

---

# Explaining abstract concepts with concrete examples – entity registry and research activity data



Cecilia Loureiro-Koechlin  
*Project Analyst, SERS,  
Bodleian Libraries,  
University of Oxford*  
Tel: 01865 280028  
E-mail: [cecilia.loureiro-koechlin@bodleian.ox.ac.uk](mailto:cecilia.loureiro-koechlin@bodleian.ox.ac.uk)

---

Sometimes when we are developing new technologies it is very difficult to explain to our users what we are trying to do and why we are doing it. This latter aspect has to do with how technologies work and how they can be used. Explaining our work to our users is a delicate task. They need to comprehend the basics and implications of it to be able to see the benefits for their own work. But what can we do if we are developing something that they cannot see or touch but which they need to comprehend in order to make the most of it?

I will attempt to answer this question by telling you about a technology-related project I recently participated in. I work as a project analyst at the systems and e-research service, Bodleian Libraries, University of Oxford. In my job I need to understand the basics of technologies – such as the semantic web, institutional repositories, software development – and the people who will use those technologies. I talk to users, ask them about what they do and what they need or think they need. I tell them what developers are doing and make conclusions about how we can help them. I pass that information to developers. In this last project I had to do this intensively as we were targeting a big, diverse audience.

## THE PROJECT

‘Building the Research Information Infrastructure’ (BRII) at the University of Oxford was a JISC

(Joint Information Systems Committee)-funded project and a collaboration between the Bodleian Libraries and the medical sciences division of the University.<sup>1</sup> The project started in September 2008 and finished in March 2010.<sup>2</sup> The aim of the project was the efficient sharing of research activity data (RAD) captured from a wide range of publicly available sources. Research activity data describe researchers and their activities, for example their research interests, projects and funders. RAD are needed in many contexts, including academic, administrative and strategic. For example a departmental administrator needs lists of experts in particular fields to write research and publicity brochures. Within the University of Oxford, RAD are available in online sources such as departmental, project and researchers’ websites, as well as departmental databases and spreadsheets. RAD about Oxford are also found in external sites such as funders’ websites, online journals and databases.

The BRII project developed a pilot of an entity registry to harvest, process, store and re-use RAD. Entities were extracted from the harvested data and divided into their constituent parts, which were then deposited into an entity store. Entities are basic types of data, for example ‘person’, ‘funder’ or ‘project’. Methods for co-referencing were incorporated to identify individuals accurately, for example to decide whether two names collected from different sources belong to the same person. Semantic web technologies (ontologies and vocabularies) were employed so that entities and the relationships between them could be described in ways that both humans and machines can understand.

The entity registry can be described as mirroring sources of data about research taking place at the University of Oxford. (See Figure 1.) This mirror has added value to the original data. It has aggregated all these data together and classified them. It has created connections between data, connections which are new, useful information to our users, for example that researcher A works in department X and participates in project M in his department and in project P in department Y. Therefore the registry has the potential to become a comprehensive and cohesive picture of research at the university, a picture that our users can be part of (if they wish to contribute their data), explore and re-use in multiple ways.<sup>3</sup>

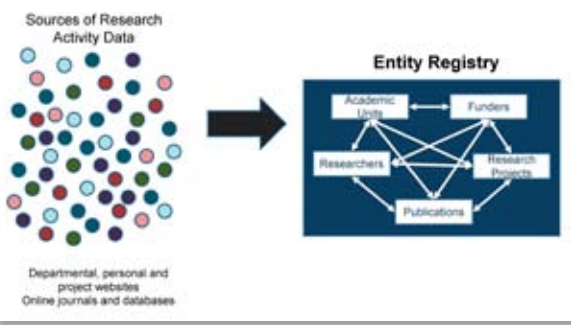


Figure 1. The BRII entity registry

The technologies used to develop the entity registry are advanced and one needs specific knowledge to understand and identify ways to use them. We cannot expect everyone to be a technology expert. BRII stakeholders have usually asked questions such as: 'What is a registry? Is it a website? Is it a database? How do I get the data? Why do we need an entity registry if we have all these online sources freely available?' The answers to those questions are not straightforward and if they are not explained carefully they could lead to confusion.

**FROM ABSTRACT CONCEPTS TO CONCRETE EXAMPLES**

Because of its nature the entity registry has become an abstract concept to most of BRII's stakeholders, a concept that we have to explain by using a variety of strategies. The central objec-

tive of BRII was the development of the registry. However, we also developed two examples to show how anyone within the university could use the registry and its data. By 'anyone' I mean both users with some technical knowledge and users with no technical knowledge at all (except from browsing the internet). I am also referring to individual users, such as academics, or users as in departments or institutes. The two exemplars are different and target different kinds of users.

The first example is a themed website and it shows how data can be re-used to create new websites for different needs. Some technical knowledge is required to assemble the website and gather data from the registry. The sample themed website uses information about research opportunities in the medical sciences division to target potential graduate students. The idea for this website arose from a strategic priority expressed by divisional staff in charge of recruitment, who had identified the needs of potential students. Its objective is to help students find information about their preferred research area, experts and supervisors. The themed website reorganises and displays sets of aggregated data collected from several departments within the medical sciences division. Little work was needed to build this website since most of its data were taken from the registry. Of course intensive work was needed initially to develop application pro-



Figure 2. The Blue Pages homepage (work-in-progress version)



Figure 3. Understanding the registry via the Blue Pages

programming interfaces (APIs) to access the registry. This website is just one example among the huge variety of possibilities for re-using RAD. We are always proposing ideas to our stakeholders – departments and individuals who may not identify immediate benefits for them – for example creating websites for multi-disciplinary collaborations where participants have different affiliations. Websites have the potential to help research dissemination and to increase visibility within and outside the university. Data can be gathered from researchers' departments, via the registry, and put together in one new website.

The second example is the 'Oxford Blue Pages'<sup>4</sup> (See Figure 2.) It is a directory of research expertise that offers several ways to view and search through RAD in the registry. It works like a search engine and can be used by anyone who knows how to surf the web. Users can browse or search through information by entering names or research keywords. The Blue Pages display the registry's entities in profiles: people (i.e. Oxford researchers), research projects (or research activities), funders (or sponsors) and academic units (e.g. departments, institutes). Data in profiles are organised in tabs, such as 'collaborators' and 'publications'. There are also links to more data in the registry (e.g. a link in a researcher's profile can take you to one of his project's profiles) or outside the registry (e.g. a link to a departmental website). The former links represent the connections made between data in the registry. These connections may have not been evident before the data were

harvested but are uncovered in the Blue Pages. For example, a researcher's biography might have been harvested from his departmental website whereas the descriptions of his projects were collected from different, disconnected departmental or project websites.

During the development of the Blue Pages we ran more than 30 tests with a variety of staff members. We asked them to perform a series of tasks in the Blue Pages and to give us their opinion afterwards. The feedback was rich both in terms of software usability and in terms of user engagement.<sup>5</sup> It was surprising to see how users gained better understanding of

the registry by testing the Blue Pages. Before the tests, users thought that the Blue Pages were the actual registry. This does not mean that users now know what the semantic web is or how the registry has been developed. What I mean is that after testing the Blue Pages they recognised that there is a store which they can use to obtain data about research in the university. (See Figure 3.) These data can be accessed via the Blue Pages or via the APIs developed by BRIL. They understood that the store was mirroring data from other sources and they could identify them by looking at the source links available in the Blue Pages. Finally, they also understood that the registry had created connections between data and that these were valuable pieces of information which could be exploited in many ways.

## CONCLUSIONS

Developing exemplars together with the registry and running user tests are by no means easy tasks. They have required an important amount of time and effort. This experience has confirmed how users can understand some technologies better by using simpler tools that keep those complex technologies behind curtains and present them in ways designed to fit into the users' work. These strategies fulfilled two purposes: they allowed us to explain our work and allowed users to understand it, and they allowed users to tell us what they wanted from the registry.

Engaging with users is critical. We have to make the effort because although we are developing

something for them we need their feedback to do our job well. We have to know our stakeholders, understand what they do and what they need. Most importantly, we need to know how our technology will help them in their own terms, not in ours. In other words, we have to identify the benefits of our work as users perceive them, because at the end of the day they are the ones who will use (or not use) what we develop. Users have the final word, and our success depends on whether they find our work easy to use and useful.

## REFERENCES

- 1 More information about the BRII project can be found at the project's website, <http://brii.bodleian.ox.ac.uk/>, and at weblog <http://brii-oxford.blogspot.com/>.
- 2 Although the BRII project has officially finished, development and user engagement work are still ongoing.
- 3 This depends on the variety and amount of data that we are able to harvest.
- 4 At the time of writing, the registry, the themed website and the Blue Pages are in a pilot stage.
- 5 If you would like to know more about the Blue Pages user tests, have a look at C. Loureiro-Koechlin, 'Uncovering user perceptions of research activity data', *Ariadne*, 62 (January 2010).