Introduction to Digital Preservation
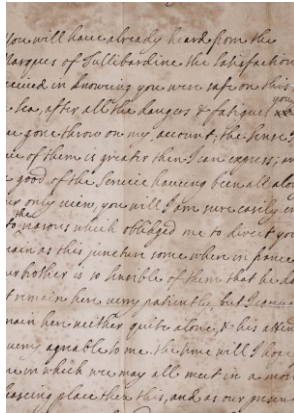
DigitalPreservationCoalition

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 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
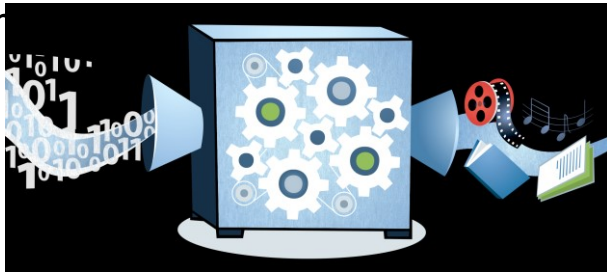
Why We Preserve….

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

A Tale of Three Models

- Three Legged Stool
- DCC Life-Cycle Model
- OAIS – Open Archive Information System

Fancy words and acronym bingo….

DP Models: Three Legged Stool

More on the Three Legged Stool

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

Organisation
- Policy
- Strategy
- Procedures
- Risks and Benefits
- Staffing

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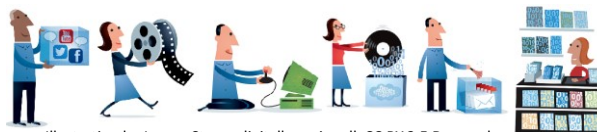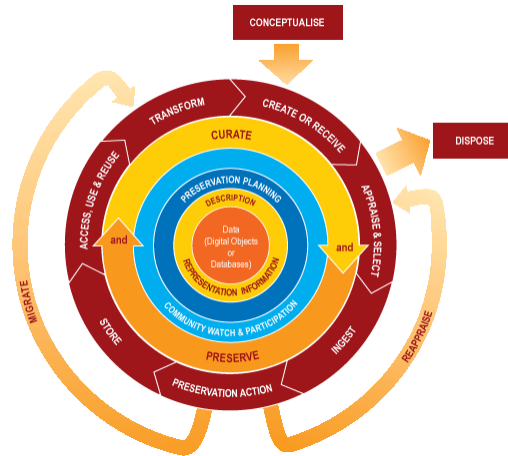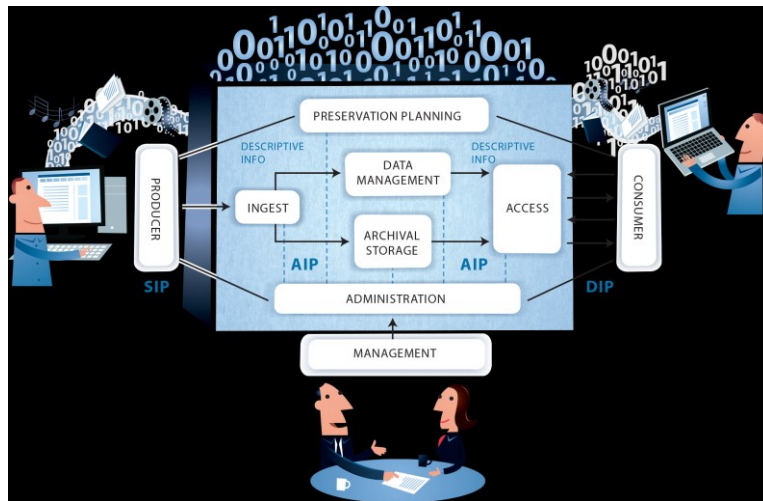- Cost modelling
- Funding
- Sustainability
- Staff skills

Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

DP Models: DCC Life-Cycle

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

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.
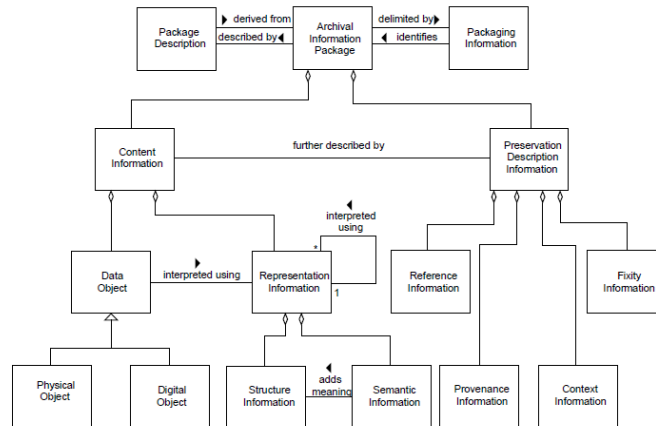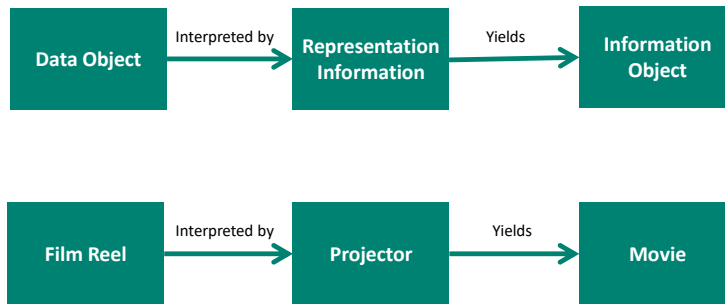
Full OAIS Standard: https://public.ccsds.org/pubs/650x0m2.pdf
Brian Lavoie's Tech Watch Report on OAIS: http://dx.doi.org/10.7207/twr14-02

Where it starts getting really scary…
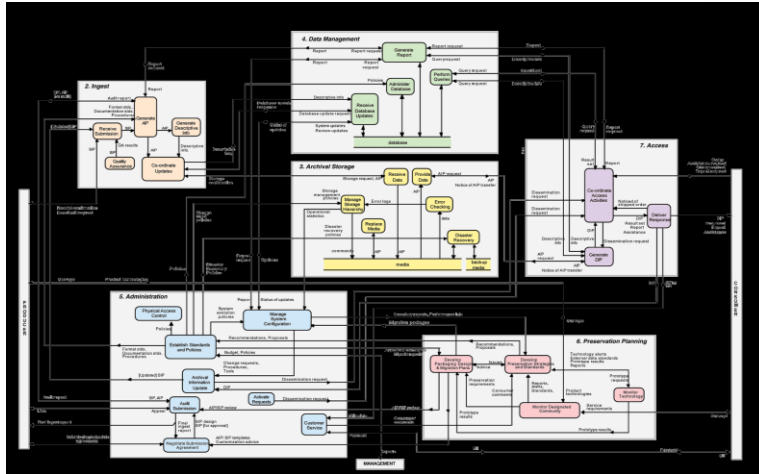
I've Always Been More Terrified Of…

Making Sense of the Information Model

| Data Object | — Interpreted by → | Representation Information | — Yields → | Information Object |

| Film Reel | — Interpreted by → | Projector | — Yields → | Movie |

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….
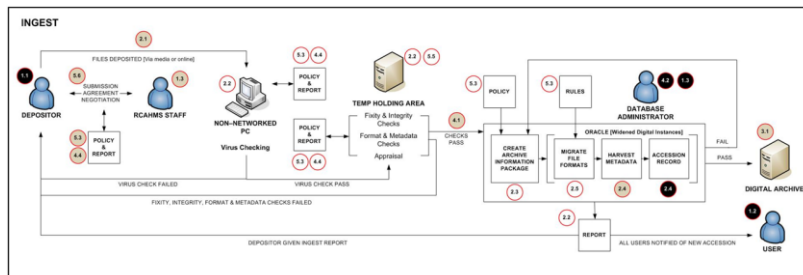
This a diagram we produced at my previous place of work, the Royal Commission on the Ancient and Historical Monuments of Scotland. The aim was to start building what a digital preservation workflow might look like then carry-out a gap analysis comparing the current systems with the OAIS functional model. 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.

Other Standards

- Metadata – METS and PREMIS
- Information Packages – E-ARK Project
- Audit and Certification – Data Seal of Approval, ISO 16363…
- Data Security
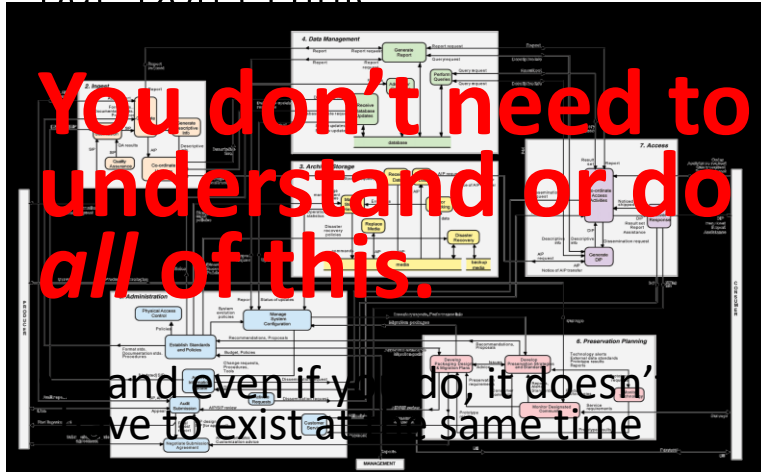- File Formats

Approaches to Preservation

- Migration
- Emulation
- Hardware Preservation
- Digital Archaeology
- Virtualisation
- etc…….

Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

But 'Don't Panic'

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

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

17

Digital Preservation Handbook

http://dpconline.org/handbook