JISC

JISC Final Report

The I-WIRE Project Final Report by Louise Harrington and Scott Hill

Louise Harrington: HarringtonLE1@cf.ac.uk

Scott Hill: HillSC@cf.ac.uk

28 February 2011

Project Acronym: The I-WIRE Project
Version: 1.0
Contact: Scott Hill
Date: 28 February 2011

Contents

Project Acronym: The I-WIRE Project
Version: 1.0
Contact: Scott Hill
Date: 28 February 2011

Acknowledgements

Executive Summary

The I-WIRE Project's mandate was to establish a quick deposit tool that would enable academic and administrative staff to deposit metadata and full text easily and swiftly into Cardiff University's Institutional Repository, Online Research @ Cardiff (ORCA), with the additional aim to make ORCA a central function in the day to day research management processes at Cardiff. The project took a user-centric approach to gathering requirements; identifying the relevant stakeholders both within academic schools and support divisions, arranging individual interviews with research administrators, and group discussions with academic authors. After the completion of the requirements capture, the project team organised a series of group sessions to analyse each documented requirement and examine its feasibility within the scope of the project. During the design and development phase, we gave regular demonstrations of the portlet to academics and research administrators, and their feedback contributed towards the enhanced development of the design. Alpha testing began in December 2010 with members of Information Services staff, and continued in January 2011 with volunteer testers from academic schools. Beta testing with selected schools will begin in March 2011, with a gradual rollout over the academic year.

The project's main output is a portal-based quick deposit tool, fully integrated into Cardiff University's Modern Working Environment, with three choices of deposit route: quick deposit, DOI deposit and Web of Science deposit. All three were developed from user needs and suggestions, as was the Selected Publications feature. What the project has developed is substantially more than a simple 'quick deposit tool', and is a step towards an institutional publication management system. Advocacy and embedding are key steps in ensuring the future sustainability and development of this service.

Project Acronym: The I-WIRE Project
Version: 1.0
Contact: Scott Hill
Date: 28 February 2011

Background

In 2008, Cardiff University completed a project to upgrade its Institutional Repository ORCA (Online Research @ Cardiff, which uses EPrints 3 software) to a fully supported service containing all the University's RAE publication data. The future role of the repository, which is seen as central to the management of all the University's research publications data, has been given added impetus and support with the University Research Committee's recent decision to approve the mandatory deposit of new publications in ORCA, "[s]ubject to detailed consideration as to the means by which staff up-load data to the repository". The Repository will be the central database of research output, and will facilitate the re-use of the content in a variety of contexts.

Aims and Objectives

The aim and objectives of the I-WIRE Project were to improve the existing repository from a "technical or process driven perspective" in the following ways:

- To provide a system that supports the management of research publication data that is agile, efficient and flexible, by providing all of the relevant stakeholders involved with access to an innovative toolset that will establish a quick and comprehensive submissions process and metadata creation workflow for the institutional repository;
- To use technology being developed in Cardiff University's MWE (Modern Working Environment) to achieve seamless interaction between the institutional repository and other relevant systems and processes in the research portal;
- To develop a system that will enable the content and metadata in the repository to be re-purposed and re-used to support the range of processes and systems used in different stages of the research lifecycle, and in various processes in the University Administration.

In meeting these objectives, the project would aim to embed the repository into the heart of the day-to-day research management process at Cardiff University.

Methodology

Project Control

We set up bi-weekly meetings of the core project team to discuss issues, plan next steps and co-ordinate activities. A project log that tracked risks and issues, lessons learned and feedback was also established. The Project Manager and Project Support Officer also posted regularly on the I-WIRE Project blog. A Wiki and team place for discussion and storing and sharing files with rest of the project group was also created.

Stakeholder Management

The project adopted a user-centric approach from the start, and began by identifying the communities that would be stakeholders in the project's outcomes.

The following communities were identified by the project team as current or potential users of ORCA, and therefore key Stakeholders in the I-WIRE project:

- Researchers and academic authors in all Schools through their need to make their research outputs available on an Open Access basis.
- Research Administrators in all Schools through their need to provide reports on the School's research outputs.
- The Repository Manager and repository team through their responsibility to ensure the quality of metadata in the repository.

Representatives of different communities (University Library Service, IT, research administrators, and academic researchers) formed the project management group, which the core project team reported to on a bi-monthly basis.

Communications

The project team arranged for a series of communications and agenda items at related meetings and committees, both within ULS and in the wider university community, to ensure awareness of the project and its objectives, and to seek support from key stakeholders where required.

User Needs

We gathered user requirements by arranging individual interviews, using a structured questionnaire, with research administrators from a wide cross-section of the 29 academic schools and departments. For help in doing this, we utilised existing structures. The subject librarians in particular played an important role in engaging stakeholders and helped the project team gain rapid and direct access to researchers, research directors and heads of school in some instances.

We also arranged group sessions with academics using LEAN methodology; these sessions captured and documented the current processes that authors follow for the management of their publications, along with associated issues. The process was looked at end-to-end from identification of research opportunities to production of reports for Schools management, in order to identify any opportunities for linking the deposit process with other processes such as Performance Management. This also gave us a baseline as-is process from which to measure the outcomes of the project.

The group then agreed and documented an enhanced and simplified future state process that the same set of authors agreed would encourage them to self-deposit in the repository, along with other requirements that they may have.

Analysis and Design

After the completion of the requirements capture phase, the project team held group sessions that analysed results of the requirements gathering, accepted and rejected findings. These were then written up as a set of user stories. The Design phase was broken into number of iterations, grouped by theme of user story. We analysed each requirement that had arisen from the interviews and group sessions to see what was feasible within the scope of the project.

Toolkit Development

During design and development, members of the project team gave demonstrations to academics and research administrators for feedback during the design stage. Early prototyping of the quick deposit tool helped us get a head start on development and feedback, and to prove the new technology.

We started Alpha testing with INSRV staff in Dec 2010 and with testers from other academic schools in Jan 2011. This allowed for a managed approach to service introduction. An online questionnaire was written to capture feedback, and other methods of feedback (i.e. email, in person) were stored on the Wiki. Beta testing is planned to start in March 2011 and will continue beyond the formal closure of the I-WIRE project, and will be managed by the recently established University Library Service Repository Working Group.

An alternative approach would have been to develop the tool in smaller iterations but it would have been difficult to secure that amount of time from researchers, due to the demand on their time. The development phase took longer than initially planned, as a crucial part of the portlet which was an outcome of the user needs analysis took some time to get right and necessitated a consultancy day with EPrints.

Implementation

Due to the late recruitment for the project manager and the project officer, the project did not get off the ground until October 2009, 6 months later than planned. Requirements gathering did not start until then and was compressed into two months rather than the nine originally planned. We aimed for as wide a cross section of schools, disciplines and support divisions as possible although it was not possible to interview from every school and division.

We started the user needs analysis phase by mapping and documenting the current state research management process, with the assistance of Cardiff University Lean team. We soon learned that processes were different enough in each school (Cardiff University had 29 schools at the time), with local variations, to make it challenging to apply LEAN thinking to a single process. However, we gained concurrence over a number of sessions on a set of common research management related tasks that could potentially be simplified by the project.

An early requirement for the project was to include a unique person identifier in the repository data that would enable other systems are Cardiff University to accurately retrieve publication lists for staff and students. It was agreed that an email address not good enough as they can be reused for multiple people and can also change. Analysis of this requirement led to us documenting the roles around repositories, and the scenarios associated with those roles, giving us a set of business rules that tested our implementation of the requirement.

The repository and portal technologies were in place before the start of the project and we had to work with what we had: the EPrints repository and the IBM WebSphere portal.

The Lead Developer attended IBM WebSphere training for the portal. While there was a requirement to fully utilise the suite of IBM applications, the Process Server element of the suite was unproven technology for Cardiff University and therefore an early decision was lobbied for to de-scope it and reduce the risk to the project's timescales.

We were faced with an integration challenge early on due to the absence of an EPrints SOAP interface. Analysis of options led us to using a combination of the SWORD deposit protocol, and the EPrints API and REST interface. Two EPrints consultancy days were needed to work out solutions for some of the more challenging requirements: the ability for a user to mark their selected publications, and a Cardiff author-only browse.

Outputs and Results

The main output of the I-WIRE Project is fully integrated portal-based quick deposit tool that presents three choices of deposit route:

- Quick Deposit: minimal data is captured from the user.
- DOI Deposit: publication data is retrieved from the CrossRef service using a DOI.
- Web of Science: publication data is retrieved from Web of Science using a simple name search using the API developed by the Bibliosight project.

The tool also shows the logged in user their publications and lets them mark their selected items, and surfaces the EPrints search and browse functions. See screen shots in appendix.

To achieve the above, the portal server has interfaces with the university's LDAP service, the local implementation of SWORD and EPrints, SHERPA's RoMEO service (via a static file), CrossRef and Web of Science.

After requirements gathering, we introduced a new 'selected publications' feature, meaning the portal is now more than just a quick deposit tool, and providing a building block from which a publication management system can be developed in the future.

Other outputs are: a documented current state research management process; documented repository roles and scenarios; documented user requirements; a formal service support structure in place, including Service Level Objective, a documented design document and a set of test stories that can be used during upgrades.

We are also looking at sustainability of the project's outcomes, and this has triggered a separate Advocacy & Embedding project that has senior management approval within Information Services at Cardiff University.

Outcomes

We've been successful in making deposit quick! We have had positive feedback about the quick deposit tool, with academic testers saying they will use it. The deposit process is much simpler for researchers who are going to have to use this, though there could be delegated deposit in research offices. The trade off is the amount of effort in reviewing new deposits and adding to the minimal publication data captured during deposit, and needs resource to be looked at.

We believe we have reached the limits of the EPrints workflow in terms of the support it provides for a medium sized team working on a single review queue, and tracking individual items through various complex scenarios. To manage this, we are defining manual workflows with the team and are considering an administrator of the queue to ensure items are not double handled.

We've introduced local customisations to our implementation of EPrints that need to be carefully tracked via a source code change management system. We have possibly customised it beyond the level we are comfortable with – it is no longer 'out of the box' – as local knowledge and specialised skill around the customisations must be retained in order to support and further develop the service.

Embedding an integrated service into the portal is a very visible win, and a first for Information Services at Cardiff University. We are setting the standard for future portalised services.

During the project's life, the repository has gained increased recognition as a key tool in preparation for the REF, and is continually mentioned in REF related working groups and committees. The detail of how the repository will play its role is yet to be finalised but certainly the I-WIRE project's outputs will place another piece of the jigsaw in developing the repository to support researchers and the university as a whole.

Great news for the project is that the school of Engineering are going to Beta test the project's quick deposit tool and integrate their own research management system with the repository. Discussions with the school of Medicine are ongoing on their requirement to connect their implementation of Symplectic with the repository. Beta testing is now being started, and roll out to remainder of schools will be over-seen by the Repository Working Group.

The Cardiff Law School has recently expressed an interest in linking their own in-house developed publication management forms to the repository, to ensure new deposits are made available in the

repository while having minimal impact on the users' current process. The school has dedicated IT resource and will need to understand the scale of the technical challenge encountered by the I-WIRE project team, and the requirement to retain staff with the appropriate knowledge and skills in order to maintain the connection in the future, particularly as the repository software is upgraded.

Conclusions

Institutions need to carefully consider their repository workflows, and understand the balance that must be made between a simplified deposit tool for researchers and the amount of effort needed in enhancing the item before making it available in the repository. When a repository functions as a publication management system for the institution, data quality is all the more important.

Dedicated resource and stakeholder management are key to a successful project outcome. It was good for the I-WIRE project to have representatives of each key stakeholder community on the Project Management Group to get things moving when needed.

Controlling scope at each stage is critical. User stories help achieve this while keeping the perspective based on the user.

Scholarly communications is a complex jigsaw. There is no silver bullet, not even a quick deposit tool. You need lots of advocacy and some very simple but direct messages that target the differing needs of different members of the community, and these messages need to attempt to tackle copyright which is as big an area of concern and confusion at the end of the project as when the project started.

Not all academics are engaged with Open Access, and some of them are unlikely to ever be engaged, so you need to target the easy wins and not beat yourself up.

Implications

The quick deposit tool has been developed using portal technology standards and can be implemented by other institutions. The code is available on request. It uses SWORD for deposit, with a few amendments for authentication. Getting data out of the repository uses the EPrints API and REST interfaces, so this would need amending for other repository platforms.

A simplified deposit tool could be implemented in stand-alone repository web interfaces but consideration must be made to resource levels if deposit is made very simple for the users.

We have enough 'nice to have' requirements left over for another major project of work.
There are further integration opportunities for the repository, such as integrating with other research management systems at the institution; with external services such as Researcher ID, Scopus, and arXiv.

References

SWORD project. http://www.crosswire.org/sword/index.jsp

Bibliosight (http://bibliosightnews.wordpress.com/) for Web of Science API.

Project Acronym: The I-WIRE Project
Version: 1.0
Contact: Scott Hill
Date: 28 February 2011

Appendixes

A: Screenshots of Manage My Publications

Figure 1: Quick Deposit

Figure 2: DOI deposit

Manage My Publications

| Quick Deposit | DOI Deposit | Web of Science | My Publications | Search & Browse |

Use this form to retrieve your publication data from Web of Science. You can then deposit your publications by ticking the check box against each, adding a file of the full text to each, and clicking the Submit button once after selecting all your publications.

* Enter Name:	Weightman A	?	Restrict to :	Select	?
Date From:		?	Date to:		?
Cardiff University Records Only:	☑ ?				

Get Records

* indicates mandatory fields.

| **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Next »

☑ **Eubacterium minutum is an earlier synonym of Eubacterium tardum and has priority**
INTERNATIONAL JOURNAL OF SYSTEMATIC BACTERIOLOGY (10 1999)
Authors: WG Wade, J Downes, MA Munson, AJ Weightman,
Full Text: C:\Documents and S Browse…

☑ **Haloalkane-utilizing Rhodococcus strains isolated from geographically distinct locations possess a highly conserved gene cluster encoding haloalkane catabolism**
JOURNAL OF BACTERIOLOGY (05 2000)
Authors: GJ Poelarends, M Zandstra, T Bosma, LA Kulakov, MJ Larkin, JR Marchesi, AJ Weightman, DB Janssen,
Full Text: Browse…

☐ **Bulleidia extructa gen. nov., sp nov., isolated from the oval cavity**
INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY (05 2000)
Authors: J Downes, B Olsvik, SJ Hiom, DA Spratt, SL Cheeseman, I Olsen, AJ Weightman, WG Wade,
Full Text: Browse…

☐ **Modified primers facilitate rapid screening of 16S rRNA gene libraries**
BIOTECHNIQUES (07 2000)
Authors: JR Marchesi, AJ Weightman,
Full Text: Browse…

☐ **Depth variation of bacterial extracellular enzyme activity and population diversity in the northeastern North Atlantic Ocean**
DEEP-SEA RESEARCH PART II-TOPICAL STUDIES IN OCEANOGRAPHY (2001)
Authors: KE Davey, RR Kirby, CM Turley, AJ Weightman, JC Fry,
Full Text: Browse…

Upload Selected Records

Figure 3: Web of Science

Figure 4: My Publications

Figure 5: Search and Browse

B: Requirements Questionnaire

I-WIRE User Needs Analysis Questionnaire for Research Administrators

An overview of the I-WIRE Project and how it aligns with other Repository projects will be provided at the beginning of each interview.

Responses will be noted below each question.

1	Background and usage of ORCA	
1.1	"Managing research publications" – What does this mean to you and how does it affect your work? *Capture name and role.*	
1.2	What do you know about the Institutional Repository (IR/ORCA/Cardiff ePrints) and does it / could it play a part in your management of research publications?	
2	**What do you presently record about publications?**	
2.1	What sort of publications do you record? *Current Content Policy: The repository will accept any material that relates to resource output. Examples of material include:* • *Journal articles* • *Conference and workshop papers* • *Theses and dissertations* • *Unpublished reports and working papers* • *Books, chapters and sections* *Are there any requirements for output beyond conventional publications (grey literature)?*	
2.2	For what purpose – What is the primary purpose for recording and storing your publication data?	
2.3	Where do you store it and in what format?	
2.4	What details are recorded, and what is the minimum set of information that needs to be recorded? *Can we have a detailed list of the data structure and a copy of it.*	
2.5	Is there any publication data that you don't want made publicly available?	

	2.6	Do you submit your research publication data to any external databases or publisher websites? *What good and bad experiences?*	*Anticipate Academics only but worth asking*
3	**How and why do you do it?**		
	3.1	Describe the current workflow for entering data for new publications – the sequence of events leading up to it and the tasks involved, and who is responsible for it.	
	3.2	How often is this done and how long does it take?	
4	**How do you use your publication data?**		
	4.1	Where and when do you need to use your research publications data? *Example for Academics: Populating research grant applications.* *Example for Research Administrators: Producing reports for School Management.*	
	4.2	How do you pull it in, and into what?	
5	**Are there any problems with the way things are done presently?**		
	5.1	Is it time consuming?	
	5.2	Is there any repetition of tasks?	
	5.3	Do you need to use other tools or switch between applications and data sources?	
6	**What suggestions can you make for improving it?**		
	6.1	What do you want to be able to do that you can't do now?	
	6.2	If you could identify 3 key "value-added" features that would make you want to put your research publication data in the IR, what would they be? *E.g.:* • *Links from repository to staff/school web pages?* • *Alerts?* • *Repurpose and reuse data from the repository to populate research grant applications, create staff profiles, etc.* • *Citation data? (Authors)*	
	6.3	Are there any other points you would like to make?	

C: User Story Template

US001 Author Self Deposit - I-WIRE - MWE Wikis - Windows Internet Explorer

https://wikis.cf.ac.uk/confluence/display/iwire/US001+Author+Self+Deposit

uwic

8 Chang... | Digital inf... | Cardiff P... | http://or... | US00... ×

ID:	Title:	Priority:	Size:
US001	**Author Self Deposit**	**MUST**	**TBA**

As an... author

I want to... deposit my publications and the associated bibliographic data in the repository

So that... I can manage my publications data in one place, enabling re-use of the data and also increasing the visibility of my publications.

I - Independent
N - Negotiable
V - Valuable
E - Estimable
S - Small
T - Testable

Acceptance Criteria: | *How do I know when I have got it? How many do I want? How long can I wait?*

1. Can I deposit my publications into the repository via the MWE Portal?
2. Can I add metadata that describes my publication?
   a. Full metadata: Publication type, Title, Abstract, Authors, Version, Status, Publisher, School, Funding Body, Key words.
   b. Type: As currently defined in ORCA
   c. Status: Submitted, Approved, Published.
3. Do I have the option to upload the full text in the format of my choice?
   a. Formats: DOC, PDF, TXT
4. Can I see the publisher's copyright policy during full text deposit?
5. Do I have the option to provide a link to the full text on another web site?
6. Can I deposit with a minimal set of data?
   a. Minimal metedata: Publication Type, Title, Authors, Status, Publisher, School.
7. Can I deposit with just the dOI where available?
8. Are the names of Cardiff people auto-completed for me, with a drop down that enables selection of the right person?
9. Is the School auto-populated for Cardiff authors?
10. Is the Journal Title, Publisher and ISSN auto-populated?
11. Does the deposit include the IdMan value for the depositor and Cardiff authors?
12. Can I import my publication data from other databases?
13. Can I deposit full text from Symplectic?
   a. Web of Science
14. Can I input an embargo date?
15. Is the metadata checked automatically during input for errors?
16. Does the system provide meaningful error messages?
17. Can the system support volumes of 1 deposit every 5 minutes during business hours?
18. Can the system store 6000 publications per year with an average file size of 5MB per publication?
19. Can I use the system at any time of the day or night, on any day of the year?
20. Are the back-end systems integrated and working in real-time, and transparent to the user?
21. Is the end-to-end solution accepted into service with clearly defined Service Level Agreements?

Have you thought about:
S - Security
P - Performance
A - Availability
M - Monitor/Measure

Origin: Research Administrators, Authors

Done | Trusted sites | 100%