

Incremental

A pilot project for supporting research data management

'Start making Sense: Talking Data Management with Researchers'

Catharine Ward, Cambridge University Library

Decoding the Digital, British Library, 27th July, 2010

Outline

- Objectives
- Scoping Study
- Concerns and Issues
- Key findings
- How can we help researchers?

Objectives

- A scoping study to determine key research data management and preservation needs of researchers
- Develop a plan for addressing these issues
- Create and pilot tools and services
- Embed the resources and findings within the University departments and supporting services and feed these up to a wider audience through the DCC, DPC and JISC

Scoping Study

- Building on HATII digital preservation study
- Semi-structured interviews
- Similar departments to allow comparison



- **Archaeology**
- **Chemistry**
- **Engineering**
- **English**
- **Public Health**
- **Biology / SPRI**

Concerns and issues

- Procedures for creating and organising data
- Data storage and access
- Backup
- Data sharing and re-use
- Preservation

Creating and Organising Data

"They've always had a common staff network drive and it's always been the bane of everyone's lives to find stuff on it."

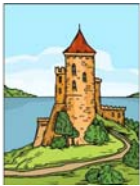


"There is a professor here who has never deleted an email message in his life."

"The volume of data produced makes maintenance a bit like drinking from a fire hose."

Data storage & access

"People will work at home, so things will be stored locally and shuffled between machines on memory sticks."



"Research groups tend to run their own little fiefdom."

"People bring in sticks with 4GB of data that simply no longer work – and nothing can be done to retrieve it."

Data backup

"I just back everything up onto data sticks. I didn't even know you could back-up to servers."



"PhD students lose material all the time...and they are exactly the people who want to be backing up. These are people who are creating data which are life and death important to them"

Data sharing and reuse

"The biggest issue to overcome is personal investment -- it's like giving away your baby!"



"It requires a lot of work to make data accessible and understandable."

Preservation of digital research data

- Current practice
- Selection
- Financing
- Feasibility

Current practice

- Do nothing! Data stored on primary storage and continually backed up.
- Other strategies include CDs, EHDs or submitted to data centre or IR
- Researchers not aware of difference between 'back up' and 'preservation'.

"I keep a defunct Macintosh under my desk solely for the purpose of being able to use file-maker pro 2.0, in order to look at a catalogue that my dead PhD supervisor created of manuscripts in the UL in 1995."

Selection

- Most researchers hope to keep everything
- Gulf in opinion between IT support and researchers

“At Cambridge, one researcher had kept every email he has sent or received, as he believes the information in them is important.”

“How do you decide what can be deleted as it’s all a record – I’m not confident to make that decision.”

Financing

- Funding bodies expectations
- Costing in preservation make grant bids look uncompetitive

“When you take a book out of the library and there are pages missing, you bring it back to the library and expect them to fix it ”.

Feasibility

- What data are worth preserving?
- What does ‘preservation’ mean?
Is it keeping the bits or retaining the usability?

“When producing something for the general public, you have to try and think ahead, but on short term funding keeping pace with technology is a challenge”.

Issues with existing guidance & training

- Difficult to find
- Difficult to use
- Don't know who to ask for assistance
- Training is often not relevant or convenient
 - need for disciplinary examples
 - online resources for anytime access

**"The whole thing
is incredibly dull."**

**"There's no point
being told all this
stuff when you're
not using it
because – I only
learn how to do
things when I need
to know."**

Key findings

- Main concerns similar between Cambridge and Glasgow - not discipline-specific - but recognise the need for disciplinary examples for training to be understood
- Simple issues often the most irksome
- Training and guidance resources must be simple, engaging and easy to access
- Points of intervention
- Language matters

How can we help researchers?

GUIDANCE:

- Make help easier to understand
 - o by producing simple guidance using language and terminology that makes sense to a researcher
- Make help easier to find
 - o by producing centralised web pages which contain or point to relevant support and guidance.



Work in progress: centralised web pages

Support for Managing Your Research Data

Creating
your data

Organising
your data

Using
Your data

Looking after
your data

How can we help researchers?

TRAINING:

- Offering practical training resources with discipline-specific examples
 - Stand alone slides/resources that can be dropped into other courses – *train the trainer*
 - Provide online training materials and tutorials



How can we help researchers?

TAILORED SUPPORT:

- Connect researchers with support staff who can offer tailored advice and partnering
 - o Raise awareness of existing support staff and services through data management web pages and outreach
 - o Offer on-to-one support to help researchers define best approach for their context

What to call this?

- Translate research data management vocabulary from specialist to non-specialist – but how?

<http://incrementalproject.wordpress.com/2010/07/14/vocabularyjargonterminology-synonyms-and-specialist-language/>

Work towards the development of a comprehensive data management infrastructure



Thanks for listening 😊

For further info:

Email: Catharine Ward - cw330@cam.ac.uk

Project website:

<http://www.lib.cam.ac.uk/preservation/incremental/>

Project blog:

<http://incrementalproject.wordpress.com/>