



BRII (Building the Research Information Infrastructure) Project: Final Report

Project Acronym	BRII		
Project Title	Building the Research Information Infrastructure		
Start Date	Oct 2008	End Date	March 2010
Lead Institution	University of Oxford		
Project Director	Sally Rumsey		
Project Manager & contact details	Sally Rumsey 01865 283860 sally.rumsey@bodleian.ox.ac.uk		
Partner Institutions	n/a		
Project Web URL	http://brii.bodleian.ox.ac.uk/		
Programme Name (and number)	Institutional Innovation 7/08 Call II		
Programme Manager	Lawrie Phipps		

Document Title	Final report		
Reporting Period	n/a		
Author(s) & project role	Sally Rumsey (Project Manager)		
Date	31 March 2010	Filename	
URL	http://brii.bodleian.ox.ac.uk/		
Access	<input checked="" type="checkbox"/> Project and JISC internal		<input checked="" type="checkbox"/> General dissemination

Version	Date	Comments
1.0	4 Feb 2010	First draft
2.0	3 March 2010	Final draft
3.0	31 March 2010	Final version



BRII (Building the Research Information Infrastructure) Project



Final Report

Sally Rumsey, Cecilia Loureiro-Koechlin, Anne Bowtell

March 2010

Table of Contents

Executive Summary	3
Background	3
Aims and Objectives	4
Methodology	4
Implementation	6
Outputs and Results	6
Outcomes	9
Conclusions	15
Implications	15
Appendixes & Glossary	16

Acknowledgements

The project was generously funded by JISC as part of the Institutional Innovation (strand 2) programme.

Contact sally.rumsey@bodleian.ox.ac.uk

Executive Summary

The aim of the project is to enable efficient sharing of research management data (ie data *about* research at the University of Oxford) from a selection of existing data stores and to create exemplar services that disseminate and re-use that data using a lightweight solution based on semantic web technologies

Objectives

1. Negotiate stakeholder input, collaboration and buy-in from a small number of internal data providers/creators, a selection of end-users and management and other stakeholders
2. Identify data sources and content for inclusion and re-use within the research information infrastructure (RII)
3. Build RII infrastructure comprising 'Combine Harvester' technology¹, instance of Fedora and web services and semantic web services
4. Provide RDF ontologies and taxonomies to define and categorise research data
5. Provide a selection of web-based applications to disseminate the research data
6. Create a University *Blue Pages* as an example of a web-based application that re-uses the data
7. Demonstrate embedding and future sustainability by data providers/creators, end users and key stakeholders

Overall approach

The project took the form of dual strands:

- 1) Stakeholder and user needs analysis and user testing. Consulting with potential users and testing pilot web interface formed the basis of the development.
- 2) Technical and web interface development informed by 1). The development was iterative i.e. it took place incrementally, each stage being informed by user feedback.

Findings

The stakeholder analysis revealed the practice and motivations of users who were categorised into three broad groups: individual researchers; administrators; disseminators and strategists. Generally users were in favour of the basic premise of the registry. Over the period of the project, conflicting attitudes towards data aggregation, dissemination and re-use became clear. Data sources were categorised with respect to their ease of harvesting and synchronising for the registry. A variety of types of data stores emerged – some were easy to harvest whilst others presented considerable problems for synchronising.

Achievements

The registry contains a good quantity of data from a number of different sources. The technical design of the registry has generated considerable interest within the developer and research information communities. There have been additional outputs to those predicted including the vocab.ox.ac.uk site for vocabularies. The stakeholder analysis provides valuable insight into stakeholder motivations and practice. The Blue Pages interface is easy to use and test users like the service.

Conclusions

Overall the project has been successful but not without problems. It met all its aims and objectives and generated much interest both internally and externally. Additional software outputs were created and released. The original intention was to make the Blue Pages freely available: consultation and feedback resulted in the resource remaining limited to within the Oxford only domain. The work of the project has been presented to high level committees of the University and a final decision regarding its future adoption as a central service has not yet been reached. The Medical Sciences Division 'themed' website (under construction) and the Blue Pages provide valuable examples of use of registry data.

Background

The project was prompted by the need to simplify deposit in ORA (Oxford University Research Archive) by obtaining existing data (author name, affiliation etc) to populate records. It transpired that obtaining this 'common' data was not an easy task at Oxford. A process to harvest the data from existing sources

¹ Created as part of the JISC funded BID (Bridging the Interoperability Divide) project

was created and plans for a 'registry' to store the data as part of OULS² digital asset management system (DAMS) was drawn up. It became clear that the data and services the registry would provide could be of great benefit to the wider Collegiate University. This realisation coincided with a call for proposals from the JISC under its 'Institutional Innovation' programme. The registry directly supports current University strategy.

The registry design uses the existing Oxford institutional repository and Oxford DAMS architecture, including a digital object store that employs semantic web technologies.

The registry enables users to:

- Re-use existing data thereby supporting increased efficiency and data consistency
- Share data using standard export methods
- Create websites that transcend department boundaries (eg cross-disciplinary)
- Search across the University's research activities
- Discover others undertaking similar or related research
- Find out who is doing what research, where and with whom (collaborators)

The design of the registry recognises and respects the devolved and disparate organisational structure of the University. Local systems retain control of their data whilst provided with a new tool that allows them to recombine and re-use their own and other Oxford data in new ways. Data complexity such as use of different forms of people's name can be dealt with and related to each other, although co-referencing techniques to ascertain if one form of name refers to the same person as another remains work in progress.

In the wider context, the technologies support data exchange with other institutions (eg. with universities and funders). This could be important for future statutory reporting.

Aims and Objectives

See above. The aims and objectives were not amended during the project.

Methodology

Overall approach

The project took the form of dual strands:

- 1) Stakeholder & user needs analysis and user testing. The basis of the development was consultation with potential users and testing pilot web interface.
- 2) Technical and web interface development was informed by 1). The development was iterative i.e. it took place incrementally, each stage being informed by user feedback.

- An iterative approach was adopted so that lessons could be learned and improvements incorporated as the project progressed. Data sources were added one by one with technical development taking place at each stage. It reduced the risk of developing something which later on required significant changes when presented to users. In order to encourage user buy-in, the stakeholder analysis attempted to discover user motivations, drivers and data sharing practices. By taking these aspects into account, the new service would be more likely to be viewed favourably.
- A comprehensive strategy was devised to interview individuals for the stakeholder analysis. They represented all dependencies of the University, including at least one department from each academic division, three offices with the University Administration and Services, two University Colleges and two other administrative areas. Additionally, staff representing a variety of roles were chosen for interviews, ranging from researchers, who are directly involved with research activities, to administrative staff who support researchers; and from departmental and divisional officers to University level administrators who need to oversee research activities across different areas. With this strategy we made sure that we were collecting information which accounted as much as possible for the variety of Oxford's administrative and research structures and cultures.
- The user testing of the Blue Pages was designed in line with the agile development strategy of the project. Tests started at the end of a development iteration and ran until feedback reached a saturation point. That is, when similar issues were being reported consistently. Testers from a

² Oxford University Library Services (OULS) was renamed The Bodleian Libraries on 2 March 2010

number of academic and administrative units were selected to represent all the variety of uses and perspectives possible. Tests involved a significant investment of time as the majority of them were carried out in testers' offices. Different scripts were designed for each iteration which reflected the changes in the tool.

- Balsamiq software (www.balsamiq.com/) was used for the first prototype mock-up of the Blue Pages interface. This worked extremely well – test users focused on the features and functions, and not on the art and design of the interface which is exactly what was required at that early stage.
- Two members of library staff were identified who were willing to test the user interface for accessibility, the first using voice recognition software and the second specialist screen reader software. This was valuable for the developers to ensure the website was accessible and to identify problems that need further attention.
- Library staff worked collaboratively with the Web Manager in the Medical Sciences Division (MSD) and who was using similar methods. It was beneficial to both parties to pool resources and expertise.
- The first data to be harvested was that which was defined as 'absolutely vital' (organisational units) followed by some 'quick wins.' Some data were included as a one-off to demonstrate the proof of concept. These decisions were taken to ensure that we had sufficient data to demonstrate the service and so that types of data did not become skewed towards a single type.
- This project has enabled technical developers to form valuable links with colleagues at the Universities of Bristol and Southampton. This has resulted in sharing expertise and solutions to problems and has proved to be valuable.
- The MSD, working with an external consultant, collaborated with the Oxford University Research Archive in undertaking further research on available ontologies for researchers and researcher outputs, resulting in a revised version of the original RES vocabulary already in use within MSD websites.
- To address the dependency of most MSD outputs (themed website etc) on available data and a stable version of the ORA API, a phased approach was taken with development (again with external consultancy). The generic framework for connecting the Plone CMS (Content Management System) to external content was built and released as Open Source. This also facilitated proof of concept work that could be used to inform ongoing requirements and could also be used as demonstrators for the Project Board and other interested parties.
- Standard semantic web technologies are used for the system.
- The technical system in use is designed to be scaleable. It has been created to be able to deal with the vast digital collections of the Bodleian Libraries.

Implementation

When the BRIL project started very little was known about the needs for research activity data across the University. Because of this the stakeholder analysis was designed as an exploratory study to try to uncover as many patterns of beliefs and behaviours as possible. The stakeholder analysis involved a wide range of areas within the University. Many emails were sent to academics and administrators across the University. Although the response rate was low, a good enough number of respondents agreed to participate in the interviews. Respondents were also helpful by suggesting further contacts for interview.

Interviews were unstructured. Questions were asked in the following three areas of enquiry: creation and management of research activity data, use of research activity data, issues and future uses of research activity data. Questions varied according to the respondents' roles and areas of expertise; sometimes covering only one or two of the areas mentioned above. Data from the interviews were transcribed and analysed soon after the interviews. These data include information which describes the stakeholders and their interests as well as their user needs. The result of this analysis can be found in the Stakeholder Analysis document³.

User tests started once the stakeholder analysis was finished. A mockup version of the Blue Pages was designed using feedback from the study. This mockup version was used in the first round of tests. Subsequent rounds used a live version of the Blue Pages. Additions and corrections in the Blue Pages

³ Available at <http://ora.ouls.ox.ac.uk/objects/uuid%3A1df69991-cd37-445b-a4c7-3573ce80c36e>

were made after each round. Work on harvesting sources of research activity data was carried out in parallel with the development of the Blue Pages.

Outputs and Results

For a full description of project outputs see 'BRIL registry and other outputs' available at <http://brii.bodleian.ox.ac.uk>. Project outputs comprise a mixture of technical items, reports and findings.

Technically focused outputs:

- Proof of concept registry containing a selection of harvested data. Constructed using the same architecture as the institutional repository and built on the Oxford DAMS (Digital Asset Management System). Entities (people, projects, funder etc) are extracted to be stored in the registry using resource maps.
- Blue Pages proof of concept directory of Oxford researchers and research activity. Available at <http://163.1.127.171/index> (Oxford only domain). The original intention was to make this website freely available, but feedback from users influenced a decision to retain it within Oxford only, until it is deemed prudent to release it more widely. A screencast demonstrating the Blue Pages is available at <http://brii.bodleian.ox.ac.uk/documentation/other-documents/BluePages.wmv>
- Example use of API demonstrating integration with a themed website. MSD data held in the registry has been re-purposed by the Division to create a cross-departmental graduate opportunities website (under construction). The need for this website was expressed by the Division.
- Collective.externalcontent has been released for general use by the Plone community (<http://svn.plone.org/svn/collective/collective.externalcontent/trunk/>). Documentation has been written for this module. Msd.rdfexport (for harvesting) and msd.ora and msd.oraexamples – building on collective.externalcontent are available for Plone users within Oxford. Screencasts of proof of concept examples can be seen at <http://webteam.medsci.ox.ac.uk/about-us/activities/directory/orainegration/>
- A freely available website was created to store and access the BRIL vocabularies. It became clear that others at Oxford working to develop services and projects using semantic technologies would want a trusted and long-lasting location for similar reasons. A request was made and granted for the website to be issued with a top level domain. It is freely available and accessible at <http://vocab.ox.ac.uk>. The software used to create the vocabulary store has been made freely available for use by other parties under an open source licence. This means that others can easily create a similar store for vocabularies. PICT (Oxford's PRAC ICT committee) has awarded funding for the vocab.ox.ac.uk site to be developed as a central mainstream service and to write a project plan for ongoing maintenance of the service for the University. This work will run April 2010 – March 2011.
- ARPFO (Academic Research Project Funding Ontology) created because there was no existing ontology to describe funders and funding activities. Available at <http://vocab.ox.ac.uk/projectfunding/>
- Webcast demonstrating example proof of concept integration of BRIL harvested data with other websites <http://webteam.medsci.ox.ac.uk/about-us/activities/directory/orainegration/>

A number of software outputs have resulted from the project:

- PairTree development⁴ for delivery of files
- Audit trails - to promote provenance, truth management and trust
- Checkm released⁵ – used to create reports on the checksums of sets of files or URLs resources for data integrity
- A library for handling digital objects with RDF metadata⁶, stored in a Pairtree structure.
- A library for handling more lightweight digital object containers in a Pairtree structure⁷ Object's metadata is serialised as JSON

⁴ <http://pypi.python.org/pypi/Pairtree>

⁵ <http://pypi.python.org/pypi/Checkm/0.3>

⁶ <http://pypi.python.org/pypi/RDFObject>

⁷ <http://pypi.python.org/pypi/RecordSilo>

- The foundation libraries for the RESTful interface to a collection of RDFObject stores, providing AMQP message passing for events⁸.

All software is released on an open source licence enabling it to be enhanced, adapted and re-used by others. Software is available at a number of locations including Python PyPi (<http://pypi.python.org/pypi>).

Software developments carried out as part of the BRIL project have generated considerable interest within the wider software developer community. This can be demonstrated by the link to the Oxford developer's instructions for use of PairTree by the California Digital Library (<https://confluence.ucop.edu/display/Curation/PairTree>) and by the endorsement of the plug-in for 4Store by EDINA (University of Edinburgh). Developers are particularly interested in the BRIL developments with PairTree, 4Store and queuing.

An algorithm to match people has been created as part of the registry. This uses the concept of SameAs i.e. this John Brown is the same person as that J. M. Brown. People are matched using a combination of last name, first name, initial, data source and affiliation. IDs are matched if available (eg RAE ID) and Oxford IDs are used if known (as on MSD people data). See 'BRIL registry and other outputs' document for more details.

APIs are available to enable web managers to extract and re-use the data held in the registry (and therefore Blue Pages). Data can be exported and re-used a number of formats including:

- JSON
- RDF
- XML
- HTML
- RSS/Atom

Future outputs will include

- Software to enable easy creation of vocabularies (taking advantage of user subject expertise and knowledge based vocabularies). This will be fully developed under the vocab.ox.ac.uk project (April 2010 – March 2011).
- Permission granted from AMS (American Mathematical Society) to create a linked data version of its Mathematical Subject Classification. www.ams.org/mathscinet/msc/msc2010.html
- If CERIF were to be adopted as the UK data exchange format, the adaptation of the BRIL registry to accept and distribute CERIF compliant data would be reasonably straightforward depending on how the national standard is set up.

Reports, publications, presentations and other works

- Rumsey, S. (2010). *A Case Analysis of Registering Research Activity for Institutional Benefit*. International Journal of Information Management Sally Rumsey DOI 10.1016/j.ijinfomgt.2009.12.004. Available at <http://ora.ouls.ox.ac.uk/objects/uuid:d71f378e-9a58-44fe-98c7-9d9eda9b0174>
- Submitted. Rumsey, S., Loureiro-Koechlin, C. 'The role of an entity registry in scholarly communication: exploring creative uses of research activity data' New Review of Academic Librarianship. Special themed issue on scholarly communication. June 2010. Proposal accepted. [Publication October 2010]
- Submitted. Rumsey, S. Paper submitted to Open Repositories 2010 conference, July 2010. 'Blurring the boundaries between an institutional repository and a research information registry: where's the join?' <http://or2010.fecyt.es/publico/Home/index.aspx>
- Work in progress. Loureiro-Koechlin, C. and Martinez-Urbe, L. *Issues and challenges relating to the use of research data and information about research. A researcher's perspective*.
- Loureiro-Koechlin, C. *Uncovering User Perceptions of Research Activity Data*. Ariadne, Issue 62. January 2010. Available at www.ariadne.ac.uk/issue62/loureiroKoechlin/
- Project website and blog available at <http://brii.bodleian.ox.ac.uk/> and <http://brii-oxford.blogspot.com>

⁸ <http://pypi.python.org/pypi/EntityStore>

- Loureiro-Koechlin, C. Stakeholder Analysis. *Exploratory Study into the requirements and uses for research activity data at the University of Oxford*. July 2009. (available at <http://ora.ouls.ox.ac.uk/objects/uuid%3A1df69991-cd37-445b-a4c7-3573ce80c36e>)
- *BRII registry and other outputs*. Available at <http://brii.bodleian.ox.ac.uk/documentation/other-documents/BRII%20registry%20-%20other%20outputs.pdf>. A comprehensive description of the technical outputs of the project. Aimed at the Project Board.
- Summative and formative evaluation reports. Project evaluation has been led and facilitated by Neil Beagrie of Charles Beagrie Limited in collaboration with the project staff. See Appendix 3
- Use cases describing scenarios where the registry and Blue Pages could be incorporated into user workflows. Available at <http://ora.ouls.ox.ac.uk/objects/uuid:3fca5ff0-30ca-47ec-9d41-b27384c8cad1>
- BRII's Stakeholder Buy-In Assembly. June 2009. Report available at the project's blog <http://brii-oxford.blogspot.com/2009/06/stakeholder-buy-in.html>
- Loureiro-Koechlin, C., Bowtell, A. *BRII Stakeholder Analysis and Sample Applications*. Making Connections JISC event. Manchester April 2009. Poster available at <http://ora.ouls.ox.ac.uk/objects/uuid%3A482c7e9a-dc55-419b-92cc-cf53f9c3cc05>
- Loureiro-Koechlin, C. *Selling an abstract concept to a practical audience*. Modular e-Administration of Teaching (MEAoT) Assembly, Cambridge University. December 2009. Presentation available at <http://ora.ouls.ox.ac.uk/objects/uuid%3Aaf26d902-3e39-46ed-8882-9cb826b91a37>
- Loureiro-Koechlin, C., *Reaching out to a big, complex university*. OR51 Conference, Warwick University, September 2009. Presentation available at <http://ora.ouls.ox.ac.uk/objects/uuid%3Aa76b8845-9599-4f96-9820-2d48c79610cf>
- Loureiro-Koechlin, C. *Building the Research Information Infrastructure (BRII)*. Inside OR, 467, 3. Short article available at <http://ora.ouls.ox.ac.uk/objects/uuid%3A9e4766cb-c972-4133-a575-eaf114230b8c>
- Table of categories of suitability of sources for harvesting and synchronising (see Appendix 1)
- List of current and planned data contributors (working document). (see Appendix 2)
- Service Usage Models (SUMs) [work in progress]
- Screencast demonstrating example proof of concept integration of BRII harvested data with other websites <http://webteam.medsci.ox.ac.uk/about-us/activities/directory/oraintegration/>
- Screencast demonstrating the Blue Pages <http://brii.bodleian.ox.ac.uk/documentation/other-documents/BluePages.wmv>
- Loureiro-Koechlin, C. *Building the Research Information Infrastructure*. Supporting research students - a unique book launch at Hull University Business School. Hull University. March 2010. Presentation available at <http://ora.ouls.ox.ac.uk/objects/uuid%3A610ccfd3-81c8-4c77-8016-ec621626e901>
- Submitted. Loureiro-Koechlin, C., 'Explaining abstract concepts with concrete examples - entity registry and research activity data.' SCOUNL Focus. March 2010.
- Forthcoming. Rumsey, S. Poster for SCOUNL conference. June 2010. Leeds.

Creation of a couple of domain maps was attempted, but it was felt that these did not really reflect the complex nature of the situation, were confusing to read and did not offer any benefits to anyone involved. They were therefore abandoned and time spent on more productive matters.

Outcomes

Project achievements against aims and objectives

The project achieved its main aim i.e. to enable efficient sharing of research management data (data about research at the University of Oxford) from a selection of existing data stores and to create exemplar services that disseminate and re-use that data using a lightweight solution based on semantic web technologies.

Outcomes listed against objectives

1) Negotiate stakeholder input and 2) Identify data sources and content for inclusion. Some groups were immediately happy to contribute data taken from their websites. This was especially the case for the MSD web pages and for one research group that is creating a publications and a project database. Others were slightly concerned because they could not commit staff time to supporting the project. Difficulties began to emerge when requesting data from a central administrative department

(Research Accounts). This data was seen by academic departments as far more sensitive than that on department websites, although comprised the same types of information. However, some project information held in Research Accounts may not be for public release.

External data providers include a couple where data is freely available to be used. Permission was granted by Cancer Research UK to use Oxford data in the way we requested. CR-UK expressed an interest in working further with us on the BRII project work. Import of this data is planned for summer 2010.

3) Build research information infrastructure. The technical development was the most straightforward part of the project, even if there were difficult challenges for developers. The biggest challenges were to persuade groups to contribute data so that it could be accessed, shared and re-used. A variety of data providers contributed data.

One problem that became apparent was that people tended to see the Blue Pages as the sole output of the project. It was clear that they found the concept an invisible registry difficult to grasp. We had to work hard to try to explain the difference between the two to enable them to understand the registry and the benefits it offers.

4) Provide RDF ontologies and taxonomies. The achievements exceeded original plans because of the development of the ARPFO ontology and the development of vocab.ox.ac.uk. Both these outputs are available for use by others. ARPFO has been discussed and extended in collaboration with colleagues at the Universities of Southampton and Bristol and is being considered for use by other individuals. Vocab.ox.ac.uk has been recognised as a valuable central resource which has been granted internal funding for further development.

5) Provide a selection of web-based applications and 6) Create a University *Blue Pages*. The Blue Pages exists as a proof of concept service. Hardware problems slowed down development of the MSD graduate opportunities website: the underlying construction has been completed and will be worked on further during April 2010.

Access to the Blue Pages

The original intention of the BRII project was to make the aggregated data stored in the registry and available via the Blue Pages freely available on the Web. The reasons for this were i) to offer wide dissemination and publicity for Oxford research and ii) to allow external groups access to that data. To avoid problems associated with sensitive data, harvested data was limited to that classed as 'publicly available data,' i.e. that which was already freely available via the Web or could be made freely available.

Despite the action taken to eliminate concerns about data being freely available, we discovered strong feelings from some people against the open nature of the registry and Blue Pages. The main concerns can be summarised as follows:

- Allowing external users access to Oxford aggregated data may i) allow other institutions to jump to incorrect conclusions about Oxford research (not knowing the background and whole story) and ii) enable the competition to spot opportunities for funding for cross- and multi-disciplinary subjects.
- Despite reassurances that we would only be using data which are already freely available or that could be made freely available, some people still did not want that data made publicly available in a single location via the Blue Pages.
- Some were concerned about sensitive data being made freely available. This includes that which is currently restricted and might be inadvertently included in the registry
- One or two individuals were worried about inadvertently sharing very sensitive data within Oxford
- One Division expressed concern about confusion arising by providing a web site (Blue Pages) that is additional to existing departmental and divisional web sites
- There were worries about data being made publicly available when being misleading and promoting an inaccurate picture of research activity if it is neither comprehensive nor accurate.
- There were a number of other issues raised such as obtaining permission from collaborators, problems arising from separating teaching and research, and the reliability of data.

Taking this feedback into account, it was clear that suggesting that data held in the registry is revealed and shared via the Blue Pages with external users is not a wholly popular development for this university at this time. The BR11 Project Team felt many of these concerns were unfounded. However, it was important not to lose user buy-in and acceptance by alienating groups of potential users by ploughing ahead whilst serious doubts were being expressed. Therefore in order to negate fears over external access to and use of registry data, it was recommended by the Project Board (at its meeting in February 2010) that a decision to freely expose the Blue Pages and registry beyond the Oxford-only domain was withheld. The data will remain accessible within the Oxford domain until such a time as is prudent to open it up on the wider Internet. This is similar to the model at the University of Melbourne where their 'Find an Expert' service remained internal only for a period of 12 months prior to public release.

Implementation of access control is a future solution for allowing access to selected data for selected users, although will take further development and was out of scope for this project. Data will remain available to Oxford users for inclusion in websites and other locations. Until access control can be implemented, data held in the registry should continue to comprise 'publicly available data.'

The problems that were encountered raise questions about what data can be shared and what can be released, what is commercially sensitive and what is deemed too sensitive for access by competitors. This goes beyond the actual content of the data and extends to the power of data aggregation which became a key factor in some areas. Because Oxford operates a devolved governance structure, there are few top-down decisions. Participation in the registry is on an opt-in not an opt-out model. It is necessary to respond to the democratic process in order to gain acceptance by the community. This makes it extremely difficult to engender institutional change, particularly rapid change both in attitudes and practice. For this reason, normal practice is to adopt a cautious approach to new development and innovation. Given more time and effort the problem may be surmounted.

One of the key lessons learned from this situation is that even the most seemingly benign assumptions made at the outset can be challenged. A second important lesson is that to gain acceptance for any major institutional change such as the registry will probably be more likely to achieve success if implemented incrementally and gradually across the divisions over a longer period.

7) Demonstrate embedding and future sustainability. This has been the most difficult part of the project. The registry will automatically be embedded as a core part of the institutional repository and so will continue beyond the project end. It will also continue to be used for the Medical Sciences web development. The project needs additional funding to continue to expand towards comprehensive coverage. Addition of more data and basic functionality is priority to continue to build the core service. Neither internal nor external project funding for continuation work has been available. The project team with support from senior members of staff has therefore been investigating the possibility of internal funding. This however, has so far been unsuccessful, hampered by the current climate of cutbacks and savings. However, funds to retain project staff over the summer 2010 have been identified whilst other options are explored. The project has benefited from members of senior staff acting as a project champion and assisting with navigating through University committees. At the time of writing this process is still in progress. The Project Board fully endorsed the project work at its final meeting in February 2010.

Project outcomes and their impact

The impact of the project outcomes are largely dependent on adoption of the registry as a central service. At the time of writing this report, this is unclear, although great efforts are being made to gain acceptance and support from key top level university committees.

- Individual researchers and other depositors in ORA will benefit in future by an easier deposit process. They may also benefit if BR11 data is used to provide contextual metadata for research data (ie output data). If the registry is adopted, then they will benefit by administrators and others being able to re-use data and not have to ask individuals for the same data numerous times.
- If the Blue Pages is adopted, central university services (such as the Press Office, ISIS Innovations and Research Services), senior management and others will benefit from easier discovery of Oxford research activities.

- Administrators will benefit from the service by being able to access and re-use data. This will be of most importance for the creation of websites, particularly cross-departmental websites. This group will also benefit from being able to pull out reports about research activities in their area.
- The *Blue Pages* online directory of researchers and research activity provides a means for Oxford to view itself in new and creative ways which has not been possible until now because of its devolved model of governance and local independence.
- The entire Oxford community could benefit from improved data quality: this is because re-use of data encourages data consistency, the facility to report errors and omissions helps identify problems and a single canonical source aids accuracy.
- The Web presence of academic and other Oxford units can be significantly enhanced using data held in the registry. Oxford groups will have the means at their disposal to be creative in how they portray themselves to the world and their colleagues for example by creating a 'themed' website.
- Aggregating Oxford's research activity data provides help to create narratives about Oxford research, and find previously undiscovered links and commonalities. Using it, members of the University can find people and research related to their own work that they never knew existed. The ability to easily identify expertise within Oxford helps the discovery of individuals or groups who might act as consultants and collaborate with external organisations.
- The design of the registry recognises and respects the devolved and disparate organisational structure of the University. Local systems retain control of their data whilst provided with a new tool that allows them to recombine and re-use their own and other Oxford data in new ways.
- The disparate nature of local systems is catered for by harmonisation of data into a single form within the registry. This means that the data can be shared, internally and externally.
- Being able to re-use existing data aids increased efficiency and encourages improved data quality.
- Over time, the registry will provide an historical record of Oxford's research activity. All data are assigned temporal and provenance information to indicate the period in which they are valid. Data will be preserved for the long-term in Bodleian Libraries' Digital Asset Management System (DAMS), a service not currently provided by existing data stores. This could be of benefit across the University.
- Undertaking this project has resulted in members of the project team being invited to participate in two groups to share expertise. These were a University research information management working group carrying out preliminary discussions to inform a new University research information management committee and the JISC research information management working group. Channels of communication have been opened up between BRII staff and those responsible for investigating a business intelligence system for the University.
- Members of the project team have collaborated with various groups which will be of benefit to all concerned and the wider research information management and semantic web communities. Developments from the BRII project have influenced many of these groups. These are: Wider work with Bristol and Southampton particularly via the ResearchRevealed project and including Readiness4REF, RKBExplorer/Dot.AC projects; CERN (want to upgrade to make an RDF version of their bibliographic data (currently in MARCXML) and has been interested to find out about BRII work including how to handle non-bibliographic data such as people); Dev8D (February 2010) including BRII work with ontologies and development of the handshake machine to record collaborations; Cornell University in the US to discuss common problems experienced by BRII and their VIVO service; Talis semantic web developers; Interest in BRII work has also been expressed by Cancer Research-UK
- The ability to record, display and report usage statistics have been created and matched between ORA and BRII and will provide an indicator of effectiveness of content.
- The MSD now has a useful set of tools to move forward in meeting its strategic recruitment and communication goals, working more efficiently with less duplication of effort. Awareness of the issues of ownership, maintenance and publication of research data has been raised, at senior management, administrative and researcher level.

Further uses within the Medical Sciences

MSD will be taking two approaches on building on the outputs of the BRII project. The first will be re-use of data. As will be illustrated in the demonstration of a theme website, MSD will be able to draw together research profiles of researchers in the same disciplines, in different departments. Work will be on-going to assist the units, consortiums and thematic collaborations to put together informative websites in the most efficient manner. There is also a growing interest within the Division to develop collaborative

Virtual Research Environments in particular disciplines and BRIL registry data can be integrated within these to provide a richer context for the collaborations taking place within them. With further work in compiling researcher profiles at department level and support from the BRIL project in drawing in other sources of researcher data, the Division could, if it wishes, look beyond the University boundaries in particular subject areas. It would be in a strong position to make an immediate contribution of data (or a subset of it) to national and international networks - such as the NIH funded VIVO project at Cornell University Library to Discover Expertise and Enable Collaborations, which has a focus on networking biomedical researchers.

The second approach will be the re-use of the technology piloted by the BRIL project. The BRIL project has the potential to provide a central, one-stop, simple to use, University-wide programming interface (API) which will enable any webmaster to access an aggregation of publicly available research information, including, of course, publication data. We would hope that ongoing work in documenting this API would facilitate its reuse. With ongoing support from the ORA registry MSD could expand its research ontologies to cover additional research 'objects' currently described on its department websites and make these available for harvesting and aggregation. These could include trials, projects, research resources, significant research output and more detailed information about research collaborations. The software which has been developed to integrate with the MSD CMS can be re-used, very efficiently, to embed this additional information in its web-delivered resources - either to enhance MSD's public facing communications or to facilitate collaboration within its research environments, or simply to provide immediate information to its staff and students.

Lessons learned applicable to other projects.

- The iterative development worked well for the web interface development. By obtaining feedback from a small number of testers, implementing changes and then re-testing on a new group, we were able to eliminate the majority focusing on the initial 'obvious' problems and gain greater depth of views.
- Use of terminology caused difficulties for clarity. For example, what one group means by the term 'project' is quite different to another and we uncovered political nuances with use of the term 'academic units.' It is important to try to resolve these difficulties if possible as it colours users' perception of the service.
- The project team was surprised by some reactions about use of already publicly available data or data which the team felt could be made available without problems. People may post material in public but they do not necessarily want it to be publicised via other channels, be re-arranged or aggregated. The original intention of the BRIL project was make the aggregated data stored in the registry and available via the Blue Pages freely available on the Web. The reasons for this were i) to offer wide dissemination and publicity for Oxford research and ii) to allow external groups access to that data. To avoid problems associated with sensitive data, harvested data was limited to that classed as 'publicly available data,' i.e. that which was already freely available via the Web or could be made freely available. Despite the action taken to eliminate concerns about the data being freely available, we have discovered concerns against the open nature of the registry and Blue Pages. This feedback has been incorporated by recommending that the data remains available within the Oxford-only domain for the foreseeable future. We have learned that it is possible to make incorrect assumptions about our users and that to gain acceptance for any major institutional change such as the registry will probably be more likely to achieve success if implemented incrementally and gradually across the divisions over a longer period.
- Problems of staff recruitment will ensure that for future funding proposals we will build in a longer period for recruitment.
- Decisions regarding embedding and future sustainability have proved to be much more complex and drawn out than anticipated. This is possibly exacerbated by the fact that the service that has been developed is attempting to offer innovation at an institutional level. The devolved structure of the University means that there are a number of hurdles to overcome. If attempting this type of project again, it would be worth considering a pilot limited to a small group within one Division or similar. Although, there would have been different problems associated with this method.
- The work involved in constructing a technical infrastructure to underpin development work encountered practical difficulties outside the control of the Project. This then impacted on the project's ability to reserve time with external consultancies.
- We realised that the Blue Pages began to be perceived as the only output of the project. The concept of the registry was difficult for many users to grasp – there is nothing as such to 'see.' The

project team therefore began to try to explain in clearer terms that the Blue Pages is only one part of the story and to ensure that users understand the concept and benefits of the registry itself.

Formal summative evaluation

A formal evaluation was undertaken by Neil Beagrie of Charles Beagrie during January and February 2010. See Appendix 3 for the full evaluation report. The key findings are as follows:

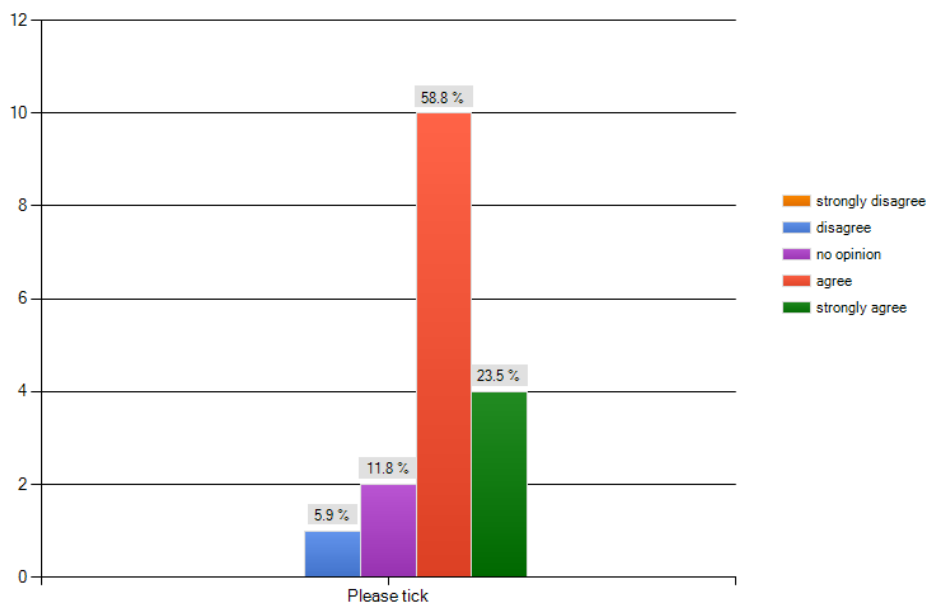
Key Findings for JISC

1. This has been a successful project which has achieved at least 11 out of 12 of its targets set out in the evaluation plan (appendix 2). The only uncertainty over a target is that it is too early to judge whether the institutional innovation has been achieved.
2. Attempting to achieve institutional innovation in a large institution in 18 months is ambitious. Technical considerations apart, in-depth interviews, user testing and consultation take considerable time. With hindsight, development would have better attempted as a three year project to allow for hiring staff, further in-depth user testing, consultation and more comprehensive data harvesting.
3. Project technical achievements have been considerable within the limited period of the project and have attracted strong interest from the UK and internationally. The project team has been pro-active in disseminating and sharing their learning and outputs with the JISC community. In addition to published articles (e.g. Ariadne January 2010) and presentations, the project has benefited from and contributed to a number of influential consultations within the UK such as the EXRI project and the JISC research information management group, via a variety of interactions. The technical outputs, stakeholder analysis and interactions and experiences of the BRII project will have significant application in other institutions.
4. Technical achievements have included creating the registry, data harvesting from a variety of sources, recording provenance, use of ontologies, and open source software outputs.
5. The project has raised a number of issues associated with research information management and exchange at the national level. How institutions share data with, for example, funders. It also encompasses strategic issues such as who creates data on research activity, who holds it and who is responsible for it. This issue should be a continuing focus for JISC dialogue and pilot projects with other national agencies and HEIs.

Key Findings for Oxford

6. The majority of user feedback on the BRII project has been very positive (see appendix 3). 82.3% of the surveyed users either agreed or strongly agreed that "The BRII project has created the foundation for an important University service to aid storage, sharing and promotion of information about Oxford research" (User Survey Q6).

How strongly do you disagree or agree with the following statement "The BRII project has created the foundation for an important University service to aid storage, sharing and promotion of information about Oxford research."



7. key benefits from implementing BRII identified in the user focus groups and online survey are seen as being the ability to quickly identify potential collaborators and related areas of work within Oxford thus improving chances of successful research funding bids (particularly for those with tight submission deadlines or inter-disciplinary opportunities); and higher visibility of research activity to potential students and researchers. There are also likely to be some time saving and efficiency gains, particularly for research support staff.
8. The ability to re-use data is a key output of the project and can be applied at both institutional and departmental levels. Easy re-use is demonstrated in a webcast showing integration of the registry within Medical Sciences. APIs are available for technical integration with the registry.
9. The BRII Project staff have been monitoring similar research activity information developments beyond the UK. The Cornell/NIH VIVOWeb work is of most interest. Originally an inhouse research activity information network at Cornell, VIVOWeb has since been adopted for local networks at other universities and institutions in the United States, Australia and China. VIVOWeb has been described as follows by Cornell:
"Before VIVO, the Cornell librarians heard a lot of frustration from faculty members who couldn't find collaborators from different disciplines across campus," Medha Devare, Cornell librarian for bioinformatics and life sciences. "The idea of VIVO was to transcend administrative divisions and create a single point of access for scholarly interaction. Now that VIVO is expanding across institutions, the biomedical community will be able to benefit from that bird's eye perspective of their research." (Cornell Press release http://www.eurekalert.org/pub_releases/2009-10/cu-vsw102109.php).
Staff at Cornell are interested in the Oxford technical model and the BRII project team have participated in a teleconference to share ideas and expertise.
10. There are doubts about inclusion of sensitive data or of the conclusions that might be drawn by others when data are aggregated in BRII. People are also concerned about comprehensiveness and accuracy of data, for example the fact that collaborations might have lapsed and which might not be indicated by the data in the registry. The BRII project team are currently considering how best to address these matters and the optimum route forward for development of the registry.

11. Adoption by the University is proving to be much more difficult than had been anticipated. Part of the problem is ownership for the ongoing work when it does not sit easily within any overall group coupled with many of the issues presented being tackled only at local level. A new university sub-committee has been set up to address research information management issues.

Conclusions

Internal concerns associated with sharing aggregated data beyond Oxford has made the project team realise that this is a highly sensitive issue for many at the university and that an even more cautious approach than had been planned needs to be taken. Whether institutional innovation has been one of the outcomes of the project remains to be seen. Despite this, the work has not been in vain. Important questions of data ownership, control and access have been raised within the University. In terms of its outputs, the project has been a resounding success and there have been many positive outcomes of the work. Much of the project development work will continue as part of the repository development. If the registry is not adopted this year, then there could be an opportunity to re-visit it in the future. This is highly likely given the direction of data exchange developments at the national level. The registry will continue to serve the MSD as, together with the Web Manager's web editing tool, it provides a service for web development that demonstrates the ease with which websites can be constructed if this method is adopted. Efforts for further development and adoption will continue as part of ORA advocacy.

Implications

There has been international interest in the work carried out as part of the project, in particular from colleagues at Cornell working on the VIVOWeb implementation. The project team intends to continue collaborations and dialogue with Cornell and with other interested groups including those at the universities in Bristol, Southampton and in Melbourne. We shall continue to contribute to the JISC research management working group and to the developer community dialogue.

The semantic web structure of the registry and its use for research information management is one which could become commonplace within the UK and elsewhere. The methods to incorporate non-bibliographic or similar entities such as funders and people together with the means to describe collaborators using semantic vocabularies could be of significant interest to others. As research information management systems and processes develop over the next few years, a number of aspects of the BRII project, together with its collaborative work with Bristol and Southampton, could become standard practice. This might include use of ontologies and vocabularies and provenance and temporal data.

There are a number of aspects of the BRII project work that could be extended:

- Controlled vocabularies and thesauri. This is one aspect of the work that we feel we would like to explore further. There are a number of developments in this field and much that the community can contribute to enable such vocabularies to enhance many services and resources. Additional vocabularies and enhancements to existing vocabularies, both structural and descriptive, would benefit the community.
- The registry contains data objects that are not publications and can express collaborations and other relationships. More work could be done in this area to add different entity types and relationships (eg geo-location data or methodologies)
- The co-referencing facility has been aided by the work carried out the Dot.AC project at Southampton. More work could (and is) being done in this area to help resolve the problem of identity and one object being the same as another.
- Collaborative work to expose data using suitable methods and the ability to re-use that data with external (non-university) partners would be a benefit to the community. For example, funding and government bodies.
- Because the registry is built using semantic web technologies and linked data, there is much that could be done with that data to present it in a variety of ways for different purposes.
- Relevant open source software will continue to be developed and contributed to the community and that already released will hopefully be taken up and adapted and used by others.

Appendixes

Glossary

BRII	Building the Research Information Infrastructure
CMS	Content Management System
DAMS	(Oxford) Digital Asset Management System
MSD	Medical Sciences Division (of the University of Oxford)
ORA	Oxford University Research Archive
PICT	Oxford's PRAC (Policy & Research Allocation Committee) ICT committee)
RDF	Resource Description Framework

Appendix 1: Categories of data sources

Type of data source	Examples	Ease of inclusion and synchronising/harvesting
1. Systems constructed to be able to 'push' data into the Registry, in batches or even as it changes on their system	e-humanities data as being developed by OeRC	Easy. Changes in data reflected quickly in registry. Registry updates and keeps an audit trail of the changes.
2. Systems constructed to allow 'pull' of data from them into Registry using atom feeds, OAI-PMH or other suitable service.	MSD websites (eg Cardiovascular)	Easy. Registry update as and when. Changes in data reflected in registry at points of updating. Only records for items that have changed, been created, or deleted will have to be updated
3. Data can be accessed somehow, but there is a way to identify what is new, changed or deleted.	NDM Pinfox	Fairly easy. More custom work required, but doable. Update on demand. Registry records can be synchronised with source records.
4. Data can be accessed, but it doesn't have history records, provenance, or obvious modification dates.	Expected to include the majority of Oxford data sources Cancer Research UK	Becoming more difficult. Need to work out a strategy for updating. It can be assisted by Last modification dates in the db/system, the presence of provenance/history data, traceable and stable identifiers for things in the system and so on. Updating realistically would be: updating existing records, getting lists of ids in the system to then work out what is new or deleted and adding/deleting those accordingly. At worst, means a total synchronise on demand, updating all records and deleting/hiding (policy) those records that no longer exist. In this case, need to work out the records replacement if it exists, or break the linkages if not.
5. One-offs of data that can change	Scraped web/home pages	Unfeasible. Essentially, identical to level 4, but the amount of work needed is unfeasible for continued updating based on the resources and staff time.
6. Static data that will not change	Philosophy RAE 2008 return	Fairly easy. Once only upload
7. Any type of data source affected by rights that restrict the upload of data no matter how accessible it.	Wellcome Trust	Requires negotiation

Appendix 2: BRIL Project Data Contributors

The following is a list of departments whose research activity data may be harvested by the RII registry during the course of the BRIL project.

Priority order	Source	Types research activity data	Location of data storage	Format	BRIL pilot	Reasons for inclusion
1 Complete	Oxford University academic divisions and departments and colleges. University's organisational chart Originate from Ilana Veitch in admin	Oxford University academic divisions and departments and colleges	PRAS website	Excel spreadsheet	Yes	Vital and quick win
2 Complete	Department of Radiation, Oncology & Biology	People, roles, topics, themes, collaborators, funders, contact details, projects(?), research group, summary	MSD web	RDF, XML	Yes	Project collaboration with MSD web manager. Data created in similar way to BRIL therefore compatible & easy to synchronise (daily)
3 Complete	EPSRC As at 11/1/10	Details of c. 400 people and 1045 Oxford grants. Department, grant (project) title, PI name, EPSRC Reference No, Co-investigator names, research contact details, Project partners (collaborators), Project date, EPSRC Research Topic Classifications; EPSRC Industrial Sector Classifications; Related Grants; project summary; Final Report Summary	ESRC website. Any changes – import to be re-run	CSV, Excel, MHTML, PDF, PSV, XML	Yes	External body. Good quantity of Live data. Mainly MPLS research. Should be re-run regularly because of updates. Needs auto alert for updates
4 Complete	Department of Cardiovascular Medicine, Centre of Research Excellence and the Cardiovascular Science theme	People, roles, topics, themes, collaborators, funders, contact details, projects(?), research group, summary	MSD web	RDF, XML	Yes	Project collaboration with MSD web manager. Data created in similar way to BRIL therefore compatible & easy to synchronise (daily)
5 Complete	Department of Physiology, Anatomy and Genetics: access to data in their website	Research themes; Research groups; Names; Collaborators; Some publications	MSD web		Yes	Complex collection of websites – personal, group etc. Project collaboration with MSD web manager. Data created in similar way to BRIL therefore compatible & easy to synchronise (daily)
6 Complete	Faculty of Philosophy: RAE 2008 return + website (needs updating because of improved website)	c. 300 records. Last Name, Initials, publications (mainly book chapters) HESA Staff Identifier, Staff Identifier, (IDs not for public view). Book editors (external departments)	Excel spreadsheet & department website		Yes	Quick win. Humanities. Good quantity of publication data. Static data which will not change (although web info has)
7 Complete	RAE 2001 return	Details of Oxford submissions – staff output & output details. Name, affiliation,	HEFCE RAE website		Yes	Wide coverage across all disciplines. Static data.

		bibliographic details, co-authors				
8 Complete	RAE 2008 return	Details of Oxford submissions – staff output & output details. Name, affiliation, bibliographic details, co-authors	HEFCE RAE website		Yes	Wide coverage across all disciplines. Static data.
9 Complete	Nuffield Laboratory of Ophthalmology	Research themes; Names; Some publications http://www.eye.ox.ac.uk/	Department website	RDF, XML	Yes	
10 Complete	HELEX Centre for Health, Law and Emerging Technologies	Research projects; Names; Collaborators; Some publications http://helex.medsci.ox.ac.uk	Department website	RDF, XML	Yes	
11 Complete	Research Accounts data for NPEU and for Humanities Division				Yes	Permission granted from depts for use of this data. Central admin department.
12 [work in progress]	National Perinatal Epidemiology Unit:	access to data in their website. Data held on two databases: People and Projects.				Example of research management system being developed for a research group
13	Balliol College	c. 65 individuals. Details of the College Fellows as displayed on the College website at http://www.balliol.ox.ac.uk/official/staff/senmem/index.asp Name; title; role; URL; college affiliation	Website CMS	Requires manual entry	Yes	Proof of concept for college. Strategically important. Encourage 1-time effort to output data in RSS/XML.
14	E-Humanities project:	People, projects	OeRC database then Oxford DAMS Excel spreadsheet and project database.		Yes	Provide back-end for e-Humanities project.
15	Cancer Research UK <i>Permission has been granted for use of data on Oxford website</i>	http://science.cancerresearchuk.org/research/loc/oxford/ research group; names; Projects	Website		Yes	External source with project data. Permission for use granted by RCUK
16	ORA	c.3000 Publications and other research outputs. People, affiliation, role, collaborations. Entities being extracted from existing records for inclusion in RII	Oxford DAMS	RDF, XML	Yes	ORA and the RII are part of same system. Work on entity extraction to separate them for RII
17	Nuffield Department of Clinical Medicine: access to publicly available data in the NDM Pinfox system	Research theme; PIs; research topic/subject; Collaborators (name, institution, country); Research group; Research institutes, centres, units; publications; titles/roles; contact details;				Large MSD department with established method for collecting research data

		keywords;				
18	Computing Laboratory - Computer Science department: access to data in their object oriented database system	TBC				Existing live data store. MPLS department
19	Oxford University Oracle database. <i>Permission sought for use of this data source</i>	Successful grants: project name, PI. Exact details to be confirmed. Humanities and NPEU data.		Export to Excel spreadsheet – CSV		Key institutional data store.
20	Nuffield Department of Anaesthetics: access to data in their website	Research topics; groups; Researcher names; roles; collaborators; Funder; publications				Small MSD department with research data. Cross department connections (eg DPAG)
	Babylab - Experimental Psychology: access to data in their website. (Note: the website is currently under development.)	Names; roles; publications				Small focused research group with cross discipline connections
	Wellcome Trust <i>Probably requires permission for use of data on Oxford website</i>	PI, Collaborators; Department; Project title; [Date]		Excel spreadsheet	Permission required	External source. Funds a large proportion of research at Oxford.
	BL/MIMAS Names project. Oxford author names. Unfortunately not constructed in an easy-to-use format. Will require extraction in somewhat unefficient method.	Author names, Names project ID		MARCXML		Large quantity of author names including non-Oxford names and additional lookup. Also national ID.
	Social Sciences Division: although the system is in the early stages of development we hope the Division will consider sharing data at the earliest opportunity.	Not yet constructed		N/A	Not yet available	Database currently under planning and at very early stage of development.

Appendix 3: Summative evaluation report



EVALUATION OF THE BRII PROJECT

Final Report - February 2010

Prepared by:

Charles Beagrie Limited

www.beagrie.com

Contents

1. Executive Summary.....	21
1.1. Introduction	21
1.2. Methodology for the Summative Evaluation.....	21
1.3. Key Findings	21
Appendix A: BRII Project Self-Evaluation Report January 2010.....	24
Appendix B: BRII Project Evaluation Plan.....	31
Appendix C: BRII Summative Evaluation User Survey Results	33

1. Executive Summary

1.1. Introduction

The BR11 project (Oct 2008 – March 2010) is based wholly at the University of Oxford. The aim of the project is to enable efficient sharing of research activity data (i.e. data *about* research at the University of Oxford) from a selection of existing data stores and to create exemplar services that disseminate and re-use that data using a lightweight solution based on semantic web technologies.

1.2. Methodology for the Summative Evaluation

The Summative evaluation has followed the framework for the project evaluation agreed with JISC (see Appendix B for the Project Evaluation Plan). The independent evaluation has been led and facilitated by Neil Beagrie of Charles Beagrie Limited. The evaluation was conducted in two parts with an interim (formative) evaluation during August/September 2009 and a final (summative) evaluation completed during January/February 2010.

In February 2010 the project team updated their self-evaluation document for the final summative evaluation. The final self-evaluation report (see Appendix A) covered issues listed in evaluation plan (see Appendix B) and the aims/objectives in the original project plan and captured comments on achievements and challenges in the project. Additional documents seen by the evaluator included the Stakeholder Analysis, and controlled terms used (all resources online and linked in the self-evaluation report Appendix A).

During January and February 2010, the evaluators completed an independent survey of BR11 users as part of the final evaluation. The survey was made available on SurveyMonkey (an online survey website) and 31 users within Oxford who had seen demonstrations or undertaken testing on the BR11 “Blue Pages” were invited via email to participate in the survey (details of the results, and any comments for each question are provided in Appendix C).

Finally in late February 2010, the results of the user survey were discussed with a focus group drawn from 3 departments and from Research Services in Oxford to confirm findings and provide further qualitative feedback. The draft self-evaluation report and user survey results were also discussed in meetings with the BR11 project board and the BR11 project team prior to finalising the evaluation report.

1.3. Key Findings

Key Findings for JISC

1. This has been a successful project which has achieved at least 11 out of 12 of its targets set out in the evaluation plan (Appendix B). The only uncertainty over a target is that it is too early to judge whether the institutional innovation has been achieved.
2. Attempting to achieve institutional innovation in a large institution in 18 months is ambitious. Technical considerations apart, in-depth interviews, user testing and consultation take considerable time. With hindsight, development would have better attempted as a three year project to allow for hiring staff, further in-depth user testing, consultation and more comprehensive data harvesting.
3. Project technical achievements have been considerable within the limited period of the project and have attracted strong interest from the UK and internationally. The project team has been pro-active in disseminating and sharing their learning and outputs with the JISC community. In addition to published articles (e.g. Ariadne January 2010) and presentations, the project has benefited from and contributed to a number of influential consultations within the UK such as the EXRI project and the

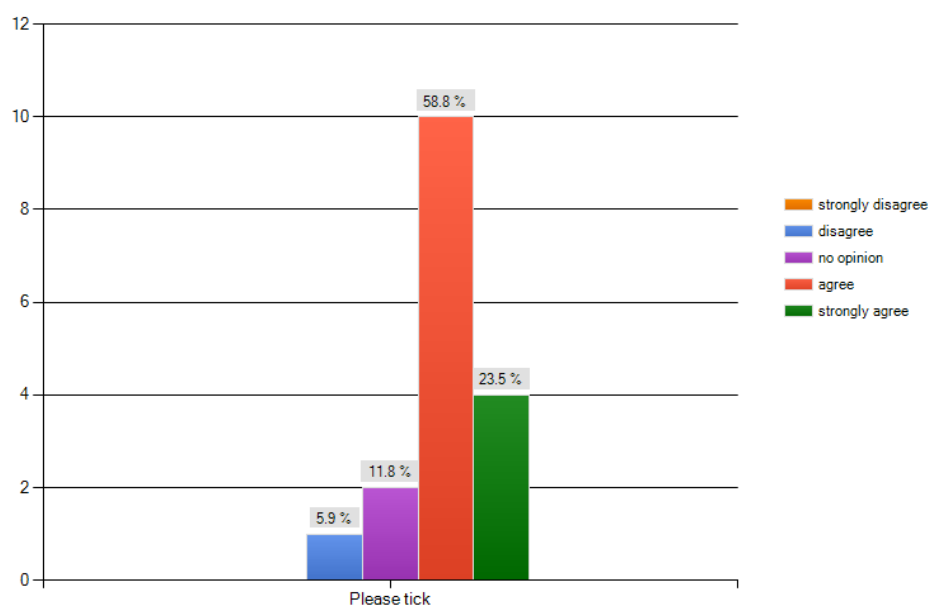
JISC research information management group, via a variety of interactions. The technical outputs, stakeholder analysis and interactions and experiences of the BR11 project will have significant application in other institutions.

4. Technical achievements have included creating the registry, data harvesting from a variety of sources, recording provenance, use of ontologies, and open source software outputs.
5. The project has raised a number of issues associated with research information management and exchange at the national level. How institutions share data with, for example, funders. It also encompasses strategic issues such as who creates data on research activity, who holds it and who is responsible for it. This issue should be a continuing focus for JISC dialogue and pilot projects with other national agencies and HEIs.

Key Findings for Oxford

6. The majority of user feedback on the BR11 project has been very positive (see appendix 3). 82.3% of the surveyed users either agreed or strongly agreed that “The BR11 project has created the foundation for an important University service to aid storage, sharing and promotion of information about Oxford research” (User Survey Q6).

How strongly do you disagree or agree with the following statement “The BR11 project has created the foundation for an important University service to aid storage, sharing and promotion of information about Oxford research.”



7. Key benefits from implementing BR11 identified in the user focus groups and online survey are seen as being the ability to quickly identify potential collaborators and related areas of work within Oxford thus improving chances of successful research funding bids (particularly for those with tight submission deadlines or inter-disciplinary opportunities); and higher visibility of research activity to potential students and researchers. There are also likely to be some time saving and efficiency gains, particularly for research support staff.
8. The ability to re-use data is a key output of the project and can be applied at both institutional and departmental levels. Easy re-use is demonstrated in a webcast

showing integration of the registry within Medical Sciences. APIs are available for technical integration with the registry.

9. The BRII Project staff have been monitoring similar research activity information developments beyond the UK. The Cornell/NIH VIVOWeb work is of most interest. Originally an in-house research activity information network at Cornell, VIVOWeb has since been adopted for local networks at other universities and institutions in the United States, Australia and China. VIVOWeb has been described as follows by Cornell:
"Before VIVO, the Cornell librarians heard a lot of frustration from faculty members who couldn't find collaborators from different disciplines across campus," Medha Devare, Cornell librarian for bioinformatics and life sciences. "The idea of VIVO was to transcend administrative divisions and create a single point of access for scholarly interaction. Now that VIVO is expanding across institutions, the biomedical community will be able to benefit from that bird's eye perspective of their research."
(Cornell Press release http://www.eurekalert.org/pub_releases/2009-10/cu-vsw102109.php).
Staff at Cornell are interested in the Oxford technical model and the BRII project team have participated in a teleconference to share ideas and expertise.
10. There are doubts about inclusion of sensitive data or of the conclusions that might be drawn by others when data are aggregated in BRII. People are also concerned about comprehensiveness and accuracy of data, for example the fact that collaborations might have lapsed and which might not be indicated by the data in the registry. The BRII project team are currently considering how best to address these matters and the optimum route forward for development of the registry.
11. Adoption by the University is proving to be much more difficult than had been anticipated. Part of the problem is ownership for the ongoing work when it does not sit easily within any overall group coupled with many of the issues presented being tackled only at local level. A new university sub-committee has been set up to address research information management issues.

Appendix A: BRII Project Self-Evaluation Report February 2010

1. Project progress

The project has mostly followed the planned schedule and, most importantly, is on target to deliver what it set out to in the original project plan. This is despite staffing difficulties and the knock-on effect on the budget. A number of additional outputs to those listed in the project plan have resulted from the work.

A comprehensive stakeholder analysis completed during the first half of the project contains valuable information about registry users, contributors and their views on research information data and its use. The pilot data registry is running and contains data from a variety of sources, both internal and external (see pages Appendix 2 above). The Blue Pages is live (within the ox.ac.uk domain only) and development has been informed by extensive user testing. The exemplar themed website has not yet been developed, but work has started on this output.

A number of additional software outputs have resulted from the project. They include:

- PairTree development⁹ for delivery of files
- Audit trails - to promote provenance, truth management and trust
- Checkm released¹⁰ – used to create reports on the checksums of sets of files or URLs resources for data integrity

All software is released on an open source licence enabling it to be enhanced, adapted and re-used by others.

A new vocabulary, ARPFO (Academic Research Project Funding Ontology), for describing the structure and relationships between organisations and individuals involved in providing funding or receiving funds for academic research or projects has been developed and made freely available at <http://vocab.ox.ac.uk/projectfunding/projectfunding>

Ontologies and vocabularies describe the data in a structured way. A freely available website was created to store and access the BRII vocabularies and make them available for others who want to re-use them. A request was made and granted for the website to be issued with a top level domain so that others at Oxford working to develop services and projects using semantic technologies have a trusted and long-lasting location for their vocabularies. It is freely available and accessible at <http://vocab.ox.ac.uk>. PICT (Oxford's Planning & Resources ICT subcommittee) has awarded funding for the vocab.ox.ac.uk site to be fully developed as a central mainstream service for the University. The work will run April 2010 – March 2011. This fits well with the project's overall aim to promote institutional innovation. The software used to create the vocabulary store has been made freely available for use by other parties under an open source licence. This means that others can easily create a similar store for vocabularies.

Two exemplar services are being developed to demonstrate use of the data. The Oxford Blue Pages is an online directory of research and researchers at the university. As such it acts as an aggregator and distributor of data about people and their research, and facilitates the discovery and display of connections between people and research activities. In these capacities the Blue Pages can be used to find information that cannot be found elsewhere and enables users to find information about research across the university. If adopted, this is another example of institutional innovation. The ultimate aim is for the Blue Pages to be available for searching and browsing to anyone with Internet access.

The data held in the BRII registry is being used to demonstrate an example of API integration with a themed website. A themed website is a means to aggregate disparate data around a common selected theme. It presents a subject-focused overview of research in areas which

⁹ <http://pypi.python.org/pypi/Pairtree>

¹⁰ <http://pypi.python.org/pypi/Checkm/0.3>

cross the more formal boundaries of departments or methodologies, for instance, cancer, neuroscience or climate change. The results could highlight potential collaboration and other connections across the University. Data may also be combined with other information not held in the registry. For the purpose of demonstration the Medical Sciences Division is constructing a website integrating information about researchers across the Division, with information about opportunities for graduate studies

The document '*BRII Registry and other outputs*' gives details of outputs from the project. The paper is available at <http://brii.bodleian.ox.ac.uk>.

At its meeting in December 2009 the BRII Project Board recommended that work on the registry should continue beyond project end. Internal funds have been identified to retain project staff from April 2010 to end of August 2010. During this period, work will continue to ascertain whether or not the University wishes to develop the service further and provide resources for extension and enhancement. At the suggestion of the Board, strategies to inform those who need to know and make top level decisions about future support for the registry, in particular the University's new Research Information Management Sub-Committee (RIMSC), are currently being acted upon.

At its recent meeting RIMSC agreed that it would provide a statement broadly endorsing the registry but with a proviso (final wording to be agreed). The statement will be to the effect that the committee agrees that the Research Registry has the potential to improve elements of the research landscape at Oxford. However the university needs to develop its information strategy before it can be clear where the Registry fits into that strategy. This statement could be used to support a request for funding with other parties.

The BRII Project Board will decide when and if the Blue Pages should be launched and how widely they should be made available. There are concerns about launching the service before it is suitably mature, or before potential users fully understand its purpose and its pilot state, giving rise to inflated expectations and possible rejection. A similar scenario at the University of Melbourne is being used as a model for how to handle the release to the community.

The BRII project has been funded under the JISC Institutional Innovation programme. It should be noted that in such a large and devolved organisation as Oxford, implementing institutional change is a monumental challenge. There has been some success in achieving institutional innovation as a result of the project. At Oxford there are few top level mandates: each organisational unit acts as a separate entity. It is therefore felt that one way to encourage change is by demonstrating the benefits of the service to early adopters. By creating as comprehensive data coverage as possible using a variety of data sources, units will be encouraged to participate rather than be left out. The understandable concerns of data sharing within some groups should not be underestimated and these need to be carefully considered over the coming months as more data are gathered.

BRII project staff have been collaborating closely with colleagues at Bristol University (ResearchRevealed and EXRI-UK projects) and Southampton University (Dot.AC project). This has been to the benefit of all parties. There has been considerable sharing of technical and other expertise between these groups.

A number of publications and other outputs have been produced. A full list can be found in the project final report and in '*BRII registry and other outputs*' document which are available at <http://brii.bodleian.ox.ac.uk/brii-papers-reports-and-presentations>

2. User engagement

During the course of the project BRII staff have contacted, spoken to and formally interviewed a large number of people from across the University. They include members of academic departments, administrative departments, senior management, administrators, researchers and technology support. Interest in the work of the project has been considerable. Most interest in potential use of the registry has been from those described as administrative staff,

and strategists & disseminators in the stakeholder analysis. Researchers themselves are not usually concerned with the 'broad brush' picture of research across the university.

An online survey was conducted as part of the final project evaluation (see Appendix 3). Respondents were selected from the stakeholder analysis and user tests groups. The registry and its APIs being rather abstract concepts, the project team considered only respondents who had had interaction with the team and a good understanding of the project and the products developed. Most survey respondents had seen a demonstration or run a test of the Blue Pages.

The scope of the project means that contact has been made with a selection of stakeholders across the University. We attempted to include a wide cross section of potential users and contributors from the academic divisions, one college and several central university services, as well as business and research consultants (ISIS Innovation) interested in commercialising intellectual property resulting from research (see stakeholder grid at <http://brii.bodleian.ox.ac.uk/brii-papers-reports-and-presentations>).

As part of the development of the registry and the Oxford Blue Pages a series of user tests were run to obtain feedback from potential users. The information obtained from these tests was useful to inform improvement of the usability of the Blue Pages and to assess the data harvested in the registry. However, most importantly, tests proved to be extremely helpful at engaging with users and assessing their perceptions of usefulness of the Blue Pages and the registry. Users' perceptions after seeing the tool supported and added to the data obtained in the stakeholder analysis. Administrators, disseminators and strategists were enthusiastic about the tool, particularly its ability to aggregate data from multiple sources in one screen and the information about collaborations. Researchers stated that research activity data and the Blue Pages could be useful for administrators and university services. They said though it would be of little use for them as it was not academic enough. However, they identified some potential uses particularly when (thinking of) starting projects and collaborations. The Blue Pages would be useful to find people and/or resources in specific fields within the University. Some academics also stated that they would value the Blue Pages and the registry more if they could obtain information from them which they cannot get from other sources. An example given was a network or map of subject fields and people in Oxford.

After experiencing the Blue Pages all testers were able to understand the mechanics of data harvesting and updating as well as the difference between the data held in original sources and the registry. Issues were raised by some of them regarding frequency of the updating processes, scope of data harvested and permissions. Users gave their opinions and suggested further avenues for exploration. In addition to user tests, software demonstrations were also conducted. These sessions, although not as exhaustive as tests, allowed participants to learn about the Blue Pages and to share their points of view.

In the Medical Sciences Division a Graduate Opportunities website is being developed which will incorporate use of the API, the specific tools developed for the MSD content management system and data from the registry. The theme for the website has emerged from the Division's strategic aims in graduate recruitment in the Division. The website will facilitate the discovery of relevant information on projects and research opportunities in subject areas across the Division. Work on this website is being carried out in close collaboration with these stakeholders.

Although reactions to the registry and other aspects of BR11 have been mainly positive, this has not been universal. There are doubts about inclusion of sensitive data or of the conclusions that might be drawn by others when data are aggregated. People are also concerned about comprehensiveness and accuracy of data, for example the fact that collaborations might have lapsed and which might not be indicated by the data in the registry. We are currently considering how best to address these matters and the optimum route forward for development of the registry.

3. Controlled terms

Vocabularies, ontologies, taxonomies and thesauri (for brevity the term 'vocabulary' is used to refer to all these types and others such as dictionaries and classifications) are used to categorise and classify data. Vocabularies help describe the structure and context of the data, as well as describing the data itself. The vocabularies used in the registry can be divided into two main types: 'structural' and 'subject descriptive.'

1. **Structural vocabularies** comprise those which enable terms to be structurally related, and relate to the architecture of the entities harvested into the registry. These tend to be ontologies or other modelling specifications describing an environment and how each 'thing' within that environment links and relates to other 'things' within the same environment (for example, SKOS).
2. **Subject descriptive vocabularies** attempt to describe a topic using a controlled scheme where terms relate to each other. These types of vocabulary include thesauri, taxonomies, dictionaries and subject classification schemes. At some stage the registry will include the LCSH as linked data. We have obtained permission from the AMS (American Mathematical Society) to produce a linked data form of the AMS Mathematical Subject Classification. We have permission to use the HASSET (social sciences) thesaurus but permission has not been granted to create a linked data form.

Vocabularies are used within the registry to describe terms. They are also published in such a way at a persistent namespace (vocab.ox.ac.uk - Oxford's new persistent namespace) that enables communities to understand how the terms have been used in each instance and to promote their re-use.

Existing vocabularies are used wherever possible. Amendments are suggested to vocabulary developers if appropriate. We have, for example, suggested additional terms which have been accepted to the standard AIISO roles ontology.

The Medical Sciences Division RES vocabulary, describing researcher activity, has been revised and a second version published. In collaboration with ORA/BR11, and after further research and investigation of currently available vocabularies, some terms have been simplified or existing terms from other vocabularies substituted.

There has been considerable success in incorporating structural ontologies, however we have not included as much work on subject vocabularies as had originally been hoped. There is scope for further work in this area.

4. Data harvesting

Data held in the registry has been drawn from a variety of sources, both internal and external. The project achieved its original aim to include data from different academic Divisions, a College, and has permission to include data from a central administrative department (although this proved more difficult than anticipated). Data has also been included from an external source (EPSRC). See Appendix 2 for details

One output of the project is a table of categories of data sources and their ease of harvesting and synchronising (see BR11 outputs document in Appendix 2). It has been discovered that data sources vary considerably in their suitability for harvesting and synchronising with the registry and maintaining their accuracy. There are a number of databases and other data management systems being planned at Oxford and the BR11 project has been timely in that we have been in close contact with departments planning new systems to discuss the possible role of the BR11 registry in their plans.

Although the project originally aimed for depth of coverage, it was decided that, having achieved this with the first few sets of data, we should also include as much data as possible within the period of the project. This is to demonstrate a more (if not completely) comprehensive system. A comprehensive system was indicated by testers to be of great importance to potential users. As a result work will focus on broad coverage and quantity of data during the next phase.

The BRII Project Board was clear that the registry does not own (and therefore does not create any new) data. Data quality (and to a large extent, quantity) is therefore governed by data providers. Data providers are clearly indicated on the Blue Pages as is the means to add, edit or report incorrect data.

All harvested data have temporal data and provenance (ie validity) information assigned. This helps indicate the trustworthiness of the data, as it is clear when it was known to be true.

Co-reference of data is being handled by a combination of pre-set rules, heuristics or user-driven choices, and informed by work by researchers at Southampton (see <http://eprints.ecs.soton.ac.uk/17587/>). There is more work needed in this area.

The MSD has focussed on delivering RDF of researchers on a small subset of its websites. RDF delivery is now being 'switched on' across a broader range of sites.

5. Project management

The project has run relatively smoothly over its course. There was an initial delay in recruitment caused by a mixture of late notification of success of the bid and problems recruiting. The staff that were appointed have proved to be of high calibre and most importantly, have worked extremely well as a team amongst themselves and with existing staff. We lost one member of staff who went on maternity leave, but a replacement was appointed without much delay. The member of staff who left the group had developed the Blue Pages web interface to an advanced state before she departed, which left the project in a strong position.

The project team has benefited from a strong lead by the Project Board under the chairmanship of Prof Paul Jeffreys. Board members have been highly supportive, have taken a keen interest in the work of the project, and have given freely of their time to prepare for and attend board meetings. The Board has offered constructive criticism and advice to the project team and has taken responsibility for key decisions during the project period. Members of the board have been pro-active in assisting with dissemination and embedding of the project across the University.

Members of the project team, i.e. those with direct responsibility for running the project day-to-day, have met fortnightly since the start of the project (Oct 09) with few breaks. This has proved beneficial ensuring all members of the team have the opportunity to share experiences and developments, especially because of the iterative nature of the development. Responsibilities such as attendance at JISC events have been shared.

There was some concern during the project that staffing problems might cause the project to be considerably underspent. Because we had two 0.5FTE developers (instead of a single 1.0FTE) there was flexibility in re-assigning staff. These two staff were therefore instructed to spend full time on the BRII project from late 2009 to make up for lost time keeping the project on track, and to ensure spend of the DI staffing budget. Relevant conferences and other events were not available at suitable times during the project: dissemination therefore was arranged through other fora (papers and reports). The conference budget was then reassigned to other costs (including additional MSD staff time).

It became clear that there were few real commonalities between the BRII project and other JISC projects funded in the same strand. The BRII project has forged strong links with other, highly relevant groups, both running JISC projects: Bristol University (ResearchRevealed and EXRI-UK) and Southampton University (Dot.AC and RKB Explorer). Through the ResearchRevealed project workshops, the BRII project has been in contact with other relevant projects such as Readiness4REF at King's and Southampton. The sharing of technical information between these groups has been invaluable. The BRII project has also been involved at the national level with input to the EXRI-UK project and to the JISC meetings about Research Information Management for high level stakeholders.

The expectation at the time of writing is that the BRII project will complete on time, on budget and with all deliverables complete. The project can demonstrate additional outputs, not forecast at the outset.

6. Learning

We have learned about how to employ user testing and software demos as tools to disseminate the project and its outputs. These are useful to inform potential users about what we are doing and the benefits offered to them, but also to garner feedback from them which we can use to improve our products. We are aware of the snowball effect of our activities. Although these occur at a small scale, we are starting to reach more people by word of mouth. We found it beneficial at the start of the development process to working steadily taking time over interviews. The team needed time to get the technological foundations ready, develop the tools and tune them according to feedback. We have learned that this is easier if we reach a small number of testers but engage with them at a deep level. In this way we can produce a working beta version tool which can then be used for dissemination to wider audiences.

The registry and research information infrastructure including the harvesting processes and the APIs are rather abstract concepts which are difficult to grasp by common users. Because of this we have used the Blue Pages tests as a gateway to testing perceptions of the registry. Although some testers were concerned only about the Blue Pages some were interested in the registry and research activity data held there. They asked questions about further uses for these data which could be beneficial to them. From these we have gathered valuable information for (possible) future developments. In the MSD, it was felt important to demonstrate the use of the API, with proof of concept visualizations and we could, potentially, have devoted more time at the outset to constructing these.

We have learned much from networking with groups in other institutions. The BRII team has been collaborating closely with colleagues using similar technical models. There has been significant sharing of technical expertise and experience. We have also been involved in high level discussions about research information data exchange (JISC organised) and other research information management topics at a national level. There has been much to learn from the other participants of these discussions. We benefited from the assembly event we held (user buy-in) and to a lesser extent from other programme events.

Throughout the project and from contact with influential stakeholders we have learned that the process of embedding of the registry and its services will take time and thorough work. For that we are exploring avenues to approach this from a number of different perspectives. We are aware that the embedding of the registry will have to be carried out alongside promotion and work on the tools to improve its acceptance by users.

A response received as part of the user survey stated that *"The project has raised very interesting, and extremely important issues about research information such as what information is useful to capture, how to capture it, where to store, can it be made accessible, who is responsible for it, etc. I strongly believe this project has made an important contribution towards answering these and many other questions in Oxford and other HEIs."* This comment summarises many of the discoveries made during the project and touches on the attitudes of users. For example, one group was particularly sensitive about being compared with other units in respect of the primary data held on local websites. We did not expect the attitudes of some users towards availability of data over the Web (even data which is already freely available).

The issue of multiple requests for the same data is one that frustrates many users. We have learnt that although this is a key functionality that the registry, that message is not always easy to convey nor are all users convinced. On reflection we perhaps should have made more use of the webcast demonstrating re-use of data created by Anne Bowtell in MSD. Many users latched on solely to the Blue Pages as 'the' output of the project, however that only demonstrates part of the story.

Project Acronym: BRII
Version: 1.0
Contact: sally.rumsey@ouls.ox.ac.uk
Date: February 2010

Adoption by the University is proving to be much more difficult than had been anticipated. Part of the problem is ownership for the ongoing work when it does not sit easily within any overall group coupled with many of the issues presented being tackled only at local level. Oxford presents an extreme case of devolved governance.

It has become clear that to achieve institutional innovation within 18 months for a service which is not anchored within a single administrative unit is an impossible task. On reflection, this is more the type of project that needs up to 3 years to achieve such significant acceptance at a large institution.

BRII Project Team (PI Sally Rumsey) February 2010

Appendix B: BR11 Project Evaluation Plan

The evaluation will fit with the development cycles. Four formal evaluation strands will be undertaken:

1. Iterative development (throughout project)
2. Formative – mid-way (after first iteration c. summer 2009)
3. Summative – towards end of project (c. January 2010)
4. Project Management, lessons learnt – towards end of project (c. January 2010)

An external evaluator (Neil Beagrie of Charles Beagrie Ltd.) will run evaluation activities and provide reports for the project.

Timing	Factor to Evaluate	Questions to Address	Method(s)	Measure of Success
A. Throughout as part of iterative development (undertaken by Project Action Group)	1. Meeting the requirements of groups of end users	Do the R11 controlled terms meet the group's requirements	Interviews and close collaboration with end users throughout the process	Compare user needs as identified in requirements to outcome
B. Formative mid-way (after first iteration c. early summer 2009). Self-evaluation with external review and report)	2. Meeting aims and objectives	Has the project progressed adequately towards meeting its aims and objectives as would be expected at this stage?	Review (read & analyse) documentation (inc progress report) and questionnaire for self-evaluation by project staff to provide evidence	Meeting milestones and comparing progress to completed workpackage tasks
	3. First user needs analysis	Has the first user needs analysis been successfully completed?	Review (read & analyse) documentation	Have all relevant stakeholders been questioned and results recorded?
	4. Controlled terms for first iteration	Has a suitable vocabulary (taxonomy and controlled terms) & any mapping to existing taxonomies if relevant been achieved?	Review vocabularies and compare with documented user feedback	Outcome matches or exceeds user requirements
	5. Data harvesting	Is the data being successfully harvested from the first data source?	Focus group (project team)	Relevant data has been harvested completely and without corruption
	6. Method for core iterative process 7. Progress so far	Has the first implementation been successful? Has the work so far been successful?	Compare outcome with initial user needs Project team focus group with report	Outcome matches or exceeds initial user requirements Provide evidence
C. Summative towards end of project (c. Jan 2010).	8. Aims and objectives	Has the project achieved its aims and objectives?	Compare outputs and outcomes (read & analyse) with aims and objectives	Outputs and outcomes matches or exceed aims and objectives
	9. User satisfaction	Are users satisfied with the project outputs and outcomes?	Review user feedback (focus groups)	80% of user groups indicate satisfaction with the outputs and outcomes

	10. Infrastructure benefits	Are the project outputs and outcomes of benefit to end users and stakeholders?	User questionnaire or interview (asking about benefits eg Multi- and cross-disciplinary information discovery, easier research management, Easier research information discovery, Time saving, increased efficiency)	80% of end users and stakeholders agree that the project outputs and outcomes are of benefit to them
	11. Embedding	Will the outputs of the project continue to be used post-project?	Focus group or questionnaire	80% of users say they will continue to use the service
	12. Institutional innovation	Are there indications of institutional change as a result of this project?	Focus group or questionnaire with senior stakeholders	90% of senior stakeholders indicate that the work should be expanded and continued
D. Additional Summative question (Project Manager)	13. Project management	What lessons have been learnt about the successful execution of a project such as BRII?	Interview with Project Manager and other staff	Lessons learned about project execution have been documented

Appendix C: BRII Summative Evaluation User Survey Results

Introduction

BRII ([Building the Research Information Infrastructure](#)) is a JISC-funded pilot project at the University of Oxford started in Oct 2008 and completing in March 2010. The aim of the project is to enable efficient sharing of research activity data (i.e. data *about* research at the University of Oxford) from a selection of existing data stores and to create exemplar services that disseminate and re-use that data using a lightweight solution based on semantic web technologies. Half the project is devoted to stakeholder input, collaboration and 'buy-in' aimed at evolving current work practices and processes. A user survey and user focus group in the final stages of the pilot formed part of the evaluation plan for the project.

Summary of Survey Results

In total 20 users completed the survey (65% of those invited to do so) by the deadline: a very high participation rate for this form of survey, which also suggests a high level of interest in the project. In addition, one user responded after the deadline with email comments which have been included after free-text comments to question 7 in section 4 below. Although it has a high participation rate, this online survey group is still to some extent self-selecting (those who chose to respond) so careful attention needs to be given to the free-text comments both negative and positive as indicators of constructive criticism which might otherwise be under-represented.

The survey was designed to provide broad user feedback on the project and address the following issues and assess the achievement of associated project targets. Outcomes were as follows:

User satisfaction - Are users satisfied with the project outputs and outcomes?

- Target: 80% of user groups indicate satisfaction with the outputs and outcomes;
- **Outcome: 80% of BRII users in the survey were either satisfied (75%) or very satisfied (5%) with the outcomes of the project. 10% had no opinion, 5% were dis-satisfied, and 5% were very dis-satisfied** (all 20 users responded to this question see Q1 responses in section 4).

Infrastructure benefits - Are the project outputs and outcomes of benefit to end users and stakeholders?

- Target: 80% of end users and stakeholders agree that the project outputs and outcomes are of benefit to them;
- **Outcome: 68.4% of BRII users in the survey stated that the Blue Pages and pilot data store of preserved research activity data available for re-use would be of future benefit to them. A further 31.6% thought they would possibly be of benefit to them. 0% users who responded thought they could not possibly be of any future benefit to them** (19 out of 20 users responded to this question – see Q2 responses in section 4).
- When users stated these outputs would be of benefit to them they were prompted to choose from a set of pre-defined potential benefits for the Blue Pages and pilot data store (13 out of 20 users responded to this question – see Q3 responses in section 4). The aggregate responses were as follows:
 - Easier research information discovery [100%]
 - Multi- and cross-disciplinary information discovery [69.2%]
 - Time saving and increased efficiency [53.8%]
 - Easier research management [30.8%]

- Other [23.1%]

Embedding - Will the outputs of the project continue to be used post-project?

- Target: 80% of users say they will continue to use the service;
- Outcome: **47.4% of BRII users in the survey said they will use the key BRII outputs after the end of the formal project in March 2010. A further 36.8% thought they might possibly use them, 10.5% did not know, and only 5.6% stated they would definitely not use them.** (19 out of 20 users responded to this question – see Q4 responses in section 4).

Methodology

A survey questionnaire was developed by the evaluators in conjunction with the project team. The survey was made available on SurveyMonkey (an online survey website) and 31 users within Oxford who had seen demonstrations or undertaken testing on the BRII “Blue Pages” were invited via email to participate in the survey. The email invitation was as follows:

Dear Colleagues,

We are nearing the end of the formal [BRII \(Building the Research Information Infrastructure\) project](#). To gather evidence of perceptions of the outputs of the project I should like to ask you as individuals who have had some contact with the work of the project and who are potential end users of the registry and Blue Pages, if you would be willing to complete a very brief survey about the project. I should be extremely grateful if you would consider participating.

There are 6 very short questions so your time commitment should be minimal: around 5-10minutes in total.

The questions focus primarily on the prototype version of the “Blue Pages”. Should you need to look at this prior to completing the questions see the prototype version of the [“Blue Pages”](#) (available from within the Oxford network only – if development work is in progress [screenshots of the Blue Pages](#) are also available).

The survey can be accessed online at <http://www.surveymonkey.com/s/BRIIsurvey> and will be available from today to 6pm Friday 5th February 2010. Responses to the questionnaire are requested by Friday 5th February 2010 at the latest.

Your help with this is much appreciated

Sally Rumsey

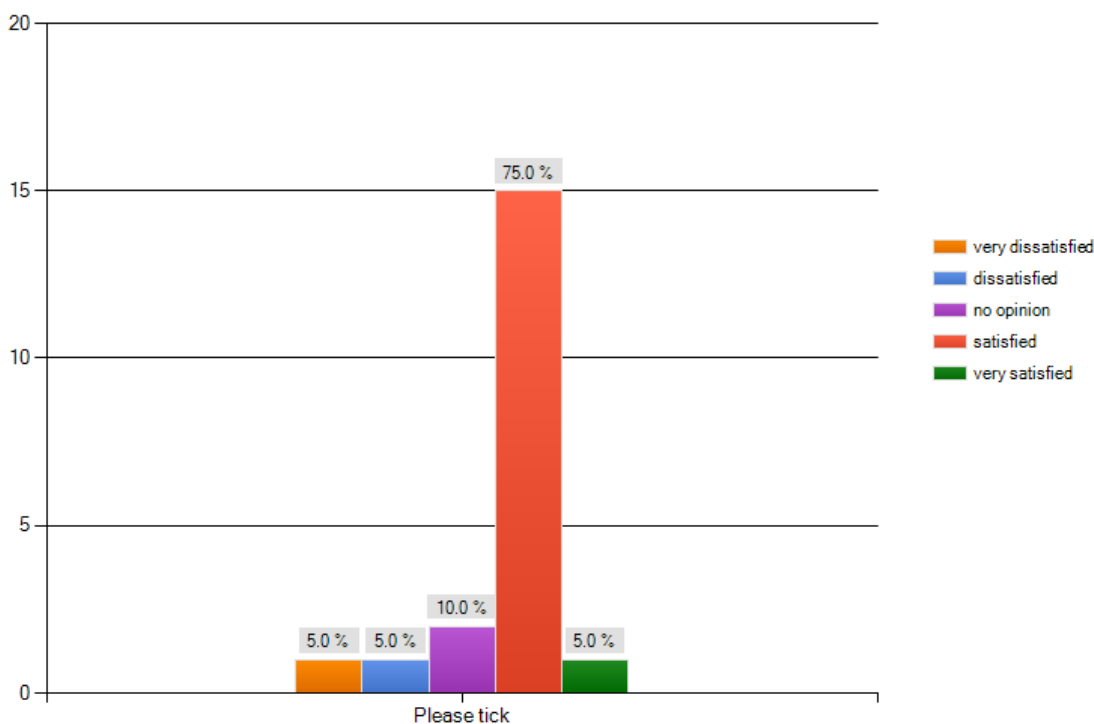
The questionnaire was available online from Monday 18 January 2010 to 6pm Friday 5th February 2010. Responses to the questionnaire were requested by Friday 5th February 2010 at the latest.

Individual questions and results

The results of the survey including any feedback in the comments are used in aggregate only and all personal data has been treated as confidential. The majority of questions were designed for easy completion and later quantitative analysis but users were invited to make use of the provision of free text fields associated with each question for additional comment and further details as needed. Details of the response rate, results, and any comments for each question are provided below.

Q1. answered question 20; skipped question 0

How satisfied or otherwise are you with the key outputs from the current phase of the BR11 project? (Note: the key outputs are the prototype version of the "Blue Pages" and the underlying data store of preserved research activity data that is available for re-use).



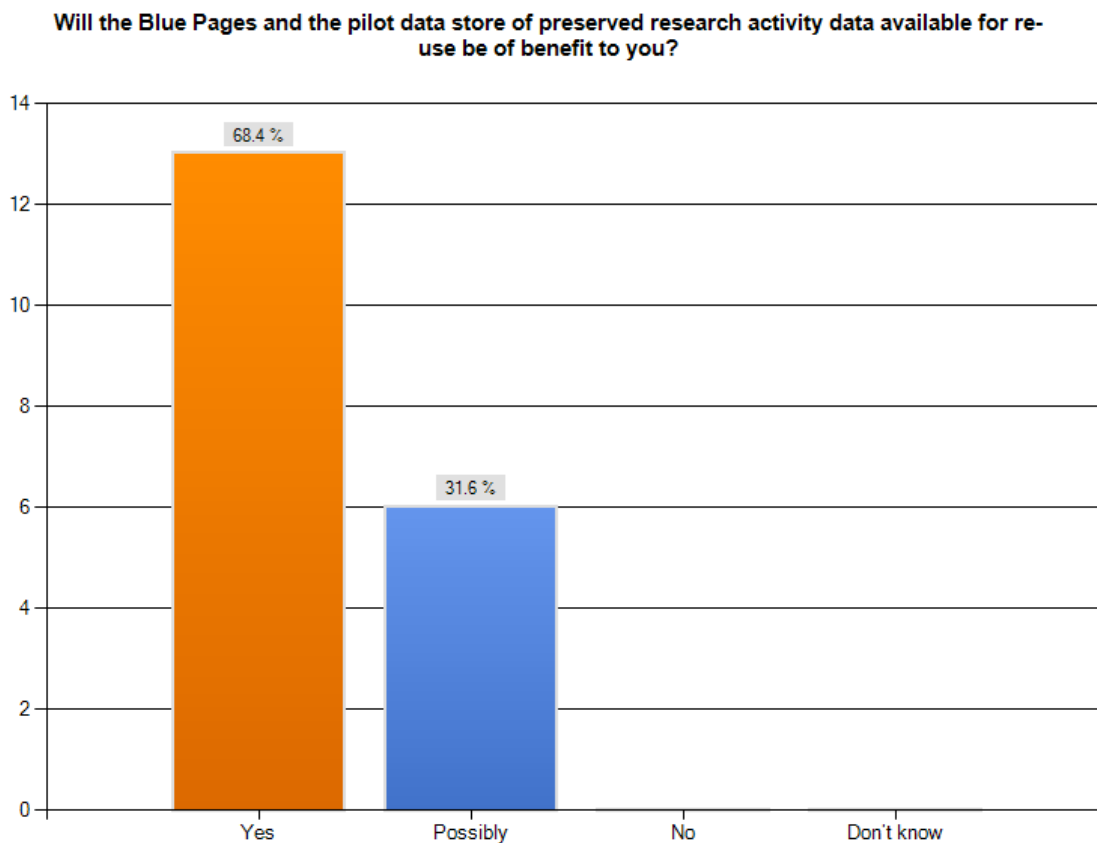
Comments?

"Recommended some changes to the evaluators"

"I feel constrained by the wording of the question and response options. The outputs are probably satisfactory within the expectations of BR11, but there seem to be questions about the utility of the outputs, who needs them, why, and their priority in relation to other demands on scarce resources."

“With my ORA hat on, I think the “Blue Pages” may be useful for checking where an academic unit sits within the University structure. Discovering collaborations between academics might prove useful too.”

Q2. answered question 19; skipped question 1



Comments?

“although the data for my department is not currently present, some useful data is already present (and clearly presented within the Blue Pages).”

“Would be cautious of replicating what is already done for research funders.”

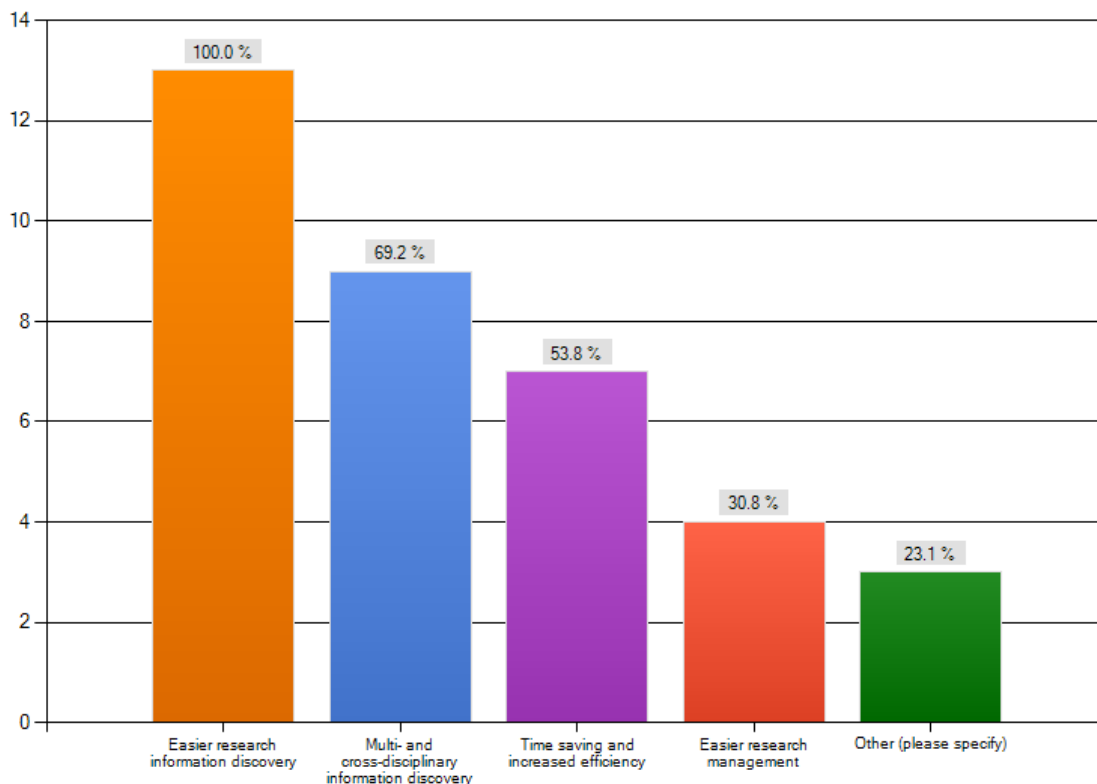
“Yes, but we need to be sure of it's reliability, accuracy and completeness.”

“Possibly limited benefit, but there are other currently unmet needs that may have higher priority.”

Q3. answered question 13; skipped question 7

If yes,

in which ways will these outputs benefit you?



Other (please specify):

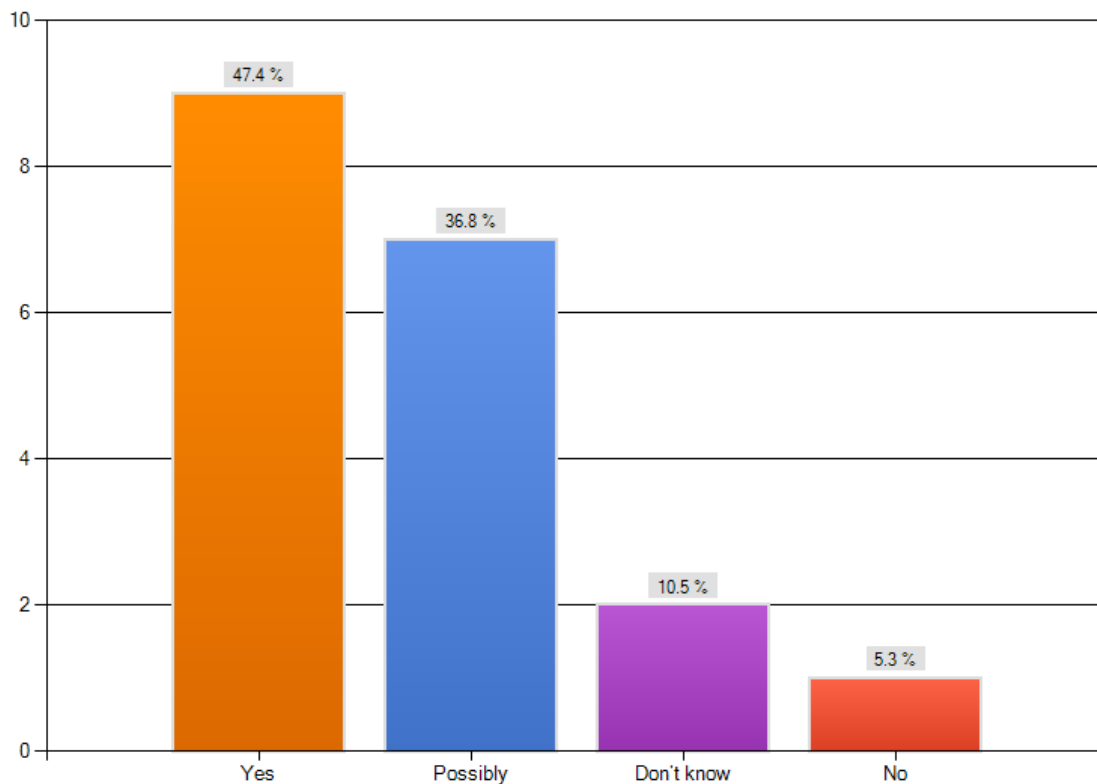
“A better understanding of what is needed in terms of research information infrastructure in a UK HEI.”

“allowing sources of outdated information to be identified and giving an idea of the visibility of the department within the university.”

“It will be easier to think of new ideas that could get funding by existing funders once we know what they are funding, the sort of projects they fund, and linking up with grant holders.”

Q4. answered question 19; skipped question 1

Will you use the key BR11 outputs after the end of the formal project (March 2010)?



Comments?

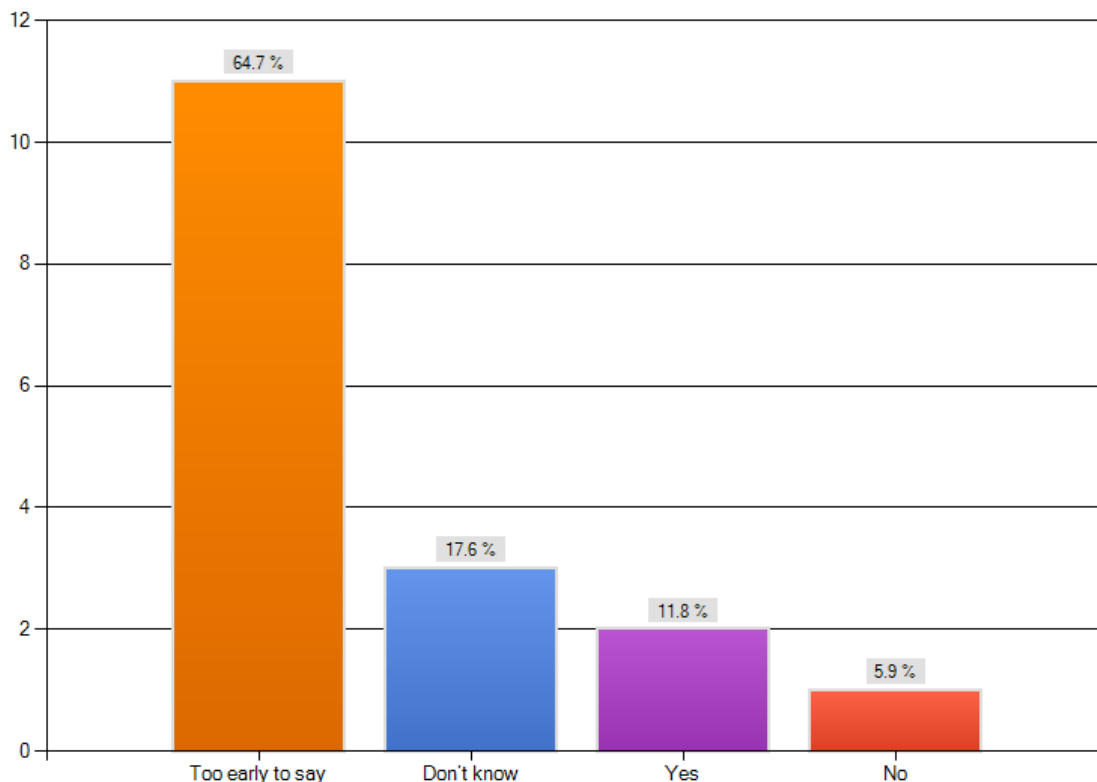
"if BR11 outputs refer to articles, it will be very good to access them and include them in teaching or further research activities."

"I hope the outputs will last beyond the formal project end."

"Probably not, as they do not relate to my area."

Q5. answered question 17; skipped question 3

Are there indications of institutional change as a result of the BR11 project?



Comments?

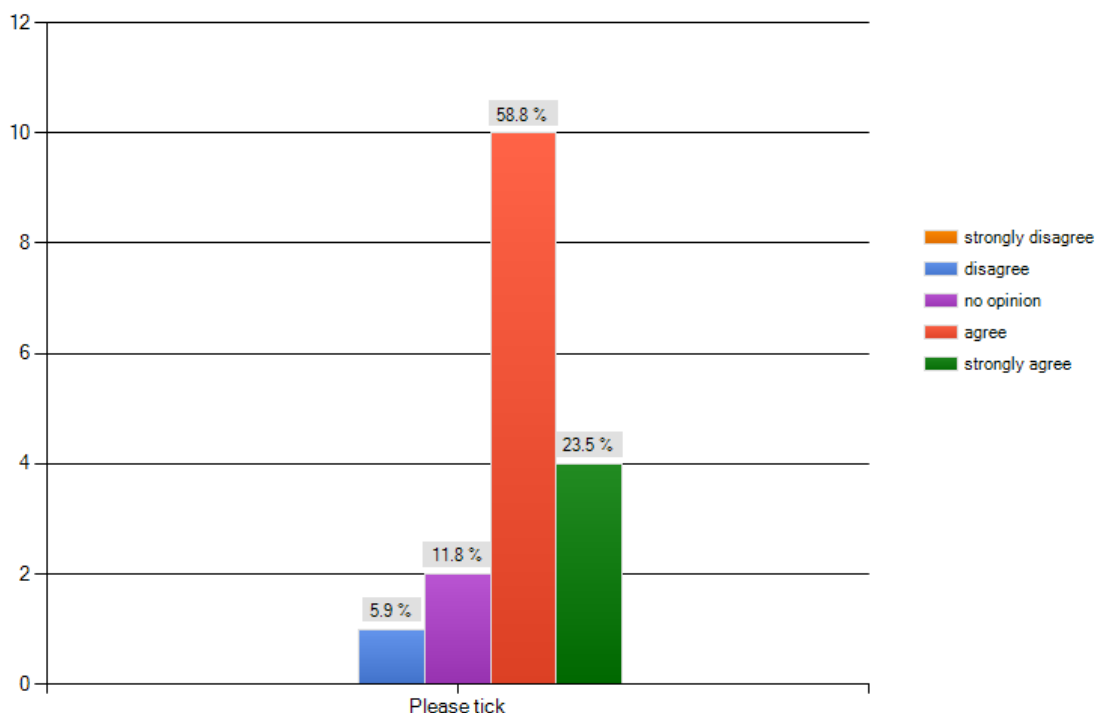
"Need to ensure that the data is constantly updated without users having to get involved. Content of this type is very dynamic, and a static instance would deem it useless within a short space of time."

"The 'local' staff web pages of units or departments normally contain people's profiles, publications and research. This information should be then reviewed to avoid duplication. My guess is that then BR11 blue pages will have to manage it and continuously update it."

"Not that I know of."

Q6. answered question 17; skipped question 3

How strongly do you disagree or agree with the following statement “The BR11 project has created the foundation for an important University service to aid storage, sharing and promotion of information about Oxford research.”



Comments?

“Data will only be as good as the ability of the originators to provide it. Need to consider also the root incentives to them enabling the availability of the data in the 1st place - a challenge that all departments still battle with!”

“I think in places that by harvesting information BR11 may be diverting resources that could be better spent in improving the data collection at the base level. Improving this foundation would be a better use of resources.”

“I don't have enough information to answer this question. It may have but I don't know.”

Q7. answered question 5; skipped question 15

As a user, are there any other comments you would like to add on the BR11 project as part of its evaluation?

Comments?

“The project has raised very interesting, and extremely important issues about research information such as what information is useful to capture, how to capture it, where to store, can it be made accesible, who is responsible for it,etc. I strongly believe this project has made an important contribution towards answering these and many other questions in Oxford and other HEIs.”

“External use is also an important part, meaning access for researchers from outside oxford. I wonder if an additional service of registering some interest in collaboration could be stored in the pages so that internal Oxford users get access to this information.”

“The blue pages could be enormously useful, but need high visibility. Could they be added to the staff gateway at www.ox.ac.uk/staff ?”

“I think it would greatly benefit the project if more consultation were carried out with academics regarding the intended usage of the BluePages.”

“Please remove the html tags in the returns of the search.”

Additional late comments outside survey via email for one user:

“ I like the new front page of the Blue pages, it's much clearer and more attractive than the previous version. I'm glad to see you changed the js that blanks out the default text in the search box to only respond to the default text, and not all text.

Just a little curious as to why the 'lt IE 7' conditional comments are in the body, and not the head, that's just a small technical note.

The revised presentation and addition of more graphics overall helps to present a clearer interface, which not only improves usability, but also confidence in the system, as it looks more professional.

I personally feel the breadcrumbs could be a little larger - I didn't notice it at first.

- The 'Research Interests' content for 'People' (e.g.

[http://163.1.127.171/search_result?search=The+Nuffield+Department+of+Cli](http://163.1.127.171/search_result?search=The+Nuffield+Department+of+Clinical+Medicine)

nical+Medicine) has it's html encoded. I'm assuming this is still in progress, but I thought I'd point it out.

From the look of it, HTML would probably break the page in this area, so I'd suggest stripping out all the HTML tags before the preview content is truncated and displayed.

I'd consider adding a max-width to the overall container. I have a large widescreen monitor which makes line lengths very long on a fluid layout without max-width unless I actually shrink my browser, which is a pain. This is a small usability issue.”