

How to get theses displaying correctly in Trove: A guide for the research community

In late 2010, the Australasian Digital Theses program will be retired, and those wishing to view Australian-produced theses will be directed to Trove. The following instructions are designed to explain the steps that institutional repositories need to take to ensure that their records are correctly included within such a search.

This guide assumes familiarity with the Australian Research Online contributor's guide. A copy can be downloaded at <http://research.nla.gov.au/main/about>.

Steps to get your repository ready

1. Contribute records to Australian Research Online (or another service feeding in to Trove)

Firstly, your theses records must be contributed to Trove. **All records contributed to Australian Research Online are also contributed to Trove**, and for the purposes of the rest of these instructions, it will be assumed that this is the method by which your records will be contributed to Trove. Please note, however, that records contributed to the Australian National Bibliographic Database will also be included within this search, providing the appropriate metadata standards are met. *If you wish to contribute records this way, please contact Alison Dellit at the NLA directly – the rest of these instructions will be different!*

2. Establish a single OAI collection point for all your Australian Research Online records

It is important that all the records you contribute to Australian Research Online are harvested by the NLA through a single OAI process. This could be by harvesting one set, or by harvesting everything in your repository. This is a change from how harvesting worked with the Australasian Digital Theses program, where a specific set for theses was harvested. ARO requires that your theses records are made available for harvesting in the **same** OAI set or collection as the rest of the records that you contribute to Australian Research Online.

3. Let us know an appropriate National Union Catalogue (NUC) code for your collection

Currently, when an institution contributes data to Trove via Australian Research Online, the name of that organisation does not clearly display in Trove. We will fix this before the Australasian Digital Theses is switched off. However, in order to make the name display properly, and also to allow, into the future, searches across an institution's records, we will need to identify each contributor using a NUC symbol. All Australian universities have a NUC symbol already, and most large organisations do as well. You can tell if your organisation has a NUC symbol by visiting <http://www.nla.gov.au/apps/ilrs> and doing a search.

We will automatically use the NUC symbol we have recorded for your institution unless you let us know you want to use something else. (For example, if you want your records to appear as "Smith University Research Repository" rather than "Smith University", you will need to use a separate NUC symbol. You may already have one, if not, we can set one up for you, if you let Alison know. (NOTE: If you do not use the NUC for your institution, then in the future, a user searching for records from "Smith University" will not see your repository records in the results).

4. Provide appropriate metadata to identify theses

Trove “decides” whether an item is a thesis or not based on what it finds in the Dublin Core type (dc.type) element from your harvest. Ideally, that element would simply contain “thesis”, as well as an additional dc.type element giving more information about the type of thesis. (You can include multiple instances of the dc.type field in your OAI export). Many repositories are already doing this.

We recognise, however, that many institutions cannot provide more than one dc.type element in an OAI record. In that case, as long as the word “thesis” is in the element, we can provide a local mapping to add the “thesis” type to all such records. For example, if you designate your records “thesis(masters)” and “thesis(phd)” we will ensure that all such types are included in the Trove theses search. If you wish to use a term such as “dissertation” or “doctorate” as the signifier, this can be negotiated by contacting Alison directly.

We do not insist on getting more granular information either, if all your theses are simply designated as “thesis” they will appear in the search. HOWEVER, it will mean we cannot provide a research theses facet (see note below).

5. Let Alison know that you are ready!

Even if you don’t contact the NLA, your records will appear if you have met all the above steps. However, it is helpful for us to know which universities are ready and which are not in order to plan timing. It also means we can cross-check what you’ve done, and make sure there are no hidden glitches.

An important note on research theses:

Trove accepts all kinds of theses, including honours theses and masters by coursework theses. A “thesis” search will include all such material. A core premise of Trove is that the more scholarly work that is findable, the better for our end-users.

The Trove team is open to the idea of allowing a specific **separate** facet-based search across “research” theses, but to do so, we will need **all** research contributors to be able to clearly identify such theses using the dc.type field. For example, by using terms such as thesis(phd) or thesis(research master) in the dc.type field or, if easier, a single type such as thesis (phd or research masters). As long as there is a clear way of separating the research theses from other material based on the wording in the dc.type field, we can do a mapping at the National Library to provide a single search point.

To emphasise, we will not provide such a search if it only works on some material. If we can’t get all the institutions to clearly identify the research theses, we will not provide a facet, because it will create a false impression. If you believe you can distinguish your data in this way, please send email Alison with the relevant types involved. These will be collated, and if we can make this comprehensive, we will progress from there.

Of course, any data you enter in the type field will be findable via a keyword search. So a keyword search for “phd thesis” will return all theses which have that entry in the type field, whether or not we provide a specific research thesis facet.

How to see what data you currently send us

In the Australian production service, you see the data after several transformations and we also use the existing ADT service to identify which items are theses. In many cases, this means that when the ADT service ceases, items previously identified as theses, will no longer be identifiable as such. **This means you may think you are including thesis as a resource type in your OAI feed, but it is not actually coming through that way. There are 7 universities we currently have no theses data at all from, and a further 5 where the total seems suspiciously low 😊**

In preparation for this process, we have changed the Australian Research Online test system so that it now displays only the raw data we get from each institutions OAI response, not the data that we have applied our mappings and 'normalisation' rules to. If this data does not clearly identify your theses, then you need to implement Step 4 above.

You can view this system at http://arrow-test.nla.gov.au/ui/main/admin_statistics. Use the drop down boxes on the top right to choose your institution, and leave "type" selected as the field. The tag cloud display on the left will show you all the data in the type field. The larger the word/phrase, the more often it appears. To see how many records have that type, simply click the word/phrase, and you will see a search restricted to that term.

Alternatively, you can go straight to <http://arrow-test.nla.gov.au/ui/>, click on your institution's name, and view the type facet on the right hand side of the results page.

Give it a go, and if you are having trouble, give Alison a call and she'll talk you through it 😊

Contact:

Alison Dellit
National Library of Australia
adellit@nla.gov.au
(02) 6262 1219.