CRIS 2014

DSpace-CRIS@HKU: Achieving Visibility With a CERIF Compliant Open Source System

David T. Palmer[a]*, Andrea Bollini[b], Susanna Mornati[b], Michele Mennielli[b]

a) The University of Hong Kong, Pokfulam, Hong Kong
b) Cineca, Via Magnanelli 6/3, 40033 Casalecchio di Reno

Abstract

The CRIS now at the University of Hong Kong, "HKU Scholars Hub" (the Hub), began as a means to provide visibility to HKU research, people, and their expertise. It now also acts as a tool for management, evaluation and visualization of HKU research. The expected addition of CERIF compliancy to DSpace-CRIS in May 2014, will further enhance the Hub's utility for visibility, collaboration, and benchmarking.

Keywords: HKU; DSpace; DSpace-CRIS; Cineca; Open Source

---

* Corresponding author.
  E-mail address: dtpalmer@hku.hk

1. Introduction

The history behind the development of a CRIS at the University of Hong Kong (HKU) is also the history of the open source project "DSpace-CRIS", developed by Cineca to fulfill HKU's need of extending the information exposed by their Hub. The Hub existed prior to 2009 as a traditional IR managed by the Libraries, and hosted in DSpace. In this year a new 5-year HKU strategic plan began that positioned Knowledge Exchange in equal importance to university Research and Teaching. The HKU Office of Knowledge Exchange (KEO) then accepted a grant proposal from the Libraries to enlarge the scope of the Hub and augment it with new record types and functions to share knowledge and expertise within and outside the University. The Libraries chose the Italian Interuniversity Consortium, Cineca, for its development partner in this project.

Cineca provides CRIS systems to more than 50 Italian universities and institutions and uses DSpace as one of its main solution technologies. The Consortium has been actively supporting the DSpace community for the last seven years and today, two of the 25 DSpace Committers are Cineca employees. For all of these reasons Cineca was able to provide HKU with the know-how and expertise required to implement the project.

The HKU Libraries realized that HKU's research, people and expertise are described and contextualized by more than publications, and that new entities and their attributes would be needed in the Hub. The Libraries wrote specifications on these entities and attributes, and their desired functions and integration. Cineca architected a solution that extended the data model of DSpace with new relational tables for people, grants, and other objects. These sit in parallel to DSpace's traditional one of publications. A customized user interface then allows search and retrieval across all such tables, and serves up an integrated mash-up display to the user.

2. HKU Processes

After authentication with HKU's Single Sign-on, a data owner can edit a few data elements, but can hide from the public any data deemed inappropriate. An authority control system enables disambiguated attribution to authors, PIs and Co-Is, inventors, etc., although most of these individuals are cited by multiple forms of name, in English and Chinese (漢字).

Data for the Hub's displays on publications, people, grants, patents, student supervision, bibliometrics, etc, is regularly extracted from several silos internal to HKU, and external. New functions exist to automate these extracts, load, de-duplicate or merge, automatically link to corresponding Hub records, and build visualizations and reports.

2.1. Visualization & Reports

Using the JavaScript InfoVis Toolkit library (http://philogb.github.io/jit/), Cineca built user interactive visualizations showing up to three degrees of separation for co-authors, co-investigators, co-inventors, co-winners, as well as researchers using same keywords in publications, grant proposals, patents, etc. Similar visualizations exist for organizational units.
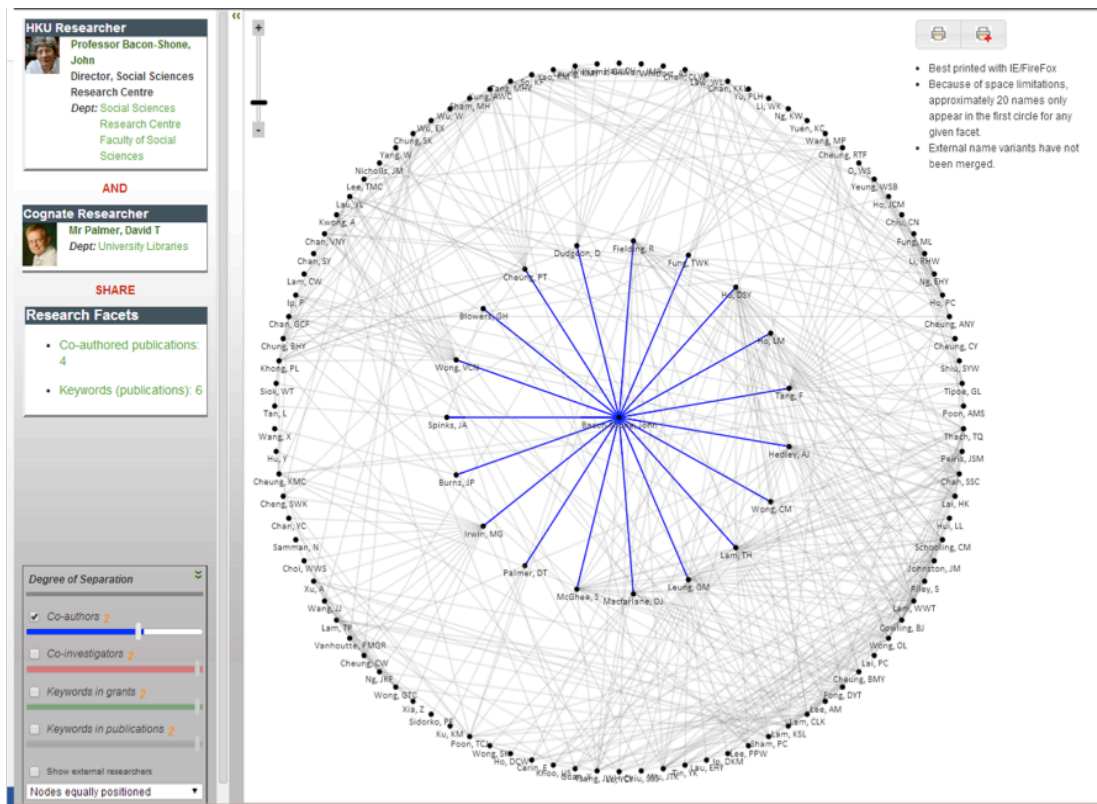
Fig. 1 Researcher Collaborations Visualized

The Hub allows users to set up email alerts on publications added to a collection, department, or author. Users can also set up emailed reports on collections, departments, authors, publications, grants, etc. showing accrual of statistics (daily, weekly or monthly). These statistics can be download and view counts, as well as citation counts from Scopus, WoS, PubMed, Google Scholar Citation, Microsoft Academic Search, etc. Users could further set these reports to show multiple collections, departments, authors, etc, and thereby benchmark the statistics of a desired object against its peers or top performers. All analyzed researchers in the Hub receive this report on their own statistics monthly, and are then free to change the frequency, or add benchmarking objects to this report.

External Citations ResearcherPage: **5**

No.	ResearcherPage	Total Counts / % Yesterday increase				
		Scopus	WoS	PubMed	Google Scholars Citations	Microsoft Academic Search
1	Luk, John Moon Ching	4541 / 0%	3196 / 0%	1358 / 0%	6650 / 0%	36 / 0%
2	Palmer, David Alexander	5 / 0%	0 / -	0 / -	253 / 0%	/ -
3	Palmer, David T	1 / 0%	0 / -	0 / -	24 / 0%	/ -
4	Sidorko, Peter Edward	21 / 0%	3 / 0%	0 / -	56 / 0%	5 / 0%
5	Tam, Paul Kwong Hang	14387 / 0%	9670 / 0%	3933 / 0.05%	/ -	2835 / 0%

Internal Metrics Items: **6**

No.	Items	Total Counts / % Yesterday increase		
		RP Views	Item Views	Item Downloads
1	Luk, John Moon Ching	7954 / 0.05%	48381 / 0.12%	2345 / 0.85%
2	Palmer, David Alexander	2766 / 0.25%	4784 / 0.17%	270 / 10.37%
3	Palmer, David T	15682 / 0.20%	16664 / 0.05%	8044 / 0.11%
4	Sidorko, Peter Edward	7006 / 0.23%	10249 / 0.07%	6271 / 0.14%
5	Tam, Paul Kwong Hang	29501 / 0.13%	122461 / 0.09%	9625 / 0.26%
6	Centre for Comparative and Public Law	7553 / 0.07%	8907 / 0.12%	305 / 0.66%

Fig. 2 Daily Emailed Statistical Report

Recently organizational units, as a new top level object, were added to the Hub. This allows displays and reports cumulated to the centre, department, and faculty. Some of these reports are,

- List of journals used in the department by item count
- List of departmental researchers by publication and grant counts
- List of departmental grants by money amount, year, etc.
- List of departmental publications by citation count from Scopus, WoS, & PubMed
- List of departmental publications & theses by download and view counts
- Departmental researcher interest inventory – sortable & searchable lists of research topics

The Hub regularly sends similar reports to the HKU Management Information Unit, which provides decision support for HKU's Senior Management Team. We are able also to meet ad hoc demands, as in the current Research Assessment Exercise, for such reports by the various individuals, departments and faculties.

2.2. Collective Intelligence

With data extracted from several sources, curated by librarians, and edited by researchers and departments, and with new functions and displays built on top of this data, it is a system of collective intelligence with an impact much greater than the sum of its parts. Besides IR and CRIS, it is also a system for reputation, identity and impact management, research networking and profiling. Further evidence of the above is that the Hub now creates, populates and updates ORCiD records for the HKU professoriate staff, with data previously curated into the Hub.

The Hub has achieved its goal of making HKU research and researchers highly visible. Editorships, contract researches, and keynote speeches have been awarded to HKU researchers because of information discovered in the Hub. The recent edition of Webometrics "Ranking Web of Repositories"[i] saw the Hub move from position 169 in July 2013 to 57 in February 2014, which is the top position in Hong Kong, China, and in Asia.

3. DSpace-CRIS

Pleased with results so far, and with the goal of knowledge exchange, KEO accepted a grant proposal from the Libraries, to fund the modularization and further development of this work for the benefit of the open source community. Known as "DSpace-CRIS", it was released in November 2012 as an open source solution to enrich DSpace 1.8.2.

DSpace-CRIS, now available as an add-on to DSpace 4.x, is compatible and compliant with the CERIF standard, with an export feature in CERIF XML.. Notably, key components of the CERIF data model are supported natively: UUID, timestamped relation, and semantic characterization. Furthermore, although DSpace-CRIS comes with a few predefined entities, the dynamic and flexible (not hardcoded) DSpace-CRIS data model makes it very easy to create new entities and configure them to be compliant with CERIF.

There are several advantages that DSpace-CRIS brings to IRs and to the DSpace community overall:

- CRIS entities apply authority control over an Item thru metadata values;

- DSpace Items can be linked and displayed in the detail page of any CRIS entity;

- Ability to create and display "selected publications" (or a highlighted subset of any other related entity) in the researcher profile;

- Analyzed statistics for each CRIS entity;

- Global & Top related CERIF Entity views & downloads referencing the CRIS entity (projects for researchers, researchers for OrgUnits, etc.);

- Global & Top item views & downloads referencing the CRIS entity;

- email and RSS alerts;

- Article level metrics for PubMed (extensible)

  o   Cited-by count in the item page

     o   Number of articles for researcher

     o   Total citations for researcher (only items in local DSpace database will be counted)

The open source approach of DSpace-CRIS allows Cineca to collaborate with a wide global community. Thanks to the contributions of this community, Cineca can continuously update DSpace-CRIS.

4. Conclusions

Development began on the Hub in 2009 for the purpose of creating higher visibility on HKU research. The ensuing five years of continued development, and ever more visibility on previously dark-archived research, allowed HKU researchers and research administrators to appreciate unexpected CRIS outcomes of this project; such as research evaluation, benchmarking, and academic search engine optimization (ASEO). The addition of CERIF compliancy in May 2014 holds promise to further increase utility in these regards. Open development continues on DSpace-CRIS, now by many players around the world, mutually benefitting from each other success.

5. References

1) Palmer, DT (2011): Knowledge Exchange: a strategy for open access success at The University of Hong Kong. *Chinese Journal of Library and Information Science*. 4 (1), 81-94. Available at http://hub.hku.hk/handle/10722/143266

2) Palmer, DT (2010): The HKU Scholars Hub; unlocking collective intelligence. Proceedings of the 14th International Conference on Electronic Publishing. Hanken School of Economics, Helsinki, June 16-18, 2010, 205-216. Available at http://hub.hku.hk/handle/10722/57674

3) Bollini, Andrea (2013): DSpace-CRIS Project Roadmap. Available at http://cineca.github.io/dspace-cris/roadmap.html

4) Bollini, Andrea (2013): First release of the DSpace CRIS module for DSpace 3. Available at http://dspace.2283337.n4.nabble.com/First-release-of-the-DSpace-CRIS-module-for-DSpace-3-td4666236.html

5) Palmer, DT (2012): DSpace-CRIS, Alpha release available. Available at http://www.duraspace.org/dspace-cris-alpha-release-available

6) Bollini, Andrea; Lo, Chun-yue; Mornati, Susanna; Palmer, DT; Pascarelli, Andrea (2012): Moving DSpace to a fully-featured CRIS System. Available at http://cineca.github.io/dspace-cris/doc/dspace-cris-poster.pdf

7) Bollini Andrea; Mornati, Susanna (2013): Integrate external services in DSpace submission process. Available at http://or2013.net/sites/or2013.net/files/slides/OR2013-Bollini.ppt

8) Mennielli, Michele (2013): DSpace-CRIS: An Open Source Solution. Available at http://www.duraspace.org/sites/duraspace.org/files/12_20_PS_Michelle%20Mennieli_0.pdf

---

[i] http://repositories.webometrics.info/en/Asia