



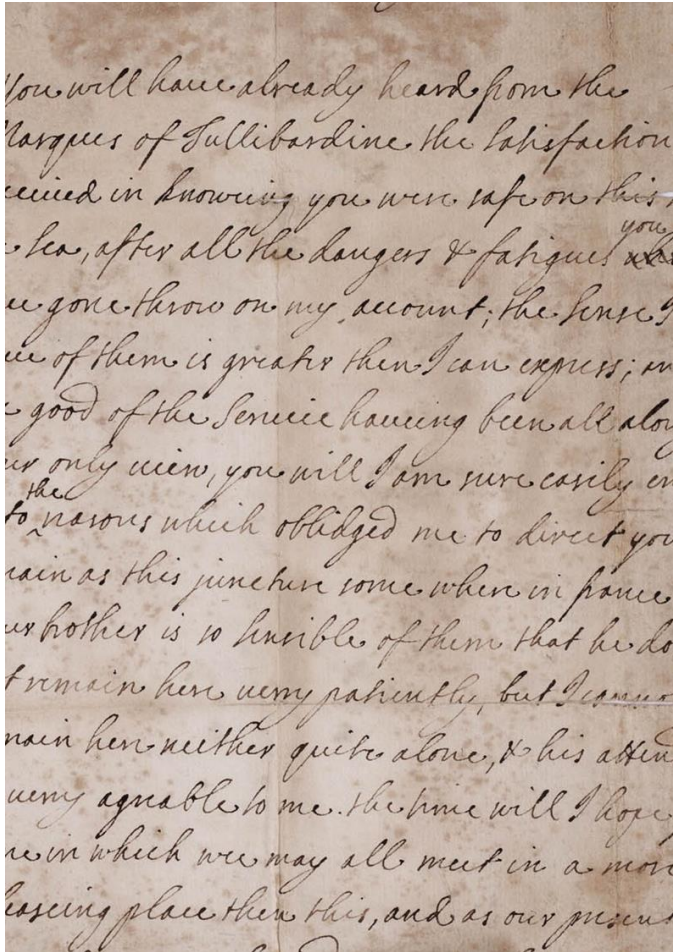
Digital**Preservation**Coalition

An Introduction to Digital Preservation

Sharon McMeekin,
Head of Training and Skills, DPC
ARA Core Training: Digital Preservation
Liverpool, 30th October 2014



Traditional Media



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

Digital Information

- Ephemeral
- Obsolescence
 - Media
 - Formats
 - Documentation
- Rights
 - Copyright
 - Moral
 - Data Protection
- Skills
- Sustainability





The Outlook is Bleak....

- Not without melodrama - ‘Digital Dark Ages’
- *“...of all the web links cited in answers to parliamentary questions 1997-2006, 40 percent are now broken”* (Spencer et al 2009)
- No easy off-the-shelf solution
 - Although services providers are now emerging



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
- We can't do it 'just because' - must explain benefits!

Seven basic challenges and some ways to address them:

- Tools
- Services
- Standards
- Fancy words and acronym bingo

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

- SPRUCE - Business Case Toolkit
 - [http://wiki.dpconline.org/index.php?title=Digital Preservation Business Case Toolkit](http://wiki.dpconline.org/index.php?title=Digital_Preservation_Business_Case_Toolkit)
- APARSEN - Example Business Cases
 - http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2014/02/APARSEN-REP-D36_2-01-1_0.pdf
- Keeping Research Data Safe - Benefits Framework
 - <http://www.beagrie.com/krds.php>

- APARSEN Recommendations for DP Policies
 - <http://www.alliancepermanentaccess.org/wp-content/plugins/download-monitor/download.php?id=D35.1+Exemplar+good+governance+structure+and+data+policies>
- PLATTER
 - <http://www.digitalpreservationeurope.eu/platter/>
- National Archives
 - <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>



OAIS Information Model

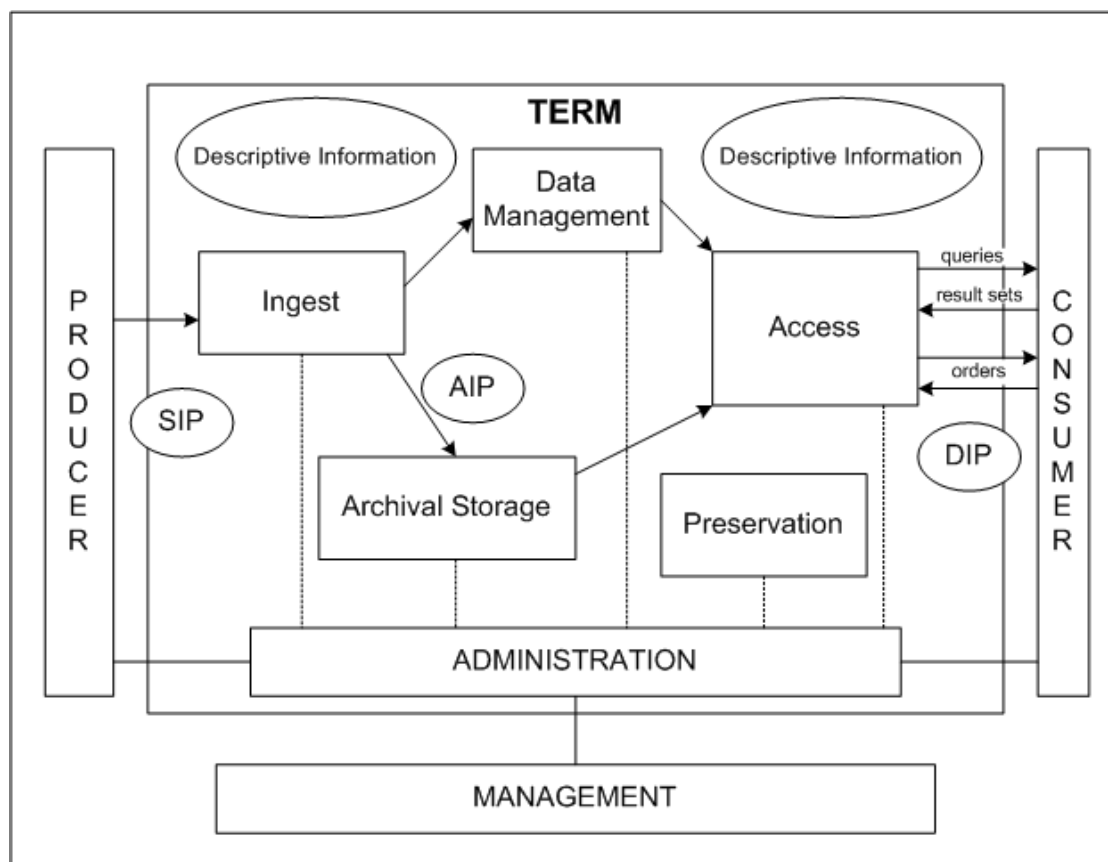


Fig. 1. Major functions of the OAIS Reference Model from Consultative Committee for Space Data Systems (CCSDS), CCSDS 650.0-W-1, Producer-Archive Interface Methodology Abstract Standard, (OAIS), White Book, Issue 1, Draft Recommendation for Space Data System Standards.



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

Challenge Two

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.



Essential to capture sufficient documentation/
metadata/representation info

- PREMIS Data Dictionary
 - <http://www.loc.gov/standards/premis/>
- METS for wrapping data
 - <http://www.loc.gov/standards/mets/>
- ISAD(G), Dublin Core.....



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>

Challenge Three

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.

Preservation Actions

1. Migration

Changing the format of a file to ensure the information content can be read

2. Emulation

Intervening in the operating system to ensure that old software can function and information content can be read

3. Hardware Preservation

Maintaining access to data and processes by maintaining the physical computing environment including hardware and peripherals.

4. etc.

Research and development field, new solutions and new approaches continue to emerge, e.g. virtualisation for preservation



Preservation Planning

- OAIS Planning Functions
- SCOUT and PLATO
 - <http://www.scape-project.eu/tools>
- SCIDIP-ES Preservation Services
 - <http://www.scidip-es.eu/preservation/>

Challenge Four

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



Digital **Preservation** Coalition

Storage and Refreshment

Different Solutions:

- Multiple media
- Controlled storage
- Self-reporting media
- 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

Beware: proliferation can become a problem

Challenge Five

Digital preservation systems are subject to the same obsolescence as the objects they safeguard...

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



Digital **Preservation** Coalition

Challenge Four: OAIS

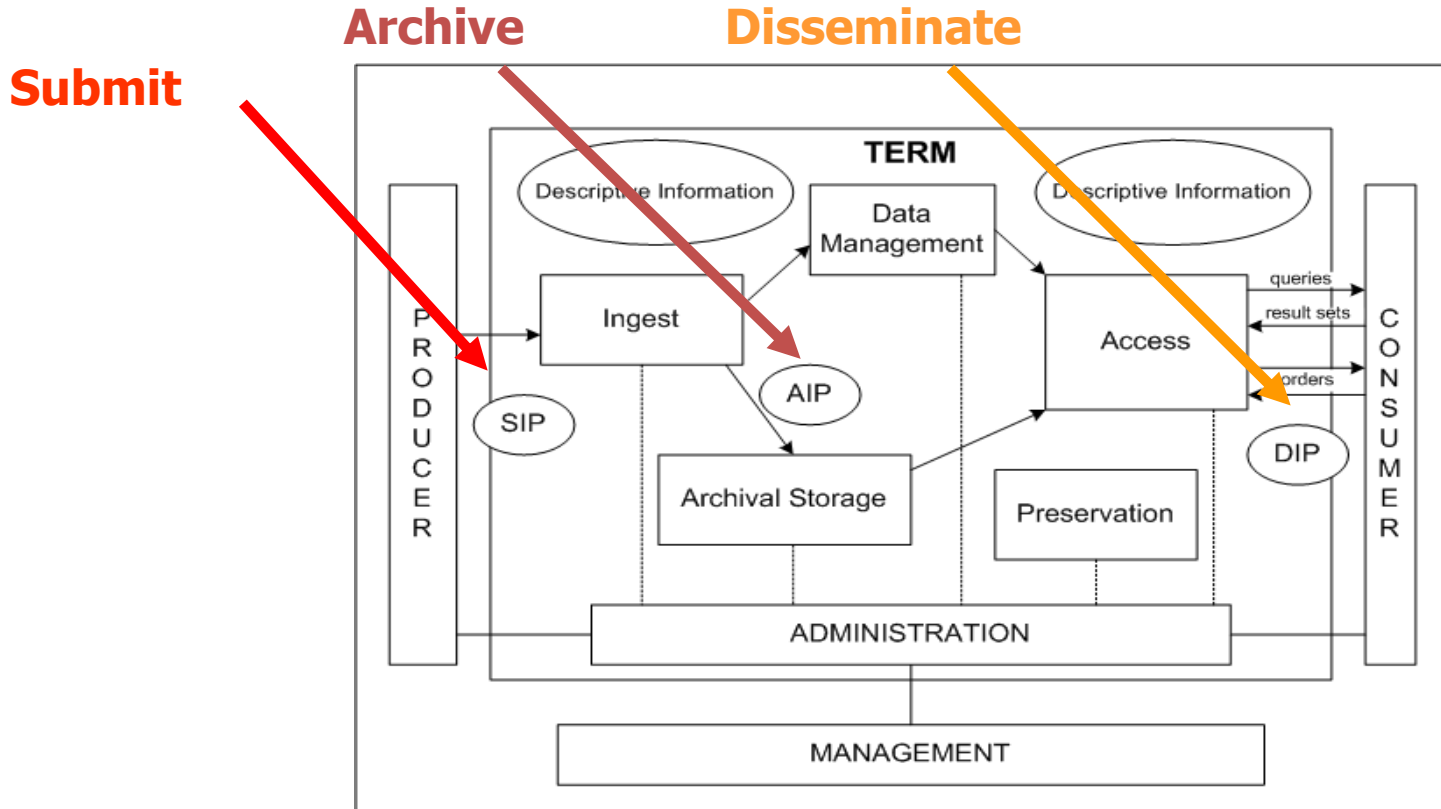
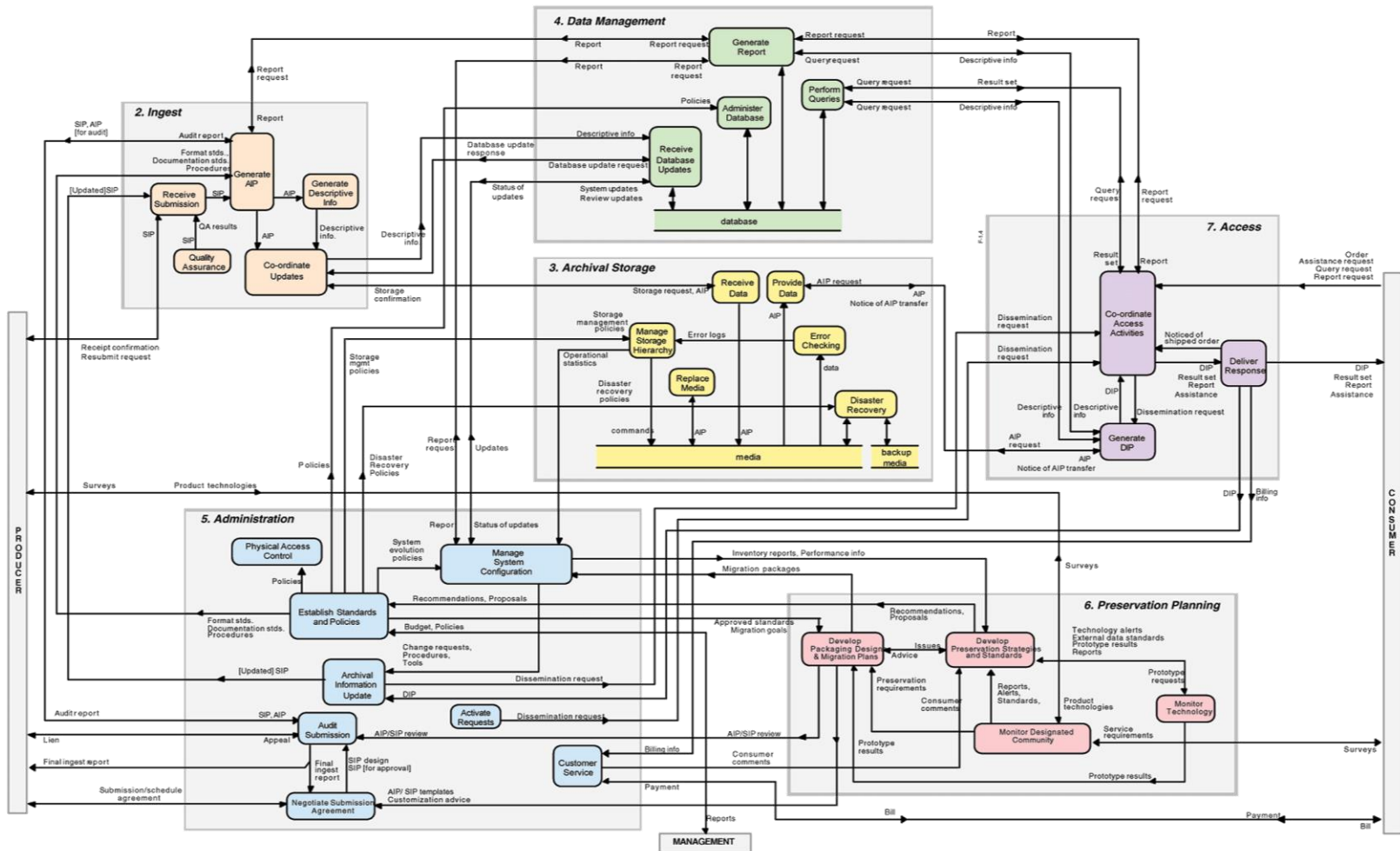


Fig. 1. Major functions of the OAIS Reference Model from Consultative Committee for Space Data Systems (CCSDS), CCSDS 650.0-W-1, *Producer-Archive Interface Methodology Abstract Standard*, (OAIS), White Book, Issue 1, Draft Recommendation for Space Data System Standards.



Digital Preservation Coalition

Where it gets really scary....

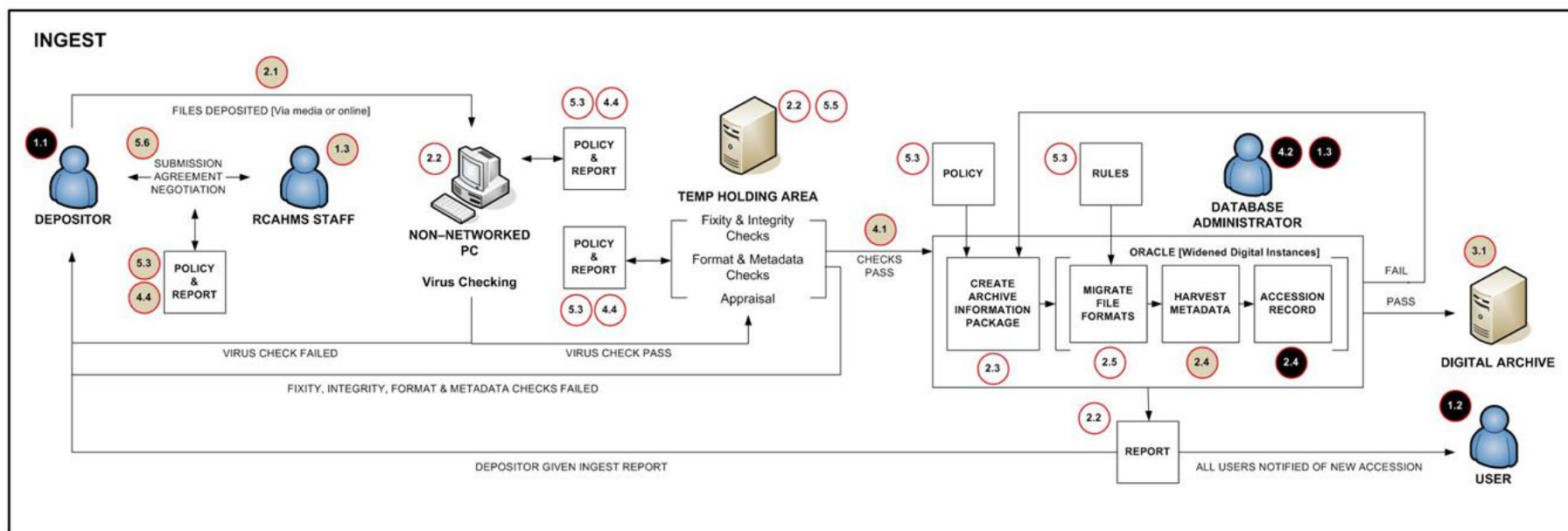


Our digital memory accessible tomorrow

Sharon McMeekin
sharon@dpconline.org
@SharonMcMeekin

www.dpconline.org

But still useful....



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. Lends itself to risk management approaches



Different responses:

- Intervene early in lifecycle
- Transferable Archive Information Packages
 - 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)

Challenge Seven

Resources can be corrupted or tampered without trace...

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



Fixity and Authenticity

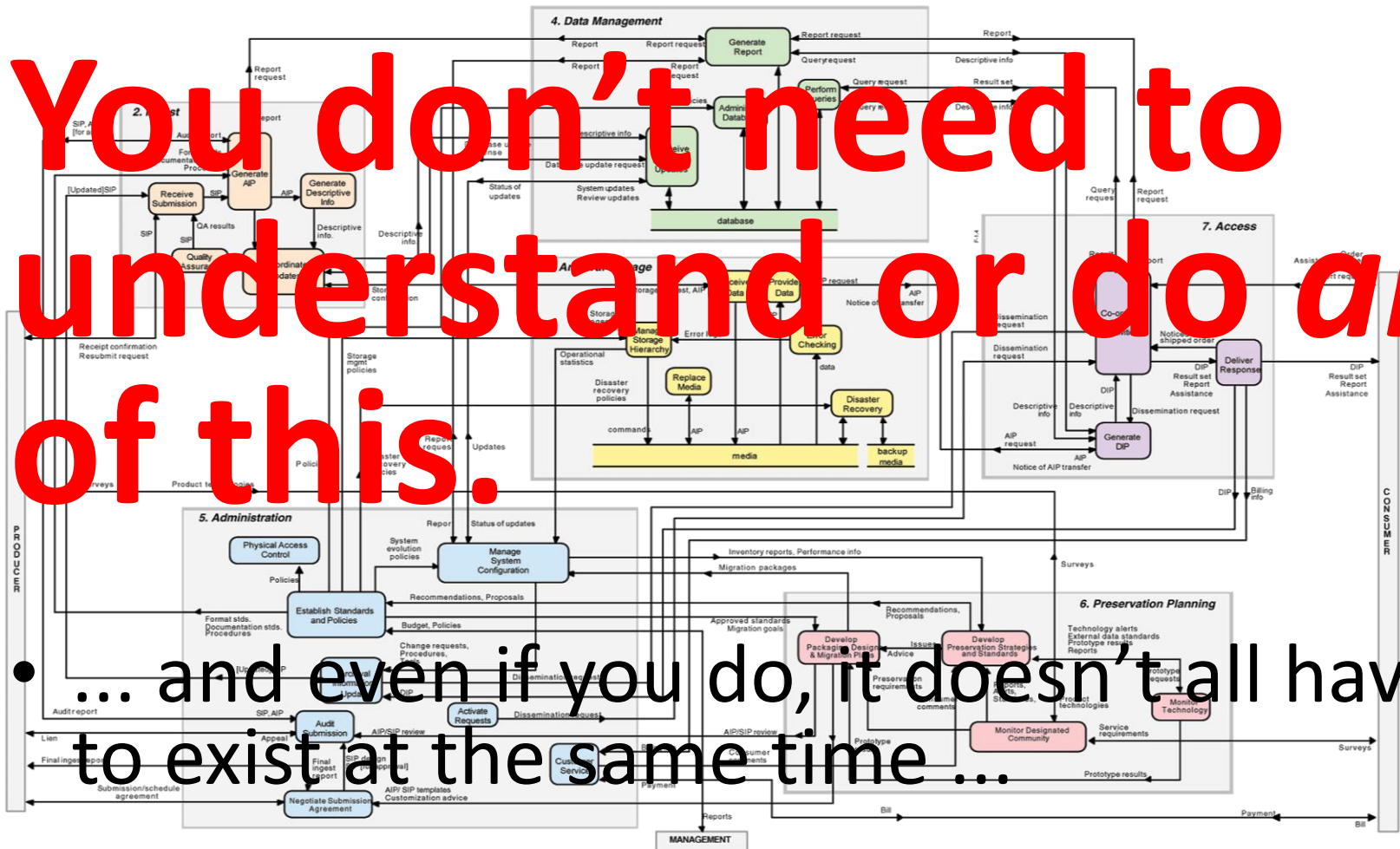
A variety of solutions:

- Checksums
- Forensic tools
 - <http://dx.doi.org/10.7207/twr12-03>
- Authenticity Evidence Records
 - http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2014/06/APARSEN-REP-D24_1-01-2_5_incURN.pdf
- Data security protocols
 - (ISO 27000 series)



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 ...



www.dpconline.org

Our digital memory accessible tomorrow

- DigCurV Framework (<http://www.digcur-education.org/>)
- Training Courses
 - Getting Started (<http://www.dpconline.org/events>)
 - Digital Preservation Training Programme (<http://www.dptp.org/>)
 - Online Training via APARSEN Virtual Centre of Excellence (<http://www.alliancepermanentaccess.org/index.php/training/courses-available/>)
- Twitter
 - #digitalpreservation
 - @WilliamKilbride



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
 - Briefing Days
 - Webinars
- Digital Preservation Handbook
(<http://www.dpconline.org/publications/digital-preservation-handbook>)
- Leadership Programme
- Member Support
- Advocacy



What Do I Need to Know?

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

And.....

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

But.....

- You already have many of the skills you need!



Digital**Preservation**Coalition

It's Not All About Digital Preservation

- Might 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/>



Digital**Preservation**Coalition

Nobody Has All the Answers

- Collaboration is key
 - Lots of relevant disciplines
 - Learn to talk the talk
- Learn from others
 - Social media, especially Twitter
 - Ask to visit like-minded organisations
 - Networking at events

Get Your Hands Dirty

- Don't be put off by job adverts
- 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



Digital**Preservation**Coalition

How to Find Me....

sharon@dpconline.org

@SharonMcMeekin

<http://www.dpconline.org>