



Digital Preservation Coalition

An Introduction to Digital Preservation

Sharon McMeekin,
Head of Training and Skills, DPC

Our digital memory accessible tomorrow

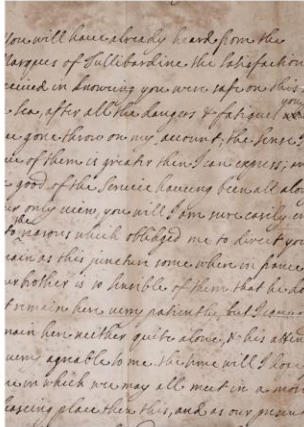
Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org



Digital Preservation Coalition

Traditional Media



- Robust
- Tangible
- Independently understandable
- Experienced in assigning value

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org

- 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 Preservation Coalition

Digital Information

- Ephemeral
- Obsolescence
 - Media
 - Formats
 - Documentation
- Rights
 - Copyright
 - Moral
 - Data Protection
- New Skills
- Benefits - accessibility



Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org

- Digital objects are ephemeral by their very nature
- They 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.



Digital Preservation Coalition

What's the Problem?

- Digital data (images, documents etc.) have value
- They can create opportunities
- ...but...
- Access depends on software hardware and people
- Technology and people change
- ...therefore...
- Technology can create barriers to reuse
- Managing data in the long term protects and creates opportunities

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org

“Digital Dark Ages”



Digital Preservation Coalition

In a Bit More Detail....

- Seven basic challenges
- Ways to address them:
 - Tools
 - Services
 - Standards
- Fancy words and acronym bingo

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org



Digital Preservation Coalition

Challenge One

Starting a digital preservation programme seems expensive and it's difficult to know where to start...

...but there's plenty of help on how to build a business case and tools to aid with planning

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org



Digital Preservation Coalition

A Solid Foundation

- **Research and Training** – don't get bogged down
- Understand **Aims and Goals**
- Identify **Benefits and Risks**
- Write a solid **Business Case**
- Relevant and clear **Digital Preservation Policy**
- Work towards **Sustainability** – embedding DP

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org



Digital Preservation Coalition

Challenge Two

Digital preservation systems can be complex and are subject to the same obsolescence as the objects they safeguard...

... so we need systems which are resilient,
based on standards and which can be tested
on an on-going basis

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org



Digital Preservation Coalition

OAIS Functional Model

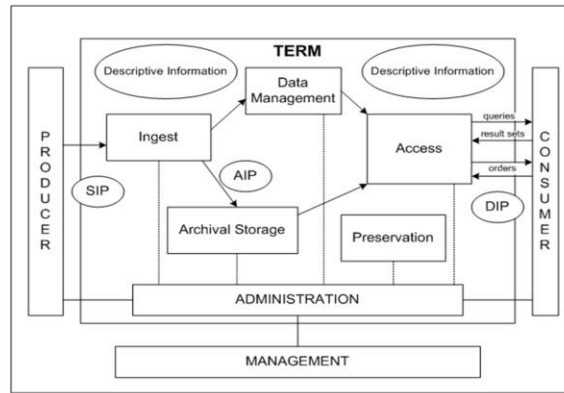


Fig. 1. Major functions of the OAIS Reference Model from Consultative Committee for Space Data Systems (CCSDS), CCSDS 650.0-W-1, Producer-Archive Interface Methodology Abstract Standard (OAIS) White Book, Issue 1, Draft Recommendation for Space Data System Standards.

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org

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.

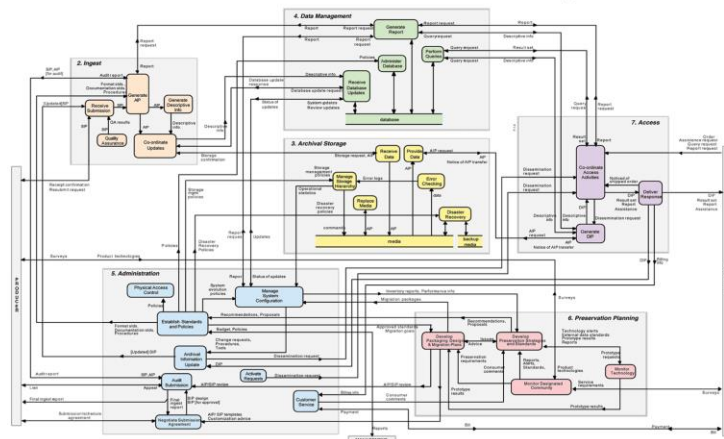
OAIS is an internal standard originally developed by the space data discipline and our Executive Director at the DPC likes to say by comparing your library or archive to this,

you're comparing yourselves to NASA. This makes a little more sense when you start delving a little further into the standard and things like the full functional model.....



Digital Preservation Coalition

Where it gets really scary....



Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

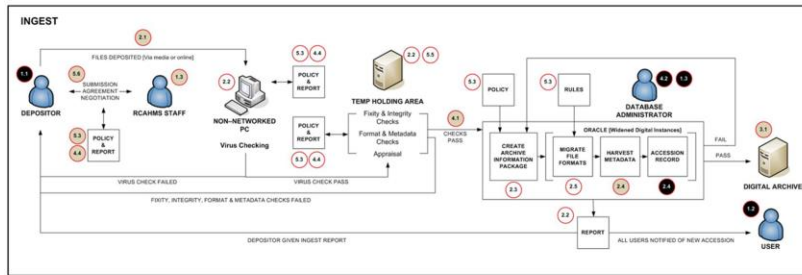
www.dpconline.org

.....although there are more accessible ways in. The DPC has published a very accessible Tech Watch Report on OAIS for which there is also an accompanying webinar you can watch on our website. The standard is also currently up for review and we'd encourage anyone with an interest to take part, more details are available on the DPC site.



Digital Preservation Coalition

But still useful....



Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org

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.



Digital Preservation Coalition

Audit and Certification

European Framework for Audit and Certification

1. Basic (Self/Peer-Assessment)

- Data Seal of Approval (<http://www.datasealofapproval.org/en/>)

2. Extended (Self-Assessment)

- ISO 16363 /DIN 31644 (<http://www.iso16363.org/>)

3. Formal (Full External Audit)

- ISO 16363 /DIN 31644

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
[@SharonMcMeekin](https://twitter.com/SharonMcMeekin)

www.dpconline.org



Digital Preservation Coalition

Challenge Three

Access and long term use depends on the constant configuration of hardware, software, data and the capacity of the operator....

... so we need to capture information on this configuration and use it to enable access.

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org



Digital Preservation Coalition

Use Standards and Tools

- Essential to capture sufficient documentation
 - Metadata Standards
 - OAIS Representation Information
- Tools
 - Registry Services
 - Characterisation
- Collection Audit

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org



Digital Preservation Coalition

Challenge Four

Technology continues to change creating
the conditions for obsolescence...

... so we need to plan and take action
accordingly, expecting that our current
plans may need to change.

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org



Digital Preservation Coalition

Active Preservation

- OAIS Preservation Planning functions
- Preservation Planning tools
- Different Approaches
 - Migration
 - Emulation
 - Hardware Preservation
 - Virtualisation
 - Etc.....

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org

- Migration – from one file format to another, or one version of a format to a newer version (i.e. .doc to .docx)
- Emulation – creating a programme that emulates the original hardware or software environment the object was created in. Allows an authentic experience although they are rarely perfectly



Digital Preservation Coalition

Challenge Five

Storage media fail, have a short life and storage devices are subject to obsolescence.

... so we need a storage strategy which includes error checking and refreshment

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org



Digital Preservation Coalition

Storage and Refreshment

Different Solutions:

- Multiple media
- Controlled storage
- Self-reporting media
- Lots Of Copies Keeps Stuff Safe
- Cloud storage

Beware: proliferation can become a problem

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org

Be aware of the storage plan for your organisation and make sure digital material well identified and is stored where it is secure and will be backed-up.



Digital Preservation Coalition

Challenge Six

Digital resources are intolerant of gaps in preservation...

...We need to act early and we need
to act on an on-going basis.

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org



Digital Preservation Coalition

Ongoing Preservation

Plan your responses:

- Intervene early in lifecycle
- Build in Sustainability
 - Transferable Archive Information Packages
- Risk management approach

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org



Digital Preservation Coalition

Challenge Seven

Resources can be corrupted or tampered
without trace...

...Need to carry out fixity and authenticity
checks on your data

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org



Digital Preservation Coalition

Fixity and Authenticity

A variety of solutions:

- Checksums
- Forensic tools
- Authenticity Evidence Records
- Data security protocols
 - (ISO 27000 series)

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org

You don't need to understand or do *all* of this.

- ... and even if you do, it doesn't all have to exist at the same time ...



Digital Preservation Coalition

The Reality...

- It won't go away.....

And.....

- It won't fix itself.....

But.....

- You already have many of the skills you need!

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org



Digital Preservation Coalition

It's Not All About Digital Preservation and Technology

- Likely to only be one aspect of your job
- A wide range of skills are required
 - Project Management
 - Risk Management
 - Communication
 - User Analysis
- DigCurV Curriculum Framework
 - <http://www.digcurv.gla.ac.uk/>

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org



Digital Preservation Coalition

In the End Nobody Has All the Answers!

- Collaboration is key
- Learn from others – Twitter, conferences.....
- Digital preservation is an active process
- There are no perfect solutions, they're still evolving
- We learn as much from our failures as our successes
- Don't be put off by job adverts!

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org



Digital Preservation Coalition

Some Resources from the DPC

- Tech Watch reports
(<http://www.dpconline.org/publications/technology-watch-reports>)
- Events (<http://www.dpconline.org/events>)
 - Getting Started and Making Progress
 - Briefing Days on a variety of themes
 - Webinars
- Digital Preservation Handbook
(<http://www.dpconline.org/publications/digital-preservation-handbook>)
- Leadership Programme - scholarships
- Member Support
- Advocacy

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
[@SharonMcMeekin](https://twitter.com/SharonMcMeekin)

www.dpconline.org



Digital Preservation Coalition

How to Find Me....

sharon@dpconline.org

@SharonMcMeekin

<http://www.dpconline.org>

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
[@SharonMcMeekin](https://twitter.com/SharonMcMeekin)

www.dpconline.org



Digital Preservation Coalition

Resources (1)

SPRUCE - Business Case Toolkit

- http://wiki.dpconline.org/index.php?title=Digital_Preservation_Business_Case_Toolkit

Keeping Research Data Safe - Benefits Framework

- <http://www.beagrie.com/krds.php>

National Archives Policy Guidance

- <http://www.nationalarchives.gov.uk/documents/information-management/digital-preservation-policies-guidance-draft-v4.2.pdf>

Jisc Digital Preservation Policies Study

- <http://www.jisc.ac.uk/publications/reports/2008/jiscpolicyfinalreport.aspx>

PREMIS Data Dictionary

- <http://www.loc.gov/standards/premis/>

METS for wrapping data

- <http://www.loc.gov/standards/mets/>

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org



Digital Preservation Coalition

Resources (2)

Registry Services and Characterisation

- PRONOM and DROID
 - <http://apps.nationalarchives.gov.uk/PRONOM/Default.aspx>
- JHOVE - <http://jhove.sourceforge.net/>
- C3PO - <http://www.scape-project.eu/tools>
- FITS - <http://fitstool.org>

Data Asset Framework - <http://www.data-audit.eu/>

SCOUT and PLATO

- <http://www.scape-project.eu/tools>

SCIDIP-ES Preservation Services

- <http://www.scidip-es.eu/preservation/>

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
[@SharonMcMeekin](https://twitter.com/SharonMcMeekin)

www.dpconline.org



Digital Preservation Coalition

Resources (3)

Lots Of Copies Keeps Stuff Safe

- <http://www.lockss.org/>

Cloud storage

- http://www.archives.org.uk/images/documents/Cloud_computing_report_final-1.pdf

E-ARK Project (<http://www.eark-project.com/>)

Risk management approach

- DRAMBORA (<http://www.repositoryaudit.eu/>)
- SPOT (www.dlib.org/dlib/september12/vermaaten/09vermaaten.html)
- TNA (www.nationalarchives.gov.uk/information-management/our-services/dc-risk-opportunities.htm)
- Forensic tools - <http://dx.doi.org/10.7207/twr12-03>

Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
[@SharonMcMeekin](https://twitter.com/SharonMcMeekin)

www.dpconline.org