

## ARDC Party Infrastructure

### Who should read this?

---

This document is intended for potential contributors to Research Data Australia.

### Why do we need information infrastructure to identify Researchers and Research Groups?

---

Keeping track of the identity of individual researchers presents many challenges. A researcher may use different variations of their name; several researchers may have the same name; confusion arises when a researcher moves from one institution to another; an institution may even have the same researcher listed more than once in their own records. Similar problems emerge with groups of researchers. To allow the discovery of research datasets that share a common researcher or research group, a common public identifier is needed for referencing.

ANDS has articulated a Data Connections Strategy to enable linkage across the Australian Research Data Commons by promoting “common” approaches to specifying ‘concepts’ and ‘entities’ which are related to datasets. This includes the researchers, groups and organisations responsible for the data.

### Which public identifier?

---

This same problem has existed in the scholarly information system for a long time. The National Library of Australia (NLA) maintains an Australian Name Authority file with a unique reference for Australians as authors and subjects of published works. This is now implemented as an online service, Trove – People and Organisations<sup>1</sup>, which supports machine access as well as public search. The public identifier used to reference these parties is called the NLA Party Identifier. ANDS decided to work with the NLA to develop this identifier system because

- most Australian researchers are publishers of scholarly information and many already have an NLA Party Identifier,
- the NLA has an established service for harvesting party information from various contributors and an identity matching facility for grouping descriptions of the same party by different contributors, and
- the NLA is committed to maintaining this identifier into the future.

Some scholarly publishers have addressed the same problem by implementing researcher identity systems of their own, such as the Scopus Author Identifier<sup>2</sup> and Thomson Reuter’s ResearcherID<sup>3</sup>. These systems have incomplete coverage and are not open for machine access. They are important identifiers though, and the NLA Party Identifier will integrate with global researcher identity systems such as these and the new ORCID service<sup>4</sup> when established.

### How will it work?

---

ANDS has funded the NLA to extend its existing infrastructure to harvest party information (people and groups) from Australian universities and other research Institutions and to improve both their automatic and manual identity matching services. The first phase of this project involved community consultation and the development

---

<sup>1</sup> <http://trove.nla.gov.au/general/aboutPeople>

<sup>2</sup> <http://www.info.sciverse.com/scopus/scopus-in-detail/tools/authoridentifier>

<sup>3</sup> <http://www.researcherid.com/>

<sup>4</sup> <http://www.orcid.org/>



of a detailed description of how this infrastructure will operate. The implementation phase is due for completion in July 2011.

The party information provided by each institution must include a local unique identifier for each researcher or group. Institutional systems can query Trove with this identifier to discover the assigned NLA Party Identifier. All other party information provided by the institution is considered to be part of the public profile for this researcher, and will be displayed in both Trove and Research Data Australia. This information would normally be available as well through the institutional website. All of the descriptions displayed in Trove for the same person or group have the contributor identified, and any reported corrections are relayed to the contributor.

## What is required of a contributor?

---

ANDS encourages all contributors to Research Data Australia to include descriptions of the parties responsible for the datasets they are describing. Once your researchers and research groups have records in Research Data Australia you can include them more easily in your descriptions of data collections. Initially these party records will not have been assigned an NLA Party Identifier.

### Direct contribution to Trove

Contributors may contribute their party information directly to the NLA through their own OAI repository. This is the preferred method. Records may either be in EAC-CPF format or RIF-CS format. Richer information can be provided in EAC-CPF<sup>5</sup> which helps improve automatic identity matching rates and improves the information displayed in Trove. Local systems can retrieve the assigned NLA party identifier for each researcher or group and use this as the key when connecting researchers or groups to other records in their feed to ANDS.

### Contributing to Trove through ANDS

Alternatively, contributors who have supplied party records in their ANDS feeds can indicate they wish these records to be uploaded to the NLA Identity Matching service by checking the '*Party records to NLA?*' checkbox in their Data Source Account 'Edit' form. An ISIL identifier will also need to be provided (see next section). All party records associated with the Data Source will then become available for harvesting by the NLA; there is no capability for harvesting selected records only.

Trove will then harvest this party information and, using automatic matching rules, will match as many records as possible against existing Trove Identities, or mint new identities with the supplied information. Unmatched identities will be available to the contributing institution via the Trove Identities Manager (TIM) web utility that allows them to easily compare their researcher with possible identities in Trove and manually match or create new identities. Only identities that have been matched, either automatically or manually using the Trove Identities Manager, will be displayed in Trove.

Contributors will then be expected to include the NLA Party Identifier as the related party in Collection, Service and Activity records. If they wish to continue to provide their own party records in Research Data Australia and use their own party keys in related objects, then the NLA Party Identifier should be included in an identifier element in their own party records so that the records about the same identity can be connected.

It is important that names supplied in party records are broken into parts, i.e. family name, given name(s), title etc. This will ensure better automatic matching rates. Party records can include other identifiers (e.g. Scopus AuthorID), research areas, publications, and biographical information. All these items assist contributors to match identities as quickly as possible, and some can be used to improve automatic matching rates.

Please contact your ANDS Liaison Officer or [services@ands.org.au](mailto:services@ands.org.au) if you wish to commence obtaining persistent public identifiers via the ARDC Party Infrastructure.

## What is an ISIL?

---

An ISIL is an International Standard Identifier for Libraries and Related Organisations (see <http://bibstandard.dk/isil/>). It is a unique number that can be assigned to every library and related organisation in

---

<sup>5</sup> <http://eac.staatsbibliothek-berlin.de/>



the world. An ISIL is alphanumeric and takes the form of a country code prefix, then a hyphen, and then an identifier issued by that country's national library authority.

The National Library of Australia has the authority to assign ISILs to libraries and related organisations in Australia. These take the format of country code 'AU' for 'Australia' then a hyphen (-) and then an identifier. The identifier issued by the National Library is a National Union Catalogue (NUC) symbol. NUC symbols are widely used by the Australian library and research community for identification and communication within the national and regional networks. For example, the ISIL for Monash University is AU-VMOU where AU is the country code and VMOU is the NUC symbol.

All of the party information you contribute to Trove will have your organisation (as identified by the ISIL) nominated as the contributor in case an alteration is requested.

To check or obtain an ISIL for your library or organisation:

1. Go to the Australian Libraries Gateway website <http://www.nla.gov.au/libraries/> and use the search options to check if your organisation has a NUC symbol.
2. If your organisation has a NUC symbol, simply add the prefix AU then a hyphen (-) to formulate your ISIL. For example, the NUC symbol for Griffith University is QGU. The ISIL for Griffith University is 'AU-QGU'.
3. If you can't find an entry for your organisation in Australian Libraries Gateway, or you are unsure of your NUC symbol, please contact the Trove team at the National Library with your query. Use the 'contact us' link from the Trove homepage at <http://trove.nla.gov.au/>

## Using the Trove Identities Manager

---

When your party records have been harvested by the NLA Identity Matching service (either through direct harvest from you or through the ANDS Registry) you are able to start manually matching any of your identities who have not been automatically matched. If you do not already have a Trove sign-in (available to anyone) please register at <http://trove.nla.gov.au/signup>.

Your ANDS Liaison Officer will arrange for your permission to match your own party records (those with your ISIL attached) using your Trove logon. They will also arrange for training in using the service and in Name Authority Matching best practice if you are not already familiar with this.

## What are the benefits?

---

Use of the NLA Party Identifier will result in improved discovery of Australian research data and research publications by enabling the linking of all research outputs of an individual researcher or group over time. By assigning a common identifier to researchers, institutional systems (such as the institutional repository or research management system) will be able to exchange information with each other and with external systems. This will enhance each institution's reputation by giving better visibility to researchers and their outputs as the records from Research Data Australia and Trove are syndicated by other larger scholarly information services (e.g. Google Scholar).

## Further Information

---

ANDS guides and other Resources: [www.ands.org.au/guides](http://www.ands.org.au/guides)

ANDS Data Connections Strategy <http://ands.org.au/dataconnections.pdf>

ARDC Party Infrastructure Project Wiki <http://wiki.nla.gov.au/display/ARDCPIP/>

Detailed description of the ARDC Party Infrastructure

[http://wiki.nla.gov.au/download/attachments/16515340/ARDCPartyInfrastructureConsultationDraftv1\\_0.pdf](http://wiki.nla.gov.au/download/attachments/16515340/ARDCPartyInfrastructureConsultationDraftv1_0.pdf)

Martin Fenner, Author Identifier Overview <http://blogs.plos.org/mfenner/author-identifier-overview/>

