

# How a CRIS can drive improvements in information management

Anna Clements, Helen Reddy  
University of St Andrews

## Summary

This paper describes how the introduction of a CRIS (Current Research Information System) at the University of St Andrews and its use in UK research assessment has highlighted inefficiencies in information management and has led to a programme of improvement based on enterprise architecture principles and practices at the University. The conclusion from our experiences so far is that because the CRIS connects information from many areas of the University (Researchers, Human Resources, Student Records, Finance, Projects, Library, Knowledge Transfer) both culture and organisational structures are as important as technical infrastructure in ensuring the full benefits of a CRIS are realised.

## 1 Background

### 1.1 Research Assessment in the UK

The first Research Assessment Exercise (RAE)<sup>1</sup> was undertaken in 1986. For the first time it introduced an explicit and formalised assessment process of the quality of research. Further exercises followed in 1989, 1992, 1996, 2001, and the latest, in 2008.

The main purpose of the Research Assessment Exercise in the UK is to enable the higher education funding bodies to distribute public funds for research selectively on the basis of quality. Institutions conducting the highest quality research, as judged by peer review, receive a larger proportion of the available funding so that the infrastructure for the highest level of research in the UK is protected and developed. This quality related funding is thus the bedrock of our research funding which then allows researchers to go out and win additional funding from competitive sources.

In the most recent exercise RAE 2008 research was assessed in subject areas, or Units of Assessment<sup>2</sup> (UoA), of which there were 67 grouped together under 15 main panels to ensure consistency across similar disciplines. The results of RAE 2008 took the form of UoA quality profiles for each institutional submission based on weighting the main elements of research assessed (Table 1).

---

<sup>1</sup> Research Assessment Exercise, *Higher Education Funding Council for England, Scottish Funding Council, Higher Education Funding Council for Wales, Department for Employment and Learning (Northern Ireland)*  
<http://www.rae.ac.uk/>

<sup>2</sup> RAE2008 Units of Assessment <http://www.rae.ac.uk/aboutus/uoa.asp>

| RAE 2008 Research Activity 2001 – 2007                                         |                                                                                                 |                                                                                   |
|--------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|
| Research Outputs                                                               | Research Environment                                                                            | Esteem                                                                            |
| Submission of Research Outputs for Peer Review. Up to 4 outputs per Researcher | Submission of information on, for example, Research Strategy, Students, Income & Infrastructure | Submission of Indicators of Esteem<br><i>e.g.</i> Membership of Learned Societies |
| Weighting between 50-90%                                                       | Weighting between 5-25%                                                                         | Weighting between 5-25%                                                           |

**Table 1** *Illustrates the information required for RAE 2008 and the range of weightings allowed for each element. Subject panels developed their own assessment criteria and used appropriate weightings for their discipline.*

Following publication of the results from RAE 2008 the profiles were incorporated into the funding models for distribution of quality related funding for research from the higher education institutional Funding Councils to institutions. The Higher Education Funding Council for England HEFCE is distributing £1.6 billion<sup>3</sup> to higher education institutions across England in the year 2010/11 using the quality profiles from RAE 2008 to inform the allocation.

As well as the monetary value from the RAE, institutions can also make reputational gains by performing well in these exercises. RAE can help to establish that an institution is carrying out significant ‘world leading’ research. The results from the assessment exercises are incorporated into many UK league tables hereby influencing the resulting ranking of institutions (*e.g.* The Times Good University Guide<sup>2</sup>).

## 1.2 Moving towards the Research Excellence Framework

The UK is now moving to a new assessment framework, called the Research Excellence Framework (REF)<sup>5</sup>. The first assessment under this new scheme is due to take place in 2012, with results published in 2013 and it is proposed to inform funding using the results in 2014/15. A change in UK government may postpone the REF by a year or so.

<sup>3</sup> HEFCE announces funding <http://www.hefce.ac.uk/news/HEFCE/2010/grant1011/announce.htm>

<sup>4</sup> Times Good University Guide [http://extras.timesonline.co.uk/tol\\_gug/gooduniversityguide.php](http://extras.timesonline.co.uk/tol_gug/gooduniversityguide.php)

<sup>5</sup> Research Excellence Framework <http://www.hefce.ac.uk/Research/ref/>

## HOW A CRIS CAN DRIVE IMPROVEMENTS IN INFORMATION MANAGEMENT

Guidance on submission to the REF is due to be published late 2010 however all stakeholders are expecting similar data requirements to those for the RAE plus two new requirements : (i) citation data on research outputs where appropriate and (ii) an impact statement and case studies to demonstrate the impact of their research (Table 2).

| REF Research Activity 2008 – 2012                                                                                            |                                                       |                                                                                                 |
|------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------|-------------------------------------------------------------------------------------------------|
| Research Outputs                                                                                                             | Research Impact                                       | Research Environment                                                                            |
| Submission of Research Outputs for Expert review informed by citation data where appropriate<br><br>3 or 4 Outputs per staff | Submission of Impact Statement<br>Impact Case Studies | Submission of information on, for example, Research Strategy, Students, Income & Infrastructure |
| Weighting 60%                                                                                                                | Weighting 25%                                         | Weighting 15%                                                                                   |

**Table 2** *Illustrates the proposed information required for the REF and the relative weightings currently proposed for each element.*

Comparing Tables 1 and 2 it can be seen that much, if not all, of the data required for RAE 2008 is likely to be required for REF; in practice, with the introduction of citation data and the assessment of impact stream, even more information may be needed for the REF submission to the funding councils.

In summary, the need to collect, maintain, quality control and understand research information will continue to be a major requirement for UK higher education institutions. With the introduction of citation analysis and gathering evidence of the impact of research in the REF, the potential burden on these institutions is likely to be even greater and the innovative use of IT to minimise this increase in burden is a high priority in the sector.

## 2 Collecting data for RAE 2008

The project to start collecting research data at the University of St Andrews for submission to RAE 2008 started in 2005 and was declared, by the Deputy Principal, to be *the* priority across the University with resources made available to ensure the data could be collected, checked, corrected and re-checked. The core team was made up of 4 people from the Research Policy Office and Business Improvements, with additional resource called in from the key data providers of Human Resources, Finance and the Registry (for Student records) as needed. At one stage temporary staff were employed to assist retrieval and electronic registering of historic paper data records. Individ-

ual researchers and school administrators and managers were involved in entering data on publications and esteem factors, and checking centrally held data on staff, students and income. In addition, in the final months prior to submission a team of staff from the university Library were tasked with checking the bibliographic data of all 2000 outputs submitted, against the source item (journal article, book, conference proceeding, etc). These full-text source items had to be physically or electronically submitted to the RAE peer review panels.

The data collection process was managed centrally with considerable use made of an in house Research Expertise Database to collect, select and check publications. We adopted the principle of 'enter once; re-use many' at the start of the process in order to ensure errors or omissions in data were corrected at source. However, although the Research Expertise database had links to the other central systems such links were either manual or inflexible and they were not able to provide information in the required formats for the RAE decision-making at the University, nor submission to the RAE. This meant much of the process was supported by a standalone MS Access database and the use of spreadsheets passed between central units and academic schools.

## **2.1 Issues**

The main problems we faced in trying to collect the data together were :

- Silos of (virtually) inaccessible data held in each central administrative system (Human Resources, Finance, Student Records). Data was made available but, mostly, via emailing spreadsheets which had to be prepared each time. A resource intensive procedure for both data provider and user. If corrections were then needed, these would be emailed back for keying in by central administrative staff with the School left with no way of verifying the changes without setting the whole cycle off again.
- Missing/incomplete data. In some cases, for example with research students going back to 2000, we did not have electronic data on award dates and funders. This meant employing temporary staff to root out the original paperwork in order to find the information and enter it into the system.
- Incompatible data sets. For example in the case of funding for research grants and research students we have two similar, but different, lists of funders. Each was maintained by a different section of our Finance department, working with two different systems. Each, therefore, had to be mapped to the RAE definitions for funding sources.

## **2.2 Lessons learnt**

While there were some improvements to research data quality and information management as a result of RAE 2008 data collection exercise, most notably ensuring data was collected once and re-used where possible, the senior management at the University recognized that considerable work still needed to be done to embed these processes within central units so that managers and schools could be confident that centrally maintained data are correct, complete, up to date and accessible.

## HOW A CRIS CAN DRIVE IMPROVEMENTS IN INFORMATION MANAGEMENT

With so much resource spent in cleaning up our research information data for the RAE we did not want to revert to the ‘dark ages’ of data silos. But, after the data were submitted to RAE 2008 there was a real danger that the old issues would gradually return.

Similar concerns were voiced by other institutions and there was a groundswell of opinion, particularly within Scotland where the research manager network had strengthened during the RAE 2008 process, that a collaborative way forward should be developed.

At St Andrews a decision was made in June 2008 to invest in a state-of-the-art CRIS either through improving our in house Research Expertise database or implementing a third party solution. Despite its short-comings, the Research Expertise Database had provided undeniable evidence of the benefits of joining up research information, entering data once/re-using many and making the information accessible. The Institution recognised continued investment was required to embed and sustain the work that had already been done.

After evaluating the market and the existing in house option, we chose to replace the Research Expertise Database with a CERIF-based CRIS, the PURE<sup>3</sup> product from Atira, in a joint project with the University of Aberdeen.

For the medium to longer-term, a programme to improve information management across the Institution has been put in place with the aim of managing all administrative data as an institutional asset and moving away from the traditional siloed approach. Although technology will act as an enabler here, as the CERIF-CRIS has for research information management, it is the governance and procedural changes that will come from this that are, arguably, more important.

### 3 Managing data as an Institutional Asset

As a first step, the post of Data Architect was created with responsibility for developing an institutional-wide view of data and working with others across the institution to improve data management.

#### 3.1 Information Management Principles

As part of this process a set of Information Management principles (*Table 3*) based on The Open Group Architecture Framework (TOGAF<sup>4</sup>) has been drawn up. TOGAF is the Enterprise Architecture framework which is currently being explored by several Institutions in the UK on a JISC-

---

<sup>3</sup> PURE standard CRIS and repository, Atira A/S, Aalborg, Denmark <http://www.atira.dk/en/pure/>

<sup>4</sup> The Open Group Architecture Framework [www.theopengroup.org](http://www.theopengroup.org)

funded Early Adopter Scheme<sup>5</sup>. In fact the principles in Table 3 are based on those developed by Roehampton University, a member of the Early Adopter Scheme. St Andrews is not a member of the pilot group but we are part of the wider Enterprise Architecture Practice Group<sup>6</sup>, as *Explorers*, defined as an Institution that *has decided to research, investigate EA and develop a case*.

| 1 Business Principles                                |                                                                                                                            |                                                                                                                                                                    |                                                                                                                                                                                              |
|------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Principle                                            | Summary                                                                                                                    | Rationale                                                                                                                                                          | Implications                                                                                                                                                                                 |
| <b>Primacy of Principles</b>                         | All parts of the University must abide by the principles                                                                   | The only way to provide consistent and quality information                                                                                                         | Developments / changes will be evaluated against this principle                                                                                                                              |
| <b>Maximise the benefit to the University</b>        | Processes and systems must maximise the benefit to the University as a whole. Corporate requirements drive IT strategy.    | Recognition of the needs of the University above individual schools and departments.                                                                               | Change in working practices. Possibly conceding own preferences. Development work is within the "plan". A forum must exist to take these decisions i.e. is this work in the "plan" or not.   |
| <b>Business must continue</b>                        | Key central operations must be maintained in spite of system interruptions.                                                | We have huge dependence on systems to carry out the University business. Alternative operation must be implemented.                                                | Management of the risks of business interruption. Recovery and business continuity must be addressed.                                                                                        |
| <b>Common use applications</b>                       | Exploiting the functionality of one system is preferable to developing separate systems                                    | We must avoid duplication of data and effort                                                                                                                       | Data needs to be standardised and understood. Default position is that central systems should be used in preference to local systems.                                                        |
| <b>Information Management is everyone's business</b> | All parts of the University participate in decisions about information management required to deliver business objectives. | Major IT projects, other than those that are purely infrastructural, should be led by the business process owner e.g. Unit or School or PO, not the IT department. | Acceptance of responsibility. Commitment of resources. Clearly defined owners of business processes. Clearly defined, communicated and enforced policies relating to information management. |
| 2 Data Principles                                    |                                                                                                                            |                                                                                                                                                                    |                                                                                                                                                                                              |
| Principle                                            | Summary                                                                                                                    | Rationale                                                                                                                                                          | Implications                                                                                                                                                                                 |
| <b>Data as Asset</b>                                 | Data are a university asset and must be managed accordingly                                                                | Accurate, timely data assists accurate, timely decision                                                                                                            | Staff must accept responsibility for their data. Data checking                                                                                                                               |

<sup>5</sup> JISC Enterprise Architectures Group Pilot

<http://www.jisc.ac.uk/whatwedo/programmes/elearningcapital/enterprisearchitectures.aspx>

<sup>6</sup> JISC Enterprise Architectures and Shared Services

<http://www.jisc.ac.uk/whatwedo/programmes/flexibleservicedelivery/enterprisearchitecture.aspx>

## HOW A CRIS CAN DRIVE IMPROVEMENTS IN INFORMATION MANAGEMENT

|                           |                                                                                 |                                                                                                                                                           |                                                                                                                                                                                                                                                                                                                              |
|---------------------------|---------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                           |                                                                                 | making.                                                                                                                                                   | procedures should be in place. Responsibility for data "trusteeship" must be set at the highest level. A forum with comprehensive University-wide representation should decide on process changes suggested by the trustee. Significant cultural change                                                                      |
| <b>Data is Shared</b>     | Users must have acceptable, justifiable and protected access to the data needed | Quality and efficiency in data access and use. "Shared" from a single source with limited sources to ensure consistency and remove duplication of effort. | Expanded use of data warehouse to assert common standards and definitions and act as institutional data catalogue. Interoperability requirements key to new system implementations.                                                                                                                                          |
| <b>Data is Accessible</b> | Data is useless unless it is accessible where and when needed                   | Efficiency, effectiveness in processes and decision making. Lack of easy access to information creates additional work in discovery and aggregation.      | Common web-enabled BI/Reporting tool<br>Policies and procedures for data security and access.<br>Electronic Records Management and retention policies, procedures and systems need to be defined and implemented.                                                                                                            |
| <b>Data has Trustee</b>   | This is the person responsible for data quality in the central units            | Users of the data must trust that the central data is complete, correct, up to date and suitable for their needs.                                         | Need to understand the data sources and changing requirements of users/customers of the data<br>Captured once and validated as close to source as possible [cut out form-filling/rekeying]<br>Trustee must be able to generate confidence in the data                                                                        |
| <b>Data Security</b>      | Data must be protected from unauthorised exposure and use                       | Legal and reputational imperatives                                                                                                                        | Data shared between systems may be subject to differing security arrangements within applications so security needs to be thought about at item level.<br>These concerns need to be addressed at the very beginning of a development.<br>Data access from outside institution on non University hardware needs to be secure. |
| <b>Common Vocabulary</b>  | Data defined once and definition is understandable by end users.                | Common vocabulary assists communication and assists in data integration / exchange                                                                        | Many people need to understand this not just the chosen few. Parochial definitions must be substituted by institutional definitions.<br>Make use of industry [community] standard definitions e.g. ISO, HESA, CERIF                                                                                                          |

*Table 3 : University of St Andrews Information Management Principles*

We are at the early stages of the programme but the framework provided by these principles is already acting to cut through the ‘walls’ between the silos of information held in the different central administrative units and academic schools. These ‘walls’ are technical, with poor interoperability functionality between administrative systems, and organisational/cultural, with central units historically inward looking. There is no formal governance procedure in place to force adherence to these principles but they are to become part of the process that prioritises projects at the University level and projects will be expected to demonstrate compliance to these principles.

## 4 Conclusion – CERIF-CRIS as an exemplar and enabler

The use of both the in house Research Expertise database during RAE 2008 and the new CERIF-CRIS has demonstrated many of the principles listed in Table 3. Data are shared, accessible, quality-controlled, secure, in the CERIF-CRIS uses an internationally recognized standard format and re-uses data from third party sources, such as Sherpa-Romeo<sup>7</sup>, Web of Science<sup>8</sup>, PubMed<sup>9</sup> and arXiv<sup>10</sup>. Most importantly data are managed as a cross-institutional asset with all stakeholders – researchers, central administrative units, library, school and university management – having access to and responsibility for the data in their domain.

This has led to an increased awareness within the central units of how they ‘fit into the bigger picture’ of research within the University. In practical terms this has resulted in the recruitment of an increasing number of highly data literate staff within the units and a redefinition of some existing roles in order to build a team of data managers embedded within the units responsible for data quality. This team works with the Data Architect in order to ensure co-ordination across the Institution.

Beyond the boundaries of the University, the CERIF-CRIS is part of a JISC-funded pilot, CRIS-Pool : Using CERIF-XML to integrate heterogeneous research information from several institutions into a single portal. Partners include the University of St Andrews (lead), University of Edinburgh, University of Glasgow, SUPA<sup>11</sup> (the Scottish Universities Physics Alliance), euro-CRIS and Atira. This project has been funded as part of the JISC Information Environment Programme<sup>12</sup> looking specifically at research information management and the use of CERIF-XML as a common exchange format across the UK research sector.

---

<sup>7</sup> Sherpa Romeo database of publishers copyright policies

<http://www.sherpa.ac.uk/romeo/>

<sup>8</sup> Thomson Reuters Web of Science

[http://thomsonreuters.com/products\\_services/science/science\\_products/a-z/web\\_of\\_science](http://thomsonreuters.com/products_services/science/science_products/a-z/web_of_science)

<sup>9</sup> Pubmed <http://www.ncbi.nlm.nih.gov/pubmed/>

<sup>10</sup> arXiv, Cornell University <http://arxiv.org/>

<sup>11</sup> Scottish Universities Physics Alliance <http://www.supa.ac.uk/>

<sup>12</sup> JISC Information Environment Programme : Research Information

## HOW A CRIS CAN DRIVE IMPROVEMENTS IN INFORMATION MANAGEMENT

### Contact Information

Anna Clements  
Data Architect  
Business Improvements  
University of St Andrews  
KY16 9AL  
United Kingdom

Email : [akc@st-andrews.ac.uk](mailto:akc@st-andrews.ac.uk)  
Tel : +44 334 462761

Dr Helen Reddy  
Senior Research Policy Officer  
Office of the Principal  
University of St Andrews  
KY16 9AJ  
United Kingdom

Email : [hr2@st-andrews.ac.uk](mailto:hr2@st-andrews.ac.uk)  
Tel : +44 334 462429