



DCH RP

Infrastructure Concertation Meeting 23/04/2014

If you have a map then you need a scale?

- What's the big picture?

The things that maps won't tell you?

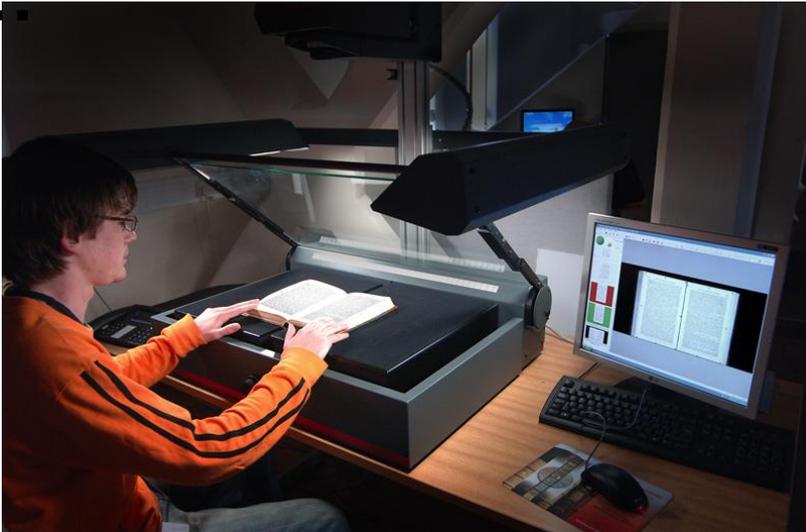
- Effort (ie cost)

What's going to stop us today:

- (Money?)
- Technology?
- People?

Shameless advert

Digital preservation typically makes bleak reading



When asked about how long their digitised resources would be available for, JISC-funded projects said ...

‘In perpetuity’

‘Indefinitely’

‘50 years’

‘10 years then

elsewhere’

‘until 2014’

‘forever or for three

years’

DPC/Portico/ULCC 2010

www.dpconline.org



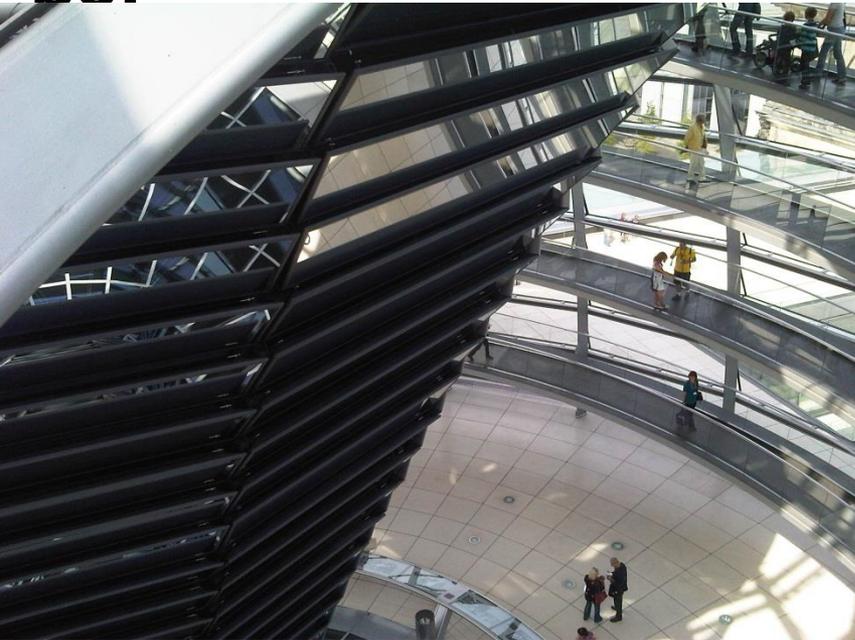
What's the problem?

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



Digital **Preservation** Coalition

**We do preservation
because we want to
be:**



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. scientific value, access to heritage, value of social knowledge

4. Wealthier

e.g. efficient business, management of IP, employment, planning, creative

5. Healthier

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

6. Greener

e.g. evidence-based policy development, efficient data retention

And because of



1. Legal Compliance

e.g. Sarbanes-Oxley, Data Protection

2. Regulatory Compliance

e.g. power generation, aviation,
banking

3. Legal protection

e.g. patents, mis-selling, detection,
audit

4. Unanticipated exploitation

e.g. petro-chemical, music,
pharmaceuticals

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 2008
www.dpconline.org



Digital **Preservation** Coalition

Digital preservation is not just about 'data':
Digital preservation is not just about 'access':
Digital preservation is not just about 'tools':

it's about

people

opportunity

A bright yellow, multi-pointed starburst graphic with a black outline, positioned over the word 'people' and partially overlapping the word 'opportunity'.

***Start
Here!***



Safer

Smarter

Healthier

Harmonisation of data storage and preservation?
Improved interoperability?
Fostering the conditions for cross-sector integration?
Governance models for infrastructure integration?

But there are no destinations.
They are means to an end

Wealthier

Greener

Fairer



Digital Preservation Coalition



Why this matters?

Money turns out to be the major problem facing the future of our digital heritage' (Rosenthal 2012)

Data volumes increasing 60% pa

Data storage increasing 25% pa

Data budgets increasing 2% pa



CHEAPER?



LOWER BARRIERS?

Does DCH-RP help this?

Maybe!

‘existing e-infrastructures are **efficient** channels for the delivery of advanced services’

‘it’s possible to establish **common** policies and processes’



CHEAPER?



LOWER BARRIERS?

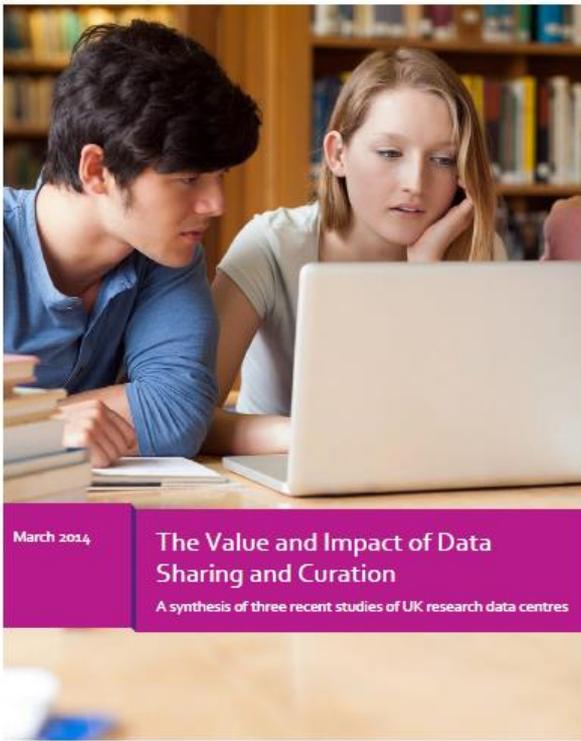
Does DCH-RP help this?

PROVE IT!

Are these e-infrastructures actually more efficient? What about other market models? Why not commercial? National? Sectoral?

Is it actually cheaper to get people to work together? What is DCH? Do they have much in common with each other?

Jisc



Does DCH-RP help?

The roadmap needs to consider how implementation of its four basic components deliver financial advantage

...

Four things to examine

Return on Investment from Data Centres

Collaboration to clarify the costs of curation

Commercial providers

Workforce developmet



We can tell a good story about costs!

quantitative analysis indicates that: ...

- The value to users exceeds the investment made in data sharing and curation via the centres in all three cases – with the benefits from 2.2 to 2.7 times the costs;
- Very significant increases in work efficiency are realised by users as a result of their use of the data centres – with efficiency gains from 2 to 20 times the costs; and
- By facilitating additional use, the data centres significantly increase the returns on investment in the creation/collection of the data hosted – with increases in returns from 2 to 12 times the costs.

<http://blog.beagrie.com/>



Digital**Preservation**Coalition

What's going to stop us today

- Money
- Technology
- People



<http://blog.beagrie.com/>



The technology? Some observations that impact on the roadmap

1. No longer clear what 'data' is
2. What are the implications for cybersecurity and acceptance testing if you need to store and execute an archived technology stack?
3. Sensitivity review (redaction) and compliance?
4. Balancing the need and practicality of metadata gathering and management?
5. Scaling up while managing (reducing ?) costs?
6. Integrate with existing infrastructure





**Specific implications for the roadmap
arise in sections 5.2.1 and 5.2.2**

5.2.1 Harmonise data storage and
preservation

5.2.4 Enhance interoperability



5.2.1 Harmonise data storage and preservation

(Is this roadmap actually a roadmap about harmonising data storage and preservation?)

Digitisation workflows should be a lot simpler to harmonise than collection of born digital collections?

e.g. metadata capture

Need to get best (any!) DP tools embedded in mainstream software

Preservation from the outset



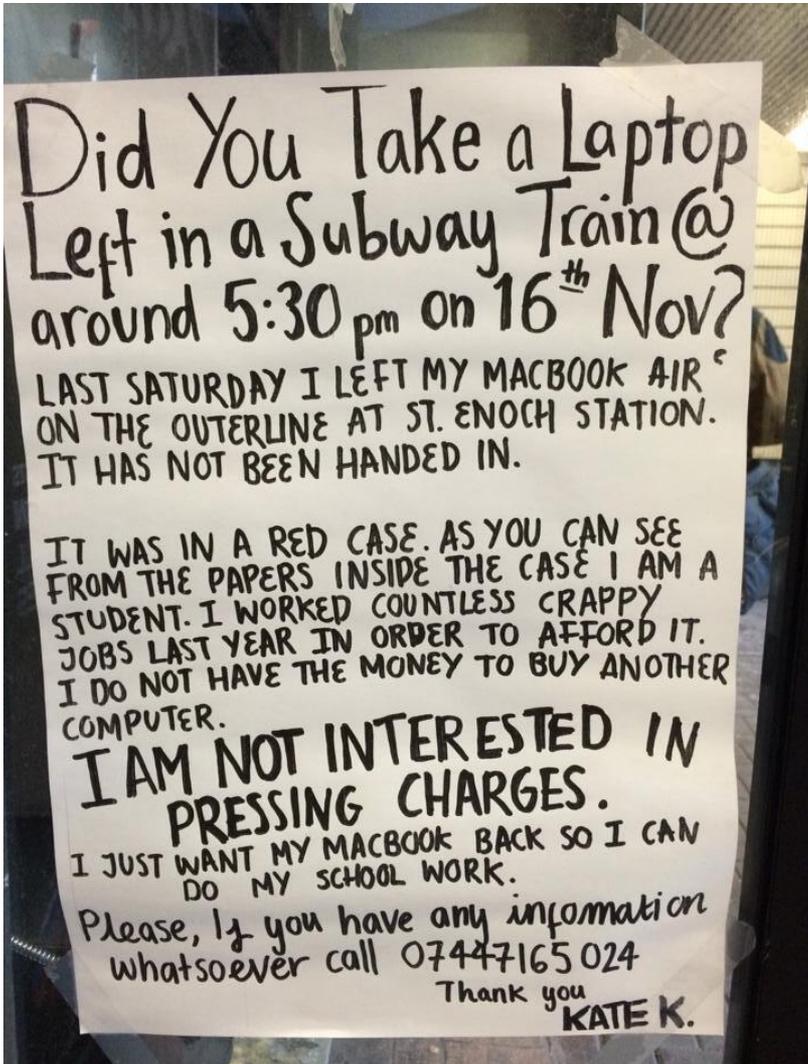
5.2.4 Enhance interoperability

Is this really just about digitisation?

Interoperability of workflows – is this realistic given diverse mission of DCH institutions?

What *are* DCH institutions? Libraries? Museums? Galleries? Publishers? Archives?

Naming things is very controversial in cultural institutions: vocabulary ontology building is tough



The people? Some observations which can mostly be substantiated

1. Community building is harder than technology building
2. There is significant distrust of (opposition to) cloud providers in the cultural heritage sector and it's not just about the technology
3. The majority of DCH institutions are small, making them hard to reach
4. There is a significant skills gap
5. Staff involved in preservation are not numerous, have numerous other responsibilities and are scattered
6. The digital preservation community is highly fragmented



Digital**Preservation**Coalition

Specific implications for sections 5.2.3 and 5.2.4

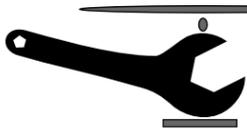
5.2.3 Establish conditions for cross
sector integration

5.2.4 Establish a governance model for
infrastructure integration

Connecting digital preservation expertise

Home · Blogs · Paul Wheatley's blog

COPTR tools registry beta launch



Almost a year ago, I presented a proposal to the Aligning National Approaches to Digital Preservation (ANADP) group to create a community tool registry. I was frustrated by the profusion of tool registries and the lack of coordination between them. Pooling the knowledge in one place would result in a far better resource. It would be easier to discover new tools, to share experience in using them and to help avoid the tool



development duplication we've seen so much of in the past. As ANADPII kicks off today in Barcelona, I'm very pleased to announce the beta launch of COPTR: the Community Owned digital Preservation Tool Registry.

We've been working to collate, combine, de-duplicate and align the contents of 5 existing tool registries from: The Open Planets Foundation (OPF), The National Digital Stewardship Alliance (NDSA), The Digital Curation Centre (DCC), The Digital Curation Exchange (DCE) and the Digital POWRR Project. There were of course quite a few duplicates to weed out but the scope and depth of COPTR now supercedes anything out there that I've seen previously, albeit with an inconsistency of depth between resulting tool registry entries. Each source registry had it's own differing characteristics. At one end of the scale the DCC registry had really strong detail but coverage of well under a hundred tools. The DCE registry included over three hundred tools but with each tool described in far less detail.

5.2.3 Establish conditions for cross sector integration

- Assumption that this is not already happening
- Model existing good practice
- Derive maximum benefit from existing community endeavours:
 - Registry services
 - Workforce development
 - Policy co-ordination



5.2.4 Establish a governance model for infrastructure integration

- Assumes that there are not already organisations that could provide such governance ... RDA, OPF, DPC, nestor, NCDD, APARSEN, Archivemata, Digital Repository of Ireland
- Etc.
- How do I get my data out of the Cloud?
- Focus on 'Trusted Digital Repository' seems redundant (and I wish you would say this explicitly!)
- IaaS changes the market place





And is there a converse side to this whole discussion

- Ask not if DCH institutions can use the cloud for preservation ...
- But ask if DCH institutions can collect and preserve the cloud?
- (Might all preservation ultimately look a bit like a web archiving workflow?)





Digital**Preservation**Coalition



5 Awards

- Research and Innovation
- Training and communications
- Student work
- Safeguarding digital legacy
- Industry

<http://www.dpconline.org/advocacy/awards/digital-preservation-awards-2014>



DCH RP

Infrastructure Concertation Meeting 23/04/2014

If you have a map then you need a scale?

- What's the big picture?

The things that maps won't tell you?

- Effort (ie cost)

What's going to stop us today:

- (Money?)
- Technology?
- People?

Shameless advert