



LIBER Case Study:

Research Data Services at University College London

AUTHOR: Dr Max Wilkinson, University College London, max.wilkinson@ucl.ac.uk

KEYWORDS: generic, institutional

1 What was the starting point?

At University College London (UCL) we have approached the design and implementation of a Research Data Service using six components (workstreams) of a strategic roadmap.

1. Organisation: Forming and embedding an organisational unit to build the services. This was primarily a targeted recruitment of higher degree graduates with excellent data and technology skills. The ongoing support for this team was key to building momentum for this work package.
2. Policy: In order to support and inform UCL researchers of their obligation to undertake high-quality research using best practice across all disciplines, UCL ratified a 'Research Data Policy' that laid out the responsibilities of all stakeholders in the Research Data Management life cycle. This policy serves as a basis which is supported further with guidelines and current service provisions. The Policy is reviewed annually.
3. Infrastructure: This workstream described the procurement, installation and commissioning of hardware and software components required for a suite of research data services.
4. Data Management: This workstream sought to define the role a centrally supported infrastructure could play in supporting data management by the researcher. There needed to be a balance between providing infrastructure over which data management can be facilitated and supporting an application layer that provides data management tools. At UCL we favour supporting infrastructure over which data management can be facilitated.
5. Service Integration: This workstream defined the existing services that would either be dependencies for the Data Management infrastructure (e.g. identity management) or services that data management would add value to, e.g. research reporting, institutional publication repositories, etc.
6. Business Analysis: This workstream is constructing sustainability models to inform service development and communicate realistic costs.
7. Communication Planning: This workstream is building a rich and comprehensive communication plan to engage uptake and feedback feature requests, etc.

2 What kind of research data is targeted?

UCL is a large and research-intensive university. UCL Research Data Services is being designed to support all forms of research across all disciplines. As such we do not make selection and/or validation judgements, nor do we have a particular focus or format prescription. Research Data are what the researcher defines and makes judgments on.



June 2014



3 What is the organisational framework?

Roles and responsibilities

Roles and responsibilities are described in the institutional research data policy and cover the following (compare UCL Research Data Policy, version 1.0, http://www.ucl.ac.uk/isd/staff/research_services/research-data/#rdpolicy):

- Data Creators (students, supervisors and researchers)
- UCL Research Data and Network Services Executive
- Director of UCL Library Services and UCL Records Manager
- RIISG
- Vice Provost (Research)
- Provost

It is the responsibility of the data creators to ensure good data management practices (incl. data management plans) are followed. Whilst it is the responsibility of the Data and Network service to provide storage and guidance for data management, the Library is responsible for curating smaller datasets and for providing guidance and advocacy for research data management.

Policies

The purpose of UCL's Research Data Policy is to provide a framework to define the responsibilities of all UCL members and to guide researchers and students as to how they should manage the data, thus enabling research data to be maintained and preserved as a first-class research object and made available to the widest possible audience for maximum impact.

The policy is intended to ensure that research data created as part of the research process are:

- Accurate, complete, authentic and reliable;
- Attributable and citable;
- Identifiable, retrievable and available with minimal barriers;
- Secure from loss and degradation;
- Retained for a minimum of ten years after publication or public release;
- Compliant with legal obligations, ethical responsibilities and the rules of funding bodies.

The policy also states that, after primary use, research data should be made available in as open a manner as possible.

4 What kind of support services are provided to researchers?

We are putting together some training courses to help researchers understand how our services can ease the burden of SOME of their responsibilities. The course design is being driven by pilot and beta user feedback as well as by analysing user behaviour with the current services. It is anticipated that our second service, UCL Research Data Archive, will require significant investment in training as this is where research data are made ready for archiving and referencing.

In the meantime, a short tutorial is sufficient to enable users to use our primary service, UCL Research Data Storage.

5 What kind of infrastructure is provided?

We are (or soon will be) providing three services at UCL Research Data Services

1. UCL Research Data Storage: A high-capacity storage service for UCL researchers, their partners and collaborators.



2. UCL Research Data Archive: A data preservation environment where researchers can 'publish' their data for others to reference and access.

3. UCL Research Data Access: A search and presentation portal to promote and access UCL Research Data assets.

We are using two technologies on our storage service, a traditional high-performance file system and an object store. The file system can provide the performance that high-data users require while the object store can scale much better than a file system.

The presentation layer consists of an SSH/SCP access feature and cloud client that permits a 'dropbox' style interface.

The service is currently available to all researchers with UCL identity and access management credentials. We are working towards being able to provide the service to external partners and collaborators as well.

The services are provided to projects rather than individuals. We collect a small amount of metadata about the project during registration.

6 What have you learned so far? What's next?

- People do not trust new services: Pilot users are generally using the service to 'back up' their data. This is slowly starting to change.
- There is a large skills variance that needs to be addressed. This often requires more one-to-one consultations, which are resource intensive.
- There is a wide variety of user/researcher needs and requirements.
- Technology, particularly infrastructures, are changing rapidly, as are the costs.
- Neither technology nor culture change will work independently. There needs to be a merging of the two and this represents a new professional space. Skilled and enthusiastic people are extremely hard to find and often attract a premium cost. Institutes should not ignore this.

Further information

WEBSITE: http://www.ucl.ac.uk/isd/staff/research_services/research-data

CONTACT: Dr Max Wilkinson, Head of Research Data Service, University College London, max.wilkinson@ucl.ac.uk