DigitalPreservationCoalition

The Digital Preservation Decade:

what we knew then, what we know now, and what I wish someone had told me ten years ago

It won't do itself
It won't go away
Don't wait for perfection

What is the question?
6 basic challenges and skills
3 lessons from experience
1 Business Trend

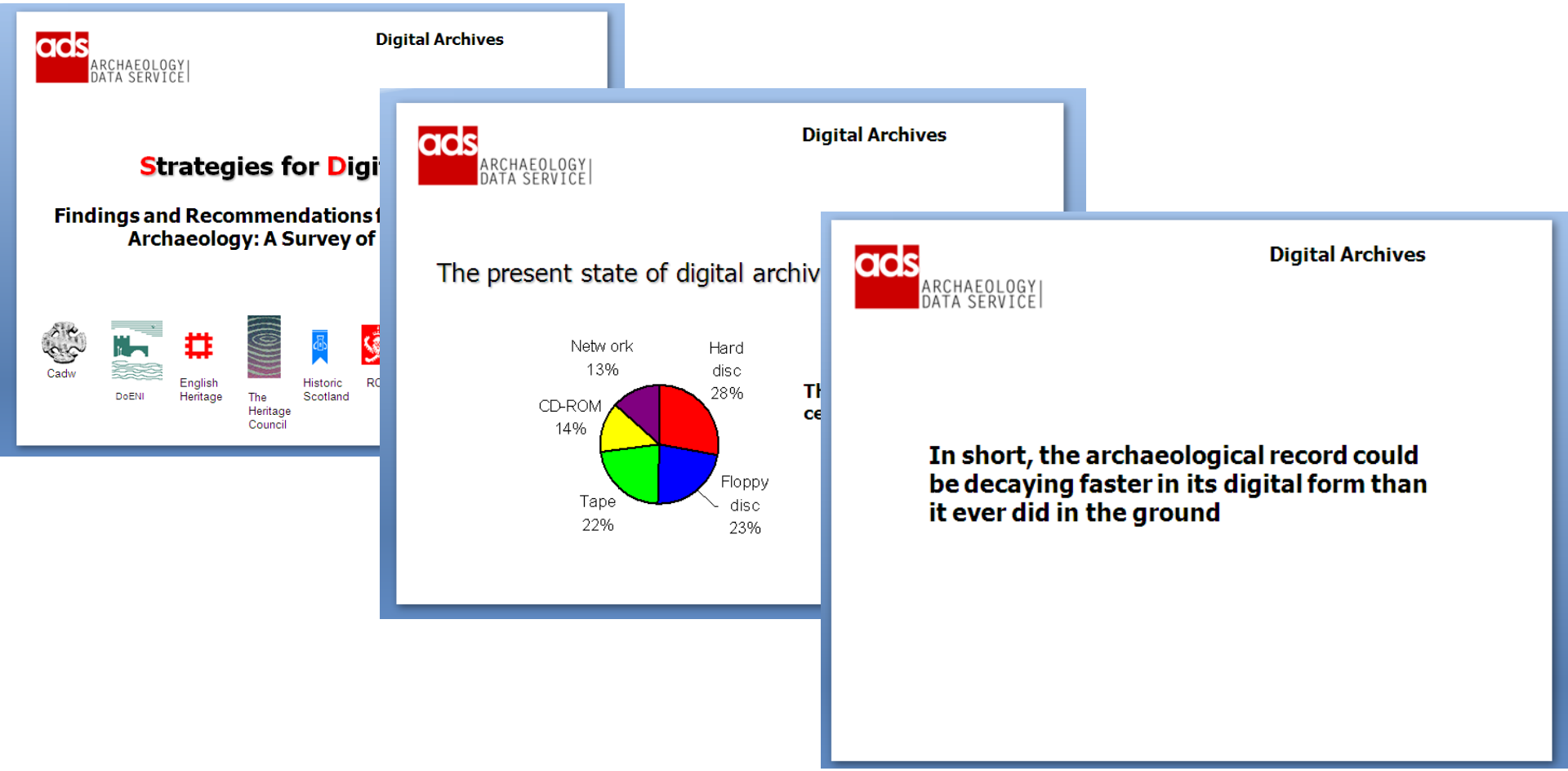
william@dpconline.org
@williamkilbride

DigitalPreservationCoalition

But first … what has brought you here?

DigitalPreservationCoalition

Digital preservation makes bleak reading …



www.dpconline.org

DigitalPreservationCoalition

Digital preservation typically makes bleak reading 2

<Enter your details here>

……

……

……

DigitalPreservationCoalition

Let's restate the problem …

- Digital stuff has value. It is an asset.
- It has potential and creates new opportunities.
- Use gives rise to direct and indirect outcomes.

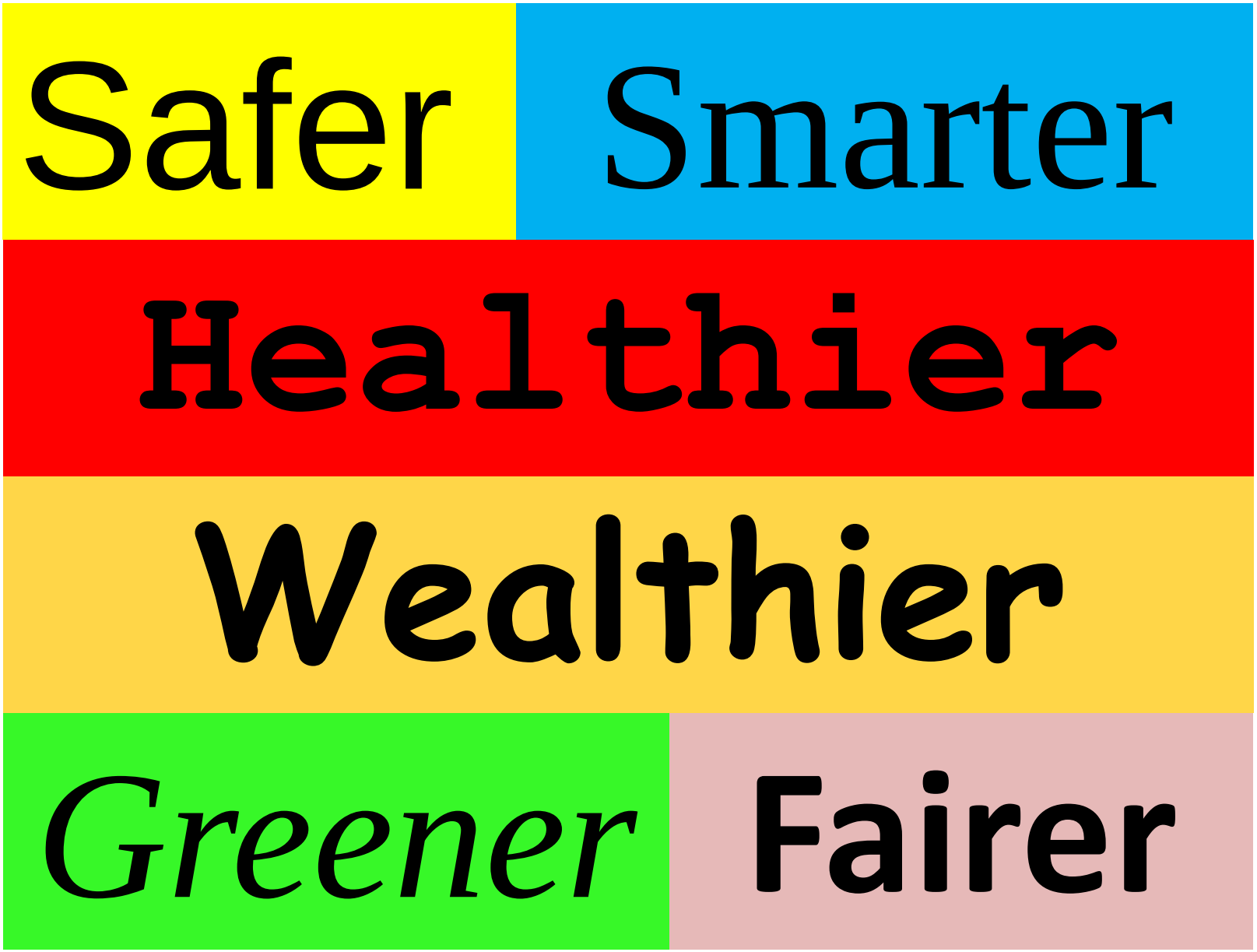
...but...

- Deployment depends on software, hardware and people.
- Software, hardware and people change.

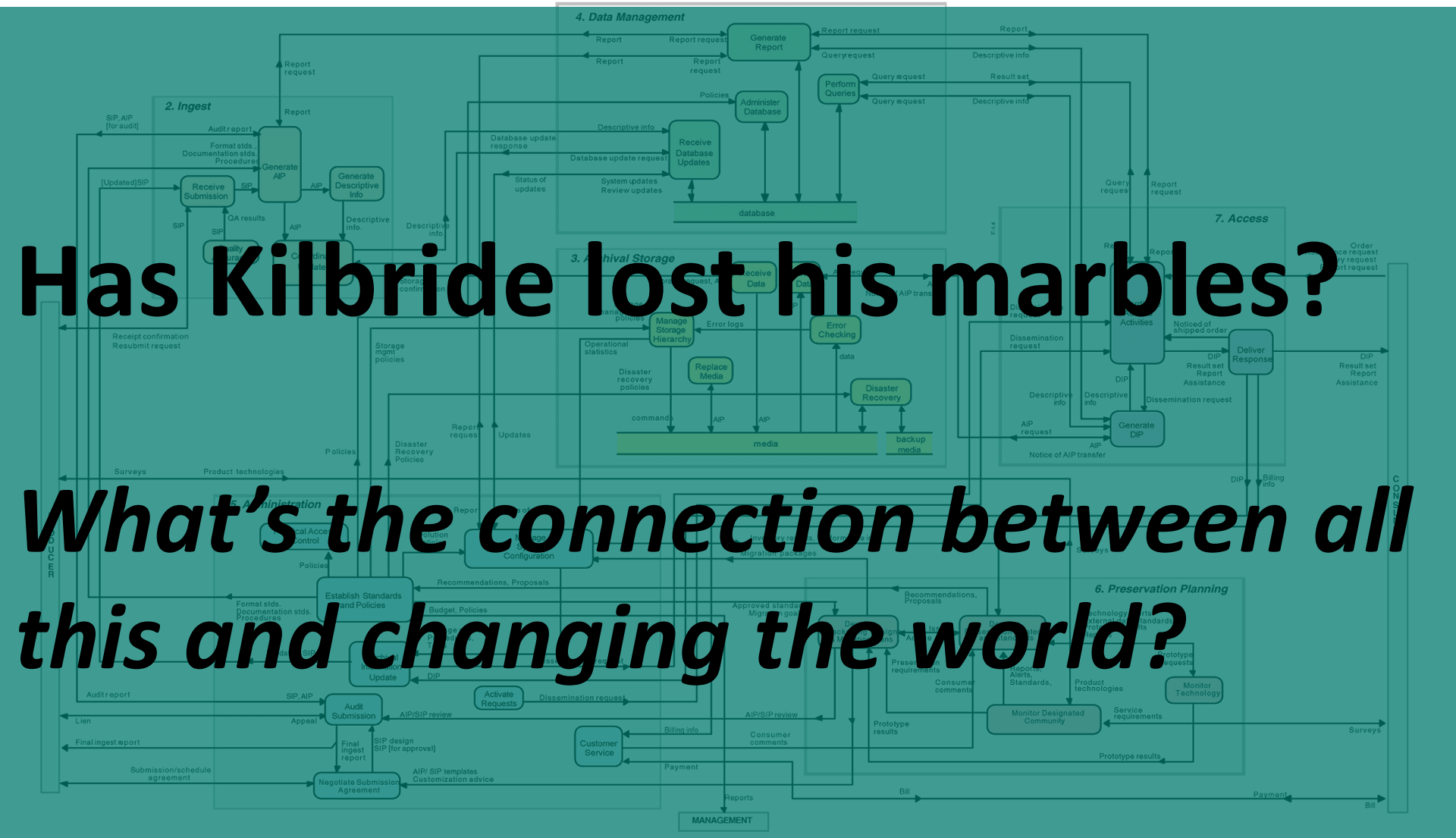
...therefore...

- Access is not guaranteed without (some) action
- Value, opportunity, impact not guaranteed

DigitalPreservationCoalition



DigitalPreservationCoalition



Has Kilbride lost his marbles?

What's the connection between all this and changing the world?

www.dpconline.org

DigitalPreservationCoalition

Why preserve in the public sector …

1. Transparent

e.g. Data Protection, Freedom of Information … childcare, human tissue

2. Safer

e.g. preparedness, detection, disaster, recovery, audit

3. Smarter

e.g. access to heritage, social knowledge, innovation, research, connected

4. Wealthier

e.g. more efficient, exploitation of IP, skills, surrogate access

5. Healthier

e.g. managed life history, research and safe innovation, identity, heritage

6. Greener

e.g. environmental policy development, efficient retention

DigitalPreservationCoalition

Why preserve for business?

1. Legal Compliance
e.g. Sarbanes-Oxley, Data Protection

2. Regulatory Compliance
e.g. power generation, aviation, banking, pharmaceuticals

3. Legal protection
e.g. patents, mis-selling, detection, audit

4. Unanticipated exploitation
e.g. petro-chemical, music, broadcast

5. Business continuity and improvement
e.g. product recall, disaster recovery

6. Business Value
e.g. getting the right information to the right people at the right time in a format they can use

DigitalPreservationCoalition

Key responses

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, eg virtualisation for preservation

DigitalPreservationCoalition

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

Change is not a bug.

DigitalPreservationCoalition

Technology continues to change creating the conditions for obsolescence.

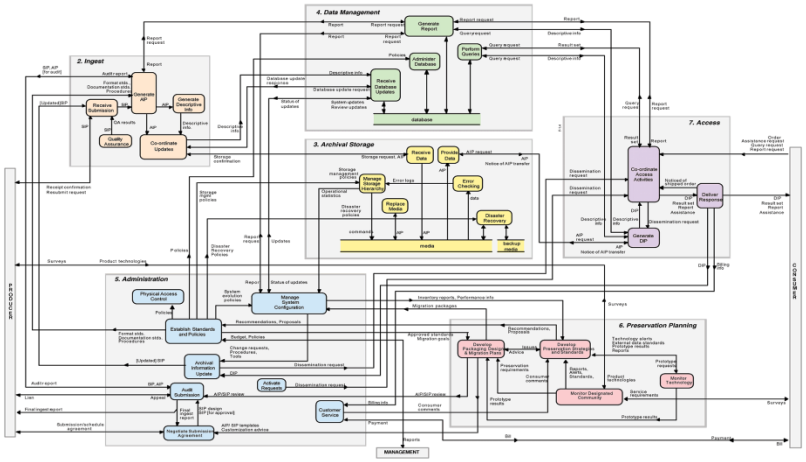
Need to become a learning institution

DigitalPreservationCoalition

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

Be mobile and format neutral

DigitalPreservationCoalition



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

Standards and modularity

DigitalPreservationCoalition

Digital resources are intolerant of gaps in preservation.

Ongoing process

DigitalPreservationCoalition

Different strategies for different types of user or collection or interaction

Find meaningful answers for how to preserve in your institutional mission

DigitalPreservationCoalition

The problems are more subtle than we realised a decade ago…

e.g. file format obsolescence

Changing file formats?
Conformant containers?
Units of information?

DigitalPreservationCoalition

How to pick a winner ...

Adoption
Dependency
Disclosure
Transparency
Metadata support
Interoperability
Complexity
Stability
Rights management

Todd, M 2009 'File formats for preservation', DPC Technology Watch Report 02/09, online at http://www.dpconline.org/advice/technology-watch-reports.html

DigitalPreservationCoalition

How to pick a winner ...

beyond and **potentially over-writing the criteria** ... repository managers should **align** the recognition and weighting of criteria with a **clear preservation strategy** that articulates the **purpose** of the repository and the **needs of its designated community**;

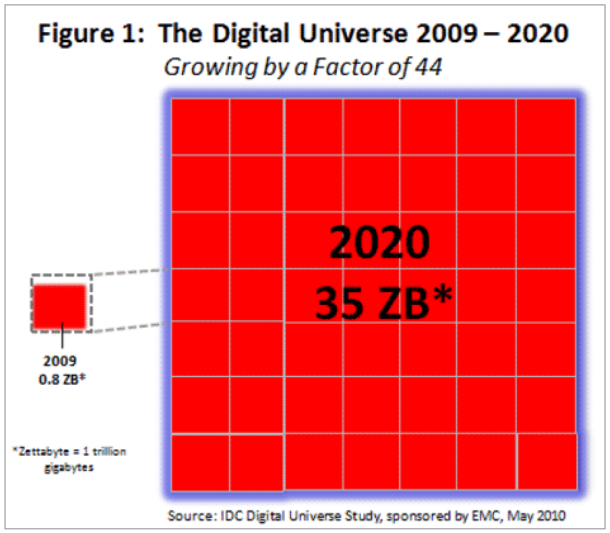
Todd, M 2009 'File formats for preservation', DPC Technology Watch Report 02/09, online at http://www.dpconline.org/advice/technology-watch-reports.html

DigitalPreservationCoalition

How to pick a winner ...

‘Digital Universe’ Nears A Zettabyte

May 4th, 2010 : Rich Miller



The Great Recession hasn’t slowed the breakneck growth of the Digital Universe. In 2010 the volume of digital information created and duplicated in a year will reach 1.2 zettabytes, according to new data from IDC.

You ain’t seen nothing yet

Data growth on 3 axes

- volume
- complexity
- expectation

... it’s not going to be about obsolescence so much as workflow and capacity

DigitalPreservationCoalition

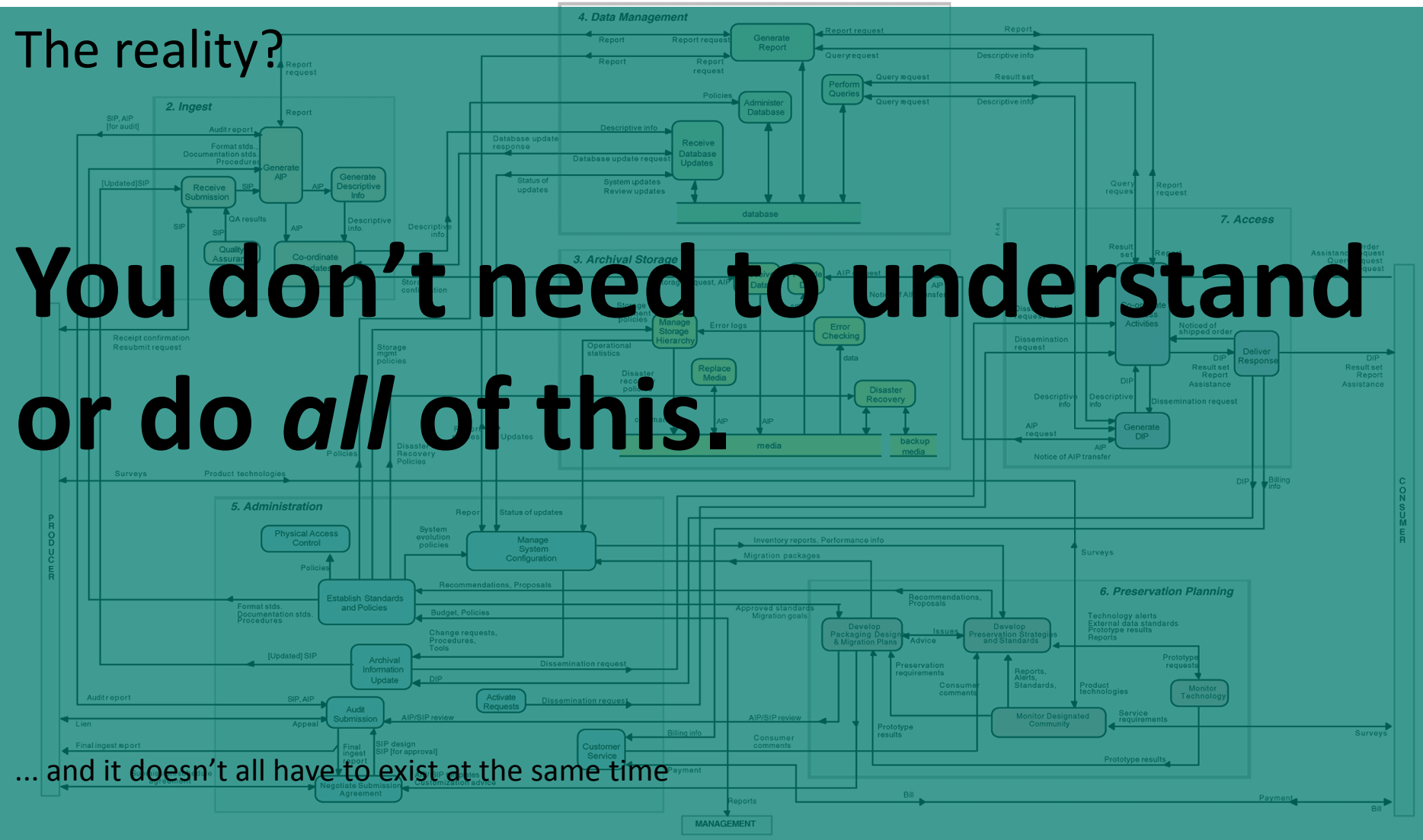
Digital Preservation as a ‘discipline’

Daunting challenge
Decade of research and development
Replete with jargon and acronyms
Turf war between professions?
A whole new barrier

The last decade has shown definitively that using fancy words are not the same as solving problems

DigitalPreservationCoalition

The reality?



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

... and it doesn't all have to exist at the same time

DigitalPreservationCoalition

The reality?

Get started now not later

DigitalPreservationCoalition

DPC's five point challenge:

Do we know which data sets from the last decade are going to be valuable in the next?

Do we have robust plans for the long-term exploitation to business-critical, high-value data?

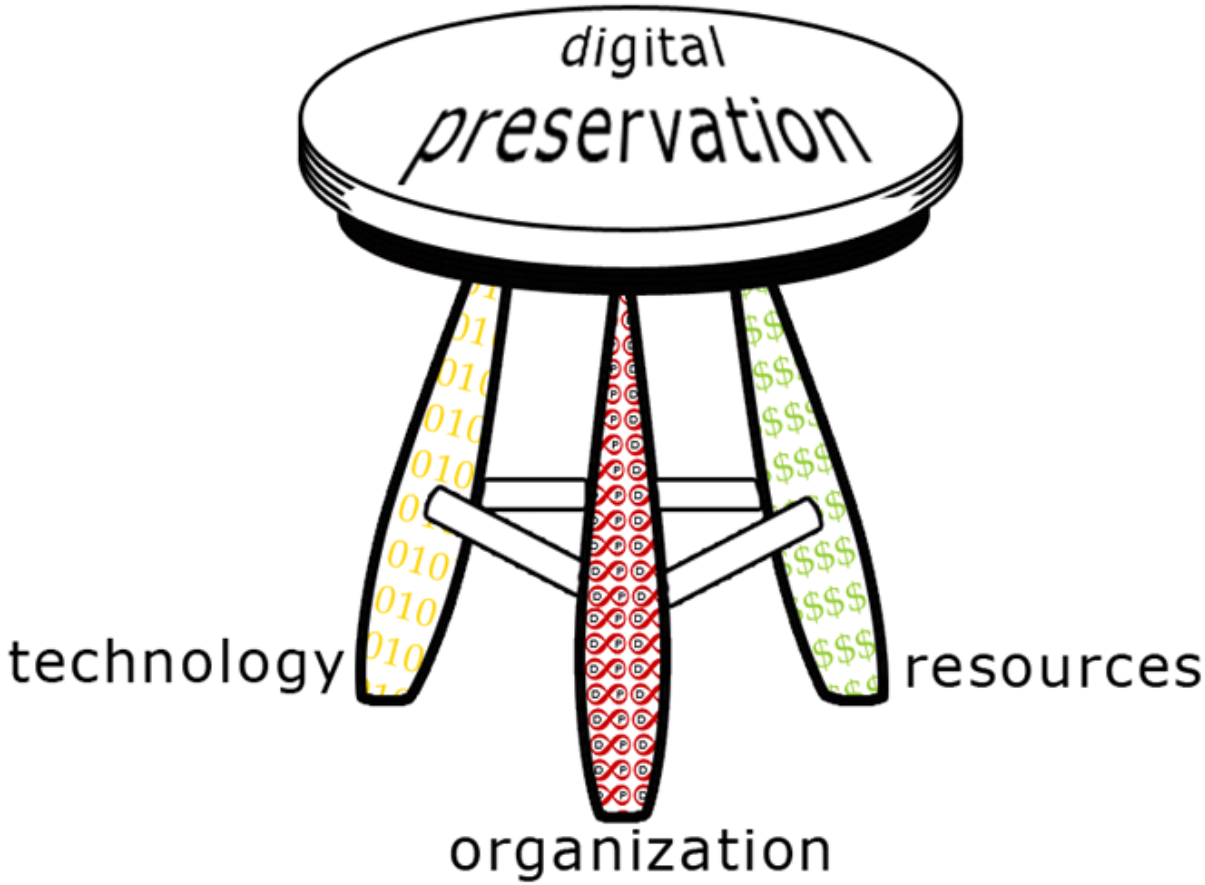
Do we have robust preservation plans to ensure long-term access to data?

How are we going to recruit or train staff with skills in digital preservation needs?

How can we collaborate more closely to meet the challenge of digital preservation?

DigitalPreservationCoalition

Getting started…



DigitalPreservationCoalition

Five step maturity model (Kenney and McGovern 2003)

Acknowledge:
Understanding that digital preservation is an issue
Act:
Initiating digital preservation projects;
Consolidate:
Seguing from projects to programs;
Institutionalize:
Incorporating the larger environment;
Externalize:
Embracing inter-institutional collaboration and dependency.

http://quod.lib.umich.edu/s/spobooks/bbv9812.0001.001/1:11?rgn=div1;view=fulltext

DigitalPreservationCoalition

JISC ULCC

AIDA – Assessing Institutional Digital Assets (ULCC 2009)

Elements	Preparedness	Levels	Institutional Scorecard
Technology (11 parts) Organisation (11 parts) Resources (9 parts)	Acknowledge Act Consolidate Institutionalize Externalize	Institutional + Departmental	NOT AN AUDIT measurement v improvement Get people talking!

http://aida.jiscinvolve.org/wp/toolkit/

DigitalPreservationCoalition

A little bit of self- assessment is a good place to get started

… it's very easy to get lost in all the other detail

DigitalPreservationCoalition

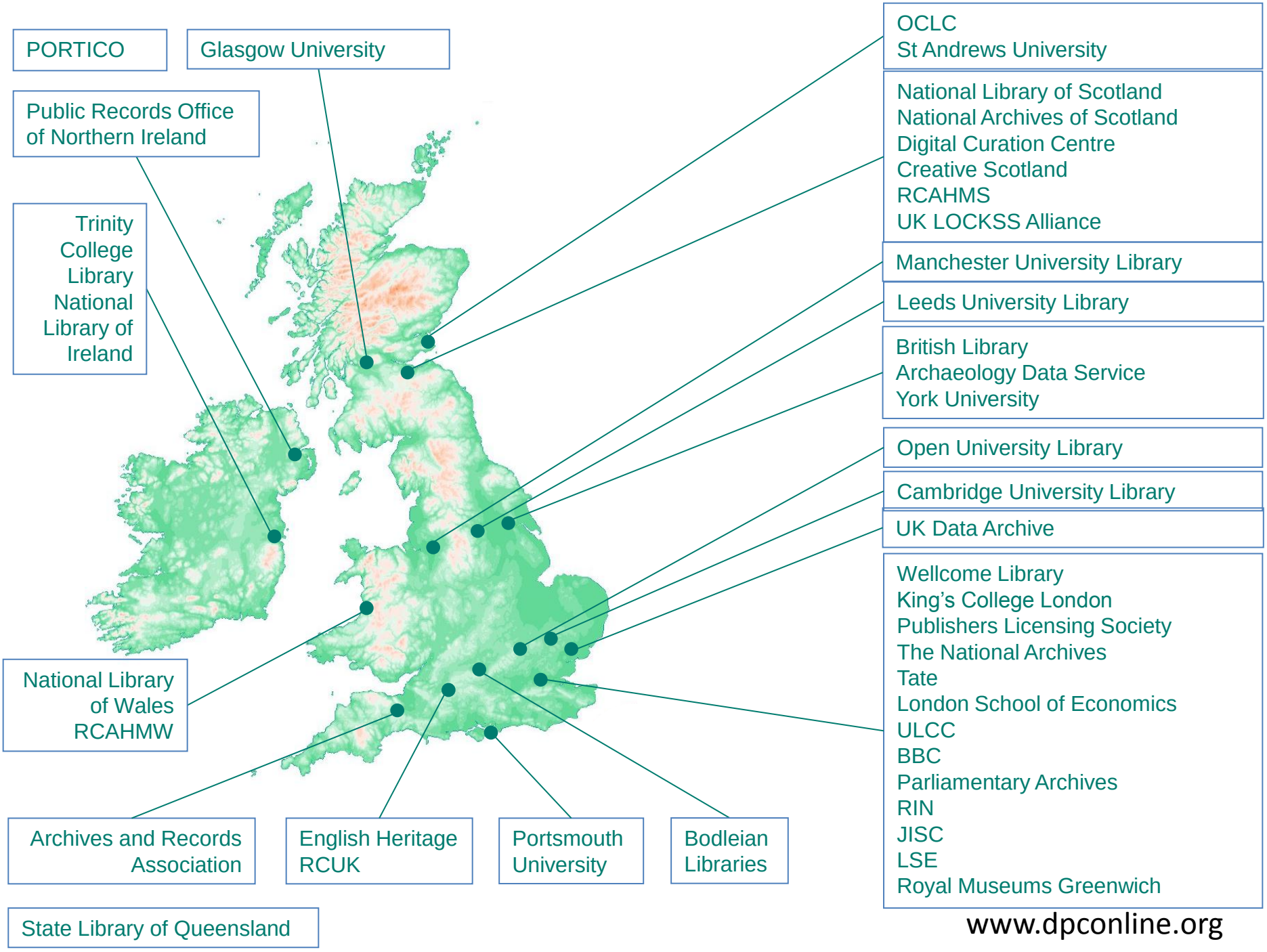
The Digital Preservation Coalition

...to make our digital memory accessible tomorrow ...

- Workforce development
- Advocacy
- Knowledge Exchange
- Assurance and Practice
- Partnership

UKLA

www.dpconline.org


PORTICO
Glasgow University
Public Records Office of Northern Ireland
Trinity College Library
National Library of Ireland
National Library of Wales
RCAHMW
Archives and Records Association
English Heritage
RCUK
Portsmouth University
Bodleian Libraries
State Library of Queensland
OCLC
St Andrews University
National Library of Scotland
National Archives of Scotland
Digital Curation Centre
Creative Scotland
RCAHMS
UK LOCKSS Alliance
Manchester University Library
Leeds University Library
British Library
Archaeology Data Service
York University
Open University Library
Cambridge University Library
UK Data Archive
Wellcome Library
King's College London
Publishers Licensing Society
The National Archives
Tate
London School of Economics
ULCC
BBC
Parliamentary Archives
RIN
JISC
LSE
Royal Museums Greenwich
www.dpconline.org

DigitalPreservationCoalition

DPC as a partner and friend in Digital Preservation

(join us)

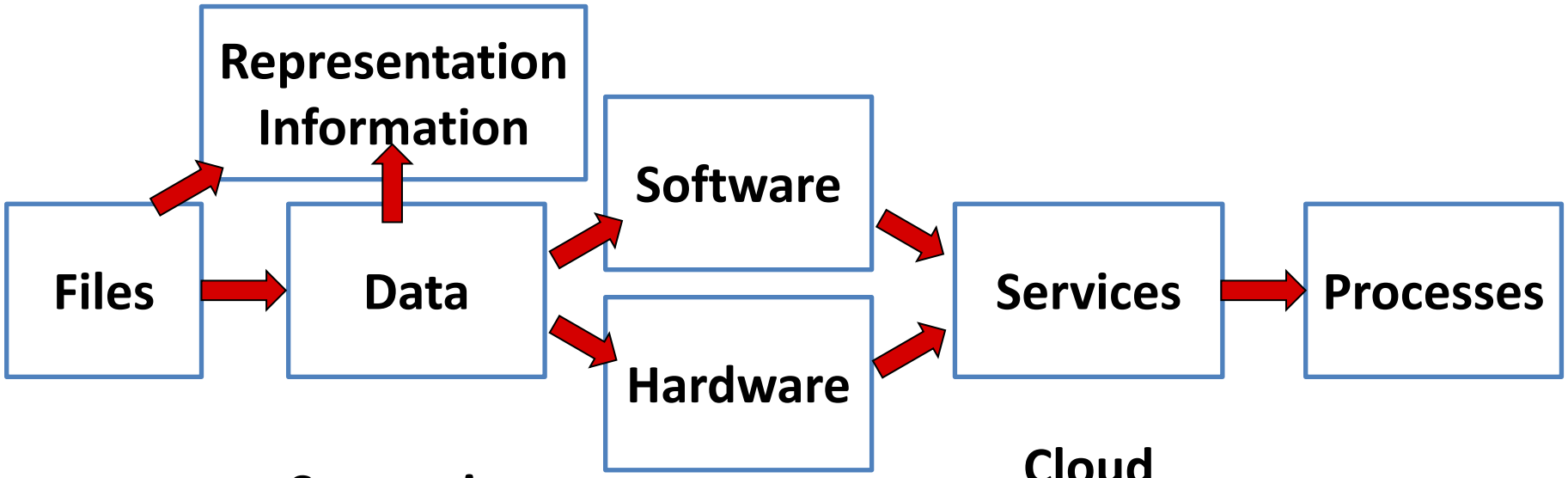
www.dpconline.org

TIMBUS

TIMELESS BUSINESS

Digital Preservation

Risk and Business Continuity Management



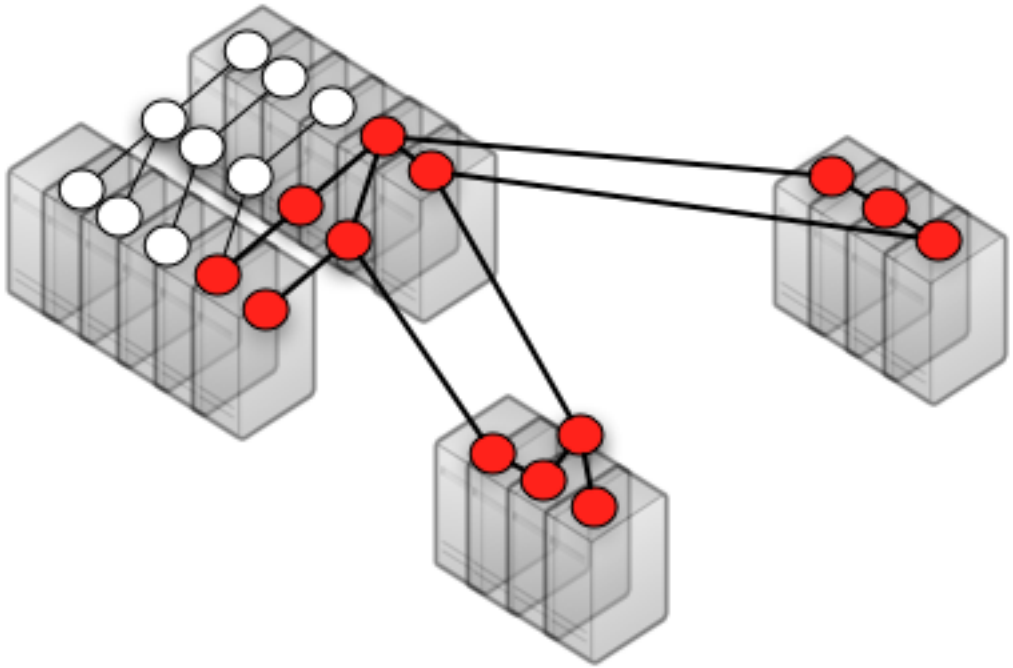
SEVENTH FRAMEWORK PROGRAMME

DigitalPreservationCoalition

TIMBUS

TIMELESS BUSINESS

1. Related services in a Business Process
2. Distributed – Public Cloud / Private Cloud / On Premise etc.



3. Risk analysis determines subset of business process must be available in >= 30 years
4. TIMBUS methods and tools used to preserve business process

SEVENTH FRAMEWORK PROGRAMME

www.dpconline.org

DigitalPreservationCoalition

TIMBUS

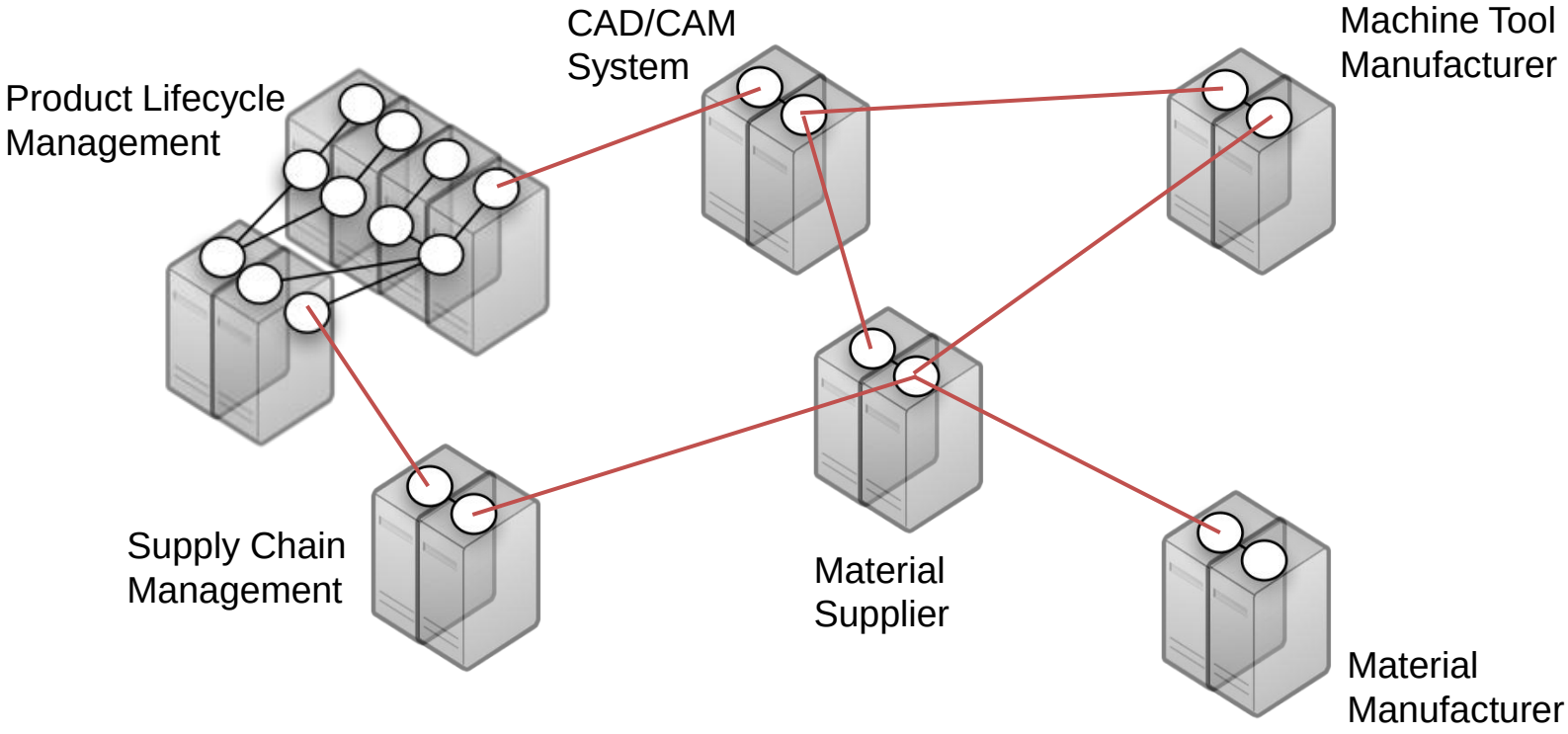
TIMELESS BUSINESS

1. Aerospace company BigAirTransportation (BAT) is involved in a project with an estimated lifecycle of 50 years.
2. This spans the initial vision through to product decommissioning and recycling.
3. The vision is to produce a lighter passenger jet that will use approximately 20% less fuel than their current product and use key components (e.g. wings, fuselage, and empennage) that require 10% less maintenance inspection.
4. After initial design and cost engineering the passenger jet will be constructed from 52% reinforced plastic composite.

SEVENTH FRAMEWORK PROGRAMME

Concrete TIMBUS Example

DigitalPreservationCoalition



DigitalPreservationCoalition

TIMBUS

TIMELESS BUSINESS

Planning

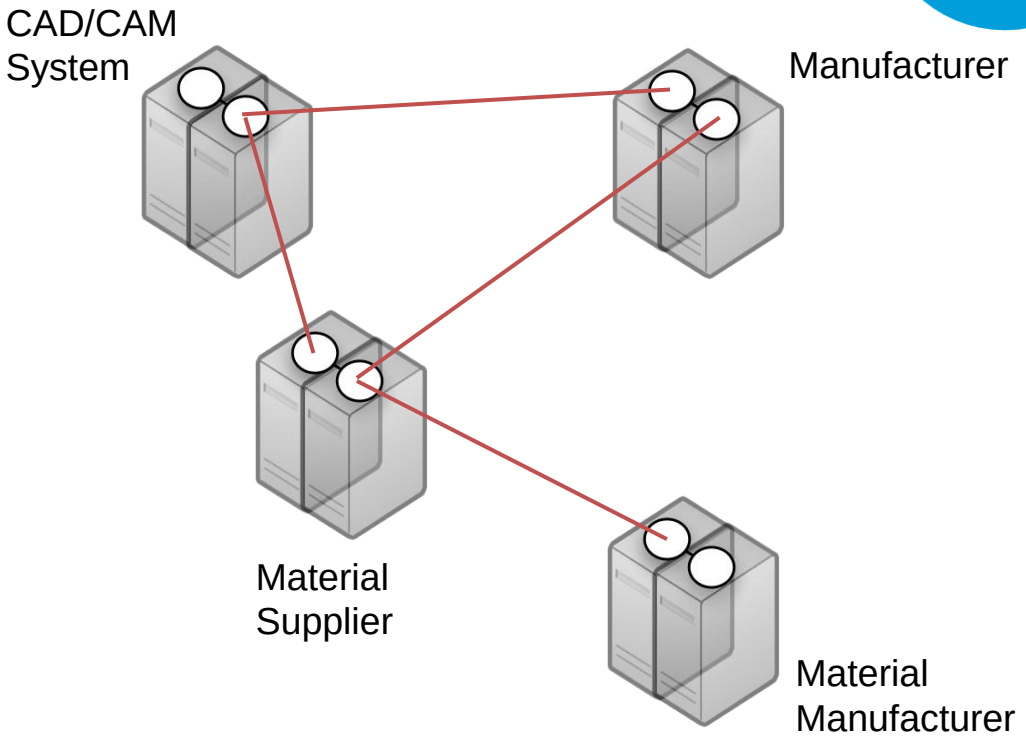
- Risk analysis performed
- PLM -> SCM *not* deemed expedient for preservation
- CAD/CAM -> Material Supplier -> Material Manufacturer -> Machine Tool Manufacturer *critically* expedient for preservation

Preserve – *performing the preservation of business processes*

- Legalities Lifecycle Management
- Software Service Engineering for Preservation
- Business Process Virtualization and Storage
- Processes and Standards for Digital Preservation of Business Processes

Re-deploy– *reruning/extending a business process at a future date*

- Business Process Exhumation and Integration Support (with Future Simulated Test Bed)



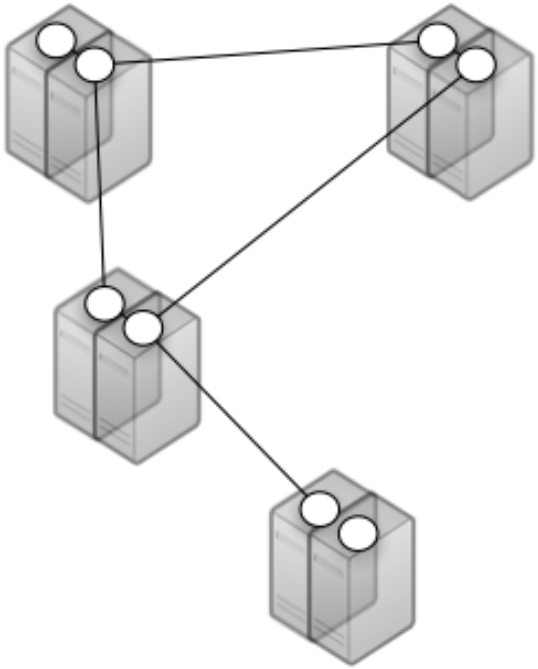
SEVENTH FRAMEWORK PROGRAMME

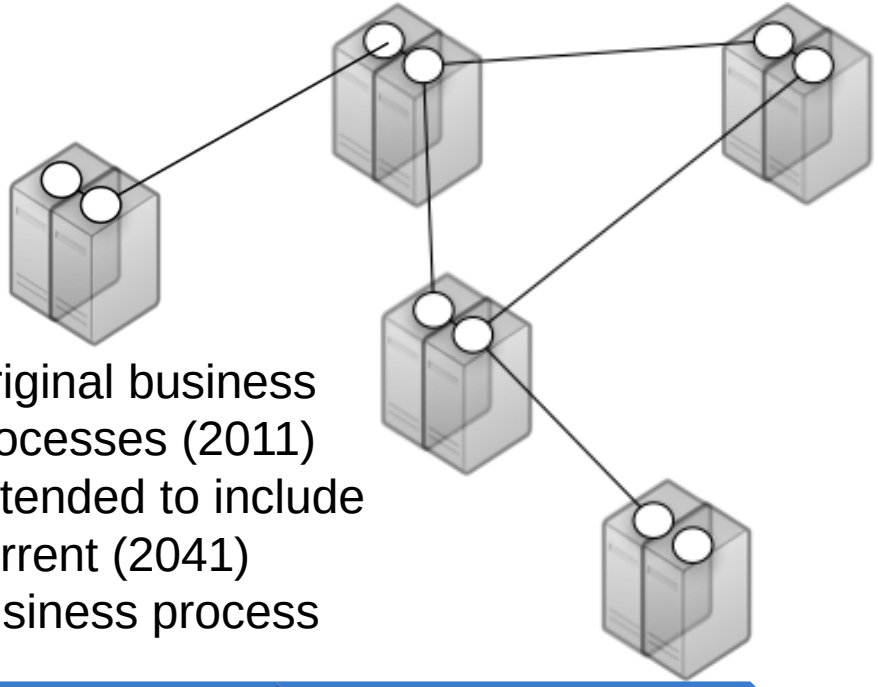
www.dpconline.org

DigitalPreservationCoalition

TIMBUS

TIMELESS BUSINESS





Original business processes (2011) extended to include current (2041) business process

2011 ... 2041

SEVENTH FRAMEWORK PROGRAMME

www.dpconline.org

TIMBUS

TIMELESS BUSINESS

- Planning Innovations
  - Service Dependency Analysis
  - Business Process Context Capture
- Preservation Innovations
  - Legalities Lifecycle Management
  - Software Service Engineering for Preservation
  - Business Process Virtualization and Storage
  - Processes and Standards for Digital Preservation of Business Processes
- Redeploy Innovations
  - Business Process Exhumation and Integration Support (with Future Simulated Test Bed)

SEVENTH FRAMEWORK PROGRAMME

DigitalPreservationCoalition

‘Digital Preservation: what I wish I knew before I started’

It won’t do itself
It won’t go away
Don’t wait for perfection

What is the question?
6 basic challenges and skills
Three lessons from experience