's collections and resources, but the processes he has developed have a wide applicability to any personal or small collections. In the fuller article (which can be downloaded as a PDF below), Wilson narrates his experiences to provide insights into the practical outcomes of using published guidelines and tools for preservation planning. Individuals and small organisations will be able to replicate those actions described here that are relevant to their own situations. They will also be able to compare their own collections and circumstances with those in this case study in order to assess common conditions and challenges. All of the documents, as well as blank templates, are available below.

- [Download the Full Text of Case Note](#)
- [Download the slides from Paul Wilson's Webinar: 'Preservation Planning for Personal Digital Archives' \(pdf\)](#)
- [Read more to access the Toolkit](#)

## Toolset: Preservation Planning for Personal Digital Collections

These documents comprise a *Toolset* derived from Paul Wilson's trials to discover a preservation planning workflow applicable to his personal collections.

### Download the Full Text PDF

Preservation Planning for Personal Digital Collections by Paul Wilson

### Related Resources

[DPC 12-Step Questionnaire template](#)

### Documents from the First Trial - PERS Collection

The following documents (suitably anonymised) were produced in the course of the first trial:

# Networking



- Important part of all our events
- Peer to peer learning
- Planning Day
- Counting the Bits

## Webinars



Covering a wide range of digital preservation issues

- Tool demos
- Research project updates
- Case studies
- Important resources

## Research Projects



Collaboration to Clarify  
the Costs of Curation



E-ARK: <http://eak-project.com/>

VeraPDF: <http://verapdf.org/>

TIMBUS: <http://timbusproject.net/>

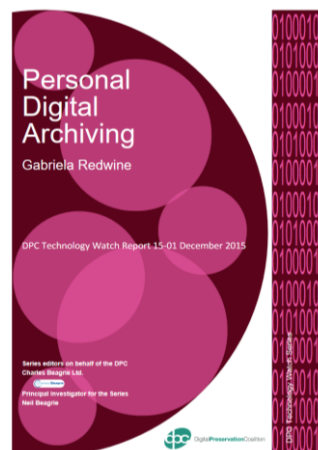
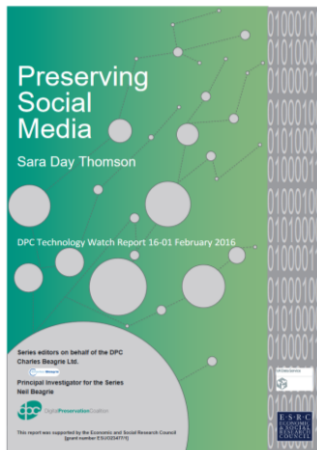
4C: <http://www.4cproject.eu/>

APARSEN: <http://www.alliancepermanentaccess.org/index.php/about-aparsen/aparsen-deliverables/>

SPRUCE: <http://wiki.opf-labs.org/display/SPR/Home>



## Tech Watch Reports



<http://www.dpconline.org/knowledge-base/tech-watch-reports>

## Training Workshops



Getting Started and Making Progress: <http://www.dpconline.org/events>

Short videos of GS online via the Handbook: <http://dpconline.org/handbook/getting-started>

# Digital Preservation Handbook



Digital Preservation **Handbook**

## Explore the Handbook

- Home
- Contents
- Introduction
- Digital preservation briefing
- Getting started
- Institutional strategies
- Organisational activities
- Technical solutions and tools
- Content-specific preservation
- Glossary



## Digital Preservation Handbook



Illustration by Jørgen Stamp, digitalisering.dk, CC BY 2.5 Denmark

Welcome to the revised 2nd edition of the Digital Preservation Handbook. A key knowledge base for digital preservation, peer-reviewed and freely accessible to all.

\*Overall the improvements to the Handbook make it, in my opinion, one of the more useful and flexible tools for identifying, understanding and getting to grips with practical approaches to the varying challenges of digital preservation. It uses approachable language, clear terminology and provides useful links to case studies and further reading which will be of

<http://dpconline.org/handbook>

sharon@dpconline.org

@SharonMcMeekin

<http://www.dpconline.org>



Digital**Preservation**Coalition