

**A project to identify successful models  
for embedding repositories in research  
management systems and processes  
within higher education institutions**

<b>Submitted to:</b>	Neil Jacob JOINT INFORMATION SYSTEMS COMMITTEE
<b>Version:</b>	Draft Final Report
<b>Date:</b>	6 June 2009

**Rightscom Ltd**  
**Linton House**  
**164-1809 Union Street**  
**London SE1 0LH**  
**UNITED KINGDOM**  
**Tel: +44 20 7620 4433**  
[www.rightscom.com](http://www.rightscom.com)

## Contents

1	Introduction .....	3
1.1	The overall picture .....	3
1.2	Institutional repositories and publications management systems.....	3
1.2.1	Publications databases and specialised publication management systems	3
1.2.2	Institutional Repositories .....	4
1.2.3	Manual or no systems .....	4
1.3	Management aspects: tracking proposals and projects .....	4
2	Case studies .....	5
2.1	Outline case-studies .....	5
2.1.1	Model 1: well-developed RMS and IR (partially integrated).....	5
2.1.2	Model 2: well-developed IR and limited RMS.....	5
2.1.3	Model 3: Highly distributed information .....	6
2.1.4	Model 4: aspiring CRIS.....	6
2.1.5	Model 5: Fully integrated IR-RMS system (Norway) .....	6
2.2	Detailed cases studies: University of Southampton and Imperial College.....	7
2.2.1	University of Southampton.....	7
2.3	Imperial College .....	12
	<i>Introduction.....</i>	12
	<i>Context: the drivers and business case for the development and its history .....</i>	13
	<i>Strategic and managerial support .....</i>	13
	<i>Planning and implementation .....</i>	14
	<i>Usage and user responses .....</i>	14
	<i>Technical challenges identified and overcome .....</i>	14
	<i>Encouraging usage and responding to user feedback .....</i>	14
	<i>Legal and rights issues.....</i>	14
	<i>Future development plans .....</i>	15
	<i>Lessons for other projects .....</i>	15
3	Drivers for linking IR and RMS.....	17
4	Conclusions .....	24
4.1	Organisational issues .....	24
4.2	Possible operations efficiencies for research managers that could result from an integrated RMS .....	26
4.3	The role JISC might play in supporting RMS .....	26
4.4	Continue to encourage adoption of Cerif .....	26
4.5	Encourage publicly accessible Cerif-complaint CRIS.....	26
4.6	Co-ordinate/stimulate activities with Research Councils and other funders	26
4.7	Investigate feasibility of shared service CRIS model .....	27

# A project to identify successful models for embedding repositories in research management systems and processes within higher education institutions

## 1 Introduction

The original objective of the project was to locate case studies that demonstrate the value of embedding repositories in research management systems. We did not find fully-fledged examples: we found some cases where people would like to develop fully-integrated systems, but have not done so yet – partly because they have not been given the resources to do so. It also seems that very few repositories have, in themselves, attracted enough content to be functionally useful in such an embedded system even if it has been achieved technically.

### 1.1 The overall picture

The main finding from our research is the variety of disparate “systems” (i.e. both technology and process) in use. Management information is derived from many different systems and processes: HR; student administration (e.g. SITS); finance; research support; and information services/library.

A number of universities use business intelligence software such as Cognos or Oracle Discoverer for monitoring activity and performance, and for resource allocation and forecasting. However, some rely on tools as simple as a series of complex interlinked spreadsheets.

The role of institutional repositories is still small. IRs are near-universal, but mainly lack critical mass of content. Publications and research expertise databases are widespread and are the main research assessment management tool for many HEIs: few of these connect to the IR, although in a small number of cases the two are linked.

We could find no implementations of a research management system with embedded repository; further, where both RMS and IRs do exist within the same HEI they are not well-integrated.

### 1.2 Institutional repositories and publications management systems

A key driver for the development of research management systems has been the need to identify research outputs to be submitted for research assessment exercises. The RAE2008 led to the formalisation or development of publications databases, and although the nature of the next assessment, the Research Excellence Framework, is not yet finalised, institutions are clearly expecting published outputs to be a vital component. However, HEIs themselves submitted metadata rather than complete outputs, and the HEFCE team responsible for the RAE arranged for publishers’ versions to be made available to the assessors.

#### 1.2.1 *Publications databases and specialised publication management systems*

These are mainly department or school-based (38%) but collated centrally in over one-quarter of HEIs surveyed by Kings College<sup>1</sup>. Centralised publications

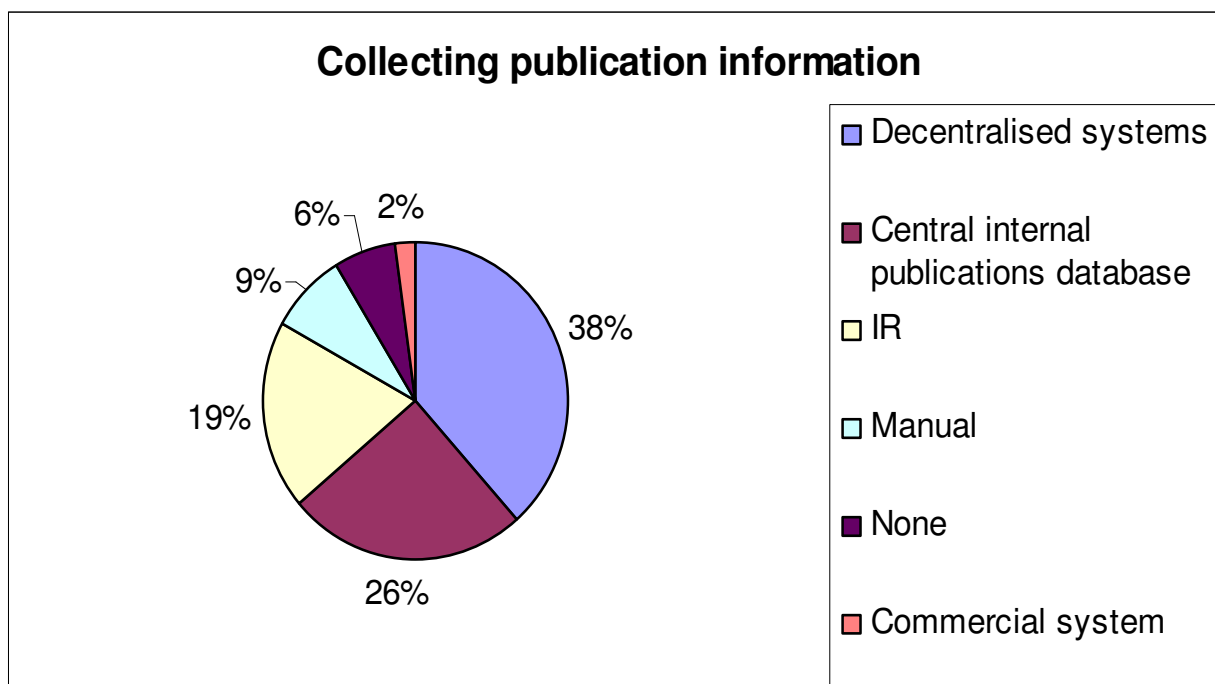
---

<sup>1</sup> <http://www.kcl.ac.uk/iss/research/ref/towards.html>

management uses a variety of tools, including general purpose applications such as Access and Excel EndNote, Reference Manager

The trend seems to favour such central publications databases, built using a variety of tools but largely managed by the research office. Some of these access external bibliographic information as well as internally generated data. This approach allows research offices to identify possible outputs in external databases, and then ask researchers for verification and additional input. Of the commercial systems, Symplectic seems to have the most mindshare.

**Figure 1 Processes for collecting publications information**



Source: survey of 49 HEIs by Kings College ISS

### 1.2.2 Institutional Repositories

The relationship between IRs and publications databases is evolving. Some IRs now take feeds from publications databases (where IR policy accepts bibliographic data as well as full-text). For full text, IRs may be populated by administrative or library staff or rely on self-deposit by researchers (a few have mandates). Around a fifth of institutions use their IR for publications management.

### 1.2.3 Manual or no systems

According to the Kings College survey, 15% of HEIs appear to rely on either manual systems, or do not have a system of any form. Given the increasing need to track outputs for to meet funders' requirements as well as present the best face of the institution, this should be a cause for some concern but would need more detailed investigation: none of the HEIs interviewed for this project were using manual systems.

## 1.3 Management aspects: tracking proposals and projects

Evidence that tracking proposals and projects is becoming a higher priority came through in several interviews. There is some concern among administrators that they are not getting the information they need from researchers, making it hard for them

to make an assessment of the overall position of the institution and identify the areas where support is needed.

Several research-intensive HEIs have implemented systems, using InfoEd (Imperial, Warwick, Edinburgh), Oracle AS Discoverer (Manchester) or SAP (Leeds). A major driver for investment in this area is the need to forecast and report Full Economic Cost for research activities to a variety of funders.

## 2 Case studies

We have presented a number of case-studies in outline to give an overall picture of the level of integration current achieved and different approaches taken to achieving it. These are based on the institutions interviewed and on further research. They are presented anonymously in most cases as this was the preference of the institution concerned for this stage of the research.

However, we have also included two detailed case studies as these illustrate many of the drivers, challenges and strategies found across the HEIs studied.

### 2.1 Outline case-studies

#### 2.1.1 Model 1: well-developed RMS and IR (partially integrated)

Profile: research-intensive institution with particularly strong non-QR research funding (RCs, medical charities)

##### *RMS elements*

The institutions does not use its RMS to monitor publications as **outputs** from projects, but is heavily focussed on proposal development and tracking (using InfoEd). It includes a database of information about the university's researchers: this has information on publications and peer esteem as well as CVs. It is automatically updated by checking personal web pages on the university's website; it also incorporates external trawling of databases (Symplectic)

Grants are managed post-award (using Oracle Grants) primarily to track the financial progress of the project.

The RMS also includes and FEC costing tool (an additional module for the InfoEd system).

##### *IR elements*

Post-award publications tracking belongs with the IR, rather than the RMS. The IR also acts as the publications database. It is basically divided into two sections:

- An "administrative repository", which contains the full text and licence information for articles.
- A "public repository", which provides public access.

Academics' personal web pages are also identified as part of the total output picture.

#### 2.1.2 Model 2: well-developed IR and limited RMS

Profile: Research-intensive institution

##### *RMS*

The RMS was built in-house, and is based on Oracle/SQL. Functionally, it is oriented to the RAE2008, though this will evolve. Its primary role is to pull together data from IR and from finance, HR and student (postgraduate) record systems.

The RMS plays no current role in grant administration.

##### *IR*

The institution maintains a well populated IR: the repository benefits from the existence of an institutional mandate for the deposit of journal articles and conference papers. It takes feeds from departmental publications databases as well as having direct input.

The IR holds both bibliographic data and full-text, and was used for preparing and delivering the RAE2008 submission.

### *2.1.3 Model 3: Highly distributed information*

Profile: decentralised institution (college-based)

This HEI has a newly-developed central IR that has not had sufficient time to become a major resource. The repository team is investigating other methods of gathering content to increase content and buy-in, but this is work in progress. There are also a number of other distributed information stores. These include:

- Publications
- Research management information
- Departmental and personal web pages

None of these are well-connected with one another. The repository team is planning a demonstrator using a web services approach to integration, and it is hoped that this will inform both the researchers and the HEI's management of the opportunities. The team is also running events on scholarly communications to help its community identify the benefits.

### *2.1.4 Model 4: aspiring CRIS*

Profile: a smaller research-intensive institution

*RMS*

This is a long-standing publications database. It helps populate some departmental websites, and formed part of the SRIS initiative. RAE2008 reinforced importance of tracking publications. The use of external ISI data is being considered, as in some areas academics are not keeping the databases updated.

A Research Expertise database is also a priority.

*IR*

There is a mandate for theses to be deposited in the IR, but senior management has not indicated a wider commitment to open access. However, it does want a 'dashboard' that will help monitor activity and performance and which might even drive staff reviews.

The funders' mandates for deposit of outputs are a new factor, but in common with other HEIs this has not yet had an immediate effect: projects covered by the new mandates have yet to deliver outputs.

Barriers to future development include uncertainty surrounding the REF specifications.

### *2.1.5 Model 5: Fully integrated IR-RMS system (Norway)*

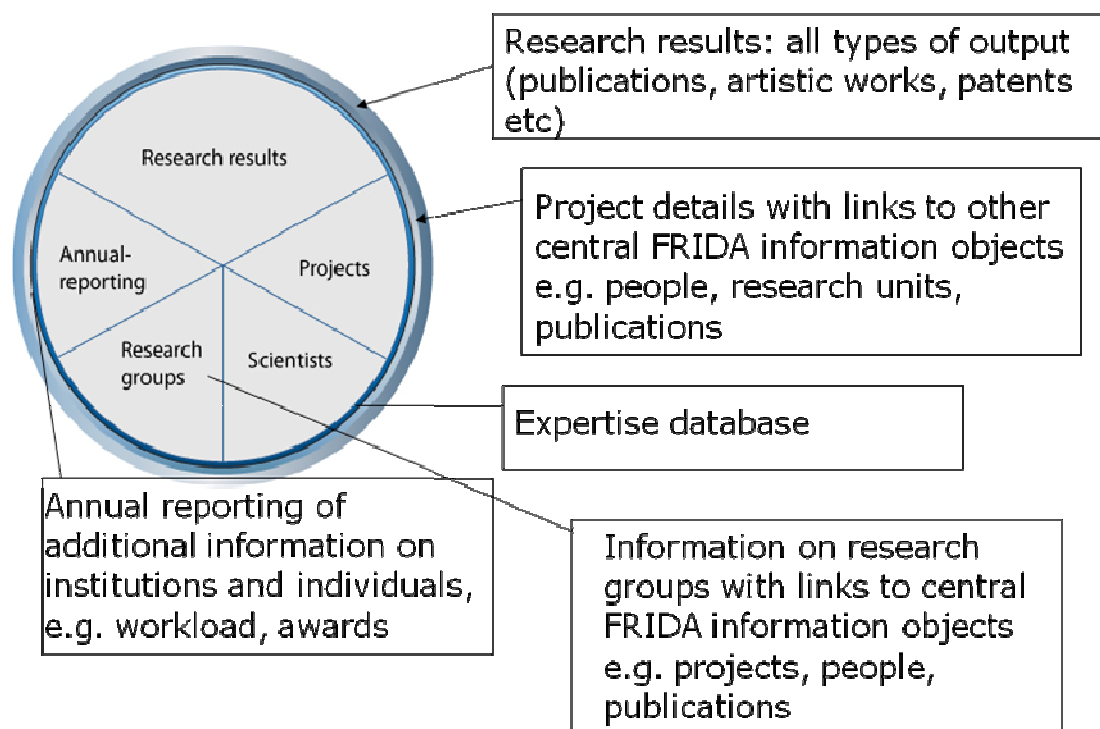
Profile: national RMS

Frida is an integrated research environment for the documentation and presentation of research activities, research results and scientific competence. Data from Frida is used to generate statistics for research activities at Norwegian universities, and

information it provides plays a major role in the annual funding of universities by the Norwegian Ministry of Education and Research<sup>2</sup>.

All institutions participating have their own full text OA repositories of research outputs. One of Frida's objectives is to act as a single data-entry point for researchers, and allow them to deposit information in a variety of institutional systems directly through Frida.

## Frida's five modules



### 2.2 Detailed cases studies: University of Southampton and Imperial College

These two detailed case studies of major research-intensive institutions illustrate the range of complex issues facing HEIs with a mixture legacy systems holding relevant information in departments as well as university-wide systems.

#### 2.2.1 University of Southampton

##### Introduction

Interviewee: Wendy White, Institutional Repository Manager, University of Southampton

The University of Southampton has been a pioneer in developing an institutional repository. The case study focuses on how the repository was used to deliver Southampton's submission for the 2008 Research Assessment Exercise and its inter-relationship with the Institutional Research Management System (IRMS).

<sup>2</sup> <http://wo.uio.no/as/WebObjects/frida.woa/wa/default?la=no&inst=185>

*Description of the system(s)*

The relevant systems are:

- The Institutional Repository, using EPrints software<sup>3</sup>
- The IRMS, built in-house using Oracle and SQL
- Departmental publications databases
- University finance, HR and student records systems

The Institutional Repository holds a variety of material, including pre- and post-prints, conference proceedings, research reports, working papers and other grey literature and postgraduate theses. Both bibliographic data and, where possible, full text are held. As well as acting as the instrument for the RAE submission, the IR is used by some schools to populate their staff web pages via the Content Management System.

The role of the IRMS is to pull together data (bibliographic and contextual) from the repository, and a number of key university administrative systems (finance, HR, postgraduate student records) to run reports and support decision-making. It does not currently deal with grant administration.

The University did not possess a CRIS, but some departments have well-developed and often long standing publications databases, which provided feeds to the IR.

*Context: the drivers for the development and its history*

The University of Southampton was early in developing an Institutional Repository, inspired by the idea of Open Access. The library, in partnership with the Information Systems Service, the School of Electronics and Computer Science and the academic community, was funded under JISC's FAIR programme to implement the IR.

The most important policy development affecting the relationship between the IR and research management came after the requirements for the 2008 RAE were first specified by HEFCE in 2003. It was at that point in the development of the IR that it began to take on its role in preparing and delivering the university's submission. In contrast with some IRs, whose commitment to OA has seen them take the view that only full text is acceptable, in Southampton a more pragmatic approach has been taken to deposit, with the result that the IR was able to act as an institutional publications database, while encouraging staff to deposit the full-text where possible.

There is a mandate in operation for conference proceedings and journal article post-prints, and where the culture of disciplines and academic concerns over publishing opportunities permit, pre-prints and book chapters are also deposited. But the repository managers prefer to build it through making it as demonstrably useful in as many different ways as possible to the institution, rather than relying purely on mandates.

The RAE submission was clearly the best opportunity to demonstrate usefulness to the core research effort of the university<sup>4</sup>. In practice, this involved:

- Storing all bibliographic information in the repository, DOIs and linking (using Cross-Ref as a check);
- Taking feeds from schools with well-developed databases which they were not ready to give up: for example, the Department of Optoelectronics has a

---

<sup>3</sup> <http://eprints.soton.ac.uk/>

<sup>4</sup> <http://eprints.ecs.soton.ac.uk/15703/>

database of full text going back about 20 years. There are 23 'schools' and each took its own decision on how to provide the material: a handful had databases, a good section decided on author deposit and the others nominated administrators to do it

- For the author and administrator deposit, the timeframe was 2001 on (the RAE start date), but where there are databases going back further, everything was taken
- Some schools put in the required '4 outputs per researcher' plus a few others, but others put in a lot more material covering the whole post-2001 period, so there is some unevenness
- The IR also optionally stored measures of esteem and the RA2, the contextual information around each submission
- While some researchers provided the published version of their work for the RAE, the repository team took responsibility for sourcing this version overall

#### *Strategic and managerial support*

The IR generally has enjoyed strong support from the leadership of the university, reflected in the policy on mandates. Its role in delivering the RAE submission has placed it strategically at the very centre of the university's core activities. This led to the repository manager being a member of the RAE planning team, which is probably very unusual, but meant that the information on what was required came directly and well in advance.

The IRMS is managed by the Planning team, not by the Research Support Office, which is more involved in grant generation, legal support and advice etc., though they were very much in the communications loop. The IRMS was really RAE-specific, though will not remain so in future.

#### *Planning and implementation*

The IRMS and the IR grew up side by side, with an awareness of the roles of each<sup>5</sup>. Although the outlines of what was going to be needed were present early on, the details often took a long time to be communicated by HEFCE, so planning had to be undertaken in anticipation of requirements. For example, prior to the HEFCE agreement with the publishers, which was quite late, the team explored with CLA (Copyright Licensing Agency) the use of the digitisation licence to allow the development of a restricted 'RAE' area within the repository holding published versions of research.

Requirements were gathered among the user group (those who would have access to administer the RAE submissions) and then meetings were held to get feedback at various stages as the development progressed.

In general, because the development was driven by HEFCE's requirements, there was little room for defining the 'what' or the 'why', so planning mainly focused on the 'how'.

The IRMS is a custom-built system, based on Oracle and SQL. The reporting functions of the IR were originally developed by ourselves in a JISC-funded project called IRRA, which supported the development of DSpace and EPrints functions for the RAE.

---

<sup>5</sup> A detailed roadmap can be found at [http://eprints.soton.ac.uk/12662/01/SPARC\\_IR\\_Workshop\\_Nov\\_2004\\_Hey\\_Simpson\\_poster.pdf](http://eprints.soton.ac.uk/12662/01/SPARC_IR_Workshop_Nov_2004_Hey_Simpson_poster.pdf)

The IRMS has to draw information from a number of administrative systems as well as the IR, so there was considerable development to enable this and to bring across data from legacy systems where these had been upgraded during the period.

#### *Usage and user responses*

The management team has access to usage data for the repository, tracking traffic and where it comes from, most popular downloads and other data., but because of some problems over interpretation of the data this is not publicly accessible as yet.

It is possible to link through to the repository from the RMS, but it is unlikely they would make the reverse journey, and in any case access to the RMS requires permissions.

The usage of the IRMS was quite small as it was restricted to perhaps a maximum of 100 people at its high point. This narrowed down as the submission data approached and schools 'locked down' their submissions, and rights of access were withdrawn.

#### *Technical challenges identified and overcome*

The main technical challenges involved were:

- Hefce classification
- Improving reporting
- The tight timescale.

**Classification:** the challenges in terms of bibliographic data were in the way that Hefce labels content: e.g. 'internet publication' – rather than 'electronic journal article'. Hefce also required the exact month of publication: however, some publishers use quarterly designations such as 'spring' or 'autumn'. This meant the team had to create schemes for Hefce. It could not be left to the time when the submission was due to go and then simply have everything put into the submission form.

**Reporting:** systems are bound to be changed during an eight-year period. In this case, there were issues of managing and mapping of legacy data from an old finance system onto a new one. Also the HR system was upgraded in the year of the RAE submission.

**Timescale:** a major challenge was to keep up to date with HEFCE's requirements. If Hefce introduced a changed requirement, something had to be implemented immediately. Changes could not be delayed for three months, as might be done normally. As the development was all in-house, it meant dependence on a few people, but the university insured itself as well as it could by having a robust project plan.

There was a lot of query answering for staff, interpreting Hefce documentation for individual academics. Although FAQs were developed in the system, there were still a lot of individual responses needed. The sheer volume of information was also a challenge.

One further challenge was in the technology platform used. The repository was not able to upgrade to EPrints 3 because EPrints 2 was already being used for the RAE. The EPrints software developer was also occupied in supporting the RAE as well.

#### *Encouraging usage and responding to user feedback*

General usage of the IR is rising as more people see its usefulness and as full-text deposit is increasing, driven by the development of new uses such as staff web pages, appraisals, and six-monthly research reviews. It becomes in the researcher's own interest to ensure that their work is in the IR to meet these different requirements.

User feedback is going to be gauged as part of the process of drawing up requirements for the next phase of development, linked with the REF replacement for the RAE.

#### *Sustainability (including business models)*

The IR was given two years' pump-priming funding initially, following which it became embedded as a regular item in budgets: this is necessary for sustainability. Relying on funding on a project-by-project basis would be exposing the repository to additional risk.]. There was a contract for the IRMS, which has now finished, and the review will look at a new contract when it becomes clearer what the IRMS will need (new tools etc) for the REF.

#### *Legal and rights issues*

There were no rights issues as Hefce had cleared the use of the published versions of research submitted across the board for the RAE.

In terms of managing access, as the Hefce agreement came quite late on, the team had already developed areas of the repository which would store the published versions for the RAE with restricted access to named individuals with RAE administration responsibilities.

#### *Future development plans*

The next step is to build a robust research infrastructure, making sure that systems speak to each other properly and data can flow seamlessly between them, and that departments and other users can get information out quickly and in the form they want. Some of this happens now, but there is a need for further work if to give everyone what they want.

The system needs not only to interface properly internally, but also externally: for example to research councils' systems. The drivers are both the REF and the need for research groups and departments to have a clear idea of what is going on at every stage.

The REF is expected to give a further impetus to the IR and to the process of building an integrated IRMS, because researchers will increasingly perceive the link between OA and citations, and because whatever HEFCE decides about the mix between bibliometrics, citations, peer review, esteem measures etc in different disciplines, it will be complex to deliver.

The IRMS will be used for other tasks in relation to research administration. It was primarily designed to facilitate the RAE, but its use will change with the REF as that is somewhat broader—the intention is to link more directly with research activity across the institution. Essentially the RAE was a large one-off task to gather all the relevant information together and submit it. It is anticipated that the REF will be much more interactive over the whole period and the systems will have to change to deal with it. Though there will be a cut off point, people will be able to monitor their position using whatever metrics are decided on throughout the period.

Scoping of requirements is underway.

#### *Lessons for other projects*

The main lessons learned by the team were:

- The key point is to ensure that the repository is addressing the core needs of the university in terms of research, helping to maximise income – the IR must engage with the critical strategic interests of the university. The REF will make that link more obvious.
- A sense of combined advocacy between the academic community and the supporting teams (library, software services etc) has been developed, which

links the supporting services closely to the core academic purpose of the university; this way of working enhances that sense of common purpose. It is important to be pragmatic and flexible in encouraging the academic community to engage, rather than trying to make them feel obligated and to impose systems on them.

- Having the library delivering core services is good both for the library and for the institution and brings expertise about research activity back into the library. Links between the library and research support office and planning have been strengthened, which has been very beneficial for the strategic position of the library and to the development of new expertise to help researchers by the library
- The development means that information in the institution is being re-used efficiently e.g. feeding information from the repository out to web pages as well as for appraisals etc.
- Being early in developing a solution has advantages and disadvantages. Some things tailored to the RAE will probably have to be unpicked and re-engineered for the REF, whereas other institutions that waited can start with the new requirements, but that was inevitable. The advantage is that being in the RAE work has made the IR and its team more central strategically.
- In the view of the team, the bottom line is: institutions want to have a reference to their complete research outputs, without requiring third parties to get it.

## 2.3 Imperial College

### *Introduction*

Interviewees: Cindy Lai and Scott Rutherford, Research Office, Imperial College, London.

Imperial College has introduced a range of systems to support research funding, and these were the focus of the discussion.

### *Description of the system(s)*

The systems discussed were:

- Proposal Development and Tracking modules (using InfoEd)
- Grants Management System (using Oracle Grants) to manage the financial transactions post-award
- FEC costing tool (part of the InfoEd system)

The Proposal Development system, which has over 1,800 users, includes a database of information about the university's researchers: this has information on publications and measures of esteem (for example, board memberships or prizes). It is automatically updated by tool developed by Symplectic Ltd that scans the web for publications that may be written by one of Imperial's researchers, and then emails the researcher with candidates and a request to confirm whether they are indeed the author. If they are, details are included in the publications database and on the researcher's personal web page.

It is populated by automatically trawling bibliographic databases such as PubMed and Web of Science, looking for articles by Imperial College researchers. Academics can also add their own papers by hand if the trawl has not picked one up. The system does some automatic de-duplication (to remove multiple references to the same article).

The individual web pages are the final component of the systems, with each academic including information under a set of pre-formatted templates covering topics such as personal information, publications and honours and awards. The individual academic can make personal decisions about which information can be displayed only to members of the university and which to the public. Most of the content is provided automatically from other systems.

The Research Services department itself does not have to deal with publication information for funders – this has been delegated to the library as part of large programme of work around its repository, Spir@l (see below).

*Context: the drivers and business case for the development and its history*

Imperial College has over 4,000 active research projects and over £200 million research income.

The systems were introduced as a result of a major change management programme, the Research Administration Management Programme (RAMP). RAMP began in an effort to improve processes and systems that were not effective at managing either applying for grants or managing the grant funding when it was received. Funders were also concerned, and their views were an important driver in the decision to make radical changes to systems and processes. Before RAMP, the pre-award and post-award processes were dealt with by different departments. As a result, faculties and individual academic departments had developed their own ad-hoc solutions, and when RAMP started there were over 60 independent databases in use around the university.

As part of RAMP, the Research Services Division was created, with dedicated teams for each of the Faculties, responsible for both pre- and post-award administration.

With the introduction of FEC, it also became essential to manage the costing and pricing of research contracts using FEC methodology. This has led to more informed pricing, and to better management of under-recovery, increasing returns from the investment.

On the output side, the university has a publications database which links to the Spir@l repository (see below). The publications database held bibliographic information on all papers published by the university's researchers. It also provided some information for use in their individual web pages. However, it did not provide full text. The publications database was also used to support the university's submission to the Research Assessment Exercise (RAE).

*Strategic and managerial support*

The RAMP programme had high-level support because of its importance to the university. The individual projects were managed by IT staff, but usually with a "business" lead as project director and an academic as project champion.

There was strong support of the systems because senior management wished to see accountability and ownership of research projects to ensure that they did not drift and that their academic and financial performance was maintained throughout their duration.

The generally business-oriented culture at Imperial helped to drive implementation of the new systems: they were seen as offering ways to track performance (e.g. proposals submitted compared to funding awarded, overspend or underspend on contracts) as well as improve administrative efficiency.

### *Planning and implementation*

The systems developed under RAMP went through two stages. Shortly after the first phase was operational, the FEC was introduced so much of the systems had to be enhanced to include it.

Planning and specification was a consultative process, with stakeholders from different areas of the university putting forward their requirements and functional needs. The main requirements were for specialised reports that could be used to help manage activities, and in the end the requirements were consolidated by a separate Reporting Group and divided between operational and strategic information. However, projects under RAMP had an initial emphasis on the internal Research Services processes, and less on the relationships with departmental administrators and academics.

Some outside expertise was brought in for the development, but most of the development was drawn from in-house resource. Users were heavily involved, as one of the key objectives was to free up the time of academics from administrative tasks by providing better support.

### *Usage and user responses*

Academics and department administrators are now used to using the systems. Problems do still arise, however.

A particular problem arose from the introduction of FEC. The systems had, up to that point, largely been driven by internal requirements. FEC, however, imposed external drivers. Interpreting the meaning of FEC terms and requirements was a substantial task. In turn, this affected procedures and user behaviour because all staff involved in research funding had to understand FEC methodology and policy. They then had to apply this knowledge and information at pre- and post-award stages.

### *Technical challenges identified and overcome*

The important technical decision was to customise software packages such as InfoEd and Oracle Grants rather than to build from scratch.

Meeting the different needs of researchers and research administrators presented technical challenges, as they are not always aligned. The reports have to deliver both individual project level and aggregated information to meet these differing needs.

The main subsequent technical challenge was the rapid development of the changes required to implement FEC.

### *Encouraging usage and responding to user feedback*

A major issue for the system was to encourage academics to contribute information needed for the proposal development system. Over 50 “roadshows” were held to explain the new systems and their role to stakeholders. This was supported by a range of printed user-guides and a website, including webcasts, e-learning modules and training sessions.

Towards the end of the project, some faculty administration staff were seconded full-time to work on project activities, mainly because their knowledge and expertise in research administration was needed.

### *Legal and rights issues*

None acknowledged.

### *Future development plans*

The REF is likely to stimulate further development, but the team's view is that if the data warehousing programme currently in progress works well, this is not likely to present serious technical problems. Imperial has managed previous RAE submissions without major problems.

### *Lessons for other projects*

The main lessons that the team felt had been learnt from the project are:

- Not to lose sight of the objectives in the pursuit of getting the project done: under the pressure of deadlines, there is a risk of projects becoming dominated by finding technical solutions at the expense of providing user benefits
- When decision-making, asking the right questions of the right people is essential
- Recognise the full impact of changes of scope
- Involve the right kind of people from the start, especially those with deep working expertise of the processes
- Estimating the time that will be taken to get things just right is hard
- Plan for the resources that will be needed to support users post-implementation, as this can take almost as much effort as the development itself
- Significant post-implementation support is essential

### *Outputs and the Spir@l repository*

We were not able to interview the relevant Imperial staff concerning management of outputs, so information on the output side (especially the repository) was drawn mainly from a recent article by Ashfari and Jones<sup>6</sup> describing the objectives and development of the institution's repository.

The Publications database was originally a standalone system, but now forms part of a new repository, Spir@l (sic). This was developed to go beyond the functionality of the publications database and also to provide full-text for academics to use on their personal web pages. Spir@l is a DSpace-based repository that has taken three years from first plans to deployment.

The major components of the Spir@l systems are the publications database, the "administrative repository", which contains the full text and licence information for articles, the "public repository", which provides the public access, and lastly the academics' web pages.

The driver for this kind of information has to be from academics themselves, who will be most affected by the need to satisfy funders that articles have been published as agreed. The team building the repository established as a core objective that it should become deeply embedded in researchers' working practices: this would ensure that it would be widely-adopted by academics, and so Spir@l has higher levels of deposit than is usual with repositories. The main reasons for this are its ability to feed web pages, and the use of nominated department staff to keep the repository up-to-date without the academics themselves having to do so.

---

<sup>6</sup> Program: Electronic Library and Information Systems Vol.( 41 ) No.( 4 ) pp 338 – 352 (can be found at <http://spir@l.imperial.ac.uk/handle/10044/1/493>)

The concepts of comprehensive coverage combined with open access were central to the development. Quality control of content was important, as was compliance with copyright requirements. All types of content are stored: journal articles, book chapters and conference papers are all included. The feeds to academics' own web pages would increase the visibility and chance of the article being found in a search.

Spir@l was designed around a "seamless interface" for the publications system, which is how academics interact with it. To keep the barriers to use as low as possible, the interface was designed to present everything on one page, and with high response times. Extensive consultation with users was carried out during the design stage.

However, the repository involves the major issue of whether or not a published article can be included, depending on the agreement with the publisher. Authors can check the SHERPA/RoMEO database from within the deposit interface to see if the journal does permit deposit. The information is checked by library staff as a safeguard. Although end-users such as the researchers themselves provide information about the license that applies to a deposited article, this is also checked in each case by repository staff.

Particular attention was paid to the design and wording of information pages explaining the relevant issues clearly to researchers depositing articles, together with information on how to get questions answered.

### 3 Drivers for linking IR and RMS

This table shows the high-level drivers for development we have identified during the research, the functions being driven, where the developments need to take place (in the research management system, in the institutional repository or in the links between the two) and whether the driver is relevant to the senior management of the institution or to individual researchers.

The high-level drivers identified were:

- Satisfying HEFCE requirements (RAE, REF)  
HEFCE assessment exercises (whatever for they take) are likely to require the collection and management of information about individual outputs. This may be a selection, or all the outputs for the HEI. Many HEIs have recognised that they need to start to manage this information well before the REF.
- Satisfying Research Council (and other funders) requirements  
The need to provide information on research outputs to funders will be on a more frequent basis with mandates becoming more commonplace. During this study the need to do this regularly was mentioned frequently: this represent a change in process for many HEIs, but one that may also assist with the REF. RMS developments in this area are driven by a need to ensure that other metrics of interest to funders are also managed.
- Improving efficiency of the individual HEI  
Developments in RMS are in part driven by a straightforward need to avoid duplication of effort, rework and human error: much the same motivations for process and systems reengineering in many different types of organisation.
- Improving efficiency of the HEI sector as whole  
HEIs are increasingly involved in collaborative research. The availability of current research information and relevant outputs can both facilitate this process; at the moment availability of both types of information is highly inconsistent across HEIs.
- Increased presence for the HEI  
As competition for both money and students grows stronger, HEIs are interested in developing a strong presence on the web in particular. Accessibility of both research project information and outputs to interested users depends in part on how accessible they are to search engines as well as the HEIs own systems.

Each of these leads to a need for specific functions or activities, and these are detailed in the table. The list is selective: we have concentrated on the functions that were most mentioned either by interviewees or in relevant literature. We also identify whether they are relevant only to RMS or repository development, or whether they require a link between the two: these are the cases of most interest in this study.

Based on the interviews we carried out during the study and on the literature, we have also suggested the relative strength for two different communities: the senior management in an institution, and academic researchers themselves. It is important to understand this difference, as it is likely to condition the extent to which researchers are willing to adapt to new or merged systems if they do not have a strong motivator. In this and other surveys, it has been repeatedly pointed out to us that it is very hard for the managers of HEIs to persuade or mandate researchers to undertake

activities that do not seem to offer an immediate benefit or which demand a change of working practice.

High-level driver	Function s needed	Relevant for developments in:			Strength of drivers for institution management	Strength of driver for researchers
		Research Management Systems	Linking RMSs and IRs	Institutional Repositories		
<b>Satisfying HEFCE requirements (RAE, REF)</b>						
	Identifying & tracking prospective outputs		✓		Weak (too much investment for uncertain payoff; publications DBs already do much of the job)	Not known
	Identifying completed outputs		✓	✓	Weak (alternatives in place in many cases; investment risk/return ratio)	Not known
<b>Satisfying Research Council requirements</b>						
	Managing grants process	✓	✓		Strong	Strong
	Keeping projects on track	✓			Strong	Weak (existing local or personal systems)

High-level driver	Function s needed	Relevant for developments in:			Strength of drivers for institution management	Strength of driver for researchers
		Research Management Systems	Linking RMSs and IRs	Institutional Repositories		
	Deposit mandate		✓	✓	Strong (information to be available in other resources besides repository)	Strong (future grant funding)
	Making information available for Research Councils to harvest			✓	Weak (can be done from IRs – but enhancements needed)	Weak (can be done from IRs – but enhancements needed)
<b>Improving efficiency of the individual HEI</b>						
	Who is doing what	✓	✓		Strong	Not known
	Other people within institution working on the same thing	✓	✓		Not known	Strong
	Measuring productivity		✓	✓	Weak (outputs in IR may not be good measure of productivity)	Weak (informal, local tools or not measured)
	Adherence to plan	✓			Strong	Weak (informal local tools)

High-level driver	Function s needed	Relevant for developments in:			Strength of drivers for institution management	Strength of driver for researchers
		Research Management Systems	Linking RMSs and IRs	Institutional Repositories		
	Management information and reporting	✓	✓		Mainly weak (linking seen as bringing no advantage over publications database in most cases); strong for small number where repository also functions as publications database	Not known
	Incorporating data from Research Councils' Grants on the Web	✓			Not known (potentially very useful but not referenced by most interviewees)	Not known
<b>Improving efficiency of the HEI sector as whole</b>						
	National-level management & collaboration	✓			Medium (not expressed but featured extensively in the literature)	Not known
	Cross-institution research project collaboration	✓	✓		Strong	Weak to strong (depends on institution)

High-level driver	Function s needed	Relevant for developments in:			Strength of drivers for institution management	Strength of driver for researchers
		Research Management Systems	Linking RMSs and IRs	Institutional Repositories		
	Wide-area repositories and RMS	✓	✓	✓	Strong (in some areas: needs local initiative)	Not known
	Consistent access to relevant outputs and personal information	✓	✓	✓	Strong (access highly inconsistent across HEIs)	Not known
<b>Increased presence for the HEI</b>						
	Visibility on search engines	✓		✓	Strong (see to be important way in; strengthened by linking)	Strong (seen to be most important way in; need to link to personal data to maximise benefits)
	Easier access to research outputs			✓	Medium (expressed by some institutions; repositories may be sufficient)	Strong among committed supporters; medium elsewhere

High-level driver	Functions needed	Relevant for developments in:			Strength of drivers for institution management	Strength of driver for researchers
		Research Management Systems	Linking RMSs and IRs	Institutional Repositories		
	Easier access to information on current research	✓	✓	✓	Strong (seen to be haphazard/lacking at present; coordinated source needed)	Medium (drawn from other sources)
	Access to information on individuals	✓	✓	✓	Strong (seen to be haphazard/lacking at present; coordinated source needed)	Strong (seen to be haphazard or lacking at present)

## 4 Conclusions

Our key conclusions focus mainly on the organisational issues

### 4.1 Organisational issues

#### *Conclusion 1: managers' interest in RMS*

RMSs are of most interest to management in universities (department/school level as well as institution level) rather than actual academics. Managers are interested in “dashboard” information about overall performance, and sometimes in drilling down to get to detailed measures (e.g. knowing which high-impact journals are accepting papers from researchers; the relative performance of different researchers in bidding for grants). However, they are not usually interested in reading the actual research output nor do they necessarily see how holding the research outputs in full will help to enhance funding.

Many managers are heavily involved in submitting proposals and finding other forms of funding, and so do not generally reflect on the longer-term needs of their university.

#### *Conclusion 2: managers' view of repositories:*

For this reason, they do not see a great deal of value in repositories, but they do see a value in publication databases – especially complete ones. From their point of view, an embedded repository is of little value unless it can also act as a publications database. There is a risk that this attitude overlooks the business intelligence benefits that may come from being able to track downloads and citations to outputs directly from an institutionally-managed resource, rather than relying on third-party services.

While growing number of institutional mandates suggest that the OA agenda is of increasing interest to managers<sup>7</sup>. The relationship between repositories and the core business of growing the university's research income is not yet fully established. As a result, managers have not always given OA developments much active attention, and crucially, this means that in many institutions (though not all) there has been little pressure on researchers to self-archive, nor much in the way of resources made available for other staff to do so on their behalf. If OA is shown to offer at least a temporary advantage in terms of citations, they may be much more interested<sup>8</sup>. However, the discussions and evidence for this are regularly presented in specialist sources that most researchers will not normally see<sup>9</sup>. Further, the advantage has been debated in more general publications<sup>10</sup> as well as in the specialist literature<sup>11</sup>. It is hard to determine the extent to which the debate has influenced either managers or researchers, but levels of deposit remain quite low.

#### *Conclusion 3: where managers are interested in repositories, their interest has a particular purpose:*

---

<sup>7</sup> For example, <http://www.ucl.ac.uk/media/library/OpenAccess>

<sup>8</sup> <https://dspace.lboro.ac.uk/dspace-jspui/bitstream/2134/4083/1/The%20Citation%20Advantage%20of%20Open%20Access%20Articles.pdf>

<sup>9</sup> An extensive bibliography on citations and open access can be found at <http://opcit.eprints.org/oacitation-biblio.html>

<sup>10</sup> See, for example, [http://www.bmj.com/cgi/content/full/337/jul31\\_1/a568](http://www.bmj.com/cgi/content/full/337/jul31_1/a568) and <http://www.hprints.org/docs/00/32/62/85/PDF/oa-interaction.pdf>

<sup>11</sup> <http://ilemt.epfl.ch/repec/pdf/cemi-workingpaper-2008-007.pdf>

The “public presentation” or “showcase” aspect of repositories is more complex: interesting to managers, but in specific contexts (e.g. for presenting a good face to research funders).

The effect of research funders’ mandates could change this situation, especially when it becomes critical for institutions, in order to gain further grant funding, to provide evidence to research funders that their mandate is being applied. Increasing downloads of outputs included in repositories might also influence attitudes to deposit. If mandates begin to affect a wider range of outputs, it will be important to provide central support to assist researchers in fulfilling them. However, in some cases mandates are reported to be driving the population of subject repositories such as UK PubMed rather than IRs<sup>12</sup> Indeed, the Wellcome Foundation mandate requires deposit in UK PubMed Central<sup>13</sup>.

*Conclusion 4: libraries are not central to the RMS agenda*

The area of the RMS the library is most likely to be involved with is the publications database. We have found that these can be fragmented across schools and departments, although some HEIs have a centralising model and are trying to enforce standards across all outlets. In some cases, library involvement has been significant in the development and operation of one key part of the RMS, the publications database<sup>14</sup>; in others the publications database has been run mainly by the research office. It seems, though, that in both cases the input has come primarily from departments or individual academics (and in some cases from systems of the type described in 1.2.1).

*Conclusion 5: institutional repositories are still limited in scope, but subject repositories are making more headway*

Most IRs are still very thinly populated and have not gained the commitment of academics or management. The best opportunity that they had was to support the most recent RAE. A small number did become important during this process, but the large majority did not.

The received wisdom among librarians responsible for repositories (based on strong anecdotal evidence) is that the limited interest that academics do have in repositories is focused more on subject than institutional repositories, although we have little concrete evidence to support this assertion<sup>15</sup>, nor did we find any examples of HEIs interested in linking their RMS and subject repositories

*Conclusion 6: this situation represents a transitory phase*

The change in the nature of research assessment is going to change the roles of both repositories and RMS, as is greater understanding of how researchers want to manage their outputs. The nature of the forthcoming REF is not yet fully clear, but it would be surprising indeed if its shape did not have a significant impact on both repositories and on RMS, especially as it is very likely that it will include significant bibliometric elements.

---

<sup>12</sup> For example, a survey carried out by the White Rose consortium reported that 15% of researchers deposited outside the IR:

[http://eprints.whiterose.ac.uk/increase/Increase\\_Questionnaire\\_Findings.pdf](http://eprints.whiterose.ac.uk/increase/Increase_Questionnaire_Findings.pdf)

<sup>13</sup> <http://www.wellcome.ac.uk/About-us/Policy/Policy-and-position-statements/WTD002766.htm>

<sup>14</sup> <http://www.leeds.ac.uk/library/teams/rae/>

<sup>15</sup> For a general review, see [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1425692](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1425692)

#### **4.2 Possible operations efficiencies for research managers that could result from an integrated RMS**

The current situation suggests that there are operational efficiencies to be had from linking RMS and repositories, but that in many HEIs these are not being realised at present.

Potential benefits at institutional level in the longer term include preparedness for REF (and its successors), better-populated IRs (although may be mainly metadata-only records) and better self-service for people interested in contacting/working with the HEI.

In the near term, the efficiencies include speedier processing of grant applications and easier progress tracking during application stages and lower costs of maintaining quality.

There is a further benefit in terms of risk reduction: the “bricolage” spreadsheet-and-database systems approach has a number of obvious vulnerabilities.

Realising the benefits depends on increasing the number of well-populated IRs and linking or merging IRs and publications database. It may be that the acceptance of metadata-only records is required to increase the level at which IRs are populated, although this is controversial to some. Compliance with mandates is obviously going to be a significant contributor to realising the benefits, but it is still too early to see how effective these will be.

The adoption of a common standard for information will further help interoperability, and the Cerif standard has clearly made progress here.

Lastly, it is important to ensure senior management commitment. Start-up costs are likely to be high, and there will be ongoing personnel costs required to maintain both quality and quantity of information. The resources to manage and maintain IR and RMS are specialist rather than generic, and if there is an increased take-up of the integrated approach it is possible that demand could exceed supply.

#### **4.3 The role JISC might play in supporting RMS**

The development of RMSs and the halo of systems surrounding them would clearly benefit from support and guidance. The greatest impetus is likely to come from managers seeing the success of others, and gaining a realistic understanding of both the benefits and the challenges of enabling development and integration of RMS and IRs within their institutions. We would suggest that at this point there are a small number of highly focussed initiatives that seem appropriate in the short term:

#### **4.4 Continue to encourage adoption of Cerif**

Encouraging the adoption of Cerif would, we believe, serve two purposes: in the long term, enable and encourage interoperability of research project information; and in the shorter term help those responsible for CRIS and RMS within institutions develop a framework for their own systems and a focus for attracting support from senior managers.

#### **4.5 Encourage publicly accessible Cerif-compliant CRIS**

The more RMS and CRIS are available to a wide range of interested users outwith as well as within institutions, the more incentive there will be for managers and researchers alike to commit resources on a sustainable basis.

#### **4.6 Co-ordinate/stimulate activities with Research Councils and other funders**

In the medium term, the development of RMS and CRIS is likely to be driven by requirements of research funders and with quality assessment bodies. The

deployment of systems to meet the needs of both these stakeholder groups will probably not happen without some co-ordination and stimulus.

#### **4.7 Investigate feasibility of shared service CRIS model**

A few national services have some traction (Narcis, Metis, or Frida, for example). Such a service may offer benefits in the UK, although it should be noted that the scale is rather different. A shared service model could offer advantages for smaller HEIs, but shared service models can be difficult to bring to fruition. A feasibility study would be needed to establish the cost-benefits and likely chance of success.