

ICT Forum, Hilary Term 2011

Report from the Bodleian Libraries

SOLO

The upgrade to version 3 in Autumn 2010 brought significant functional, interface and performance enhancements. Otherwise, development effort has centred on a cluster of enrichment services. The recommender service, bX, based on data-mining journal article usage, will suggest articles of possible interest to researchers. In addition, Nearly 800,000 records in SOLO now include tables of contents and/or summary information from Nielsen BookData. This enrichment data is searchable, which is especially valuable for works with individually authored chapters. For example, a search for "Cyndi Banks" in SOLO has eight results - four of which would not be found without the enrichment data. Chapter titles and book summaries will also contain terms that are not in the title or subject headings. For example, a search for "kloosterman sums" has six results, only one of which would be found without enrichment. Being able to view tables of contents and summaries will also help readers in deciding whether they wish to request books from closed stacks.

In the coming months, we shall be working on the changes behind the scenes that will enable SOLO to interoperate with the new ILS supporting OLIS.

OLIS: The OLIS Section

The OLIS Section continues to support the current OLIS while forming part of the new ILS implementation & training team and of the BSFIS support team. We are also the programming and processing support unit for the Inventory Control (i.e. barcoding) Project, a major undertaking which updates both the OLIS database and the Depository control data while all Bodleian and RSL closed access stock is barcoded and, for the first time since the 17th Century, inventoried.

OLIS: The New ILS Project

Interim acceptance of the new ILS was provided to the vendor (Ex Libris) on the 17th December 2011 having successfully tested that:

- the data can be successfully migrated from the current system
- the software enhancements to support stack requests are complete and correct
- the system supports the workflows necessary to manage the libraries.

Though acceptance of the system is an important milestone, there is still much work to be done including:

- further ILS configuration to support each individual library's set-up
- further development on system interfaces
- preparation for the final test data migration in April
- preparation of training materials for an intensive period of staff training from the end of April through to August

- final enhancements to the hardware infrastructure to fully support the requirements of a live service

The project continues to go well and we remain on track to go live in 18 July 2011.

BSFIS Implementation Project

With the staff training complete and the Libraries' Book Storage Facility at Swindon now fully operational, the implementation project is at an end. As of 7 March, 2,305,970 books Friday 14 January, have been ingested into the BSF.

ORA (Oxford University Research Archive)

The new ORA Developer has been familiarizing herself with the system software and architecture. The system is now more stable than it has been in the recent past. We are investigating use of Hydra/Hydrangea (<https://wiki.duraspace.org/display/hydra/Hydrangea>), a new OS software to replace the unsatisfactory VALET software and which will be used for ingest, editing and other workflow processes. Use of a top level domain name for ORA has been approved (ora.ox.ac.uk), though it is not yet deployed. This is important in terms of maintaining persistent URLs for the long-term and also the perception of ORA as a central university service, managed and maintained by the Bodleian Libraries. The top level domain name is set up for DataBank and work is progressing on DataBank for the Admiral project including assigning DOIs to datasets.

Metadata and full text has been imported from the University of Kent repository for an academic who has recently joined Oxford. The bibliographic records were obtained in JSON format and the metadata mapped to ORA. The base code to convert metadata into MODS and DC has been written. This has provided valuable experience for the next tranche of imports from local Oxford ePrints repositories (eg Economics and Maths).

We have laid the former Oxford EPrints service to rest. Links are now redirected to ORA and the service has been archived. Whilst attending to this, we also added improved wording for a new 404 error message. ORA is now being harvested by DRIVER (Digital Repository Infrastructure Vision for European Research) www.driver-repository.eu/.

The vocab.ox.ac.uk user admin interface is nearing completion. User testing is planned for end of January.

ORA Help & Information website (www.bodleian.ox.ac.uk/ora) has been transferred to the new CMS. Information about ORA has also been added to the Bodleian Libraries LibGuides. To address problems of sensitive and copyright content in theses, we will be advising students on the merits of using a separate appendix for such material. It is clear that many theses contain large amounts of such content which has proved difficult to deal with both for authors and ORA staff. This solution is intended to simplify procedures. A presentation to library staff by PRAS about Symplectic has been planned for 25th January. Work is in progress to import Symplectic data to ORA. Two new research institutes have expressed a strong desire to include their publications in ORA: Reuters Institute for the Study of Journalism and the Oxford Byzantine Research Centre.

Oxford Digital Library

In addition to their everyday tasks, staff working on EEBO-TCP Phase II have been actively engaging with colleagues at JISC to assist with the development of an "Historic Books Platform" (to be based at MIMAS, in Manchester). This platform will provide users with a single interface to cross-search the holdings of EEBO, ECCO, and the British Library's public domain digital publications (e.g. newspapers, and books digitized for the Open Content Alliance). We also collaborated with JISC Collections and MIMAS on a proposal to develop the technical infrastructure required to enable academic users of the Historic Books Platform to enhance and tag the underlying OCR and SGML/XML texts; the proposal was submitted to JISC before Christmas, and the outcome should be known by the end of Q1'11.

The JISC-funded "Mapping Crime" project (<http://www.jisc.ac.uk/whatwedo/programmes/digitisation/econtent/mappingcrime.aspx>) is entering its final phase, and work has been scaled-back during a period of no-cost extension. The outputs of this project will be used to enhance the results of its predecessor, "The John Johnson Collection: An Archive of Printed Ephemera" (<http://johnjohnson.chadwyck.co.uk/home.do>) -- and will thus also be made freely available to UK-based users in educational institutions and public libraries. In recognition of the invaluable expertise that he has acquired, the Project Manager for Mapping Crime, David Tomkins, has been asked to take over the running of the Google Music Project (mentioned elsewhere in this report).

As contributing partners to the "Europeana Libraries Project", we were delighted to learn that this proposal has been approved by the EC under the ICT PSP forth call, and has entered the negotiations stage. This means that between January 2011 and December 2013 the Bodleian Libraries will be collaborating with 25 partner institutions from across the EU, to ensure that extensive metadata for selected digital collections will be integrated into, and searchable by, the tools and technologies that have been created by the European Library (Europeana).

Preliminary work on various other digitization projects has been ongoing. In particular, devising a workflow that will enable our plans to digitize up to 5,000 Oxford DPhil Theses to proceed in a timely manner, whilst also respecting IPR and third party rights, has presented some interesting challenges. We have also been scoping a series of related projects which, if successful, will result in the digitization of several major catalogues of Bodleian Library holdings that are currently only available in analogue form.

Our programme of outreach and engagement activities saw staff from the ODL representing Bodleian Libraries at such events as the recent LUNA UK Users Group meeting at Manchester University, and the concluding workshop of the AHRC-funded "e-Science and Ancient Documents" project. Presentations on our work were given at conferences hosted by CILIP and the Association of Independent Libraries, as well as a number of smaller workshops and events.

IT Strategy & Development Project Updates

ITSD has lost one of their database developers, whose contract we did not have the funding to renew. We have two full time staff—a database administrator and a system architect—beginning this term.

The EIDCSR <http://eidcsr.oucs.ox.ac.uk/> project (in collaboration with OUCS) has come to an end. Backend development for Vocab (<http://vocab.ox.ac.uk>) is nearly completed, with front end work still underway. The Google Books project is progressing with strategies for bulk downloads underway and a

small pilot using the Read and Note software to annotate the Google scans underway this term with several faculty members and students.

We have a new project, funded by Google, to scan sheet music and make it publicly available. The collections to be digitised are not currently catalogued, though, and this project will be in part an experiment in 'crowd-sourcing' the creation of metadata. We hope to work with the creators of the Zooniverse projects to find a suitable technical platform and workflow.

We currently have a bid out to tender with JISC. In partnership with the University of Cambridge, this project is to create a gateway to Islamic collections across the UK. A significant Board of Advisors and Contributors has been assembled from the across the UK. In addition to providing a point of focus to the otherwise hidden collections in Arabic scripts, this gateway would serve to showcase the metadata developments made in OCIMCO and encourage the adoption of the OCIMCO TEI format by other institutions.

Other projects underway include:

- Cultures of Knowledge www.culturesofknowledge.org
- OCIMCO <http://www.bodleian.ox.ac.uk/bodley/library/specialcollections/projects/ocimco>
- Blockbooks http://www.odl.ox.ac.uk/digitalimagelibrary/blockbooks_home.html (live)
- Bodleian Incunables
- ADMIRAL <http://imageweb.zoo.ox.ac.uk/wiki/index.php/ADMIRAL> (live)
- SUDAMIH <http://sudamih.oucs.ox.ac.uk/>
- Genizah <http://www.ochjs.ac.uk/news/newsletter/GenizahDigitizationProject.html>
- Repository Junction
http://edina.ac.uk/projects/Open_Access_Repository_Junction_summary.html
- PIRUS2 <http://www.cranfieldlibrary.cranfield.ac.uk/pirus2/tiki-index.php?page=pirus2>
- The BRII project website is still live (as requested by JISC) at <http://brii.bodleian.ox.ac.uk/>
- Vocab <http://vocab.ox.ac.uk>
- Medieval Libraries of Great Britain is now operational though with ongoing development
<http://www.history.ox.ac.uk/research/projects/MLGB3.htm>
- Impact (13th-15th Islamic Codices) has started and will be using the Cultures of Knowledge code.
- DMSTech - examining visualisation standards and tools for medieval manuscripts
<http://lib.stanford.edu/DMSTech>; http://www.stanford.edu/group/dmstech/cgi-bin/mediawiki/index.php/DMSTech:Community_Portal
-

Dave Price, SERS, 9 March 2011