Institutional Change through Exposing Data: the James Cook University Research Portfolio

David Beitey[1], Floris van der Leest[2]

[1]James Cook University, Townsville, Australia, david.beitey@jcu.edu.au
[2]James Cook University, Townsville, Australia, floris.vanderleest@jcu.edu.au

Even within a single department in an institution, research-related data exists everywhere: different formats, inconsistent databases, separate environments. In a typical situation, this data unlikely to be directly accessible, nor visible to, or perhaps even known to exist by the person or group that it pertains to. This leads to a number of problems: access, being a difficult, disjointed experience for those who seek the information; poor quality or incompleteness, as the information may be hidden and not maintained; and a potential lack of consistency across different uses of the data. Through exposing useful data from primary sources in a logical and consistent fashion, great improvements across a research institution are possible.

Prior to the creation of the Research Portfolio, the exposure of researchers at James Cook University (JCU) and their associated activities was a disjointed experience. Visitors and collaborators seeking information were faced with numerous different sources of information – many of these public-facing identities were required to be manually maintained by researchers or their administrative counterparts. In addition, external researchers seeking to collaborate, higher-degree research students, and the general public were also faced with needing to logically 'connect' information together, due to a lack of consistency and quality of information available in at one source. The result was diminished visibility of research activities, and little or no clear path for improvement across the various bespoke environments. Without a central point, attempting to better expose research activities, or improve researchers' practices, would be an impossible task.

The James Cook University (JCU) Research Portfolio (http://jcu.me) aims to provide a solution to these various issues, spurring institutional change in the process. The Research Portfolio brings together as much valuable information as possible about research activities together into a cohesive, visible, consistent fashion.

WHAT IT IS

The conceptualisation of the JCU Research Portfolio arose from the need for consistency, quality, and visibility of research data. The resulting solution provides an open-access, web-accessible research directory that aims to expose all known valuable research metadata in a systematic manner. Portfolio pages cohesively integrate the most pertinent and novel information about a researcher – including biography, interests, experience, publications, grant projects and funding, student supervision, worldwide collaborations, and dataset downloads. External systems are integrated directly into a researcher's page, such that information appears consistently, and is accessible without leaving the context of a researcher's profile. Information is sourced from existing databases, and allows individual researchers to feed improvements to their information back upstream. The resulting tool endeavours to improve communication between researchers and the world, enable collaboration opportunities, enhance media exposure, and promote open access.

CHANGE FROM DATA

Whilst one of the primary aims of developing the Research Portfolio was to enhance research visibility through contemporary web-based academic profiles, a natural development has been the improvement to all aspects of data availability and quality. The nature of a one-size-fits-all approach effectively encourages consistency and highlights the benefits of clarity in user-entered data as other profile pages can be used as exemplars on how to best promote one's research. For example, once logged in to one's own profile, a researcher is shown all potential aspects of data that may be shown against their portfolio page, even if no data is available. Helpful hints and instructions are provided explaining where each piece of information is coming from and the relevant department to contact if the data is in need of modification. With all potential data types exposed, users are able to identify what should be shown, and take steps to improve their public display.

On an institutional level, processes and policies are being refined and streamlined. Information that was once only collated on request, is now always visible, and automatically kept updated. Listings of various types of researchers and groupings that needed to be maintained manually previously are now automatically managed, and guaranteed to be canonical, being generated using the most recent information from data stores. Exposing data to the public is useful in identifying procedures that should be present and were not at

the time—for example, requirements on data storage, metadata records, and quality and consistency of descriptions. In addition, by communicating openly with original data owners with regards to this user data and interaction, the Portfolio has been designed to limit itself become 'yet another system' with a custom set of data. In doing so, it has spurred the creation of new databases and storage as a method of departments improving existing capabilities. These new stores of data are 'owned' beyond the Research Portfolio, and thus can be utilised and leveraged accordingly in day-to-day and reporting capacities. So, because of this, a significant portion of manual work across positions and departments is no longer necessary.

Since the inception for the Portfolio, more researchers are now fully knowledgeable of their data. Interestingly, in a variety of cases, individuals have suddenly discovered connections and collaborations they were not aware they had, either in different countries or at different institutions. In a similar vein, researchers and administrative staff are now more familiar with the lifecycle of the data, including a better knowledge of data ownership and awareness of data generated about activities taking place. Because data now has a clear, obvious path from paper or electronic record to display to the public on the Research Portfolio, individuals now take greater care in specifying suitability descriptive information, and ensuring source data is complete and correct. As an example, research cover sheets, which were previously seen as a formality by some, are now completed with care to ensure correct association with those involved with the grant or project. In addition, many researchers have a new-found respect for the nature of this data, and thus ensure information is sufficiently detailed, as it will eventually be displayed publicly on both their and their colleagues' profiles online. Each improvement stands as a significant benefit to all involved in each research activity; the more information that is known, the more useful the underlying data inherently will be as well.

FUTURE

The development of the Research Portfolio and the encouragement of researchers to contribute more information about themselves paves the way for future analysis of data. In addition to being able to better associate individuals within the University through capturing information not previously stored—such as past experience, honours and awards bestowed, and consistently-formatted biographies—a key factor is that researchers can connect themselves with their external identities. Due to the complexity of naming and identity management, it is difficult to connect a researcher's identity across different systems. However, by allowing researchers to validate this themselves across their various identities, such as ORCID, ResearcherID, Scopus, and more, more connections are possible to the external datasets of publications and citations within these systems. This allows for future integration of additional data directly into Research Portfolio pages, and other systems within the University, and allows for future innovation in terms of tools that build upon that data. Overall, this integration amounts to a significant saving of time and human effort in terms of manual data maintenance as data is now stored consistently, and will further improve data management processes.

In addition to improvements in underlying data, the exposure of collated information within the Research Portfolio has increased awareness of the benefits of open information and the need for research collaboration. Researchers are now more familiar with services available to them, especially those that may be new to the research sector, and how integrating their research activities with eResearch platforms assists them. In particular, researchers are showing greater interest in services for data and metadata storage, including the Australian National Data Service (ANDS) and Tropical Data Hub initiative, and investigating the potential for the other data they have. As the Portfolio continues to be refined, it is clear that many researchers are aware of the benefits of opening their activities to the wider world and to adjust their practices accordingly.

CONCLUSION

The creation of a single system that exposes numerous types of data has instigated significant change across the University in a very short time since its inception. At its most basic level, the JCU Research Portfolio has rethought how research individuals and groups are displayed to the world, ensuring each is now more visible and accessible. By providing a consistent design, and an automated approach to how system data is constructed, there is a clearer path to research activities and a far greater understanding – and appreciation – of institutional data sources. Several key institutional processes and policies have been improved through the collation of the data displayed, leading to a reduction in workloads across numerous positions in the University. As data is now automatically visible to the world, the quality of data held by corporate and research data stores continues to improve as information is easily reviewed and adjusted by researchers. On the whole, exposing useful research-centric data in a consistent, easy-to-understand format is beneficial to all stakeholders, and paves the way for future innovation in data association.

The JCU Research Portfolio is available at http://jcu.me and provides a public view to all of the research activities undertaken at the University.

ABOUT THE AUTHORS

Mr David Beitey is the Online Services Manager within the eResearch Centre at James Cook University. In addition to acting as the Lead Developer for the JCU Research Portfolio (http://jcu.me), he manages the various online services being provided to the University and wider communities. His primary interests are web service integration, online security and client-side testing, efficient Content Management System (CMS) provisioning, and all aspects of systems optimisation. As a passionate advocate of open-source development, collaboration, and programming, David is a contributor to numerous public-domain projects, and strives to improve open-source projects whenever possible. He is also the founding organiser of the Python North Queensland Users Group (http://pynq.org).

Mr Floris van der Leest is responsible for the strategic development of James Cook University's Research Information Management System (RIMS), supporting the functions of research management at the University, as well as public accountability. He manages the maintenance of this system, and plans its ongoing development, integration with peripheral systems, and external interfaces with the web. Floris is responsible for research performance analysis, benchmarking and reporting, as well as strategic analysis of accountability requirements and information needs – for example, for national research quality assessments.