



# ANDS Guide



## Data citation

**Level:** Awareness

**Last updated:** 19 January 2017

**Web link:** [www.ands.org.au/guides/data-citation-awareness](http://www.ands.org.au/guides/data-citation-awareness)

### This Guide covers:

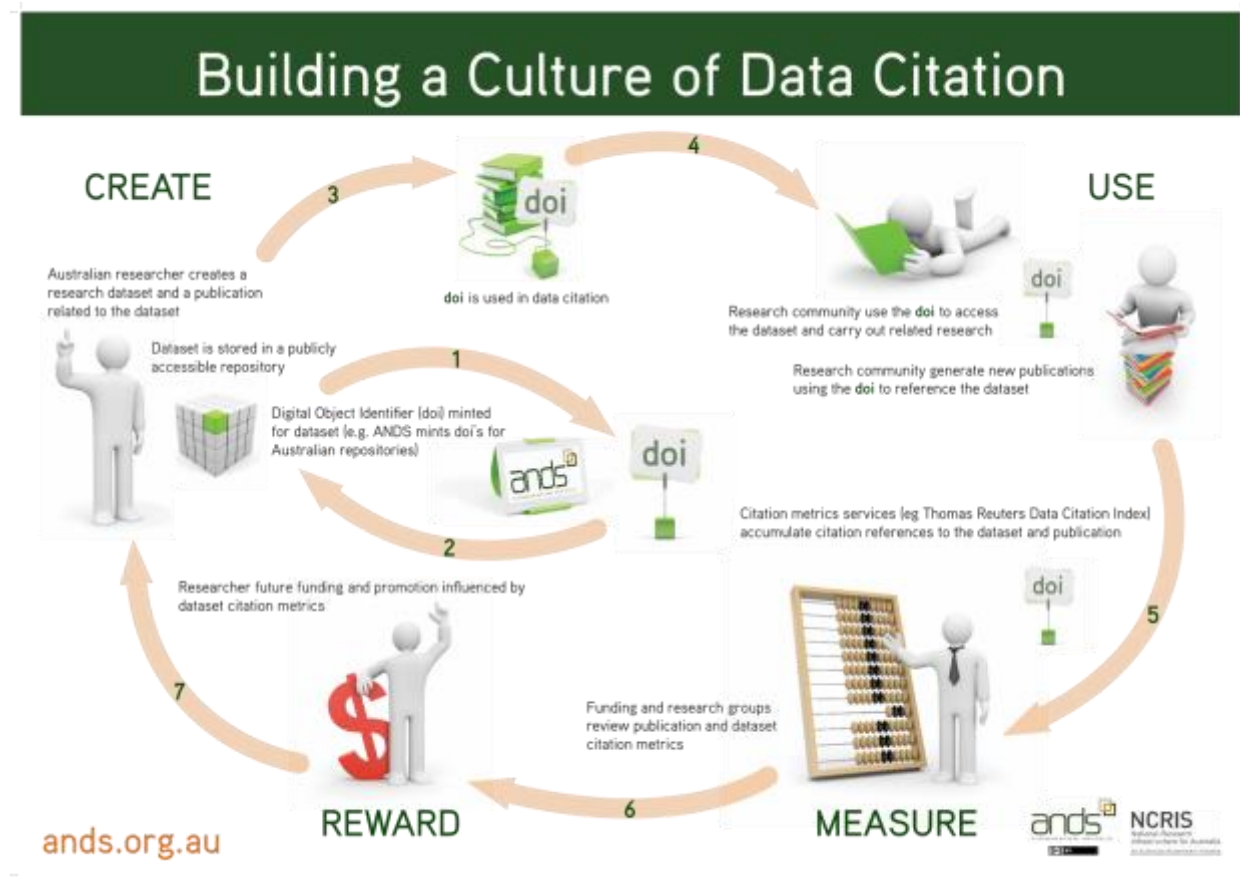
- What do we mean by data citation?
- Did you know...?
- How do you cite data?
  - Data citation styles and formats - with and without a DOI
  - Data repositories: styles and formats
  - Journals: data citation styles and formats
  - Data citation elements for repository managers
- Why include a DOI in data citations?
- How do you count data citations?
- Connecting data and related citations
- Implementing data citation at institutions
- ANDS services for data citation

### What do we mean by data citation?

Data citation refers to the practice of providing a reference to data in the same way as researchers routinely provide a bibliographic reference to other scholarly resources. While data has often been shared in the past, it was seldom cited in the same way as journal articles or other publications. This culture is, however, rapidly changing.

Driven by publisher and funder requirements for reproducible research and reusable research outputs, data has come to be regarded as a **primary** research output rather than a by-product of research. When datasets are cited,

they become discoverable, reusable and citable by others. Citation of data also enables recognition of scholarly effort in disciplines and organisations that want to acknowledge and reward data outputs.



View the poster [here](#)

### Did you know...?

- The creation of data is increasingly being recognised as a primary research output
- Many journal publishers now encourage or require citation of research data
- There is a global network of discipline and institutional data repositories where research data collections are described with a preformatted citation statement provided
- Only cited data can be counted and tracked (in a similar manner to journal articles) to measure impact
- A [Digital Object Identifier \(DOI\)](#) may be assigned to data in the same way as journal articles
- Data citation information may soon be incorporated into practices for research evaluation and reward
- Some bibliographic management systems (e.g. [EndNote](#)) now include a template for research data citations.

## How do you cite data?

Data citation styles continue to evolve and vary across disciplines and publishers: short [Overview](#) of important elements in citing data, regardless of citation style, publisher or repository guidelines.

## Data citation styles and formats

### DOIs and data citation styles and formats

DOIs are not essential, but are considered best practice for data citation. DOIs are a unique, persistent identifier that can be used to track data citation metrics and to link related outputs such as journal articles, research data and software.

- DataCite [DOI Citation Formatter](#) is a simple form based system which uses your dataset DOI to allow you to quickly format your citation in hundreds of different styles.

### Data citation without a DOI

- [Chicago-Style Citations](#): Social & Online Media, Data, Music, and Other Non-Print Materials: Datasets
- [Data Citation styles](#) for: APA, APSA Style Manual for Political Science, American Sociological Association Style Guide.

eg APA style

Pew Hispanic Center. (2008). *2007 Hispanic Healthcare Survey* [Data file and code book]. Retrieved from <http://www.pewhispanic.org/2007/09/23/2007-hispanic-healthcare-survey/>

## Data repositories: citation styles and formats

Some data repositories (e.g. ICPSR, PANGAEA) provide a recommended format for citing data from that repository eg ICPSR

Kessler, Ronald C. National Comorbidity Survey: Baseline (NCS-1), 1990-1992 (Restricted Version) [Computer file]. ICPSR25381-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2009-05-11. doi:10.3886/ICPSR2538

## Journals: data citation styles and formats

Many journal publishers also provide guidance on how to cite data. For example, the [SpringerNature data policy](#) states:

Where datasets are hosted in public repositories that provide datasets with Digital Object Identifiers (DOIs), we encourage these datasets to be formally cited in reference lists. Citations of datasets, when they appear in the reference list, should include the minimum information recommended by [DataCite](#) and follow journal style. For example:

Hao, Z., AghaKouchak, A., Nakhjiri, N., Farahmand, A. Global Integrated Drought Monitoring and Prediction System (GIDMaPS) Data sets. *figshare*. <http://dx.doi.org/10.6084/m9.figshare.853801> (2014)

## Data citation for repository managers

DataCite [DOI Citation Formatter](#) is a simple form based system which uses dataset DOIs to format data citations in hundreds of different citation styles. Do you want to integrate this service? Check the [Documentation](#).

DataCite (the agency who manages the global system for DOIs for datasets) recommends using one of the following generic formats:

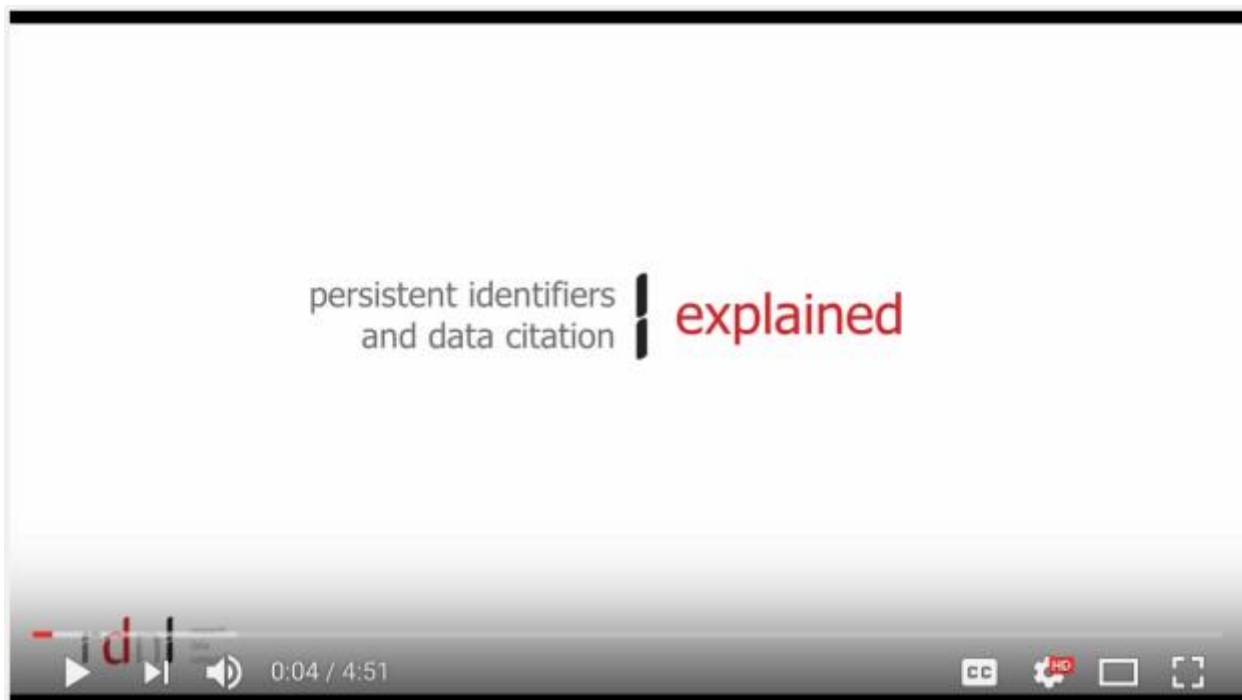
- **Mandatory citation elements only**  
Hanigan, Ivan (2012): Monthly drought data for Australia 1890-2008 using the Hutchinson Drought Index. The Australian National University Australian Data Archive.  
<http://doi.org/10.4225/13/50BBFD7E6727A>
- **Including version and resource type**  
Bradford, Matt; Murphy, Helen; Ford, Andrew; Hogan, Dominic; Metcalfe, Dan (2014): CSIRO Permanent Rainforest Plots of North Queensland. v2. CSIRO. Data Collection.  
<http://doi.org/10.4225/08/53C4CC1D94DA0>

## Why include a DOI in data citations?

While data may be cited without a DOI, assigning a [Digital Object Identifier](#) (DOI) to data facilitates data citation and is considered best practice. A DOI is a type of [persistent identifier](#) that is optimised for citation and indicates a dataset will be well managed and accessible for the long term. The [ANDS DOI Service](#) can be used by Australian publicly funded research organisations and government agencies to mint DOIs for datasets and collections that can be included in data citations.

DOIs for data enable:

- Easy and persistent access to research data available via the internet through resolving the DOI link eg <http://doi.org/10.4225/08/51145B02F20A8>
- Enhanced discovery, retrieval and management of data to enable data reuse and verification of research results
- Persistent links between related outputs: DOIs can be applied to journal articles, reports, grey literature, data and software
- Support for measuring the impact of data through:
  - [citation indexing services](#) such as [Thomson Reuters Data Citation Index](#)
  - altmetrics such as number of views, downloads and 'likes'.



Watch the video [here](#)

## How do you count data citations?

Only cited data can be counted and tracked (in a similar manner to journal articles) to measure impact. Use of persistent identifiers within the citation are a key factor in ensuring all data citations are accurately tracked and counted.

Citation indices measure the reuse of research data. Such metrics are commonly used for performance appraisal and reporting. It is also possible to measure the impact of research data through 'altmetrics' such as downloads, 'likes' and social media mentions.

Initiatives to support data citation and tracking include:

- [Thomson Reuters Data Citation Index](#) - analogous to the use of products such as Web of Science and Scopus to measure citations to journal articles and other types of scholarly publications.
- [Europe PubMed Central](#) web service
- [Scholix](#) framework developed through the Research Data Alliance

## Connecting data and related citations

The connection between research data and publications is increasingly recognised.

[How to Cite Datasets and Link to Publications](#): This practical guide from the Digital Curation Centre (UK) will help you create links between your academic publications and the underlying datasets, so that anyone viewing the publication will be able to locate the dataset and vice versa. It provides a working knowledge of the issues and challenges involved, and of how current approaches seek to address them. This guide should interest researchers and principal investigators working on data-led research, as well as the data repositories with which they work.

The [Monthly drought data for Australia 1980-2008](#) record in Research Data Australia shows:

- how the dataset should be cited: the 'Cite' button
- Citation information for related research outputs: publications, services, grants, original and derived datasets, and software

## Implementing data citation at institutions

Good planning will smooth the path for implementing data citation in your organisation. Consider key questions early, learn from what others have done and talk with ANDS staff about the services and support we offer.

To guide your planning for data citation, consider these key issues:

1. How have others (institution and repositories) approached building a culture of data citation? [Case studies](#)
2. Resource planning: have you identified the [people and skills](#), infrastructure, support and resourcing you will need to successfully implement data citation?
3. Business rules & procedures: have you determined guidelines or business rules for [identifying citable data \(PDF, 0.2 MB\)](#)?
4. Metadata framework: are you able to provide at least the minimum [metadata requirements to form a data citation](#)?
5. Data and metadata storage: are you able to guarantee that the data and metadata will be persistently [stored and managed](#) for the appropriate time for a scholarly record?

## ANDS and data citation

ANDS is engaged in activities that will make it easier to share data, to recognise the importance of making data available and to make data citation a standard practice. To this end:

- ANDS is engaging with research funding agencies to promote data publication as a primary research output and the inclusion of data in the research assessment process

- ANDS is working with Thomson Reuters [Data Citation Index](#) to track and record data citations as part of research assessment activities
- Through the [Research Data Alliance](#), ANDS is contributing to international initiatives aimed at improving data citation and tracking
- ANDS offers a [DOI Service](#) that can be used to assign DOIs to datasets and collections.

More information about ANDS services: [contact@ands.org.au](mailto:contact@ands.org.au)

## Feedback?

We welcome your feedback on this guide. Please email [contact@ands.org.au](mailto:contact@ands.org.au) with any comments or questions.

## About ANDS

**The Australian National Data Service (ANDS) makes Australia’s research data assets more valuable for researchers, research institutions and the nation.**

ANDS is a partnership led by Monash University in collaboration with the Australian National University (ANU) and the Commonwealth Scientific and Industrial Research Organisation (CSIRO). It is funded by the Australian Government through the National Collaborative Research Infrastructure Strategy (NCRIS).

This work is licensed under a [Creative Commons Attribution 4.0 International License](#). You are free to reuse and republish this work, or any part of it, with attribution to the Australian National Data Service (ANDS).