# EPrints: A Hybrid CRIS/Repository

*Leslie Carr, University of Southampton & EPrints Repository Platform*

## WORKSHOP ON CRIS, CERIF AND INSTITUTIONAL REPOSITORIES 10-11 May 2010, Rome

You must remember this: a CRIS is just a CRIS…

The fundamental things apply as time goes by.

*With apologies to* *Herman Hupfeld*.

It is a truth universally acknowledged that a repository collects and manages research papers, data, reports, patents and software (curating many kinds of research outputs for immediate access and long-term preservation) but that it is *just* *one* of the systems that a CRIS has to interact with. Specialised research databases are used not only for research outputs, but for human resources, project management, finance, grant funding and research expertise. A CRIS offers a comprehensive overview of an institution’s research activities, pulling together information from all of these research-relevant databases to improve the administrative processes of an institution.

EPrints, as a repository, supports the CERIF standard to allow it to communicate with CRIS systems, exchanging relevant information about research publications, to provide a comprehensive picture of the research aspect of an institution’s business activities.

However, in some universities a repository also provides some degree of research management reporting – publications by research group and topic, according to citation impact and other criteria. This uses information from the HR department (organizational structure in terms of research group affiliation), requiring it to be incorporated into the repository. *In doing so, the repository starts to act (in a small and incomplete way) like a CRIS.*

Furthermore, information that has previously been strictly administrative in nature (for example, grant administration) and hence subject to administrative processes and controls, is increasingly becoming relevant to the researcher as an end-user. Not only are funded project figures, titles and dates required from a grant database for CVs, but they are part of the research story-telling that underpins both impact reporting and marketing that researchers themselves have to undertake.

These new tasks create further requirements of the data collected about grants: as well as the basic facts and figures, descriptions, pictures (logos!) and press-releases are required for presentations, reports and websites. Mainly run behind the scenes in administrative offices, grant databases are not configured for descriptive, narrative and pictorial information. In fact, such information has previously been relegated to an *ad-hoc* position on departmental websites (updated if the researcher has time), but is now becoming increasingly valuable information.

*The JISC Open Impact project is developing a repository of High Impact research in Computer Science, run by the UK’s Learned Society (BCS). It helps researchers collect and collate information about their projects, research outputs, events, teams and individuals to form a coherent description of a decade or more of activity.*

Whereas grants databases are usually run as B-B services (mediating information between funding bodies and research institutions), repositories have typically adopted a B-C role, mediating research information between an institution and the thousands of grass-roots researchers who work in it (as well as the research community outside it). *Therefore it does not seem unusual for a repository to offer other researcher-serving information services.*

In a recent workshop run by the JISC *Readiness for REF* project[[1]](#footnote-0), a straw poll of the attendees indicated that they would favour their repositories taking on information management roles more usually associated with a CRIS (specifically information about projects and funding organisations).

## The Hybrid Repository

As well as accommodating the CERIF data model for interchange, EPrints is extending its internal data models to accommodate not just “publications” but “projects” and “organisations” as well. Such CERIFization requires a repository to have many separate datasets, all linked together via explicit relationships; projects and funding organisations that were just names typed into the paper’s metadata record are now objects in their own right linked to that metadata record.

Hence, data about projects, publishers, funders etc can be managed in EPrints just like publications; they can be edited, displayed, searched, exported. Because a repository is designed for engagement with end users, the user interface and capabilities may be better (richer, easier) than those associated with an administrative database, and so it may be easier for users to keep the repository up-to-date than the administrative master.

## Questions

What should the relationship be between a repository and a CRIS? Or the repository and the CRIS component systems? The kinds of research reporting that is being undertaken is much richer than previously: does the CRIS model mediating between independent systems still work?

1. http://www.kcl.ac.uk/iss/cerch/projects/portfolio/r4r.html [↑](#footnote-ref-0)