

## Research information meets research data management

By Anna Clements, Enterprise Architect, University of St Andrews, St Andrews, Scotland

### CRIS-IR from the outset

At the University of St Andrews we have had an integrated research information infrastructure since 2006. The overall architecture has remained unchanged, with a current research information system (CRIS) providing tools for managers and researchers to access all research-related institutional data from corporate systems such as Human Resources, Student Records, Research Grants and Finance. In addition the CRIS stores research outputs, outcomes, impacts and activities either via harvesting from third-party sources, such as Scopus and Web of Science, or via manual data entry by researchers.



Anna Clements

The technology has been updated over the years, with an in-house CRIS being replaced by the market leader, Pure<sup>1</sup> CRIS, in 2010.

From the outset the CRIS has been integrated with our open access (OA) institutional repository running in the DSpace platform. The CRIS is the single, “golden” data source for the research publication metadata and, where a full-text version can be made OA, these metadata are pushed through to the institutional repository (IR) together with the full text. All workflow on copyright clearance and embargo periods is done in the CRIS. Thus the IR acts as a genuine *repository* of openly accessible documents.

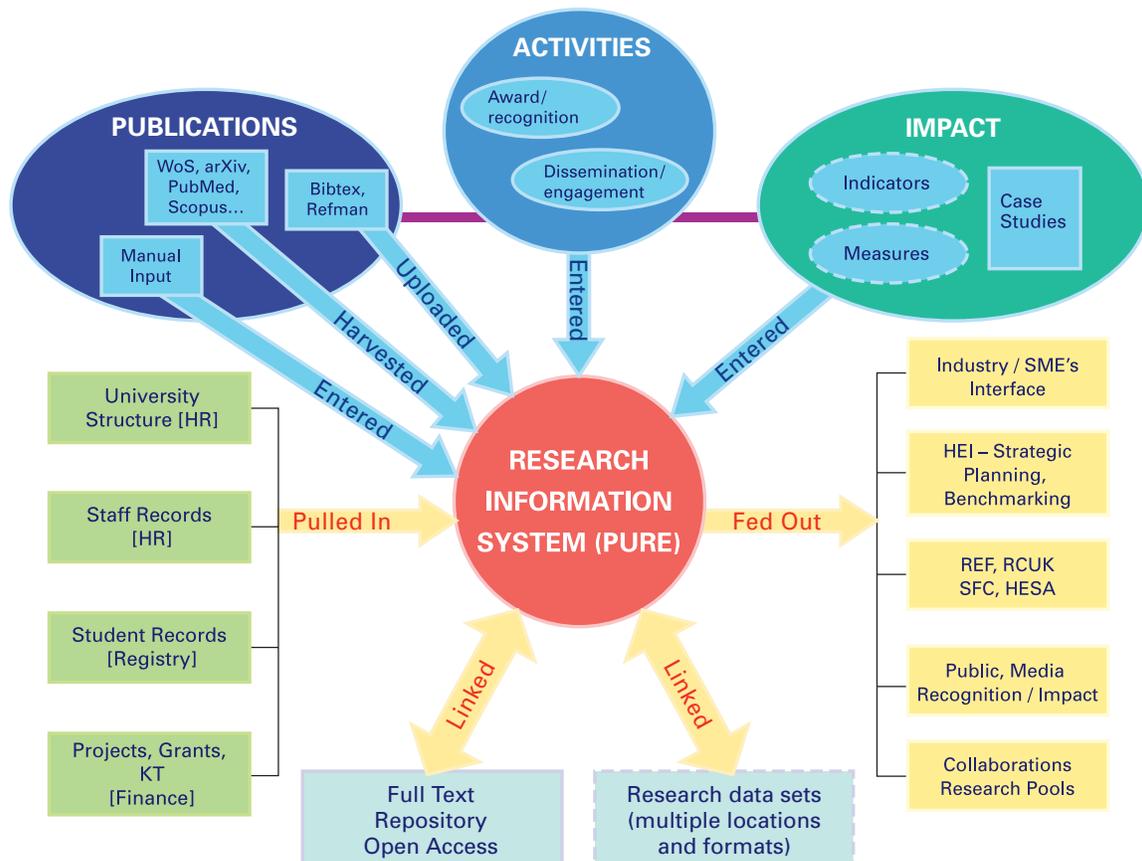
### Information management principles

The CRIS-IR is a prime example of the successful practical application of the principles of good information management:

- Data are entered once, as close to source as possible, and reused.
- Data stewards keep control of the data within their domain of expertise.
- Data are available only to those who need it, in the format needed and when needed.
- Data standards such as CERIF<sup>2</sup> and existing data sources such as Web of Knowledge<sup>3</sup> and Scopus<sup>4</sup> are used.

>> page 2

Figure 1: Existing research information architecture – indicating future link to data sets





Anna Clements << page 1

This approach has also resulted in close cooperation across several university functions, principally the Research Policy Office, the Library and IT Services. This strong sense of co-ownership of the CRIS-IR infrastructure has resulted in clear and coordinated communication to researchers about services and tools available and an effective mechanism for gathering feedback to improve these services. It has also ensured a very strong technical and support infrastructure to build on new services, such as research data management.

### Extending the CRIS-IR

In the UK, research data management is high on the agenda of funders and, therefore, institutions. A new joint OA policy from the UK's seven Research Councils (RCUK) goes into effect April 1, 2013. It will not only require the article to be published in an RCUK OA-compliant journal, but also specifies that it "must include ... a statement on how the underlying research materials — such as data, samples or models — can be assessed."

In addition, data management plans or similar are required for all RCUK grant applications. The Engineering and Physical Sciences Research Council in particular has set a May 1, 2015, deadline for institutions to have the policies, processes, infrastructure and tools in place to satisfy the main principles "that publicly funded research data should generally be made as widely and freely available as possible in a timely and responsible manner; and, secondly, that the research process should not be damaged by the inappropriate release of such data."

At first these requirements appear contradictory; the Research Councils are working with universities and other stakeholders to provide further clarification. However, whatever the interpretation

of the words, the practical consequences are far reaching — not least because universities need to know what research data its researchers generate (who funded it, formats used, quantities, how sensitive it is, where it is, whether it needs to be stored, and so on). The minimum requirement is a catalog of research data sets, and that is where our existing CRIS-IR infrastructure comes in. We can link our research data sets to the information on people, organizations, projects, funding, outputs, impacts and activities that we already have in one place. We are still at the early stages of determining where the research data are and should be stored, how much we need to keep, and for how long — but we are at least in the position to build a catalog of the data within the existing research information environment.

### CRIS-IR or IR-CRIS: Who cares?

I am often asked to talk about CRIS-IR and where the responsibility for such systems should lie within an institution — or perhaps more controversially, "Do we need an IR if we have a CRIS?" or vice versa. My answer: Stop thinking about systems and think instead about services. What people, processes, tools and standards are available internally or externally that can best deliver the necessary services now and in the future? At St Andrews we have certainly benefited from this approach and concentrated our efforts into delivering joined-up services to our researchers and research managers, whatever the system. **LC**

✉ [akc@st-andrews.ac.uk](mailto:akc@st-andrews.ac.uk)

### References

<sup>1</sup> <http://www.atira.dk/en/pure/>

<sup>2</sup> Common European Research Information Format, see [www.euroCRIS.org](http://www.euroCRIS.org)

<sup>3</sup> [http://thomsonreuters.com/products\\_services/science/science\\_products/a-z/isi\\_web\\_of\\_knowledge/](http://thomsonreuters.com/products_services/science/science_products/a-z/isi_web_of_knowledge/)

<sup>4</sup> <http://www.info.sciverse.com/scopus>