

# The Incremental approach

Practical guidance and training to facilitate research data  
management in Cambridge

Elin Stangeland  
DSpace@Cambridge Repository Manager  
Cambridge University Library

RDMF VI Leicester 5-6<sup>th</sup> of May 2011

# Outline

- Cambridge background
- Changing context
- Projects
- Current status
- Next steps



# Cambridge University background

- Annual report
  - £266.9m - £ 123.3 RCUK
  - Appx. 6000 academic staff
- Organisation
  - Decentralised IT services
  - Collegiate University
- Policy level
  - Information strategy 2009
  - Guidelines of Good Research Practice
  - UL Strategic framework 2010-2013

# DSpace@Cambridge



- University of Cambridge - Institutional repository.
- Currently holds appx. 200.000 items covering a wide range of content:
  - Research Publications
  - E-theses
  - Research data
  - Digital media
  - Cultural heritage
  - Research data



CUPID →

Incremental

DataTrain

# Incremental

---

- Scoping study
- Implementation of support activities

# Scoping study findings

- Creating and managing files
- Storage and back-up
- Preservation and re-use
- Data sharing
- Existing guidance and training
- Terminology is not straight forward

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

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

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

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

Home

- [Create](#)
- [Organise](#)
- [Access](#)
- [Look After](#)

About

- [Browse](#)
- [Glossary](#)
- [Training](#)
- [Advice & Support](#)

## Support for Managing Research Data

Research data takes many forms, ranging from measurements, numbers and images to documents and publications. These web pages will help you plan, create, organise, share, and look after your electronic research materials, in whatever form they take.



Resources for funding bids, data collection, data planning, Intellectual Property Rights and ethics.



Resources for naming files, structuring folders, version-control, and documenting your work.



Resources for accessing your own data, external data sources, collaboration, and publishing.



Resources for storing, backing up, and maintaining your data over time, archiving, and data sharing.

Home

- [Create](#)
- [Organise](#)
- [Access](#)
- [Look After](#)
- [About](#)
- [Browse](#)
- [Glossary](#)
- [Training](#)
- [Advice & Support](#)

## Data Management Planning

The University and research funders encourage researchers to plan their data management from the outset. This will often mean a formal plan for managing data (a 'data management plan').

However, even informally setting out your plans and project guidelines can make your life much easier. If you want to be able to reuse your data or manage collaboration with colleagues, it helps to plan for that from the outset. Decisions you make about which software to use, how to organise, store and manage your data, and the consent agreements you negotiate, will all affect what is possible to do in the future.

Planning ahead for your data management needs and aspirations will help ensure that:

- you have adequate technological resources (e.g. storage space, support staff time).
- your data will be robust and free from versioning errors and gaps in documentation, etc.
- your data is backed up and safe from sudden loss or corruption.
- you can meet legal and ethical requirements.
- you are able to share your finalised data publicly, if you and/or the funder desires.
- your data will remain accessible and comprehensible in the near, middle, and distant future.

### Frequently Asked Questions

Expand All

- [What do research funders expect?](#)
- [Where do I start?](#)
- [What should I cover in a full data plan?](#)
- [Who can help with data planning at the University of Cambridge?](#)
- [Further Reading](#)

### Related Links

- [Data Management for Funding Bids](#)
- [Information on Funding Council Requirements for Data Management Plans \(DCC\)](#)
- [Overview of Individual Funders' Data Policies \(DCC\)](#)
- [DMP Online \(interactive tool for creating your data plan\)](#)

### Downloads

- [Data Management Plan Template](#)
- [Managing and Sharing Data: a best practices guide for researchers \(UK Data Archive\)](#)
- [UK funding councils: Data retention and access policies](#)

# Training

- Practical training covering relevant aspects in data management
- Learning videos



What are the main things that researchers should make sure to do at the start of a project using sensitive data if they might want to share or archive it in the future?



# DataTrain

- Data management training for Department of Archaeology and Department of Social Anthropology
- Collaboration with Archaeology Data Service (ADS)
- Shareable resources – deposited with ADS and JORUM
- Status
  - Archaeology course ran over 4 sessions in March 2011 – will be permanently integrated with Research skills training
  - Social Anthropology ran yesterday – day workshop

## Data Management Plan for Post-Graduate Research Projects: Prompt Sheet

<b>Researcher:</b> Name		
<b>Project Title:</b> Provisional dissertation / thesis title		
<b>Project Duration:</b> Dates of post-graduate research project		
<b>Project Context:</b> Where is the research being carried out, and what is under study? Is the research individually based, part of a larger project, or being carried out in agreement with an institution(s), e.g. an archaeology contract unit(s), a museum(s), a state/regional authority, or community group(s), etc?		
<b>1. What Data will be Produced?</b> What physical data will you study? And what digital data will be captured/derived from these? (field notes, images, measurements, spreadsheets, survey data, etc). What data will be 'created' digitally (images, some analytical and survey data, etc)? Describe the methods/standards for data creation? What file formats and software will you use? Consider how many individual files you expect to make, anticipated file sizes, and total storage volume.		
<b>2. How will the Data be Documented and Described?</b> Think about what contextual information is required to make the data understandable to others: What standards will be used to record the data? What information on the data collection methods, standards, and context ('metadata') will be recorded for each data type/set? Where will the metadata for each data type/set be located? (e.g. within the data file and/or as separate metadata text document, and/or in method chapter/appendices in the thesis)		
<b>3. Has a 'File Structure/Naming Form' been completed?</b> (see separate form)		
Date Created:	Date Amended:	Version no.
<b>4. Deposition of E-Thesis: delete as appropriate and state reasons:</b> A. Intend to deposit e-thesis with [fill in]... with open access. B. Intend to deposit e-thesis with [fill in]... with a time-limited embargo on open access. C. Do not intend to deposit e-thesis. <b>Give Reasons:</b> Intended publication of articles or book (e.g. BAR report) after submitting (three years is the standard length of an embargo) Agreement with sponsoring body or institution (museum etc) to embargo e-thesis. <b>NR</b> If you intend to deposit your thesis with a digital repository agreement must be sought with all		

# Current status

- We have a better understanding of the provision and needs of Cambridge researchers
- We have some experiences in training postgraduate students
- Repository service review
- Repository infrastructure needs evaluation

# Areas of future activity

- DSpace@Cambridge Research Data and Digital Curation Officer
- Research office collaboration
  - Project support
- Training
- Infrastructure
- Data management community

# DSpace@Cambridge Research Data and Digital Curation Officer

- Incremental and DataTrain activities
- Institutional liaison
- Contribute to DSpace@Cambridge service and infrastructure development
- Contribute to other preservation related activities

# Research Office collaboration

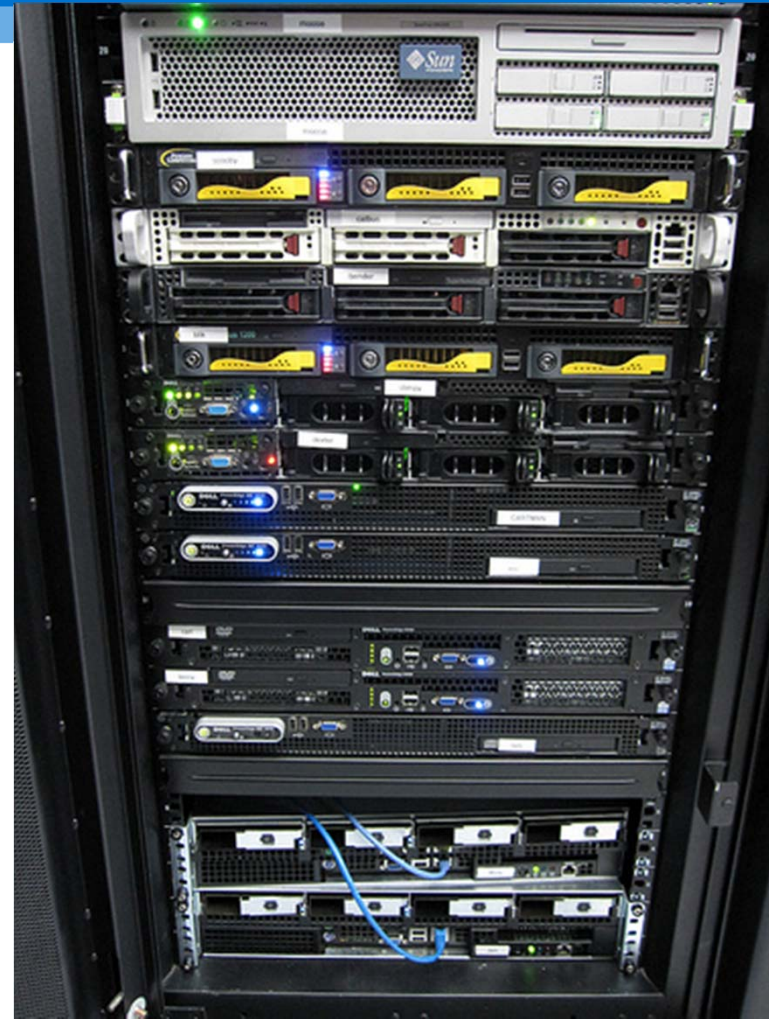
- Awareness raising – service alignment
  - Staff presentations
- Support for bid writing
- Project support
  - Guidance on complying with Funding body requirements

# Training initiatives

- Continue DataTrain initiatives
- Contribute to the organisation of data management workshops
- DSpace train the trainer sessions
- University Library Skills course
- More DC 101 courses
- DCC Roadshow – November 2011
- RDMF in Cambridge – Autumn 2012

# Infrastructure

- Continue developing DSpace@Cambridge service to better support perpetual access to research data
- Expand current provision to reference research materials created by Cambridge researchers
- Support developments of infrastructure to support management of “active” data



# Questions?

DSpace@Cambridge:

<http://www.dspace.cam.ac.uk/>

Incremental project website:

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

DataTrain project web site:

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

Email: Elin Stangeland - [es444@cam.ac.uk](mailto:es444@cam.ac.uk)