



Project Document Cover Sheet

Project Information			
Project Acronym	WRAP		
Project Title	Warwick Research Archive Project		
Start Date	1 October 2007	End Date	31 March 2009
Lead Institution	University of Warwick		
Project Director	Robin Green		
Project Manager & contact details	Jenny Delasalle University of Warwick Library, Gibbet Hill Road, Coventry CV4 7AL Email: j.delasalle@warwick.ac.uk Tel.: 024 7657 5793		
Partner Institutions			
Project Web URL	http://www2.warwick.ac.uk/services/library/main/research/instrep/erepositories/		
Programme Name (and number)	Repositories and Preservation Programme – Repositories start-up and enhancement projects		
Programme Manager	Andrew McGregor		

Document Name			
Document Title	<i>Final Report</i>		
Author(s) & project role	Robin Green, Project Director, and Jenny Delasalle, Project Manager		
Date	25 March 2009	Filename	WRAP_FinalReport_March2009.doc
URL	http://www2.warwick.ac.uk/services/library/main/research/instrep/erepositories/		
Access	<input type="checkbox"/> Project and JISC internal	<input checked="" type="checkbox"/> General dissemination	

Document History		
Version	Date	Comments
1	25 March 2009	

WRAP – Warwick Research Archive Project: Final Report

Table of Contents

Section	Page number
1. Executive summary	2
2. Background	2
3. Aims and Objectives	3
4. Methodology	4
5. Implementation	5
6. Outputs & Results	8
7. Outcomes	10
8. Conclusions	12
9. Implications	12
10. Recommendations	13
11. References	14
12. Appendices	15

Acknowledgments

External funding for the Warwick Research Archive Project (WRAP) was provided by the JISC as a repository startup project under its *Repositories and Preservation Programme*.

1. Executive summary

WRAP's aim was to capitalise on the learning from early adopter institutional repository projects and build a repository for the University of Warwick that would further develop understanding of how repositories can meet the needs of their stakeholders.

Key objectives were to implement a repository for preprints, postprints and theses with the EPrints Open Source software, using SWAP, the Scholarly Works Application Profile, and the EThOS EPrints OAI plugin; to explore the potential for interoperability of the repository with other campus systems; to develop an infrastructure to receive around 350 theses a year; to attract participation by a range of departments and researchers; and to achieve a corpus of over 1500 items by project end.

IT Services, Library and Research Support Services staff comprised the project team: the Library as lead partner, with IT Services hosting the repository software and providing the technical expertise to integrate WRAP with the campus search engine, and Research Support Services providing linkage between WRAP and its Expertise/My Profile system (which contains information about Warwick researchers, including publication details).

The project team worked with a small set of pilot academic Departments in the earlier stages of the project in order to assess effective advocacy approaches and to obtain content prior to the repository being made live. Additional Departments were included as the project progressed and awareness of WRAP increased through promotion at meetings of academic groups, University committees and internal events.

WRAP has been implemented as a full-text only repository to expose and emphasise Warwick research excellence. The project team also adopted a mediated submission process combined with creation of high quality metadata, including full Library of Congress Subject Headings, to maximise discoverability, interoperability and provide future-proofing.

Although content submission levels are very good technical delays when implementing SWAP with EPrints, combined with the time and effort required to create high quality metadata, have significantly impacted on record creation and ability to meet the initial volume target. However, as the project closes WRAP has achieved many of its aims and objectives, including its primary purpose of providing a repository service that is embedded within the institution. In particular, WRAP has implemented the SWAP metadata schema with EPrints software, obtained an institutional mandate for submission of e-theses, developed procedures for populating the repository, integrated WRAP with the campus search engine, and created a mechanism for transfer of content to and from the Expertise/My Profile system. In addition, the WRAP team has both contributed to and made good use of the range of initiatives and shared infrastructure services sponsored by JISC to support IR developments. Learning from and sharing this experience has been a large part of the project, making it a valuable initiative not only for the University of Warwick but also for the wider repositories community, both in the UK and abroad.

2. Background

2.1 Institutional

Although the University of Warwick is one of the UK's leading research institutions, prior to the Repositories Startup and Enhancement call it had not implemented an institutional repository. The JISC call was seen as an opportunity to raise the University's profile and support the wider research community by improving the availability of Warwick research. At the same time an updated University Strategy, inspired by the arrival of a new Vice-Chancellor, began a move to expand the University's research base and generate new research ventures as well as further develop international collaborative activity, with the intention of establishing Warwick as one of the top 50 institutions worldwide by 2015. The University's senior administrative and academic management viewed an institutional repository as a means to help achieve this aim.

2.1 Library

From the Library's perspective there were a number of additional drivers to implementing a repository. These included: recognition of a need on behalf of the institution to host and make available the University's research and other intellectual assets (it is estimated that well over 2000 research papers alone are published annually); support present and future RAEs and other administrative requirements such as annual review; the Library's participation in the EThOSⁱⁱ project resulting in the availability of digitised theses; and developing Warwick's digital library.

The Library had indicated in its Strategic Plan 2003-2008 its intention to take a more active role in supporting the institutional mission and to provide value-added services, particular around engagement in the research process and the development and management of the University information environment. Implementation of an institutional repository was a specific goal stated in the Strategy.

3. Aims and Objectives

The aim of the Warwick Research Archive Project was to establish, populate and promote an interoperable open access institutional repository for the University of Warwick. This repository would house, manage, preserve, expose and disseminate research outputs produced within the University in order to support the institutional strategy and increase the accessibility of research outputs for the benefit of the wider community, as well as the timeliness with which they become available.

The original intention during the project lifetime was to focus on preprints, postprints and theses. As the project team engaged with researchers in pilot and other departments it was recognised that despite increasing publisher willingness to accept deposit of preprints, researchers in many departments are unwilling to do so. This is for a number of reasons, including the likelihood of significant change between pre- and postprint (and therefore the preprint's lack of authority) and a perception that its lack of status means reduced security of IPR. This unwillingness was negatively impacting on success of advocacy activity and as a result the team concentrated on postprints and theses, except where a discipline (such as Economics with its use of Working Papers) supports the concept of preprints.

The project would also assess the potential for interoperability with local and appropriate national services such as Warwick's Expertise system (InfoEd) and the UK Federation in order to support the embedding of repositories within institutional infrastructures and as tools for promoting the UK research endeavour. Work with the Expertise system and University's search engine took place but a policy decision was made that in the interests of accessibility the repository would not introduce authentication so now activity around the Access Federation has taken place.

Specific objectives were to:

- Commission the WRAP repository service
- Acquire a critical mass of participating departments and content that would demonstrate WRAP's capability of supporting the institutional mission
- Promote self-archiving amongst participating Departments and raise awareness of its benefits across the University in order to establish culture change in management of research outputs
- Implement a local infrastructure capable of handling approximately 350 theses a year, including the introduction of an institutional requirement for submission of theses in electronic format
- Identify issues and requirements relating to interoperability with local and other systems
- Participate in the JISC-funded network of digital repositories and wider repository community
- Disseminate through the project Web site and other means the knowledge and experience gained from WRAP
- Identify the resources required to sustain continued management and operation of the repository and to extend its remit to other types of content and services.

4. Methodology

4.1 Summary

WRAP adopted EPrints as the preferred repository software tool. A repository for full-text journal articles and theses was implemented using EPrints configured with SWAP and UKETD.

A mediated submission process with a single and simple to use submission form was introduced to minimise researcher effort in depositing material in WRAP. WRAP staff expand the basic information entered by authors, check publisher restrictions on submitted works, and liaise with individuals and Departments as appropriate. High quality metadata was created for each record using SWAP's flexibility and applying Library of Congress Subject Headings.

Project staff and Library Subject Support teams worked with an initial set of 4 pilot departments to explore advocacy methods and obtain content for the repository prior to its live release. As the project progressed further departments were added.

Various methods of obtaining content were employed, including monitoring Zetoc alerts in order to contact authors who have recently published.

To obtain thesis deposit the WRAP team worked with the Board of Graduate Studies and Office and Graduate School to introduce a mandate for submission of theses in electronic format as well as print, and to devise procedures for managing such submissions.

The JISC Repositories and EPrints communities were used to both seek and provide guidance as required.

4.2 Methodology in Detail

The project was undertaken as a collaboration between three core academic support services: the Library as project lead, Research Support Services (who manage the InfoEd/Expertise system that administers grant processes and researcher profiles, including details of publications), and IT Services, who have provided hosting and technical support for the repository software and development work on interfacing with the local search service.

WRAP has used the EPrints open source software developed at the University of Southampton. EPrints was selected as it offered a mature and stable solution that would support the different types of material that would potentially be deposited in a repository, had a record of ongoing development led by Southampton (complemented by an active user community), and was the most widely-adopted institutional repository software across peer institutions to the University of Warwick within the Russell Group. It was anticipated that the EPrints community would provide advice and other support with technical issues or developments.

The adoption of appropriate standards was essential to ensure the repository was able to interface with the broader information environment. The following standards and protocols were adopted: Dublin Coreⁱⁱⁱ; OAI-PMH^{iv}; UK Metadata Core Set for ETD (UKETD) to enable interaction with EThOS, the new British Library-managed national e-theses service; and SWAP, the Scholarly Works Application Profile^v, a new FRBR-based Dublin Core application profile specifically designed via UKOLN for the description of scholarly works and research papers. This was the first implementation of SWAP in a live service. In addition, Library of Congress Subject Headings were created for each record. The use of SWAP together with controlled subject headings enabled high quality metadata creation in order to:

- Improve discoverability through increased routes to content
- Support collocation of results
- Enable resources in WRAP to be discovered by, and used in, other services, to give Warwick research the maximum possible distribution and visibility
- Ensure greater interoperability across University systems

- Proof against future requirements, e.g., presentation of content within a vertical search tool such as Innovative's 'Encore', or service migrations that we may undertake.

Evidence from many earlier repository implementations^{vi} has shown that researchers' focus is on their core activity and that any perceived administrative barriers, including length of time required to make a repository submission, will deter submissions. In order to minimise this effect - and to ensure quality control of metadata and alignment to standards - WRAP adopted a deliberately simple and heavily mediated approach to submission:

- Provision of a standard electronic submission form with the following fields: staff/student number, bibliographic data (one field, allowing for simple copy and paste, version status of the work, indication of acceptance of a submission agreement, DOI if known, and 'any further information', which included a prompt for funder and grant information where applicable)
- Checks on publisher policies on copyright and self-archiving through SHERPA ROMEO and other sources by WRAP staff
- Liaison with publishers regarding policies on behalf of authors
- Addition by WRAP staff of Library of Congress subject headings to enhance metadata.

Metadata staff working on WRAP were embedded within the Library's Data Services Team as a model for future activity - it is anticipated that recording and curation of the institution's intellectual assets will be of increasing centrality to their role.

The OpenDOAR Policies Tool was used as a basis for defining a set of policies for the repository.

The project team worked with an initial set of four pilot departments and extended to others at specific points during the project. This approach was taken in order to acquire initial content to populate the repository before it was made openly accessible, and to enable the team to gain an understanding within a 'controlled environment' of researcher perceptions of self-archiving and how best to carry out advocacy work in this area. New departments were added shortly after the repository went live to increase institutional awareness and content flow, and at later points as workflows permitted. As promotion of the repository took place submissions were encouraged from researchers across the institution in order to begin to embed the repository in their publication processes. At the same time work was undertaken with the Graduate School to implement an e-thesis submission mandate (this was introduced in October 2008) and develop procedures for submission.

To ensure sufficient critical mass could be attained if obtaining material directly from researchers proved difficult the project team identified pre-existing internal and external sources of Warwick content, including a departmental repository in the University's Centre for Scientific Computing, the arXiv and BioMed Central repositories, and so on (however, the flow of submissions has been such that these sources have not yet been drawn upon. This will happen post-project).

The project team met frequently in the early part of the project until the repository had been implemented and was stable. As planned, frequency reduced from that point. To maintain engagement at institutional level project progress and issues were reported regularly to appropriate committees and to an oversight group of administrative and academic staff chaired by a Pro Vice-Chancellor for Research.

5. Implementation

Project work was planned and mapped into work packages throughout the course of the project. A project team was established and met regularly to track the course of the project through the workpackages, and work towards the delivery of planned outputs. Throughout the course of the project it proved necessary to adapt plans in their timing and specific details, and the team's combined expertise was useful when responding to the challenges that arose.

5.1 Technical delays

Configuring the EPrints software was a much larger hurdle than expected. There were substantial delays owing to our determination to implement SWAP, which required much editing of programme

files. Whilst 'basic' EPrints implementation was perceived to be manageable, this additional requirement went beyond local technical knowledge, and there was little EPrints documentation on this area regarding which files would alter what functions of the repository. After a number of failed iterations the original intention to cope internally with the technical aspects of the implementation had to be adapted and we subscribed to the support service offered by the EPrints Services team in Southampton (EPrints has since been upgraded to make such configuration possible via the user interface rather than the command line). The repository eventually went live in July 2008, several months later than planned.

5.2 Staff recruitment is not easy

Another delay experienced was in the recruitment of metadata staff. A nationally-placed advertisement for a part-time project Metadata Librarian resulted in no suitable applications. It eventually proved possible to make internal arrangements to extend existing part-time staff hours. Internal recruitment of a Repositories Assistant was fortunately also managed internally, and as a development opportunity. Both appointments were later than originally planned owing to the delay in achieving a working repository.

5.3 Metadata quality versus speed of growth

Once metadata record creation began it became clear that the highest quality SWAP took time to produce and it became necessary to refine our processes. A significant amount of time was invested in the inclusion of bibliographic references in the SWAP metadata. Initially the Metadata Librarian re-formatted the references from every article into the Harvard style. This would be the most useful way to store the metadata, presenting consistent and high quality data to searchers and enabling further technical developments to process the data and make further use of it, such as by turning the references into links to those works online. However, the re-formatting of hundreds of references - whether done by professional or support staff - turned out to be prohibitively expensive in terms of staff time and could not be continued. As a compromise solution references are still included in WRAP metadata records, but they are left in whatever format they are presented in the original full text item.

WRAP metadata also includes Library of Congress Subject Headings for each item. The creation of these subject headings requires the involvement of professional staff who understand the schema and are able to assess a document's subject content and assign appropriate headings. This process does take a considerable amount of time, but it is the inclusion of these subject headings that will provide the flexibility to identify items in WRAP that are suitable for sharing with other central repositories, and to allow subject-based cross searching of WRAP with other quality metadata resources such as the library catalogue.

5.4 Processing submissions

Prior to creation of the metadata described above the Repositories Assistant checks submitted items against the terms of the publisher's copyright agreement to ensure that the author has sent a document version that can be included in the repository. The Assistant also checks the basic details of the metadata, and begins the record creation process.

A best practice issue investigated as part of the project was addition of cover sheets to the documents themselves. We became aware that repository documents discovered through Google are presented without their metadata, and were concerned that documents submitted to WRAP as early drafts often contained little or no information about the author, the publication, or the University of Warwick. It was decided that cover sheets were the only way to remedy this situation. The JISC-Repositories list was used to ascertain what other repositories did and why in relation to cover sheets. This informed the development of a suitable procedure for WRAP. A summary of findings was made available through the list and a briefing sheet was later produced on behalf of the Repositories Support Project (see Appendix 1 below. The published version is now available on the RSP Web site at http://www.rsp.ac.uk/pubs/cover_sheets_web_advisory_doc.pdf).

A general approach - which proved very productive - of determining community views through the JISC-Repositories list, of maintaining awareness via that list of what other repositories were doing,

and visiting the live repositories themselves, was used to investigate all the major issues that arose throughout the project.

The processing and metadata record creation time has been optimised such that circa 20 items are now fully processed regularly each week.

5.5 Advocacy

Advocacy activity necessarily had to be adapted in the light of technical delays described above. Rather than openly marketing WRAP before a live repository and robust processing and cataloguing procedures were established, staff in the pilot Departments were written to individually. This gave the control to ensure that throughput could be managed and author expectations regarding processing speed could be met, as well as handling enquiries and concerns in a way which best suited those whose co-operation the repository depends upon, i.e. at the individual and personal level.

An initial briefing was prepared for each department, and presentations were given at departmental meetings, with the key message being that authors should keep appropriate versions of their journal articles to submit them to the repository. Slightly different attitudes and issues were discovered in each of the departments, and these were responded to accordingly.

Having gained support from the Head of Department in each of the pilot Departments, time was invested in looking at information that authors had already made available on personal Web pages about their publications, and in writing to ask them for specific versions of their five most recently published journal articles (having already checked the publishers' policies for those articles). Where a final version could be deposited, authors were simply informed that we had done so on their behalf. These e-mail invitations proved very effective in prompting authors to submit not only the items asked for but also, in some cases, more of their work.

The process of checking copyright policies in advance of requesting specific items has the following drawbacks:

- Authors have not always kept appropriate versions and so could not submit them even when willing. Looking on the positive side, this was valuable in raising authors' awareness of and the benefits of retaining versions of works!
- Some authors did not respond at all, and time in checking their items could perhaps have been better spent. It is for these reasons that only the five most recently published articles were checked, and the general approach only used in selected departments.

Another means of generating content has been the use of Zetoc alerts. This mechanism is managed through members of the Library's Subject Teams and is predicated on their knowledge of the academic staff in relevant Departments. A list of authors by surname and initial is set up as alerts on Zetoc. Each member of staff monitors the names for one department only, and is therefore able to delete alerts for articles in the wrong subject area. Articles on appropriate subject areas are investigated further to check they are alerts for journal articles rather than other types of content such as book reviews, and that it is indeed the right research area for the Warwick academic of that name. Alerts for items meeting the required criteria are forwarded by the Academic Support team member to the WRAP team, who then write to the author. At this particular point in the publication cycle time authors are likely to still retain the all-important postprint version, and roughly 20% of requests resulted in a deposit in WRAP. Zetoc alert monitoring for non-pilot Departments stimulated a number of enquiries and interest in WRAP beyond article deposit.

In addition to Departmental briefings, individual invitations and Zetoc alert invitations, authors with material in WRAP were informed of the URL(s) of their item(s), and the RSS feed URL of a search by their surname, to maintain awareness of WRAP and prompt them to continue to submit content.

The WRAP submission form includes a check box for authors to select to be sent more information. Those who signed up were sent a Christmas bulletin with information about visits to WRAP and the most viewed items in the repository. The most popular authors were informed early in the new year of the number of views of their works; copies of these messages were sent to their Heads of Department.

Up to February 2009, in addition to initial pilot Departments:

- Individual authors in 4 Departments have been contacted to request postprints of their most recent journal articles; other Departments have been primed to expect this
- Zetoc alerts for 6 Departments are being monitored, with 2 further Departments identified for inclusion
- Briefings in a further 5 Departments are underway.

The project team plans to increase Subject Teams' awareness regarding the importance of the Zetoc alert monitoring, and to investigate other alerting services such as PubMed for the Medical School.

The project plan had assumed identification and early ingest of Warwick research already held in repositories such as arXiv, BioMed Central and a local service managed Centre for Scientific Computing in order to meet the content target; however, considerable resource will be required to achieve this and as the project progressed it became clear that in order to attain steady and sustainable growth WRAP had to be embedded within authors' processes. As this could best be achieved by putting resource into working with authors to acquire current material, the decision was taken to obtain material already securely available through other sources at an appropriate point in the future. This has impacted on the volume of content in WRAP (as has the major delay in delivery of content and metadata from EThOS) but the project team remains confident that this was the correct route to follow.

5.6 Theses

The University's Graduate School has agreed to add electronic versions of theses to its process for print thesis acceptance to ensure WRAP acquires a copy of every thesis at the time of submission, emphasising to students the importance of submitting a copy identical to their final print version.

A WRAP consent form has been included in this process, designed to take into account possible requirements for restricted accessibility, primarily where third party copyright content is a factor. WRAP also accepts standard University embargo periods where imposed.

Guidelines have been issued for students and supervisors regarding the new deposit requirement and dealing with copyright material in relation to WRAP, together with information of the benefits that WRAP could achieve on behalf of authors. Theses are coming to us through this traditional route, which is a slow but steady source.

6. Outputs and Results

6.1 General

A successful, fully standards-compliant live WRAP repository is the primary project output, together with the first implementation of SWAP.

By end of project WRAP had 500 records (more statistics in Appendix 3), with almost all records containing full text files. The diagram below (Figure 1), taken from the ROAR^{vii} website, shows the growth of the WRAP repository in the preceding months. Growth is now steady and consistent.

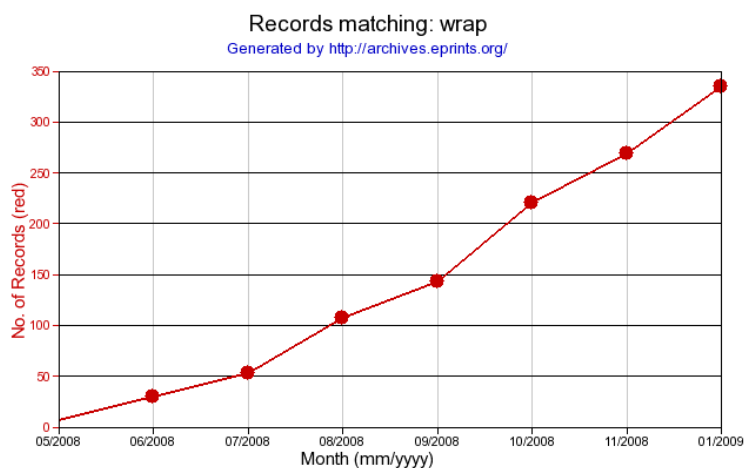


Figure 1: The growth of items held in WRAP, from May 2008 to January 2009.

Owing to technical difficulties in implementing EPrints with SWAP (detailed in the Project Progress Reports) together with the length of time taken to provide high quality metadata for each record the project target of 1500 items will not be achieved; however, WRAP is now a robust embedded institutional service.

A highly mediated workflow process has been implemented. This includes a single low-demand submission interface. Feedback from researchers has been positive.

A University mandate for submission of theses in electronic format was introduced in October 2008, and work with the Graduate School has developed a process for handling the c.350 theses submitted annually. The first submissions are being received by the Graduate School and are testing how well staff on the ground have understood this.

Integration with the University's Lucene-based search engine has been achieved, so that results from a search of the University Web site include results from WRAP – with a sometimes unexpected co-location of source types.

Exposure of the repository for harvesting has resulted in the majority of visitors to WRAP arriving via search engines: mostly from Google, with the second highest source of visits from within the Warwick Web site. From August 2008 to December 2008, there were 10,000 visits from 125 countries/territories, with just over half of the visits coming from the UK. (See Appendix 3 for more statistics.)

A bilateral workflow has been developed with the University's Expertise system. Although to date it has not been possible to fully automate this process, submissions to WRAP are passed to Expertise in spreadsheet form for easy upload, and Expertise have provided a link from their user data entry page to the WRAP submission form, so that researchers can begin a submission process from either service. Further development work remains to be undertaken in this area.

WRAP has enabled the University, as a NEREUS project partner, to contribute records to the NEEO (Network of European Economists Online^{viii}) disciplinary repository and thus take an active role in the consortium's agenda to facilitate access to the key research outputs from the top European Economics departments.

6.2 Documentation

The project has produced two Progress Reports. These detail a number of findings, issues and implications relating to the repository itself, SWAP implementation, the project methodology, impact of a mediated approach on resourcing requirements and implications for sustainability, and discipline-based advocacy with a multidisciplinary academic community.

The Project Manager produced a briefing sheet on cover sheets in repositories, and the implications of using them, on behalf of the Repositories Support Project. (Appendix 1)

A document describing the modifications necessary to implement SWAP with EPrints has been shared with those who have enquired about our implementation. (Appendix 2)

The WRAP Web site provides information for external and internal access, explaining about the project and its activities generally, but also providing FAQs and advice, e.g., regarding third party copyright, open access, journal article versions, and so on. From August to December 2008 there were around 9000 views for these pages.

6.3 Web 2.0

A WRAP project blog^{ix} was created and maintained throughout the project. There has been engagement from the community through visitor comments on the blog, and positive feedback about its usefulness.

The Project Manager has also saved numerous repository-related bookmarks on del.icio.us and faves.com, and joined the Repositories Support Project Facebook group.

6.4 Events

WRAP staff organised a JISC RPPROG meeting 3 December 2008, jointly with the University of Northampton, on the timely theme of demonstrating and exploiting repository value. Speakers were approached to cover the theme of repository statistics and their usefulness in advocacy work. Further details of the event are online^x.

A poster presentation was given at the Open Repositories 2008 and E-Prints User Group workshop, Southampton, 1 - 4 April 2008^{xi}.

6.5 Other

Experience gained from the project, and the results of our enquiries about other repositories, have been shared through the JISC mailing lists, particularly the JISC-Repositories list where the Project Manager Jenny Delasalle posted on numerous occasions, was a maternity cover list-owner of LIS-Copyseek for 12 months, and other relevant lists such as JISC-RPPROG-SUE, JISC-RPPROG and UKCORR-discussion. Team members have also contributed to the repository community through attendance at numerous events, including those offered by the RSP.

Publisher policies have been investigated throughout the course of our processing work, and details of these have been shared with SHERPA wherever possible. ROMEO is a critical service for the UK and wider; the more complete it is the more useful, and crowdsourcing contribution from the community is important.

The team has, like others, responded to direct e-mail enquiries from other repository managers, undertaken telephone interviews with publishers and UKOLN staff and facilitated visits. WRAP has also actively participated in relevant surveys such as the recent Research Information Network^{xii} exploration of how funder and grant information is handled.

7. Outcomes

WRAP has achieved most - but not all - all its stated objectives.

The repository is live with SWAP and UKETD implemented. As the first site to implement SWAP with EPrints the project has been able to provide valuable feedback to UKOLN, EPrints and the repository community.

The project has engaged with ten Departments so far and its level of internal exposure (the second highest proportion of visitors to WRAP come from within Warwick) has led to it receiving submissions from researchers from across the University. Researchers are beginning to understand its potential benefits to enhance their profile^{xiii}, and it is anticipated use of such 'best practice' and promotion of the service – e.g., dissemination of statistics, lists of 'popular' papers, etc. – will lead to a cultural shift towards self-archiving.

For WRAP, working with an initially small number of pilot Departments and expanding this set at appropriate points has proved a successful model. Preparatory work with the pilot Departments showed that tailored advocacy approaches and messages would be necessary to take disciplinary differences into account, and the perspective of each discipline on the scholarly publication process and open access publishing had to be understood prior to open discussions. The managed expansion also gave the ability to control volume of submissions when workflows were being refined once the repository was live.

An innovation part way through the project has been to use ZETOC alerts, actioned by the Library's Academic Support teams, to identify new publications by Warwick researchers in pilot Departments. If these have not already been submitted to WRAP then authors are approached for postprint versions. The closeness to publication means that authors are still likely to have these (a major challenge in trying to obtain older publications). This has proved a useful additional aid to obtaining content and raising awareness not only of WRAP but of rights transfer implications.

There is already increased researcher awareness at a personal level of the impact of rights transfer through dialogue with the WRAP and Library Academic Support teams and others over restrictions on placing their content in WRAP, and alternative approaches such as the 'author pays' publishing model. This links to other work in this area at Warwick: WRAP is acting as a catalyst in stimulating discussion around scholarly communications, funder mandates, IPR and Open Access issues. It is anticipated that an outcome will be that researchers publish more in Open Access journals, as a number of researchers have requested an open access publishing fund be instituted.

In October 2008 the University introduced a requirement that theses must be submitted in electronic as well as print versions. This will enable the University to participate fully in the national EThOS e-theses service and indicates the support WRAP has within the University.

Interoperability with other systems has been explored and WRAP can now be searched using the University's search engine. It has not proved possible so far to provide automatic updating between WRAP and the InfoEd system but structured and checked data is regularly fed through from WRAP to the team managing the My Profile module, and there is a link from the My Profile contribution page directly to WRAP to facilitate submission. Work on this will continue beyond the project as Warwick is likely to use the My Profile module as the institutional Publications Database, and outcomes will be made available to institutions using both InfoEd and EPrints.

Two factors have seriously impacted on the volume of content held in WRAP: technical difficulties in configuring SWAP with EPrints delayed the repository's launch by six months, and the combination of using SWAP and Library of Congress Subject Headings to build high quality metadata for each record means each bibliographic record takes over an hour to create. Together these have restricted the number of items added to the repository to just over 400 rather than the target of a minimum of 1500. However, this has provided an understanding of the level of resource required to sustain the quality of record that will be necessary if repositories were to become included in future research assessment frameworks.

The project has also surfaced 'hidden' institutional intellectual assets, from reports to funding bodies and other grey literature to learning objects to research data. These have various origins, e.g., they may be outputs from past projects or research studies, but are held by individuals, groups and Departments, perhaps on individual PCs, and are potentially at risk. A centrally-provided infrastructure, perhaps based on additional repository development, could offer more coherent and efficient management, preservation and promotion of such assets. This issue has been raised at University level for further consideration.

WRAP activity and advocacy has stimulated considerable discussion over scholarly communications generally and the open access publishing in particular, both in the institution's academic and administrative communities, resulting in the project team becoming involved in issues such as a requirement for a central open access publishing fund, the potential for overlay journals or a scholarly digital press, and so on. This has implications for time and effort devoted to core repository activity, the need for awareness of scholarly communications issues, and more broadly the developing role of the Library as an institutional service provider.

As noted under Outputs above, the Project Manager has made a strong contribution to the wider repository community, including production of a briefing sheet on cover sheets in repositories, and the implications of using them, on behalf of the Repositories Support Project (Appendix 1). The Project Manager has now been elected Chair of the UK-CoRR (UK Council of Research Repositories^{xiv}). The Project Director has participated in focus groups and meetings to support consultancy work to develop a JISC Repositories Recruitment Toolkit^{xv}.

The sustainability of WRAP following project end was initially addressed in a paper to the University's Steering Committee in October 2008, and funding has been allocated to recruit a Metadata Librarian when JISC funding ceases. A Working Group chaired by a Pro Vice-Chancellor for research is considering practical and resourcing requirements for WRAP together with the implications of developing a University Publications Database to support future REF requirements.

8. Conclusions

Implementing an institutional repository is not a closed activity but will surface a range of issues that have implications for institutional policy, e.g., should an institutional mandate to deposit be introduced or promulgated; is there a responsibility to host centrally, as an institutional asset, the range of different types of material created and held (and potentially at risk) at individual or departmental level; should there be institutional support for Open Access, e.g., creation of an Open Access publishing fund, and so on.

Implementing a repository will also have implications for the role of the Library and its staff. The repository cannot be embedded in the research process without the involvement of subject support (and other) staff. However, their engagement in both high-level and very practical scholarly communications matters with researchers will require new understanding of a complex and evolving area, and new skillsets. This may in turn be a driver for structural change.

There is a tension between willingness to take on the role of an early adopter, taking the implications of this on faith, and the ability or commitment to fulfil that role, with the unanticipated challenges or costs that might bring. This is not an argument against taking on that role but those doing so should consider it carefully and be prepared to make that commitment.

9. Implications

Working with researchers to implement a repository for research outputs reveals major differences in disciplinary cultures relating to scholarly publication and open access. As a result generic approaches and messages may hold little value for researchers, and may actually have a negative effect on their engagement with issues and willingness to deposit content. Advocacy material must be appropriately targeted. There are few resources available for those implementing repositories to draw on to avoid duplicating previous effort. It would be of benefit to the community if such information was collated and made widely available. This would be a useful role for the Repositories Support Project.

The decision to use the EPrints Open Source tool rather than adopt a commercial package or develop an in-house service was partly on the basis that EPrints is a well-developed and functional package that has been widely adopted and has community engagement. It was therefore expected that only a low level of in-house technical expertise would be required, and that the EPrints network would overcome any difficulties. In practice, particularly owing to first-time implementation of SWAP, greater expertise proved to be required and it proved necessary to subscribe to the EPrints support service.

Such issues could arise with any software: those implementing repositories need to have a clear understanding of the technical requirements relating to the repository tool they plan to use.

Writing to authors individually to request copies of their latest publications has proved effective as a means of generating deposits although this requires high investment of staff time. Though not easily scalable it is likely to be most cost-effective for those at an early stage of repository implementation or engagement with a Department.

WRAP's mediated deposit and metadata record creation process results in high quality metadata, but is resource intensive and slows the rate of accession of new material. Benefits of this model include:

- high interoperability of metadata records to support harvesting and discoverability by services like Intute, EThOS, library catalogues and vertical search engines such as Encore that require structured metadata
- SWAP's comprehensiveness brings flexibility if there is a requirement to export to other schema
- collocation of results is improved
- future-proofing (e.g., data in metadata records on citations could be used to link through to cited works or used to support future Research Assessment frameworks)

The main disadvantage is that the process is time-demanding (particularly regarding citation management and Library of Congress Subject Heading creation) resource-hungry, and requires skilled staff. This is difficult to argue when benefits are not immediately demonstrable, particularly as metadata records are not (currently) the primary route to discoverability of the work in the repository. Most visits to WRAP are through Google searches; Google does not index metadata records specifically but indexes them as html files and indexes the full text files.

Those planning to implement repositories should consider the current and future requirements of the repository as both institutional and wider service. This juxtaposes high quality with acceptability and affordability of costs. For example, which other repositories and services they envisage their metadata being used for? What metadata formats will they expect to be able to harvest? What other technical developments are likely to enable them to use their metadata for the benefit of their stakeholders?

WRAP's mediated approach supports the building of a discrete collection which has a cohesive identity and collection management policy, differing from other unmediated repositories which take whatever content authors wish to deposit, without mediation. A controlled collection provides a guarantee for both authors and institution that the quality of the repository content will serve to maintain a high institutional reputation, which in itself is likely to encourage submissions. Long term preservation needs can also be met more easily. However, it is apparent that there is also a need to host other types of content. WRAP does not do this but it could sit within a broader repository landscape at Warwick: different repository instances could provide such a service. Consideration needs to be given to the cost of enabling and maintaining different collections, and of appropriate submission processes.

However, defining 'University of Warwick research' for the purpose of accepting content for the repository is a potentially complex issue that may be determined in a number of ways, e.g., material only written when employed at Warwick or when the research for a work was undertaken whilst at Warwick, all publications of current Warwick staff whether written at Warwick or not, and so on. This is an issue for all repositories, and may be artificially determined by the requirements of the REF.

There is potential to build on WRAP's initial impetus by linking it with work beginning on a University Publication's Database, by creating further instances of the repository software in order to extend the types of content that is currently collected, and to integrate it further within the University's infrastructure in order to further exploit its content.

10. Recommendations

A thorough assessment of the available options together with apparent and hidden costs should be made when selecting the most appropriate way to implement a repository. Examples of this would be

the level of technical expertise required to adopt and understand an open source tool that by its nature has a different support culture from a commercial package, or the potential lack of freedom if an externally-hosted service was chosen to avoid technical support costs.

Those implementing repositories should consider carefully the long-term and institutional purpose of their repository in relation to the level of mediation they provide and the quality and level of metadata they plan to create. There are high resourcing costs in providing a mediated service and quality metadata but if the purpose of the repository is raise the institutional profile and at the same time emphasise the quality of the institution, or to future-proof the repository for embedding in other services these might be justified.

The repository should be considered as a strategic tool that, together with related developments such as Open Access publishing, management of research data, etc., can be used to move forward both Library and institution thinking.

11. References

-
- ⁱ <http://www.jisc.ac.uk/whatwedo/programmes/reppres.aspx>
 - ⁱⁱ <http://ethos.bl.uk>
 - ⁱⁱⁱ <http://dublincore.org/>
 - ^{iv} <http://www.openarchives.org/OAI/openarchivesprotocol.html>
 - ^v http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_Application_Profile
 - ^{vi} Carr, L., Swan, A., Sale, A., Oppenheim, C., Brody, T., Hitchcock, S., Hajjem, C. and Harnad, S. (2006) "Repositories for Institutional Open Access: Mandated Deposit Policies"
<http://eprints.ecs.soton.ac.uk/13099/>
 - ^{vii} <http://roar.eprints.org>
 - ^{viii} http://www.nereus4economics.info/neo_intro_press.html
 - ^{ix} <http://blogs.warwick.ac.uk/wrap/>
 - ^x http://www2.warwick.ac.uk/services/library/main/research/instrep/ir_value_event/
 - ^{xi} Details of the Open Repositories 2008 event are at <http://pubs.or08.ecs.soton.ac.uk/33/>
 - ^{xii} <http://www.rin.ac.uk>
 - ^{xiii} For example, see <http://www2.warwick.ac.uk/fac/soc/pais/staff/breslin/research>
 - ^{xiv} <http://www.ukcorr.org/index.php>
 - ^{xv} <http://www.jisc.ac.uk/whatwedo/programmes/reppres/recruitmenttoolkit.aspx>

12. Appendices

Appendix 1 – draft text of RSP briefing sheet on Cover sheets.

Cover Sheets

You may wish to consider whether it would be beneficial to add cover sheets to the full text files in your repository, or to recommend to your authors that they add them, upon self-archiving. The sections below deal with the reasons for and against using cover sheets, what information they can be used to convey, and how you might consider adding them. Whether or not you use them in your repository is likely to depend on a combination of factors, including your workflows, your resources and the main purpose of your repository.

Introducing your repository as the source of the item

Many visitors will come via a search engine to the full text file, bypassing your repository home page and metadata record, and a cover sheet is the only way to inform them of your repository's existence as the source. This is an opportunity to demonstrate the value of repositories to researchers, and a chance to market your own institution. Tying the work to the repository can also be of benefit to authors, adding the weight of your institution's reputation to the item, which postgraduates in particular may appreciate for their theses.

Consider whether your cover sheet should contain any branding such as logos or images: it might become labour intensive to replace cover sheets if your repository and/or the institution are re-branded. You can also use the cover sheet as a way to link to your repository home page, and/or the metadata record for the item in the repository. The link to the metadata record may be all you need to consider, as the other information that could be included on the cover sheet might be in your metadata record.

Meeting copyright holders' policies

If you are using cover sheets, then displaying a copyright statement is highly recommended. Some publishers allow post-prints to be made available online whilst insisting upon standard phrases being displayed along with the text: the cover sheet is one way to meet those requirements.

Even when publishers do not require such statements, a standard statement referring to copyright demonstrates respect for rights holders, just as others who are reading your content should show. A simple statement that all rights are retained by the copyright holder(s) and the terms under which you make the item available in your repository can be included on a cover sheet. Consider that your policy might change over time, so you might prefer that a standard statement refers to a policy on your website, rather than using the cover sheet to detail the policy itself.

Information about which version is presented

There is often little information in the metadata record as to which version of the article is held in the repository, even whilst the final, published version is referred to. The cover sheet is an opportunity to redress this, but even if the metadata record does make it clear which version the full text represents, of course the metadata record may not be the reader's entry point to the text.

It is often difficult to describe the version held in the repository in either a cover sheet or the metadata, owing to the confusion over terminology and authors' imprecise records. However, referring to the published version elsewhere can make it clear that yours is not the published "version of record". It may also be worth pointing out that access to the published version may require a subscription, which explains the reason for making an alternative version available.

VIF guidelines on cover sheets give more detail about version information:
<http://www.lse.ac.uk/library/vif/Framework/Object/cover.html>

Helping others to trace the published version

It is the published version, the “version of record” that anyone wishing to cite the article is likely to want to cite (See page 9 of the VERSIONS toolkit: http://www.lse.ac.uk/library/versions/VERSIONS_Toolkit_v1_final.pdf), and the cover sheet can be used to give the reader the information needed to trace and cite the published article. It would be a shame for an author to lose a potential citation because the repository version of the article itself did not contain enough information to make it easy for the reader to cite.

If possible, it would be ideal to put a full reference for the repository version and for the published version onto cover sheets, making it extremely easy for their works to be cited.

Reasons not to bother with Cover Sheets

The metadata record is the link we encourage authors to use as a handle for the items in the repository. As the pages with links to them, they are likely to appear highest in Google results lists. If your metadata records are stored as static web pages then they will be indexed by Google, even though Google don't harvest through OAI-PMH. So you may find that the metadata records are the pages that most people will come across, anyway, rather than the full text items.

Your authors may prefer to attach their own cover sheets rather than have a standard one displayed. Especially if they wish their own work to appear in any “rollover” images in the metadata record, rather than a standard cover sheet.

The cover sheet need not appear on every item or item type in your repository. Your cover sheets might be different for theses than for journal articles. You might have a policy to only attach cover sheets when authors have not supplied sufficient information in the text, to meet rights holders' requirements and enable others to trace and cite the final published work.

A cover sheet can confuse metadata extraction tools like Data Fountains, which finds things such as a rights statement, mission statement, name/url of the repository, or something else in prominent positions where it would normally find information from which it could generate a metadata record. If you plan to use such tools in your metadata creation, then you might not want cover sheets attached to your files.

How to attach/display cover sheets

A cover sheet might be page 1 of the file itself, added by authors before/whilst depositing, or by repository staff after deposit. Or the cover sheet might be generated on the fly at the point the document is requested by the reader, rather than being a part of the file itself.

An automatically generated cover sheet would not interfere with any automated metadata creation tools, nor with authors' wishes to have the front page of their work displayed in rollover images just as they intended it to look, and it would mean no extra work for depositors or repository administrators in attaching such cover sheets. Plus, any later branding or policy changes can be altered in the source information for such cover sheets.

Another way to potentially automate cover sheet creation would be to offer an “add-a-cover-sheet” option to authors as they deposit and upload the file. In this way, repository staff are not involved in lengthy processes, authors are prompted to include appropriate information in the files themselves, and authors get to choose whether they want a standard cover sheet or not.

Appendix 2 – List of EPrints3 files modified to fit SWAP

/package/eprints3/archives/test1/cfg/cfg.d/

eprint_fields.pl
eprint_render.pl
oai.pl
search.pl
user_roles.pl
views.pl

/package/eprints3/archives/test1/cfg/citations/eprint/
thread_commentary.xml

/package/eprints3/archives/test1/cfg/lang/en/phrases/
eprint_fields.xml
eprint_types.xml
views.xml

/package/eprints3/archives/test1/cfg/lang/en/static/
index.xpage
information.xpage

/package/eprints3/archives/test1/cfg/lang/en/templates/
default.xml

/package/eprints3/archives/test1/cfg/namedsets/
eprint

/package/eprints3/archives/test1/cfg/workflows/
default.xml

/package/eprints3/archives/test1/cfg/workflows/eprint/
default.xml

/package/eprints3/archives/test1/cfg/
subjects

/package/eprints3/perl_lib/EPrints/Plugin/Export/
EAP.pm

Appendix 3 - Some statistics, as reported to the WRAP team by Jenny Delasalle on 5 Feb 2009 and updated on 24 March 2009

WRAP, THE PROJECT:

To summarise where we're at, in terms of approaching departments...

We have written to all individuals from WMS, Economics, Psychology, Sociology, Education and History, requesting copies of their most recent journal articles. We continue to monitor Zetoc alerts for:

Biological Sciences
HRI
Economics
Psychology
History
Film and Television Studies.

WRAP staff have also been briefing people in departments, upon invitation:

Politics
Institute for Employment Research
School of Health and Social Studies

And Subject Librarians are initiating discussions about WRAP in their departments as well.

WRAP REPOSITORY ADMIN:

As of 24 Mar 2009, WRAP has 492 records of nearly all full text journal articles. (Excepting some experiments, basically.)

There are 26 user accounts including library staff and repository administrators, but we don't really expect people to set these up or use them in earnest: users can only save searches or set up alerts with an account on WRAP. WRAP allows anyone to create an account and there is no "are you a human being" check at the account creation stage, so we are deleting 'rubbish' accounts monthly. As the repository grows in popularity more accounts of both real and robot nature are being created, so we're watching this growth.

VISITOR INFORMATION FROM GOOGLE ANALYTICS:

WRAP has a significant number of visitors, and the number of visits/visitors is growing along with the collection of items within WRAP.

In the month 25 Dec - 25 Jan we had 3,639 visits from 101 countries/territories: 1,777 from the UK (230 were on the Warwick network), with the US 2nd on 566. They add up to 39,520 page views and 9,496 unique visitors.

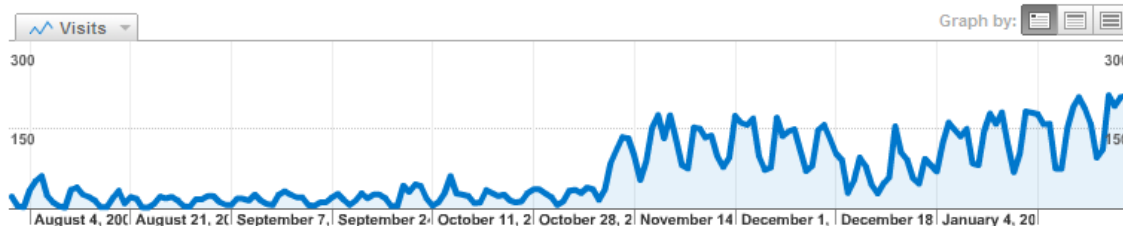
From 1 Aug 2008 until 5 Feb 2009 we had 10,833 visits from 129 countries/territories: 5,789 from the UK with the US 2nd on 1,747. Canada is 3rd on 335 and India 4th on 250.

Noticeably academic network locations of our visitors during that time are: University of Oxford, University of Manchester, LSE, University of East Anglia... and others, including the Houses of Parliament! (Also, since WRAP was launched: Southampton, Bath and Birmingham University networks).

We had a noticeable rise in visitors in November 2008 (no known reason), and visit numbers drop every weekend, which suggests that WRAP is a work-based activity, as seen on the diagram below:

Dashboard

Aug 1, 2008 - Feb 5, 2009

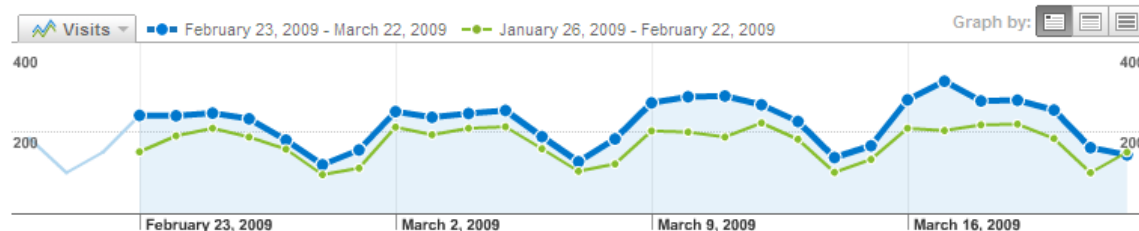


Visitor numbers continue to grow steadily, as the mapping of February to March 2009 visit numbers shows, below:

Dashboard

Feb 23, 2009 - Mar 22, 2009

Comparing to: Jan 26, 2009 - Feb 22, 2009



Most of these visitors came to WRAP via search engines and the proportion appears to be increasing, against direct traffic and referring sites. From 1 Aug 2008 to 5 Feb 2009, 76% of visitors came through search engines. For the month 23 Feb 2009 to 22 Mar 2009 the proportion of visitors from search engines rose to 88.9%. This may be due to the additional content in WRAP but also could be associated with a growing status in ranking of the WRAP website in search engine results: it is hard to tell since the ranking of WRAP in results may also be dependent on factors like the size of the website and the number of referring websites.

Keywords which appear to have brought visitors to WRAP are increasingly reflecting search terms not associated with WRAP itself (which brought us most visitors in the early days) but from within the content. Hot topics for March 2009 include "inflation and unemployment" and "educational leadership theories".

Our most popular page is the home page, followed by our departmental browse page, then advanced search.

After such general pages are accounted for, the most popular paper, with 562 pageviews from 1 Aug 08 to 23 Mar 09, is: Wolke, Dieter and Sapouna, Maria (2007) **Big men feeling small: childhood bullying experience, muscle dysmorphia and other mental health problems in bodybuilders**. Psychology of Sport and Exercise, Vol. 9. pp. 595-694. ISSN 1469-0292

...which gained early popularity from a link from a forum somewhere in Thailand (6 pageviews), and has now had many visits through a link from an Estonian body building forum (146 pageviews), but even so, the majority of visitors found it through Google (192 pageviews). It seems that Google ranks it fairly highly in its results, since there are links to it from elsewhere: the number of pageviews from the Thai and Estonian forums has altered by just one, from the end of Feb to the end of March, whilst the number of pageviews through Google in this period has soared from 192 total pageviews to 357. These figures reinforce the importance of getting links to repository content, in boosting "Google juice", which then takes on a growth of its own.

Our second "all time" most popular paper in February 2009 was (177 pageviews): Tennant, Ruth and Goens, Cristina and Barlow, Jane and Day, Crispin and Stewart-Brown, Sarah (2007) **A systematic review of reviews of interventions to promote mental health and prevent mental health problems in children and young people.** Journal of Public Mental Health, Vol.6 (No.1). pp. 25-32. ISSN 1746-5729

Most visitors came through Google, and there are no interesting link sources... except perhaps that we did set up an experimental page on Jane Barlow's Medical School web pages which linked to it: this accounted for 2 pageviews!

This remains the second most popular "all time" paper for March 2009, and the number of page views has grown as exponentially as for the most popular paper (now at 339 pageviews). This suggests that, with the importance of Google in bringing visitors to WRAP, the most popular "all time" papers are always likely to be the earliest deposited ones, as these gather a momentum of visitor growth: we will watch the long term trend of such visitor numbers to our papers.

From an advocacy point of view, it is worth regularly announcing the most popular paper for a particular month or date period, rather than the "all time" most visited paper, because otherwise the story remains the same.

SITEBUILDER PAGE STATS:

March 2009 updates have only been provided where there is a significant difference worth discussing.

Our "Repositories at Warwick" page is the one linked to from the page within the Research section of the Warwick website, and is structurally at the top of information about repositories on our site:

For the whole site, from 1 Aug 2008 till 26 Jan 09 - Total page views: 9,185. 69% of visitors were not logged in, only 2% were students and 29% were staff. 65% of visitors were from external networks. This suggests to me that the site as a whole is not of great interest to students, which is fine because it's not of great relevance to them! The page I do want some of them to look at is our "Thesis deposit information", of which there is a usage breakdown below.

For this specific page, from 1 Aug 2008 till 26 Jan 09 - Total page views: 801 from 499 distinct PCs 81% of which were from people not logged in, 67% of which were from external networks. It is one of our more external facing pages, so this is all fine. The table below shows how people got to this page for the last month alone (26 Dec 08 - 26 Jan 09):

How they got here	Hits	%
Typed in address or used bookmark	42	35
/services/library/main/research/ (One level up on the Library page)	27	23
/services/library/main/research/instrep/... (The About WRAP page)	7	5
/research/wrap/ (Also our go/repositories page)	6	5
/services/library/main/research/instrep/... (Using the WRAP repository)	3	2
/ (Not sure: we don't have a link from the home page!)	3	2
http://text.www2.warwick.ac.uk/ (Not sure: we don't have a link from the home page!)	3	2
/services/library/main/research/instrep/... (FAQs page)	3	2
From SiteBuilder search: (keywords: <i>repositories, online submission jisc</i>)	2	1
/services/library/main/research/instrep/... (Apparently from the Using the WRAP repository page, which is the same as the page listed above)	2	1

Apart from the mysteries which intrigue me, this tells me that a fair few people are reaching our repositories page through browsing the library's website, and by following the link from the page that we set up for the JISC (About WRAP, the project). No problems or surprises there.

Here is a breakdown of pageviews of the top pages on our site:

Title of page	Pageviews	Description of page
Submit your written work to the WRAP ins...	1,406	Submission form
About WRAP, the project	1,154	Page linked to from JISC website
Repositories at Warwick	801	Main repositories page, linked to from Research area of Warwick website.
Thesis deposit information (students)	690	Self evident!
Demonstrating and exploiting repository ...	676	For the event held in Northampton in December
Programme	505	For the event held in Northampton in December
Repository FAQs	379	Self evident!
Using the WRAP repository	332	Instructions for creating RSS feeds and accounts in WRAP, etc.
WRAP Policies	287	Self evident!
Booking form	254	For the event held in Northampton in December

More information about our submission form:

From 1 Aug 2008 till 26 Jan 09 - Total page views: 1,406 from 394 distinct PCs. 55% of visitors were not logged in, 44% staff and just 1% were students. 48% were on external networks while were 51% were on campus.

The number of pageviews is interesting, because in this time we have had 360 successful submissions of items to the repository. It is also a page that you have to be logged in to actually see. Further investigation reveals that Sitebuilder will count the visit when someone arrives who is not logged in. They will then sign in, and be counted again as a second access to the page. So, we could theoretically halve the number of pageviews, if we could assume that most people are not logged in when they get to our page: I'm not sure how likely that is. (Since the "How people got here" Sitebuilder statistics also suggest that, in the last month, we had 65 pageviews (18%) resulting from a link on this very page, that suggests to me that many people did have to go through the login upon arrival at the page, but not half of them. Although I'm not sure that following a link from the same page is exactly a measure of the logging in process...)

Even when halving the number of pageviews to about 700, we still seem to have a significant number of people who have looked at the page and submitted nothing to us through it: this is a mild concern, because I'm not sure what is turning them away. Is it because they don't actually need that page at all, or were they thinking of submitting, but were put off by the look of our form?

This page has its own "go" URL which we give out all the time, in practically every e-mail Marie or I send! It is probably the most important page we have, excepting the repository itself. It is also linked to from the repository's "About" page, although that link is not well followed, according to Google Analytics (Content Overlay showed 0 clicks on it!). The "How people got here" Sitebuilder statistics shows that most people typed in the address or followed a bookmark in the last month, although we did have 25 pageviews from the "About" page on WRAP. Which just goes to show that Google Analytics is not 100% foolproof... either that, or I'm not the best interpreter of it!

More information about the "About WRAP" page:

From 1 Aug 2008 till 26 Jan 09 - Total page views: 1,154 from 622 distinct PCs. 82% of visitors were not logged in, 14% staff and 4% students. 71% on external networks, 25% on campus.

This page is the least well maintained or informative, and was set up to meet the requirements of the JISC funding. I suspect it gets lots of visits because of the link from the JISC website and the relatively large proportion of external networks and un-logged in visitors supports this.

More information about the Thesis deposit information page, for students:

From 1 Aug 2008 till 26 Jan 09 - Total page views: 690 from 370 distinct PCs. 82% of visitors were not logged in, 15% staff and 3% students. 75% were on external networks, 18% on campus and 7% in Residences.

Interesting to note that staff are looking at the advice for students: we also have a page for supervisors, containing largely the same information. But perhaps it is also students using staff accounts to look at the pages: some PhD students also teach so might have more than one account. Also, there are plenty of external visitors, which doesn't tell us too much as many PhD students won't be based on campus.

March update: a very similar picture for the period 1 Oct 08 to 23 Mar 09. We can only get stats for up to 180 days at a time, so I can't tell the picture since 01 Aug 08. However, with no particular spikes or troughs or variations in who is looking at the page, there is nothing more to add: interest in this page remains very steady.

Repository FAQs:

From 1 Aug 2008 till 26 Jan 09 - Total page views: 1,045 from 284 distinct PCs. 85% not logged in, 14% staff, 1% students. 82% on external computers, 18% on campus.

NB: This section of the website was originally one page, but was redesigned to include a short answer on the main page and a detailed page for every question, using the contents page type in Sitebuilder. The stats above are therefore for this page and sub-pages.

The pages are primarily aimed at staff, but there is no requirement for logging in, so the balance of staff and students and logged in people is no surprise. It is interesting to see that most people are from external networks, though.

March update: None of these proportions have altered significantly.

Looking at the sub-pages, we can see which are the most popular:

1 Aug 08 – 26 Jan 09		1 Oct 08 – 24 Mar 09	
Repository FAQs	379	Repository FAQs	434
What is the value of depositing in WRAP?	56	What is the value of depositing in WRAP?	85
Which version of my journal article should I submit to WRAP?	48	Which version of my journal article shou...	74
How do I find out about my journal publisher's policies?	41	How do I find out about my journal publi...	67
Do I need to update WRAP with my entire back catalogue of work?	41	What is open access?	61
How do I use a repository to find research on my topic?	39	Do I need to update WRAP with my entire ...	58
What is open access?	39	Why don't you just put everything I've published into WRAP?	57
What content is eligible for inclusion in the WRAP repository?	38	What content is eligible for inclusion i...	56
How do I submit to WRAP?	37	How do I use a repository to find resear...	55
Why are institutions and research funders introducing mandates?	36	How do I submit to WRAP?	54

This in no way correlates with the order they are displayed on the contents page! There are 18 FAQs in total, and the most visited question about the value is placed near the top of the list: often at the very top, although most recently created content sits above it for a while. The question "What is open access", which is linked 14th on the contents page, but is the 6th most visited FAQ until Jan and the 4th most for Oct-March. For people to visit this page, they must have browsed through the other

questions, which might be an opportunity for us. Alternatively, perhaps we should consider putting our FAQs listing in order of the most popular questions. In any case, our current FAQs listing is probably long enough at 18 questions, and were there any more we would need to group them further into themes or organise them in some way.

March thoughts: The number of visitors to the page that gives more detailed information about a question is an indicator of which questions are interesting to our audience, but since the contents page also attempts to give a brief answer, the number of visitors to a page could actually be an indication that we have failed to satisfactorily explain the answer in brief!

Appendix 5 – Comments on the Project Budget

Significant variations in the overall WRAP budget were:

Staffing

The Budget originally included additional support at Grade 2. It became clear when workflows were being designed that this role would not be required during the startup phase of the repository (though a need has been identified for such a role post-project). This post was not recruited to and the Budget amended accordingly.

The delay in implementation of the repository (as discussed in the Report) also meant a delay in recruiting to the Repositories Assistant and Metadata Librarian. As the project progressed it was determined that throughput by the Metadata Librarian should be prioritised and the underspend utilised to increase the hours of that post.

Non - Staff

Dissemination

Plans for a programme of widespread internal advocacy and promotion of WRAP were put on hold firstly owing to the delayed implementation then a requirement to match submissions to the rate of processing. As discussed in Section 5.5 above, it was felt beneficial for the longer-term success of the repository to obtain a steady flow of submissions through targeted approaches and meet author expectations through a reasonable turnaround time rather than obtain a large amount of content but have disillusioned authors. Instead of production of leaflets as envisaged, internal dissemination has primarily been carried out by email, University publications, giving presentations and in committee meetings. Externally, dissemination has primarily been through the Website and blog, attendance (and presenting) at meetings/events and active engagement with relevant email lists.

Travel and expenses

The Project Manager participated in a larger number of repository community and training activities than had initially been expected, and the Library felt this important to support.

Hardware/software

It proved possible to utilise existing equipment so less expenditure was required than had been allowed for.

Evaluation

This has been managed without cost.

Other

Owing to difficulties configuring EPrints with SWAP it was felt necessary to acquire a subscription to the EPrints Support service at a cost of £1500 excluding VAT. This expenditure was unexpected but having access to the service proved very useful.



JISC Project Final Reports Budget Template

Directly Incurred Staff	TOTAL BUDGET £	Year 2007-2008 Actual Expenditure	Year 2008-2009 Actual Expenditure	TOTAL EXPENDITURE £	TOTAL VARIANCE
Project Manager (1FTE@FA6)	£56261	£18426	£39282	£57708	£1447
Metadata Librarian (0.58FTE@FA5)	£15376	£0	£20600	£20600	£5224
Repositories Assistant (0.5FTE@FA3)	£16208	£0	£8999	£8999	£-7209
Total Directly Incurred Staff (A)	£87845	£18426	£68881	£87307	£-538
Non-Staff					
Travel and expenses	£250	£0	£1251	£1251	£1001
Hardware/software	£3900	£1600	£1600	£3200	£-700
Dissemination	£1500	£0	£0	£0	£-1500
Evaluation	£250	£0	£0	£0	£-250
Other	£600	£200	£2298	£2498	£1898
Total Directly Incurred Non-Staff (B)	£6500	£1800	£5149	£6949	£449
Directly Incurred Total (A+B=C) (C)					
	£94345	£20226	£74030	£94256	£-89

Directly Allocated					
Staff	£22252	£5246	£17006	£22252	£0
Estates	£18463	£6154	£12309	£18463	£0
Other	£0	£0	£0	£0	£0
Directly Allocated Total (D)	£40715	£11400	£29315	40715	£0
Indirect Costs (E)	£64900	£15706	£49194	£64900	£0
Total Project Cost (C+D+E)	£199960	£47332	£152539	£199871	£-89
Funds Received from JISC	£29910	£4682	£25228	£29910	£0
Institutional Contributions	£170050	£42650	£127311	£169961	£-89

Nature of Institutional Contributions

Directly Incurred Staff					
Project Manager (1FTE@FA6) Metadata Librarian (0.58FTE@FA5) Repositories Assistant (0.5FTE@FA3)	£57935	£13744	£43653	£57397	£-538
Directly Incurred Non Staff					
Hardware/Software etc.	£6500	£1800	£5149	£6949	£449
Directly Allocated					
Staff, Estates etc.	£40715	£11400	£29315	£40715	£0
Indirect Costs					
Indirect Costs	£64900	£15706	£49194	£64900	£0
Total Institutional Contributions	£170050	£42650	£127311	£169961	£-89