DSpace@Cambridge Case History

By Elin Stangeland, Repository Manager

Organisational context

The University of Cambridge is a confederation of Colleges, Faculties and other institutions and functions with a relatively small central administration, and with central bodies consisting of, and mainly elected by, the current academic personnel of the Faculties and Colleges. The University has a decentralised structure, and large parts of the University decision making is done at faculty or departmental level.

DSpace@Cambridge was established as part of a three-year project running from 2003-2005. The project was a collaboration between Cambridge University Library, Cambridge University Computing Service, and MIT Libraries, funded by a grant from the Cambridge-MIT Institute. The mandate was to implement the DSpace software at Cambridge University and to further develop this into a managed preservation and delivery service. On 1 July 2006 DSpace@Cambridge became a strategic service to the University of Cambridge, managed jointly by Cambridge University Library and the University Computing Service. From the autumn of 2007 a service team has been established with a mandate to support the service and facilitate submission of digital content to DSpace@Cambridge.

The repository's mission statement

DSpace@Cambridge, the institutional repository of the University of Cambridge, is established to support students and employees at the University in their processes of studying, teaching and conducting research. This is achieved by facilitating a managed environment where scholarly content in digital form can be stored, preserved and disseminated online available for the international community.

Overview of current contents

When DSpace@Cambridge was established a survey was conducted in order to map the target community’s requirements of an institutional repository. The results led to a decision of capturing a wide range of digital content in DSpace@Cambridge, such as scholarly papers, research data, images, various multimedia etc. The DSpace@Cambridge team are also working together with the University Copyright Officer and representatives from the Board of Graduate Studies to establish a voluntary deposit service for e-theses.

Overview of current deposit activity

In the last couple of years very little promotion of DSpace@Cambridge has taken place, therefore few departments deposit on a regular basis. Main deposits currently consist of batch uploads from digitisation projects, for example, the Scott Polar Research Institutes Freeze Frame project1. Beyond batch uploads there is little faculty engagement at the moment, mainly because the service has not been advertised beyond the initial project phase. The new team is currently running a promotional campaign targeting new departments at the University and their academics.

1 http://www.spri.cam.ac.uk/resources/freezeframe/
Developmental phases

From an organisational point of view the DSpace@Cambridge service has gone through three phases, the initial DSpace@Cambridge project, where the service was established and DSpace were adapted to University of Cambridge repository requirements. After project finalisation in 2005 the service was maintained rather than developed for about two years, before a new team was established in 2007 consisting of four people: a repository manager, a support and liaison officer, a DSpace developer and a systems manager.

Institutional embedding and service sustainability

A business model and supporting research was commissioned from Apto Consulting Ltd. in 2004. This led to DSpace@Cambridge being developed into a centrally funded service for the University of Cambridge. It has received funding from the Planning and Resources Committee for the next five years and renewed funding depends upon the project team’s achievements during those five years. Since particular attention is being given to the preservation aspect of DSpace@Cambridge, commitment from both the University Library and the University Computing service has been secured to maintain the existing content of DSpace@Cambridge independently of staff retention.

Policy formulation and licensing

Formulating policies for DSpace@Cambridge has been a main focus this winter and a new set of policies has recently been implemented locally. The intention is also to implement these in the OpenDOAR policy tool. The policies are available at the DSpace@Cambridge support web site: http://docs.repository.cam.ac.uk/

The process of developing policies for an institutional repository is always an interesting one; on the one hand you have to comply with the wishes of the University’s legal council, on the other academics who wish to open access to the content. For the time being we have decided to follow UK copyright law regulations for the DSpace@Cambridge policies and licenses. In the future the intention is to move towards a set of licenses allowing the depositors increased flexibility in what rights they choose to retain or release.

Hosting and support

The DSpace@Cambridge team is responsible for both hosting and support of the service. The team is split between two departments: the University Computing service hosts the service, this also includes technical maintenance and continued service development. Staff employed by the University Library provides support services for depositors and other users of the system.

Measuring and demonstrating success

Since DSpace@Cambridge are only funded for the next five years, being able to measure and demonstrate that the repository is an important service is an important task for the new team. We are currently in the process of mapping possible evaluation measures and setting development targets for the service. Here is an overview of some measures we consider using:
Number of new departments depositing per year

- Deposit statistics (batch and front end deposits)
- Access and download statistics
- Number of possible depositors (i.e. University of Cambridge researchers) vs. number of active DSpace@Cambridge depositors
- We are currently in the process of establishing a publications database, Citations@Cambridge where various departments can maintain their publication lists. For participating departments we will in the future be able to use this to compare number of full text deposits in DSpace@Cambridge with the actual number of publications from the various departments.
- Citation counts would allow us to compare citation rates between deposited research papers and non deposited papers. This does however depend upon us acquiring information about citation patterns for University of Cambridge researchers, information which is available, but costly.

Key challenges faced

For the University Library part of the team the main future challenge lies in engaging departments to deposit in DSpace. As mentioned above we are starting up a promotional campaign where we initially will be contacting Heads of Departments following that up with training and general support.

Another future challenge DSpace@Cambridge is faced with is that of digital preservation of the deposited content. Preservation is one of our key service offerings, but as is well known the process of conducting preservation are still in development. For now we have committed to the following:

- Items will be migrated to new file formats where necessary.
- It may not be possible to guarantee the readability of unusual or proprietary file formats.
- The repository stores and maintains its files according to current best practice.
- The original file will be retained for all items, in addition to any upgraded formats.

We also plan to develop a preservation strategy for DSpace@Cambridge, this we plan to do together with the University Library’s new Digitisation and Digital Preservation Specialist, a new University Library position planned to be advertised shortly.

Important unresolved issues

DSpace@Cambridge is currently suffering from a multitude of technical problems, most of which are related to the large number of items stored in the repository which are related to the DSpace software’s inability to handle large collections. We hope that a number of these may be resolved by the next DSpace release in addition to us maintaining local “workarounds”. If our problems continue over a longer period of time we may have to consider using another software platform for DSpace@Cambridge.

The handling of rights is another important issue. An example is repository deposits of research papers where the process of clearing copyright for self archiving is so cumbersome that it ends up as a threshold for depositors. We try to improve this by
creating awareness amongst the researchers on which online resources to use, e.g. SHERPA/RoMEO\textsuperscript{2}, but at the end of the day we know that many people do not deposit articles to DSpace@Cambridge because of the above described obstacles.

\textsuperscript{2} http://www.sherpa.ac.uk/romeo.php