

Looking deep at current research information systems: The Information Science perspective in Higher Education

Sofia Fernandes

Universidade Católica Portuguesa, Lisboa, Portugal

Abstract: The objective of this paper is to present and discuss a project that is being developed to implement a Current Research Information System (CRIS) in a Higher Education Institution (HEI). This article discusses the various steps that are being followed to design, define and implement a CRIS system, while at the same time presenting a comparative study between two commercially available CRIS solutions and a Dspace CRIS, highlighting the most common key features that should make part of such a R&D management and support system. Implementing a CRIS in a HEI requires a methodology plan that takes into consideration all the institutional, technological and organizational restrictions and requirements, and in this paper we share an Information Management perspective on this challenge, where the main concern is being able to help design and develop an optimized and efficient R&D managing system that fosters and leverages research results and their worldwide dissemination and visibility.

Keywords: Current Research Information System, Project Management, Institutional Repository, Information Systems, Information Management, Research Management

1. Introduction

Research is one of the most important undertakings of HEIs all over the world. The information that stems from research is accounted as fundamental to improve the development of knowledge.

The new information technologies that are nowadays available at universities have an important role improving the information management. However, the research lifecycle demands are progressively challenging. Given its importance, the R&D activities of a HEI increasingly require a structured and interrelated information system, which should support R&D managers and support staff in defining and monitor a R&D strategic plan, while at the same time complying with the requirements of national and international funding agencies.

Institutional repositories (IR) had answered, over the last years, some of these requirements by providing consolidated reports about institutional publications. Nonetheless, IRs still fail to properly match publications to researchers, project management, or performance evaluation, or perform more complex R&D management and reports.

As a result, a Current Research Information System (CRIS) must provide such an extended view on all the R&D activities in a HEI, and in order to do so it needs to be interconnected to the IR (where both publications and/or data are deposited), the ERP (where the institution's financial and human resources are managed), the project management system (where information about project funding, partners, etc., is stored and managed), the academic management system, as well as the staff performance evaluation system.

In addition, a CRIS should be an optimized and efficient R&D managing system that fosters and maximizes worldwide visibility and dissemination of researchers, their research, and their affiliated HEIs (Kahn, Higgs, Davidson, & Jones, 2014; Mallikarjun & Kumar H., 2015). This paves the way to a broader recognition of the R&D work done at the HEI, increasing further worldwide collaborations and partnerships, and making the institution more attractive to top-level researchers.

As Kahn et al. (2014) demonstrates, the collaboration between researchers and institutions can reduce the efforts done conducting research by sharing and reusing information. These aspects accelerate the creation and knowledge development.

An appropriate CRIS assists the compliance with national and international requirements, namely concerning funding, and improve the scholarly outputs (Lynch, 2014).

2. CRIS assets, features and specifications

A CRIS supports the information management of all R&D activity in an institution. EUROCRIS (2018) defines a CRIS as “any informational tool dedicated to provide access to and disseminate research information”.

The major benefits of a CRIS are related to the strategic and the operational needs of a HEI, and its researchers (see Fig. 1).

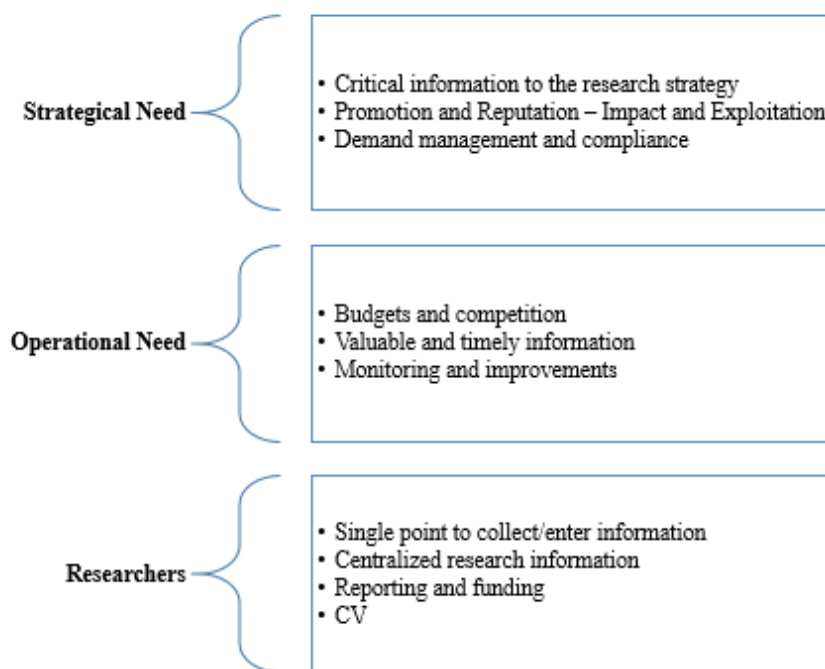


Figure 1 - The main assets provided by a CRIS system.

A CRIS can be quite helpful to most R&D management and strategic requirements by providing critical information about the research activities at the HEI, which can then be used to increase reputation and impact, supporting the compliance with all the external entities and systems. Valuable and timely information about all the R&D activities in a HEI is vital to monitoring and reporting, as well as to quickly act on any required improvements or corrective measures.

Researchers can get a single point to collect or enter research data related to publications or datasets (using ORCID for example) centralizing information that can be used in CV, reports and funding requirements. The following table provides more detailed information about these assets: people, outputs, organizations, finance, activities and impact (see Fig. 2).

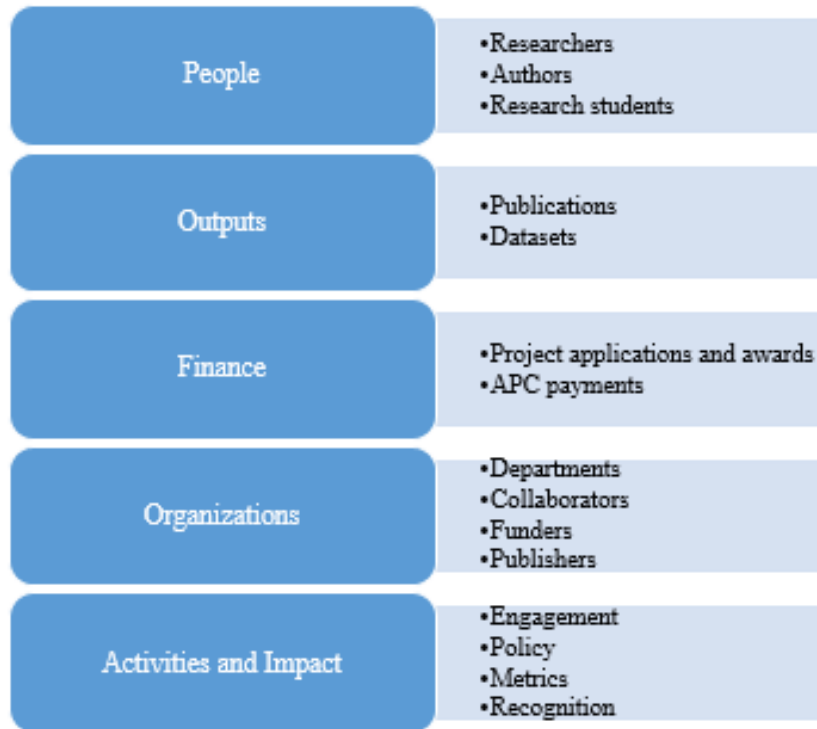


Figure 2 - Specification of the main entities and services associated when developing a CRIS.

A CRIS connects these entities: the researchers and the authors with their outputs (publications and datasets), published in the scope of projects, dealing with article processing charges (APC payments) and funders, collaborating with other departments to achieve recognition and good metrics evaluation.

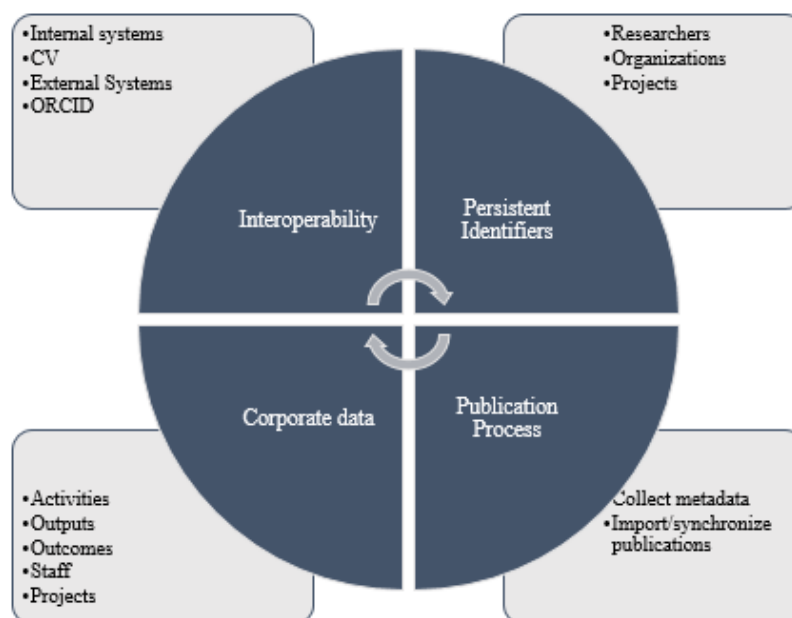


Figure 3 – Representing the key and functional domains of a CRIS.

As such, the key features of a CRIS are: interoperability, persistent identifiers, corporate data and the publication process (see Fig. 3).

It is important to guarantee interoperability between the CRIS and several internal systems of a HEI, like the academic management system for example, a researcher's CV platform (specially if an official governmental national platform exists), and any widely used and recognized external systems, such ORCID. One of the main features of a CRIS is the existence of persistent identifiers for researchers, organizations and projects. The CRIS must collect, import and synchronize publications in order to support the publication process. Finally, the outputs be connected to the activities, the outcomes, staff and projects.

3. Planning a CRIS

In this section it is presented the planning methodology to implement a CRIS based in the assets, features and domains presented during the previous section. The approach includes three main frontlines: organizational structure, systems and compliance (see Fig. 4).

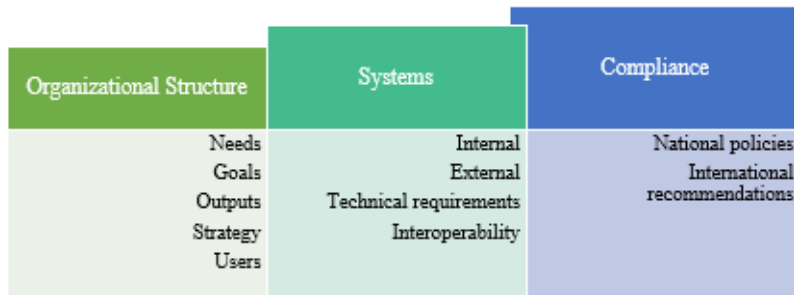


Figure 4 – Planning methodology for implement a CRIS.

Organizational structure:

- 1- Identify the needs and goals (outputs);
- 2- Determine the users;
- 3- Consider the organizational strategy;

Systems:

- 4- Recognise internal systems;
- 5- Name external systems;
- 6- Verify technical requirements;
- 7- Assure interoperability;

Compliance:

- 8- Assume national and international compliance.

Given the identified requirements, and the proposed methodology, the final step is to explore solutions to the development and implementation of a CRIS solutions. Different solutions are possible, namely: “in-house” development of a CRIS, commercial systems, or extended IRs (e.g. DSpace CRIS).

Pure (Elsevier) and Converis (Clarivate Analytics) are some of the commercial solutions.

Elsevier (2018) describes Pure as a “sophisticated research information management system – used as a system-of-record for research-related data management by universities, as a research profiling solution, as a research networking solution, as an academic activity reporting solution and as a grant management solution. To ensure maximum value as soon as possible, while accommodating specific customer needs, we offer comprehensive setup support options.”.

Clarivate Analytics (2018) set up the goals of Converis: “collect all your institutional information in a single system and be able to report on it; integrate information from all other sources”.

Analysing these two definitions, we can understand that Pure intends to offer a comprehensive research information management system that involves research profiling, record and information related and reporting. Converis proposes the impression of a simplified solution by integrating information in one place from other data sources.

In order to evaluate these two systems and DSpace CRIS, Table 1 lists the product features of each proposal.

Researcher Features	CONVERIS	PURE	DSpace CRIS
Data Source	WoS, Scopus, Internal data can be configured	Scopus, WoS, ORCID, PubMed, NIH, RePORTER, SciVal Funding, Embase, Mendeley, arXive, WorldCat, Cross Ref, Journal TOC, CAB Abstracts. Internal systems (Human Resources, Finance...) and Repositories	PubMed, Scopus, Web of Science, ORCID integration for both public and member APIs, Internal systems (Human Resources...)
Auto-ingest	Yes - can be used pre-populated and alerts recommend new data	Yes	
User Profiles	Yes - all	Yes - all	Yes - all
Updating Profiles		Yes - import form	CERIF 1.6 XML
Disambiguate Authors	Yes	Yes	Yes
Networking	Yes	Yes	Yes
Expertise vs. Funding Opportunities	In some fields	Yes - advanced search and web profile search	Yes

Table 1 – Main product features related to Researchers.

The number of data sources available is higher in Pure, that means that researchers can import their publications from many databases.

Metadata and Sources feature	CONVERIS	PURE	DSpace CRIS
Thesaurus	Yes	Yes - Fingerprint engine	Yes
Metadata Model	Dublin Core, Social Media, compatible with VIVO and others	VIVO mapping	Dublin Core
Linked Open Data	Yes - can be used to power VIVO	Yes	Yes
Institutional Systems	Yes	Web Services and local archives	SOAP Web Services for READ-ONLY access to CRIS information
External Systems	Yes - WoS, ORCID, PubMed, EuroPubmed	Yes - Scopus, WoS, ORCID, PubMed, NIH, RePORTER, SciVal Funding, Embase, Mendeley, arXive, WorldCat, Cross Ref, Journal TOC, CAB Abstracts	PubMed, Scopus, Web of Science
Integration with other products	Yes - Pentaho Research Analytics and inCites	Yes - Funding Discovery module	CKAN Integration - Research Data Management
Connectivity with other implementation	No	Yes - members VIVO can search in a single networks	
Export format types	XML, .mbb, XSL	XML, RDF, SPARQL, CSV, CERIF XML, MS WORD, XLS, PDF, ATOM/XML, Web Services, EndNote/Reference Manager, BibTex	CERIF XML 1.6

Table 2 – Main product features concerning metadata and sources.

Pure offers an interesting feature, Fingerprint, that is based in data mining and extracts keywords to the authors. These keywords provide automatic text extraction to build fields and research orientations useful to find collaboration networks.

Dublin Core metadata is used to supplement existing methods for searching and indexing Web-based metadata, regardless of whether the corresponding resource is an electronic document or a "real" physical object. Pure and DSpace CRIS have this flexible metadata model.

Web services are an important feature because they can assure interoperability with any system that also provide these web interfaces.

Converis offers the link between ORCID and Converis account: synchronize research works data among researcher's ORCID and Converis records. It collects and connect badges up and include the synchronize option. However, it is not configured to push affiliation data to research record.

Pure also links the researchers to ORCID ID – it can claim works and push affiliations to researchers ORCID records. DSpace CRIS also connects to ORCID.

The export format types available is the most wide-range, it can be helpful to extract information that can be read easily or integrated with other systems.

Hosting Features	CONVERIS	PURE	DSpace CRIS
Local Hosting	Yes	Yes - customer option	Yes
Cloud Hosting	Yes	Yes - standard option	
Dual Environment		Yes – optional diferent host for stage and production environments	
Custom		Yes - premium option	

Table 3 – Main product features about hosting options.

Pure and Converis offer cloud hosting by default, though both have the possibility to be installed in a local server.

4. Conclusions

The implementation of a CRIS system presents many challenges, both in what regards interoperability between IT systems, (re) design of processes and workflows, and changes in the organizational culture of a HEI. However, it is an unavoidable challenge towards the demands that R&D centers nowadays face, in the context of their own hosting HEI regulations, as well as in what concerns requirements imposed by national and international evaluation and funding agencies. On the other hand, institutions need to advance open access to

research carried out by linking to institutional repositories so that they can meet the new open access requirements in funding projects.

Decidedly, CRIS is one of the most important and challenging projects among HEIs nowadays.

A CRIS provides critical information that serves the institutional strategic and operational demands and helps the research lifecycle.

CRIS is all about integration: integrates systems, persons, organizations, finances, activities and impact in a relationship data model. The key features involve interoperability, persistent identifiers, corporate data and support to the publication process.

The planning methodology to implement a CRIS solution has three major milestones: organizational structure, systems and compliance. The CRIS product chosen must follow these approaches.

Analyzing the product features presented, the main highlights go to the connection to data sources and the export format types. Pure has a great connectivity and can connect with the most known databases to import data. The interoperability through web services can be a solution for the most cases. The hosting options can answer to everyone's needs.

Nonetheless, Converis main product feature is the customizing and flexibility because institutions can set up and customize on their own without being dependent on the vendor.

Finally, the main characteristic of DSpace-CRIS is its flexible data model, which allows you to collect and manage research data and information typical of a CRIS system but as an extension of the IR.

References

Eurocris (2018). Current Research Information Systems. Retrieved from: <https://www.eurocris.org/>

Dura Space (2018) Dspace Cris. Retrieved from: <https://wiki.duraspace.org/display/DSPACECRIS/DSpace-CRIS+Home>

Elsevier Research Intelligence (2017). Sistema de Informação de Pesquisa Corrente (Current Research Information System – CRIS) na USP. Retrieved from: <https://www.slideshare.net/sibiusp/sistema-de-informao-de-pesquisa-corrente-current-research-information-system-cris-na-usp>

Clarivate Analytics (2018). Converis. Retrieved from: <https://clarivate.com/products/converis>

Kahn, M., Higgs, R., Davidson, J., & Jones, S. (2014). Research data management in south africa: How we shape up. *Australian Academic & Research Libraries*, 45(4), 296-308. doi:10.1080/00048623.2014.951910

Mallikarjun, D., & Kumar H., A. (2015). (2015). Managing research data in academic institutions: Role of libraries. Paper presented at the Caliber, Shimla. Retrieved from <http://hdl.handle.net/10760/24911>

ORCID (2018). Connecting Research and Researchers. Retrieved from: <https://orcid.org/>