What can publishers do to support data?
Dryad’s perspective

STM Annual US Conference - April 22, 2015

Meredith Morovati
Executive Director
@datadryad
Dryad Digital Repository

- **Curated** resource that makes the data underlying scientific publications discoverable, freely reusable, and citable.
- Integration - support the work **without recreating infrastructure**
- All data submitted to Dryad are **associated with a publication**.

![Monthly Data Packages Released](chart.png)
Why have a policy?

Vines, T (2013, January 15) “I Want” Doesn’t Get — Just Recommending Data Archiving Isn’t Nearly as Effective as Requiring It, the scholarly kitchen http://bit.ly/1O7UCq5
An effective data policy

First the policy

• Joint Data Archiving Policy (JDAP)
• Data are important products of the scientific enterprise, and they should be preserved and usable for decades in the future
• ... Data supporting the results in the article should be deposited in an appropriate public archive
• Authors may elect to embargo...
• Embargo exceptions may be granted...

Now the reality

Integrated data submission

Matching functionality with policies
- Simple data deposition
- Permanent link to data within each article (and vice versa).
- Data can be used for review
- Data availability statements
- Embargo options

Integration options

Journals and publishers tell Dryad how they wish to configure integration to meet their needs. Key questions include:
- Whether data should be available before MS review
- Whether secure data review is needed
- Whether data embargoes are allowed

<table>
<thead>
<tr>
<th>Workflow options</th>
<th>Where author goes first</th>
<th>When author submits data</th>
<th>Metadata transfer to Dryad</th>
<th>Duplicate data entry?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dryad-first</td>
<td>Dryad</td>
<td>When MS is in preparation</td>
<td>Once, at MS acceptance</td>
<td>More</td>
</tr>
<tr>
<td>Review</td>
<td>Journal</td>
<td>Before MS review**</td>
<td>Twice, at MS submission and acceptance</td>
<td>Less</td>
</tr>
<tr>
<td>Acceptance</td>
<td>Journal</td>
<td>On MS acceptance**</td>
<td>Once, at MS acceptance</td>
<td>Less</td>
</tr>
</tbody>
</table>

**author submits via custom submission link
*also allows for secure data review
Data Availability Statement

For data deposited in a repository, which has been declared in the data availability statement, please provide a URL to the data for editorial access during review. Insert this as a Data Review URL submission item below. Our data repository integration partners facilitate this process by providing authors with the DOI and private reviewer URL at the point of data deposition. This ensures that authors have all the data access information needed when submitting their manuscript. Please see the Data Reporting Guidelines for more detail: [http://www.plosone.org/static/guidelines#accession](http://www.plosone.org/static/guidelines#accession)

Enter a Description and specify your External Resource. Then click Attach This Information.

**Description**: Data Review URL

**URL/URI/External Resource Location**: http://

Please select http://, https:// or ftp://, then type in the remainder of the web address (e.g. www.targetsite.edu/datasets/dataset1) and click the 'Preview' button to ensure that the address you entered navigates to the correct location. Once you are satisfied with your entry, click the 'Attach This Information' button.

**Data Availability**: The authors confirm that all data underlying the findings are fully available without restriction. The raw TPS file from this study has been deposited to Dryad ([http://dx.doi.org/10.5061/dryad.vn413](http://dx.doi.org/10.5061/dryad.vn413)).
What about REVIEW?

A. How would you expect a **published** dataset to differ from a **shared** one?

- Openly available: 68%
- Deposited in a repository: 54%
- Basis of a research paper: 43%
- Rich metadata: 39%
- Has a unique identifier: 39%
- Peer-reviewed: 28%
- Formal metadata: 25%
- Described in a data paper: 22%

B. What would you expect data **peer review** to consider?

- Methods are appropriate: 90%
- Enough metadata for replication: 80%
- Technical details check out: 70%
- Data is plausible: 61%
- Metadata properly standardized: 39%
- Novel/impactful: 22%

Dryad policies on citing data

When using this data, please cite the original publication:


Additionally, please cite the Dryad data package:


Download the data package citation in the following formats:

**RIS** (compatible with EndNote, Reference Manager, ProCite, RefWorks)

**BibTex** (compatible with BibDesk, LaTeX)
Recommendations

For publishers planning to reference datasets that have DataCite DOIs, CrossRef strongly recommends using the DataCite DOI in the citation, and also putting the citation in the reference section of the article (not just inline in the text). These practices will ensure that the data is permanently accessible via link and will make it easier to find data references within the journal and other scholarly literature.
Citing data in the literature

Novel forests maintain ecosystem processes after the decline of native tree species

Joseph Mascaro, R. Flint Hughes, and Stefan A. Schnitzer

1Department of Biological Sciences, University of Wisconsin, Milwaukee, Wisconsin 53211 USA
2Institute for Pacific Islands Forestry, USDA Forest Service, Hilo, Hawaii 96720 USA
3Smithsonian Tropical Research Institute, Apartado 2072, Balboa, Republic of Panama

Abstract. The positive relationship between species diversity (richness and evenness) and critical ecosystem functions, such as productivity, carbon storage, and nutrient cycling, is often used to predict the consequences of extinction. At regional scales, however, plant species richness is mostly increasing rather than decreasing because successful plant species introductions far outnumber extinctions. If these regional increases in richness lead to local


Dryad’s experience

Featured Data: Avoidable errors in the modeling of outbreaks of emerging pathogens... dx.doi.org/10.5061/dryad.

Featured Data: Recovery of a mini damaged stream ecosystem dx.doi.org/10.5061/dryad....

Featured Data: Towards a framework for agent-based image analysis of remote-sensing data dx.doi.org/10.5061/dryad....
Recommendations

Box 2. Recommendations for Publishers to Increase Access to Data

1. Establish and enforce a mandatory data availability policy.
2. Contribute to establishing community standards for data management and sharing.
3. Contribute to establishing community standards for data preservation in trusted repositories.
4. Provide formal channels to share data.
5. Work with repositories to streamline data submission.
6. Require appropriate citation to all data associated with a publication—both produced and used.
7. Develop and report indicators that will support data as a first-class scholarly output.
8. Incentivize data sharing by promoting the value of data sharing.


Thank you

Dryad Community Meeting
Taking a closer look at data

Datadryad.org

27 May 2015. Washington, DC
8:30AM - 4:00PM

Repository home: http://datadryad.org
News: http://blog.datadryad.org
Twitter: @datadryad

Meredith Morovati, mmorovati@datadryad.org
For more information

Data Policies

- Vines, T (2013, January 15) “I Want” doesn’t get — Just recommending data archiving isn’t nearly as effective as requiring It, the scholarly kitchen [http://bit.ly/1O7UCq5](http://bit.ly/1O7UCq5)


Data Review


Data Citation

- Data Citation Synthesis Group: Joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego CA: FORCE11; 2014 [https://www.force11.org/datacitation](https://www.force11.org/datacitation)

- [http://www.crossref.org/10quarterly/quarterly_jan12.html](http://www.crossref.org/10quarterly/quarterly_jan12.html)