ANSTO, the Australian Synchrotron, metadata catalogues and the Australian National Data Service

Nick Hauser[1], Uli Felzmann[2], Saliya Wimalaratne[1]

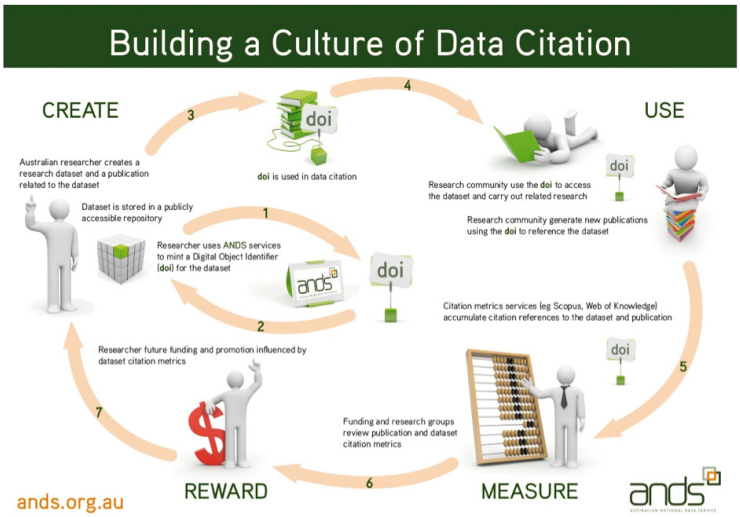
[1]ANSTO, Lucas Heights, NSW 2234. Australia
[2]Australian Synchrotron, Clayton, Vic 3168. Australia

Why share scientific data?

ANSTO and the Australian Synchrotron host experiments for domestic and international science community over a wide range of scientific disciplines. Data citation, management and discovery are important to the facilities and the scientists that use them. Gone are the days when raw data is written to a removable media and subsequently lost or locked away in a desk to collect dust.

The metadata catalogue Tardis is being used by both ANSTO and the Australian Synchrotron. Metadata is harvested from the neutron beam and X-ray instruments raw experimental files and catalogued in databases that are local to the facilities. The data is accessible via a web portal. Data policies are applied to embargo data prior to placing data in the public domain. Public domain data is published to the Australian Research Data Commons using the OAI-PMH standard. The Commons is run by the Australian National Data Service (ANDS), who was the project sponsor. The Commons is a web robot friendly site.

ANDS also sponsors digital object identifiers (DOI) for deposited datasets, which allows raw data to now be a first class research output, allowing scientists that collect data to gain recognition in the same way as those who publish journal articles. Raw data is increasingly required by journals to allow for more rigorous referring of publications. Data is being discovered, cited, reused and collaborations initiated through the Commons.



Management of Research Data

Under the responsibilities of institutions, the Code requires;

"Provide secure research data storage…"

"Identify ownership of research data…"

"Ensure security and confidentiality of research data…"

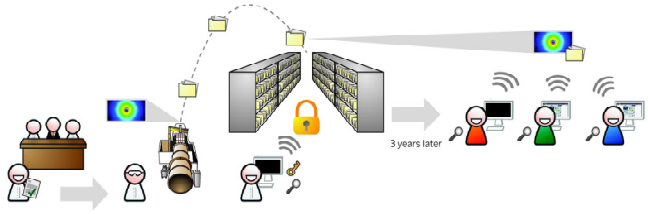
The Code requires that data must be in a;

" durable, indexed and retrievable form"

" catalogue … in an accessible form"

"A policy is required that cover the secure and safe disposal…"

.



Proposal Acquire data Embargoed data Public data with DOI

How to share data. mytardis (metadata catalogue)

mytardis harvests metadata from NeXus files and the proposal management system, to provide users with a summary of their experiments.

Scientists can download their data files easily via a web browser.

An embargo feature ensures that the data is accessed only by the scientists. The data is made public after the embargo, allowing scientists time to publish their results. Shown in the graphic above. Once public, the data is assigned a digital object identifier (DOI) which scientists can use to cite the data.

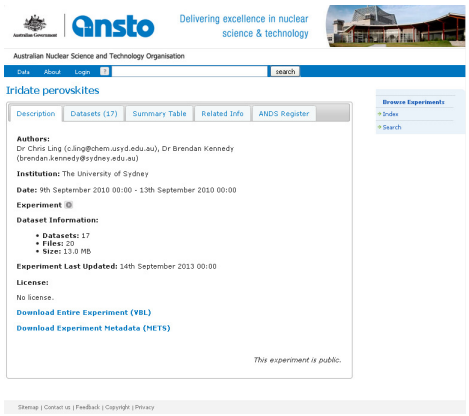
The webpage to the right shows the summary table view in mytardis, where all the datasets for an experiment can be viewed and sorted.

Web services allow applications to query the database and download data files.

mytardis is being used by ANSTO and the Australian Synchrotron.

Open source under the New BSD license.

http://code.google.com/p/mytardis/



Tardis web interface

Conclusion

Using the metadata catalogue Tardis, ANSTO and the Australian Synchrotron are providing services allowing the researchers that use these facilities to fulfil the Australian Code for the Responsible Conduct of Research in managing research data. Confidentiality and intellectual property rights are protected, whilst providing a mechanism to enhance scientific collaboration and data reuse. Raw data is accessible, searchable and citable.

Acknowledgement

ANDS is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy Program and the Education Investment Fund (EIF) Super Science Initiative