International Standard Name Identifier (ISNI) identifiers for one university's researchers: what, why, how

Roderick Sadler
La Trobe University, Bundoora, VIC Australia. R.Sadler@latrobe.edu.au

Simon Huggard
La Trobe University, Bundoora, VIC Australia. S.Huggard@latrobe.edu.au

THE NEED FOR NAME IDENTIFIERS

Research institutions face a task that is often very time-consuming and inefficient in compiling information about the scholarly outputs of their researchers. They also face the problem of ensuring that those scholarly outputs are correctly attributed and that the organisational affiliations of the creators are recognised externally. It is no longer sufficient for institutions to have complete and correct data to use for purposes such as performance appraisals or institutional reporting requirements if the data available to external agencies, such as those performing institutional ranking tasks, are inaccurate or incomplete. It is not enough for institutions simply to be able to measure their own scholarly outputs. Other agencies are also engaged in assessing the value and impact of those outputs, but without the benefit of having a direct relationship with the researchers involved to assist in gathering accurate and complete information; the potential for a significant disparity between what is measured internally and what may be measured by external agencies is clearly apparent. The obvious element missing is an unambiguous persistent identifier for researchers that could be employed from within the institution while also being recognised externally, particularly by organisations engaged in the work of attributing authorship of research publications to researchers. Concomitantly, use of such an identifier offers considerable time-saving and efficiency gains internally for institutions in their data-gathering work. To begin working to this end, in late 2013 La Trobe University agreed with the International Standard Name Identifier (ISNI) Implementation Agency (OCLC) to act as a test site for a service being developed to accept a file of data about researchers from an institution with the intention of adding high quality data to the ISNI database and in turn to receive back ISNI identifiers where data has been incorporated into an assigned ISNI record.

LIBRARIES AND NAME IDENTIFIERS

Libraries and library agencies have worked for many years associating the works of particular creators with distinctive and consistent forms of name for those individuals to try to ensure that the right works, and only the right works, appear in connection with those forms of name. This authority work (meaning work relating to authors) has led to the creation of many database records representing creators and sketching out a kind of publishing "identity" for them. What many researchers may not realise is how large the infrastructure behind this work has become. They may be identified in many national and international library systems and in systems in other parts of the international bibliographic industry, in records which attempt to establish their publishing identity on the basis of works which they have published. As a result, many researchers will already have multiple identifiers associated with them.

THE ROLE OF ISNI – EXTENDING AND CONNECTING

While previously this work has largely been confined to assisting the operations of libraries in their direct provision of services, this name identification work is now being extended through the ISNI infrastructure to support other goals and functions, as the ISNI system also brings in data from copyright agencies and other organisations. This further extends the scope and potential effectiveness of the ISNI system throughout the various sectors relating to intellectual and artistic creation. It is a name identification system based on the processing of industry or agency data rather than directly incorporating data provided by individuals to describe or represent themselves.

HOW ISNI WORKS

The International Standard Name Identifier (ISNI) system (based on ISO standard 27729) has been implemented in a way that builds on the vast amount of data gathered and created internationally by libraries and other agencies about author identities and their publications, and transforms that data into a comprehensive system of name identification. An ISNI identifier is created after disparate sources of data about authors and publications have been processed and matched with consideration to the content of the data (for example, consistency between forms of name used by authors in publishing, uniqueness of a name, consistency between works associated with authors, other authors associated with authors and organisations associated with authors), and found to comprise a logically consistent and sufficiently distinct identity. (Many, many names in the data submitted to ISNI from various sources are not given an ISNI identifier because they do not meet these requirements.)

The ISNI system has been setup to work meaningfully with as much existing reliable data as possible in allocating name identifiers, and it is a system based on processing data rather than on allowing individuals to simply ask the system to give them a name identifier. For us, this has several implications. Many of our researchers will already have an ISNI identifier, whether they know it or not. Where our researchers do not have an ISNI identifier it would not be an easy task to get one assigned.

LA TROBE UNIVERSITY'S WORK WITH ISNI

We have provided ISNI with a data file of names of La Trobe University researchers matched with publications as reported in the university's research management system, and this has progressed from their "acceptance" environment to the production environment for processing. We should receive the results of this soon. We have incorporated the same local identifiers for researchers in this file as we have used in populating the "Mint" system in our ReDBox/Mint research data management configuration. This has been done to maintain consistency in researcher identification across all of our relevant systems and in all the systems to which we contribute such data (including the Trove Party Infrastructure). We did not use Mint system-generated identifiers for our researchers but instead loaded identifiers based on those used in the university's research management system, for consistency.

As an ISNI data contributor we can work manually in the ISNI system to carry out various tasks on our batch-contributed records and to contribute new records. Where records have not matched a record with an assigned ISNI and have not met the requirements to be assigned a new identifier automatically, we will need to work manually on disambiguating or enriching those records until they meet the requirements for ISNI identifier assignment. Preliminary indications are that about two-thirds of our contributed records will require manual work before they are assigned an ISNI or are matched to an existing ISNI. The matching rate achieved has been considerably higher than expected. Maintaining currency in the ISNI system will require the addition of about 150 new records per year, as new research staff join the university. (For many who are already published authors an assigned ISNI will already exist.) Library staff will have responsibility for the ongoing work in the ISNI system, as they have had for the initial batch load, and they will need to coordinate with data in the university's research system.

The approach taken in this project was to contribute to ISNI only data that is already public, and thus to exclude dates of birth and other private data. What we were relying on in the process of building up researcher identities for ISNI processing was the authoritative association of researchers with as extensive sets of publications as we were able to assemble. The organisation employing the researchers involved is well placed to do this, especially with access to the annual publication data reported directed by researchers to the university.

Once we have retrieved a significant set of ISNI identifiers for our researchers we intend to promote their use to our researchers for inclusion alongside their names when publishing. We expect that eventually we will be able to use ISNI identifiers to assist in automated harvesting of research output data from indexing services and elsewhere.

References

1. Data contributors, available at http://www.isni.org/content/data-contributors
2. How ISNI works, available at http://www.isni.org/how-isni-works

About the Authors

Roderick Sadler has been employed at La Trobe University since 2004 and worked at several libraries previously. He has worked primarily in roles focused on metadata creation and management and in the maintenance and configuration of systems working with bibliographic metadata. He leads the team responsible for the university's institutional repository, La Trobe University Research Online.

Simon Huggard joined La Trobe University in June 2012 and has previously worked at the State Library of Victoria and Monash University. He manages a team of staff responsible for the Library's Research Online Repository, search, linking and discovery systems, library management system, image repository, and copyright and audio-visual requests.

Simon has written many articles and presentations relating to Library systems and information access and is a strong advocate for open access and the provision of smart systems for accessing digital library content.