DATA RETENTION AND STORAGE

POST PUBLICATION

Advice to Research Staff and Students

The following advice provides options for you to meet your obligations for data retention and storage under the ***Australia Code for the Responsible Conduct of Research***, post publication.

Advice on support for meeting your obligations for data retention and storage are still under development as part of the review of the UC *Management of Research Data and Primary Materials* policy.

Background

Policies on data retention and storage are required to address the ownership of research materials and data, their storage, their retention beyond the end of the project, and appropriate access to them by the research community.

A central aim of such policies is that sufficient materials and data are retained to justify the outcomes of the research and to defend them if they are challenged.

Research data should be made available for use by other researchers unless this is prevented by ethical, privacy or confidentiality matters.

Considerable advantages derive from arrangements where the data is held in a repository that is public or widely available, or in some other way independent of the researcher, should there be a dispute.

Researchers should retain research data and primary materials for sufficient time to allow reference to them by other researchers and interested parties. This may be for as long as interest and discussion persist following publication.

Options Available

1. ***Supplementary Materials***

   Most journals will have generous limits on the inclusion of supplementary materials associated with publication, and so provide an avenue for the inclusion of raw data tables, spreadsheets, images and even videos. This may be the option of choice for researchers whose data are relatively limited in size.

2. ***Dryad***

   The Dryad Digital Repository, available at http://datadryad.org/ is a curated resource that makes the data underlying scientific publications discoverable, freely reusable, and citable. Dryad provides a general-purpose home for a wide diversity of datatypes.

   It is likely that 90% of obligations of researchers for data retention and access, post publication, can be met by the Dryad facility. Submitting to Dryad better connects data with the papers arising from those data, and helps researchers get more credit for their work by increasing the usability and citability of their data. Dryad does charge for service, and can be prohibitively expensive for large datasets (>10 Gb). See

http://datadryad.org/pages/payment for more information on charging. We assume researchers have avenues for meeting the cost of dryad submission through their faculty or research institute.

3. ***GigaScience and GigaDB***

This initiative aims to become a home to research from the growing number of biological and biomedical fields handling "big-data", GigaScience is a new type of open access, open data journal that provides standard scientific publishing linked directly to a database that hosts its relevant data. The associated GigaDB database, launched last year, provides a home to all of the supporting data and tools associated with research, overcoming one of the biggest challenges holding back reproducible research and research integrity. More information on GigaScience can be found at http://gigascience.biomedcentral.com/.

4. ***Medical and Health Sciences***

NCRIS and other national research support agencies have established a national facility to provide petabyte-scale research data storage, and related high-speed networked computational services, to Australian medical and health research organisations. This can be accessed through the med.data.edu.au site.

5. ***Other Disciplines***

There are a range of discipline-specific options for data retention, including for social sciences and humanities. In particular, social science researchers are directed to the Australian Data Archive (ADA, http://www.ada.edu.au/) which provides a national service for the collection and preservation of digital research data and makes these data available for secondary analysis by academic researchers and other users.

6. ***University of Canberra Research Repository***

Staff are encouraged to lodge a reference to digital data held in public repositories such as Dryad and GigaScience with the University of Canberra Research Repository in the form attached, as with other citable publications. This will both serve to showcase our research, and provide an alternate avenue for researchers to discover and cite your work.

Note that the University of Canberra Research Repository does not have the capacity to hold the data themselves.

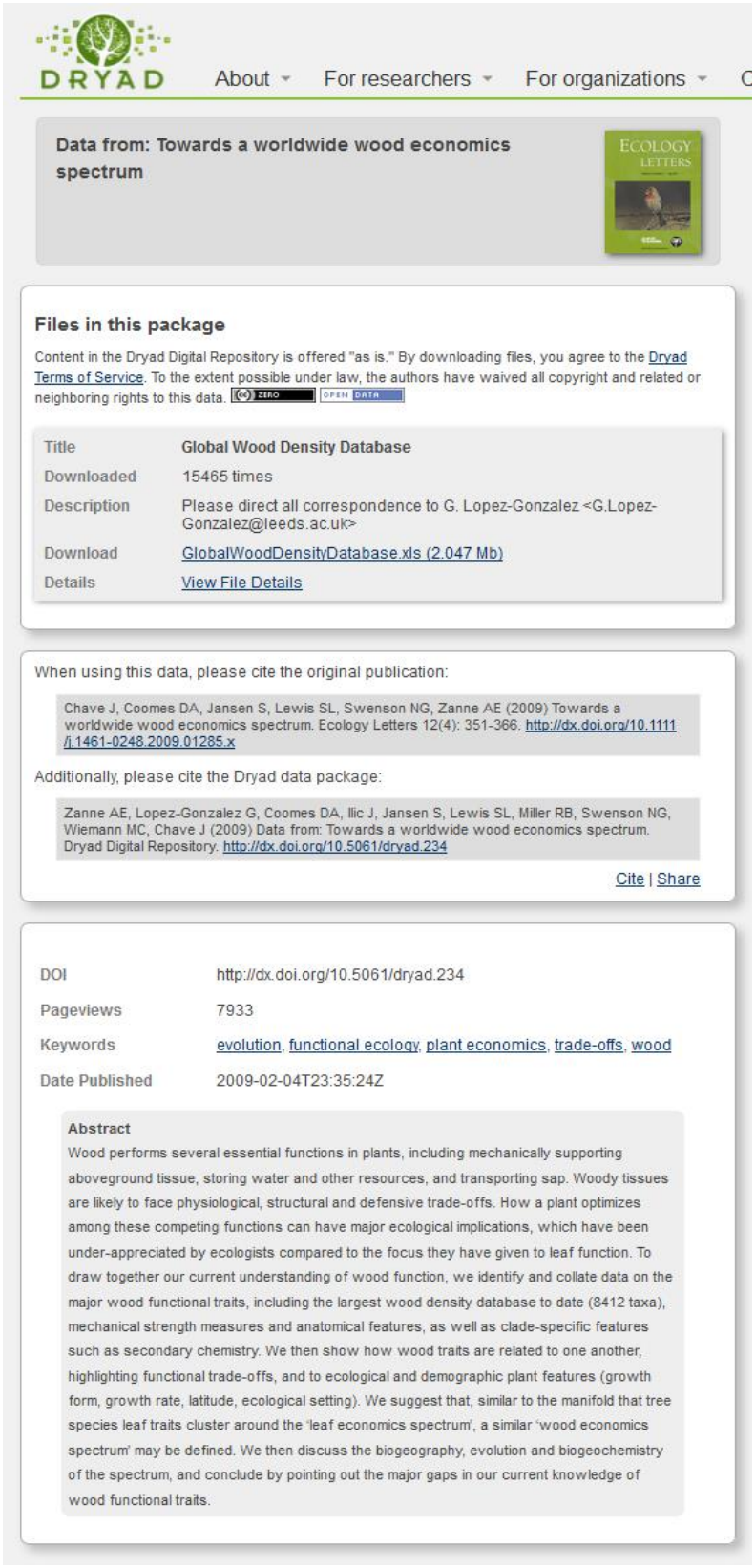
7. ***Confidential data***

Other arrangements may be required for the storage of confidential data post publication. The decision to store sensitive data in any environment involves assessing both the likelihood of unauthorised access to data and the severity of the consequences if a breach were to occur. Since no two datasets are completely alike, there can be no standard which applies in every case and as such a risk assessment should be considered. A great amount of information relevant to this matter has been collected on the med.data website

(http://med.data.edu.au/). As both the domain expert and data steward/custodian you will know your data best and have the best understanding of the protocols that should be observed in its handling. There is a great deal of difference, for instance, between information about an individual that can be considered sensitive (e.g. health or financial records) but has been anonymised and that which has not. Before transferring identifiable, sensitive data to any shared system, you should consider these matters, including with regard to any requirements outlined in ethics protocols governing this project, as well as principles for maintaining high standards in the responsible conduct of research, as articulated in the *Australian Code for the Responsible Conduct of Research* and internal UC policies.

Please consult with the UC eResearch Analyst, Mahin Raissi (mahin.raissi@canberra.edu.au) about available options.

Sample Dryad entry in a form that can be lodged as a pdf with the Library (as with other publications) with its associated citation details and doi. This particular example has had 15,259 downloads.



DRYAD About For researchers For organizations C

Data from: Towards a worldwide wood economics spectrum

Files in this package

Content in the Dryad Digital Repository is offered "as is." By downloading files, you agree to the Dryad Terms of Service. To the extent possible under law, the authors have waived all copyright and related or neighboring rights to this data.

Title	**Global Wood Density Database**
Downloaded	15465 times
Description	Please direct all correspondence to G. Lopez-Gonzalez <G.Lopez-Gonzalez@leeds.ac.uk>
Download	GlobalWoodDensityDatabase.xls (2.047 Mb)
Details	View File Details

When using this data, please cite the original publication:

Chave J, Coomes DA, Jansen S, Lewis SL, Swenson NG, Zanne AE (2009) Towards a worldwide wood economics spectrum. Ecology Letters 12(4): 351-366. http://dx.doi.org/10.1111/j.1461-0248.2009.01285.x

Additionally, please cite the Dryad data package:

Zanne AE, Lopez-Gonzalez G, Coomes DA, Ilic J, Jansen S, Lewis SL, Miller RB, Swenson NG, Wiemann MC, Chave J (2009) Data from: Towards a worldwide wood economics spectrum. Dryad Digital Repository. http://dx.doi.org/10.5061/dryad.234

Cite | Share

DOI	http://dx.doi.org/10.5061/dryad.234
Pageviews	7933
Keywords	evolution, functional ecology, plant economics, trade-offs, wood
Date Published	2009-02-04T23:35:24Z

Abstract

Wood performs several essential functions in plants, including mechanically supporting aboveground tissue, storing water and other resources, and transporting sap. Woody tissues are likely to face physiological, structural and defensive trade-offs. How a plant optimizes among these competing functions can have major ecological implications, which have been under-appreciated by ecologists compared to the focus they have given to leaf function. To draw together our current understanding of wood function, we identify and collate data on the major wood functional traits, including the largest wood density database to date (8412 taxa), mechanical strength measures and anatomical features, as well as clade-specific features such as secondary chemistry. We then show how wood traits are related to one another, highlighting functional trade-offs, and to ecological and demographic plant features (growth form, growth rate, latitude, ecological setting). We suggest that, similar to the manifold that tree species leaf traits cluster around the 'leaf economics spectrum', a similar 'wood economics spectrum' may be defined. We then discuss the biogeography, evolution and biogeochemistry of the spectrum, and conclude by pointing out the major gaps in our current knowledge of wood functional traits.