

Are repositories ready for REF? Institutional repositories to support periodic research assessment in the United Kingdom

The institutional repository is rapidly gaining a role as a definitive record of a higher education institution's contribution to knowledge through its research activities. The repository provides visibility for an institution's scholarly output by providing open access to self-archived material in a single location, including such important material as unpublished grey literature, and is increasingly becoming a mandatory part of the scholar's research cycle.

Where such repositories are presently often less than fully effective is in fulfilling all the requirements of a Current Research Information System (CRIS). CRISs are becoming increasingly necessary as a result of the imperatives of funding environments: fund managers, for instance, often require detailed information on research activities and their financial management to ensure the optimal use of increasingly scarce funds. They also provide valuable information for administrators to allow easy analyses of activities, particularly quantitative measurements, bibliometric and otherwise, which are increasingly required as measures of performance.

A CRIS will in most cases, therefore, require additional content to that can be provided by a repository alone, but there is a large overlap between the requirements of both, and some method of interfacing between the two will save much time and effort that would otherwise be expended in duplicating information between the two systems. The *Readiness for Ref (R4R)* project at King's College, London¹, aims to provide mechanisms for interfacing of this type by the use of the CERIF² interchange standard for research data. It will also allow repositories to fulfil some of the requirements of a CRIS: such 'enhanced' repositories will be particularly appealing to smaller institutions.

CERIF, maintained by euroCRIS³, is a long-established data model for the encoding of the information that characterises a CRIS. It includes, amongst other things, information on people, research outputs and funding, and provides extensive linking mechanisms to allow a rich web of associations to be encoded between these elements, so affording even the most complex research environment to be represented. CERIF is format-neutral, but is available as a series of XML schemas for ready interchangeability.

The R4R project looks forward to the 2012 Research Excellence Framework (REF) exercise in the United Kingdom, which, like the earlier Research Assessment Exercise (RAE) in 2008, will require higher education institutions to provide an audit of their research activities and outputs on the basis of which future research funding will be determined. The project aims particularly to provide a mechanism for interfacing data held in institutional repositories with the requirements of this exercise by producing an XML schema (CERIF4REF) which will mediate between this data, the requirements of the REF, and CERIF. Information encoded in this schema will be used to generate, through XSLT stylesheets, instances of the REF's own XML schema and CERIF XML files.

A series of plugins will be produced as part of the project to extract data from *Dspace*, *Fedora* and *ePrints* repositories into CERIF4REF, from where it will, with any necessary added data dictated by the requirements of the REF, be ready for incorporation into REF or CERIF. The project should, by these mechanisms, render the whole REF exercise much less burdensome, as well as, for the first time, making it possible to present information on a large portion of the UK's research activities available for the first time in the CERIF format. With such a critical mass of information in CERIF, future systems should more readily be built on interchangeable and malleable standard.

1 <http://www.kcl.ac.uk/iss/cerch/projects/portfolio/r4r.html>

2 <http://www.eurocris.org/cerif/introduction/>

3 <http://www.eurocris.org/>

