

---

# Developing the Digital Research Repository at the University of St Andrews



Janet Aucock  
*Bibliographic Data Services  
Manager,  
University of St Andrews*  
Tel: 01334 462299  
E-mail: [ja@st-andrews.ac.uk](mailto:ja@st-andrews.ac.uk)

---

## INTRODUCTION

The move to develop this new service was primarily driven by St Andrews University Library's planning and policy. This clearly identified a need for improved service provision to the local, national and international research community by providing digital delivery mechanisms. The university's strategic planning with regard to making research outputs available to a wider audience was also a major driver to the development. The deposit and delivery of theses in electronic format was identified as the primary service aim of the University of St Andrews Digital Research Repository development.

Closely allied to this is the development of a pilot service to enable the deposit of full-text published research outputs from the university's academics. The drivers here are to develop the Digital Research Repository in parallel with the university's research publications database and information system. The aims are to establish good communication and practice in the centralised and coordinated development of both databases and to embed the deposit, storage and dissemination of research-profile data and publications into centralised workflows.

## THE DIGITAL RESEARCH REPOSITORY

### *Background*

The University of St Andrews Digital Research Repository<sup>1</sup> is the successor to the St Andrews eprints service, which was established as a test project under the JISC-funded HAIRST (Harvest-

ing Institutional Resources in Scotland) project.<sup>2</sup> The eprints service is due to be decommissioned in 2008/2009 and relevant content will migrate to the Digital Research Repository. The Digital Research Repository service began in late 2006 and up until now has been primarily focused upon the deposit and delivery of full-text electronic theses. An additional pilot service is currently in place to demonstrate the new functionality, which allows the full text of research papers to be submitted to the repository via the Research Expertise Database,<sup>3</sup> the university's publications database, which was intensively used during the latest research assessment exercise.

### *Technical information*

The Digital Research Repository uses DSpace software.<sup>4</sup> DSpace is one of the primary open-source solutions for accessing, managing and preserving scholarly works. There is a strong DSpace community commitment to development, with regular new software releases, and DSpace is becoming more formalised as a community and an organization. DSpace supports the Open Archives Initiative's Protocol for Metadata Harvesting (OAI-PMH) v2.0 as a data-provider, which exposes data in local repositories to a world-wide audience through standard search engines. Our installation of DSpace is managed on our behalf by the Scottish Digital Library Consortium (SDLC).<sup>5</sup> Hardware and technical and development expertise are provided by the SDLC as part of our contract with them. Administration of the system is carried out locally by repository staff based in the university library.

### *Structure and metadata*

The content structure for the Digital Research Repository mirrors the university's school and departmental structure and also that of the research centres and institutes. This allows content to be placed into the relevant thesis or research collection for each school and research centre. Content can be mapped across into relevant collections if a researcher's primary affiliation is with both a school and a research centre. Each community and collection in the DRR links out to the relevant institutional web pages and can be branded with school and departmental logos, text and so on. Related research centres and institutes are listed and linked to from each school collection page. The repository offers standard database search features and the metadata is created and stored in Dublin Core format. PDF file format is encouraged for both thesis and research output deposit. Permanent URLs are provided for all full-text deposits.

## **ELECTRONIC THESIS CONTENT AND WORKFLOW**

The deposit of theses in electronic format was mandated by the university from the academic session 2006/7. There are now almost 250 theses available. This policy change was made in the context of national and international developments to facilitate access to theses, and many other individual academic institutions have now adopted a similar policy. Postgraduate students deliver one hard copy and one CD copy of their final completed thesis to the university's academic management and support office. Once they are receipted and all the relevant administrative paperwork has been completed, the copies are passed on to repository staff in the library to complete the deposit process. Postgraduates are invited to register on the repository to provide metadata, including the abstract, and to grant the necessary licences to complete the deposit process.

### *Policy issues*

Policy issues that have been addressed during the development of the service are: restriction/embargo framework, updating of the Postgraduate Code of Practice, deposit and end-user licences, author copyright statements and third-party copyright, file formats and conversion of data to PDF. The major stakeholders in this process and the issues that had to be resolved are described in more detail in a poster presented to the ETD (Electronic Theses and Dissertations) conference held in Aberdeen in June 2008.<sup>6</sup>

### *Support, training and publicity*

The University's Gradskills programme is a primary delivery platform to train and inform postgraduate students.<sup>7</sup> Courses on 'How to deposit your electronic thesis' are held each semester and provide support for students on the actual deposit process and the preparations they need to make prior to deposit. Extra training is scheduled for the start of the session 2008/9 with the introduction of specific short courses on electronic theses and copyright which will be held in individual schools. These will be joint presentations by the repository manager and the University's copyright officer. Further support is provided by the comprehensive information, factsheets, relevant documents and guidelines provided on the library's thesis web pages,<sup>8</sup> with links to the relevant pages on the university's copyright site.<sup>9</sup>

### *Retrospective conversion and digitization of theses*

Local arrangements are now also in place for handling requests for the digitisation of pre-current

St Andrews theses, and workflows have been developed to deal with these requests. Full details of procedures and contact details are provided for users on the library's thesis web pages. For each request, the author of the thesis is asked for permission to digitise it and to make it available in the repository. We handle specific requests and in addition we are choosing to digitise the most heavily consulted older material. We are also subscribing to the EThOS (Electronic Theses Online Service) project<sup>10</sup> which is a national initiative:

'to deliver a fully operational, easily scaleable and financially viable prototype UK online electronic theses service ... to enable students, researchers and in fact anyone else with an interest in post graduate research theses material, to search and access, from the desktop, the full text, in secure format ...'.

St Andrews has already contributed about 40 of our most heavily used theses for digitisation as part of the project. Content will be delivered by the EThOS site which is scheduled for full launch in early 2009, with initial kick-start thesis content from UK institutions of around 20,000 full-text theses. Digitised copies will also be returned back to the local institution for delivery through the local repository.

#### **RESEARCH PUBLICATIONS: PUBLISHED RESEARCH OUTPUTS**

The full text of published and peer-reviewed research outputs can now be deposited into the Digital Research Repository via the Research Expertise Database (the university's publications database). This is still a pilot service and feedback is being sought from the academic community. Once a publication's bibliographical details are deposited in the Research Expertise Database the interface allows the upload of the full text and the abstract of the publication. The depositor is asked to identify the version of the full text (i.e. publisher's PDF or author's final version) in order to provide information that can assist repository administrators with the copyright clearance of the publication. Open access to the full text of publications in institutional repositories is very much governed by individual publishers' policies. The SHERPA service,<sup>11</sup> supported by JISC (Joint Information Systems Committee)<sup>12</sup> and RSP (Repositories Support Programme),<sup>13</sup> provides crucial services to assist with copyright clearance. The ROMEO database, which lists publishers' copyright policies with regard to self-archiving, provides a summary of permissions that are normally given as part of each publisher's copyright

transfer agreement.<sup>14</sup> Repository staff will check each submission against the ROMEO database and contact depositors if there are queries about their deposit. SHERPA also provides another useful service, JULIET,<sup>15</sup> which lists the research funders' open access policies and a summary of policies given by research funders as part of their grant awards.

A more detailed description of the aims of this pilot project and a description of the workflows can be found in the poster presentation I made to the 2008 Open Repositories Conference.<sup>16</sup>

#### **CONCLUSION AND FUTURE DEVELOPMENTS**

The electronic theses service has been successfully implemented and has transformed the previous delivery method of print copies of theses – stored in closed access in library stacks – to free online delivery in an open-access institutional repository. The success of the current service is very dependent on key partnerships which have been established between relevant agencies within the institution. Communication and establishment of workflows between the library and the academic management and support office are crucial, as are the partnerships between repository staff, copyright staff and supervisors, to provide support and training to postgraduate students. There have also been significant training and reskilling issues for university staff working closely with the service, and many opportunities for contact with other relevant projects and external support agencies. The process has been one of constant re-evaluation of institutional and user needs. Relevant institutional policies have been developed and the framework put in place to embed the process in institutional practice. This has resulted in the provision of a high-quality service.

We seek to encourage the reuse of Digital Research Repository data. For example, lists of current theses for specific schools in the university can now be linked to and can provide instant lists for the schools' web pages, with access to the full text. Links to the full text of research publications are retained in the Research Expertise Database and are used in publication lists, which can also be reused on the schools' web pages. The biology and modern languages staff pages are good examples of this practice.

The pilot project for published research outputs is still ongoing. This too has demonstrated a real need for key partnerships and effective communication between stakeholders. Strong links have

now been developed between the library, the business improvements unit and the university's research office. The aims of this project are still very centred on data reuse, dataflow and integration and the repurposing of data. Efficiency in gathering, updating and preserving the university's research information and staff publications is a primary goal. Acquiring the full text of published research outputs and making them available for open access is another focus. A more significant driver is the current emphasis on research assessment and the need for academic institutions to prepare effectively to support the next research assessment exercise, the Research Excellence Framework (REF).<sup>17</sup> Many current initiatives are highlighting the need for CRIS (Current Research Information Systems)-type systems<sup>18</sup> to be put into place at institutional and national levels. The future role of the Digital Research Repository will be to support, and to be interoperable with, any such system developments that may take place within the institution.

Another recent strand of national development has been a growing interest in the development of policies and services to support research datasets and the archival preservation of this material. The UKRDS (UK Research Data Service) project<sup>19</sup> will also impact on the development of the Digital Research Repository and the services it may provide in the future. Defining and redefining a content policy for the repository is an ongoing task which is influenced by both external and internal factors. Considerations for future content might include datasets, digital objects, content from digitisation projects and in-house journals.

The development of the University of St Andrews Digital Research Repository has already made significant changes in the scholarly communication process for all St Andrews research outputs, and current emphasis on the use, dissemination and preservation of these outputs indicates that its contribution will continue to be relevant. (Contact me, at the e-mail address above, for more information.)

#### REFERENCES

- 1 See <http://research-repository.st-andrews.ac.uk/>
- 2 HAIRST project website: <http://hairst.cdlr.strath.ac.uk/>
- 3 See <http://portal.st-andrews.ac.uk/research-expertise/>
- 4 DSpace home page: <http://www.dspace.org/>
- 5 See <http://www.nls.uk/sdlc/>
- 6 See Janet Aucock, 'Electronic theses at the University of St Andrews: institutional infrastructure, policy and support to establish an electronic theses service', poster presentation from ETD 2008: 'Spreading the light', 11th International Symposium on Electronic Theses and Dissertations, Robert Gordon University, Aberdeen, June 2008 (<http://hdl.handle.net/10023/513>)
- 7 Gradskills programme: <http://www.st-andrews.ac.uk/GRADskills/>
- 8 University of St Andrews library web pages: <http://www.standrews.ac.uk/library/Collections/Theses/>
- 9 See <https://www.st-andrews.ac.uk/students/policy/Copyright/copyrightthesis/>
- 10 See <http://www.ethos.ac.uk/> and live service at <http://ethos.bl.uk/>
- 11 <http://www.sherpa.ac.uk/>
- 12 <http://www.jisc.ac.uk/>
- 13 <http://www.rsp.ac.uk/>
- 14 See <http://www.sherpa.ac.uk/romeo/>
- 15 See <http://www.sherpa.ac.uk/juliet/index.php>
- 16 Janet Aucock, 'Project to integrate research publication deposit through a centralised deposit workflow: optimising the relationship and functionality of the St Andrews Research Expertise Database and the St Andrews Digital Research Repository', Third International Conference on Open Repositories, 1-4 April 2008, Southampton, United Kingdom, available at <http://pubs.or08.ecs.soton.ac.uk/65/>
- 17 See <http://www.hefce.ac.uk/Research/ref/>
- 18 See <http://www.eurocris.org/>
- 19 See <http://www.ukrds.ac.uk/>