



Facilitating data publishing through journal integration

Elizabeth Hull¹, Jamie Lamkin², Todd J. Vision³

Dryad Digital Repository

Dryad is a **curated** resource that makes the data underlying scientific publications **discoverable, freely reusable, and citable**.

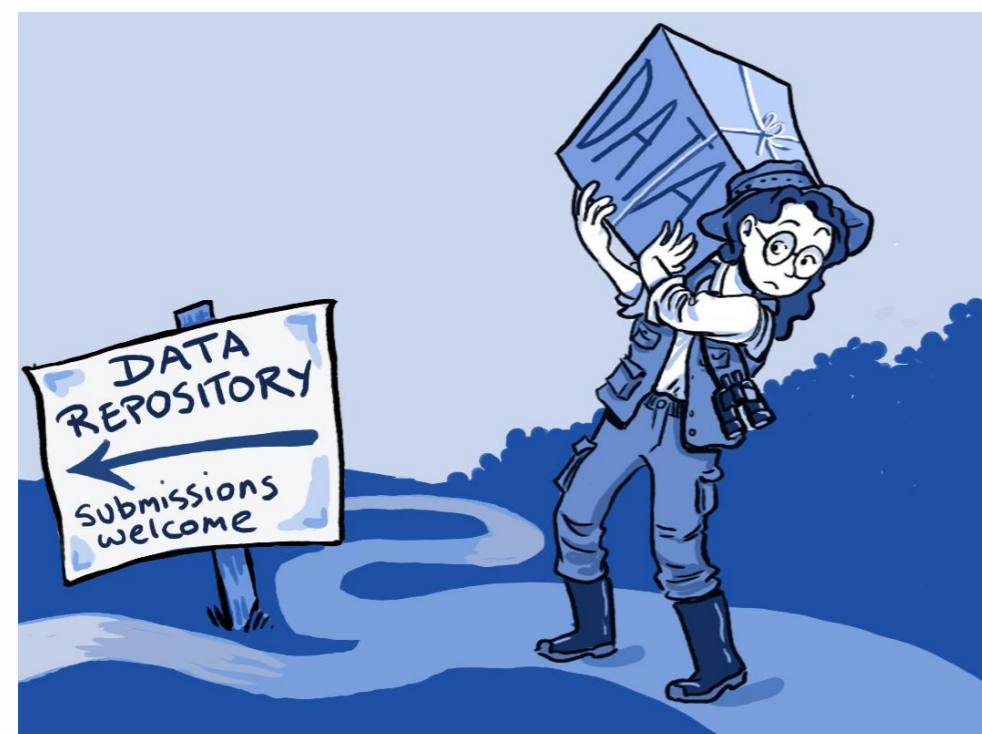


Illustration credit: Ainsley Seago.

Dryad provides a **general-purpose** home for a wide diversity of data types.

All data submissions to Dryad **must be associated with a publication**.

What is submission integration?

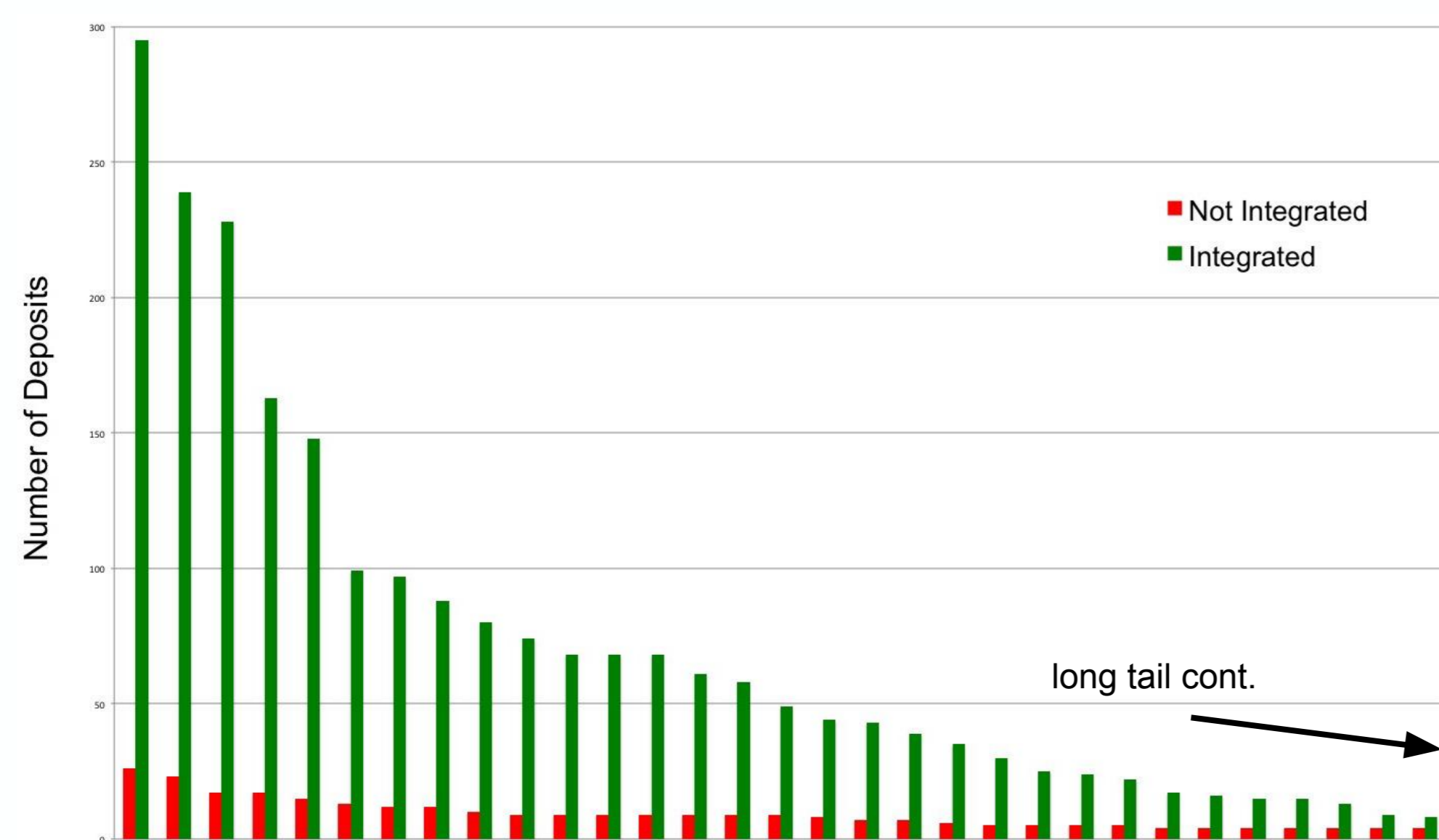
A free service that allows publishers to coordinate manuscript submission with data submission by setting up automatic notifications between journals and Dryad. Benefits:

- Simplify the process of data submission for authors.
- Reduce the rate of noncompliance with journal data policy.
- Ensure bidirectional links between the article and the data and increased visibility for both.

67 journals are currently integrated with the Dryad data submission system.

365 journals are currently represented in Dryad.

Integrated journals account for the vast majority of deposits:

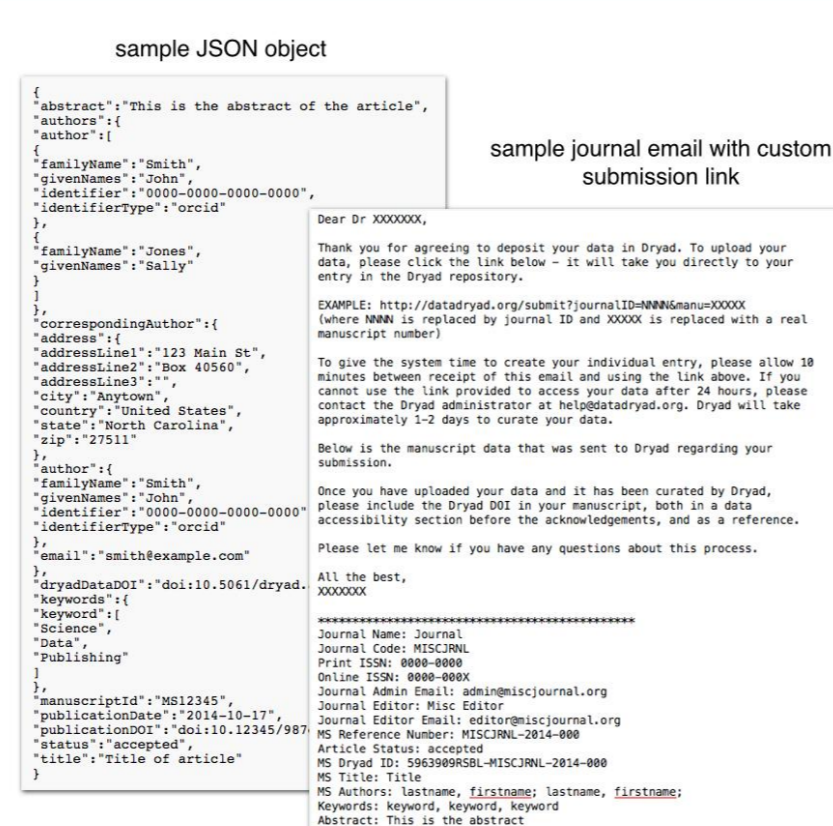


2014 Dryad deposits by journal. Only journals with the most deposits are displayed.

Metadata transfer

Article/manuscript descriptions from publishers can come to Dryad via:

1. **Email:** notifications sent by the journal system are automatically processed upon receipt.
2. **REST API:** journal sends API calls to Dryad consisting of a structured JSON object containing manuscript metadata.



Integration options

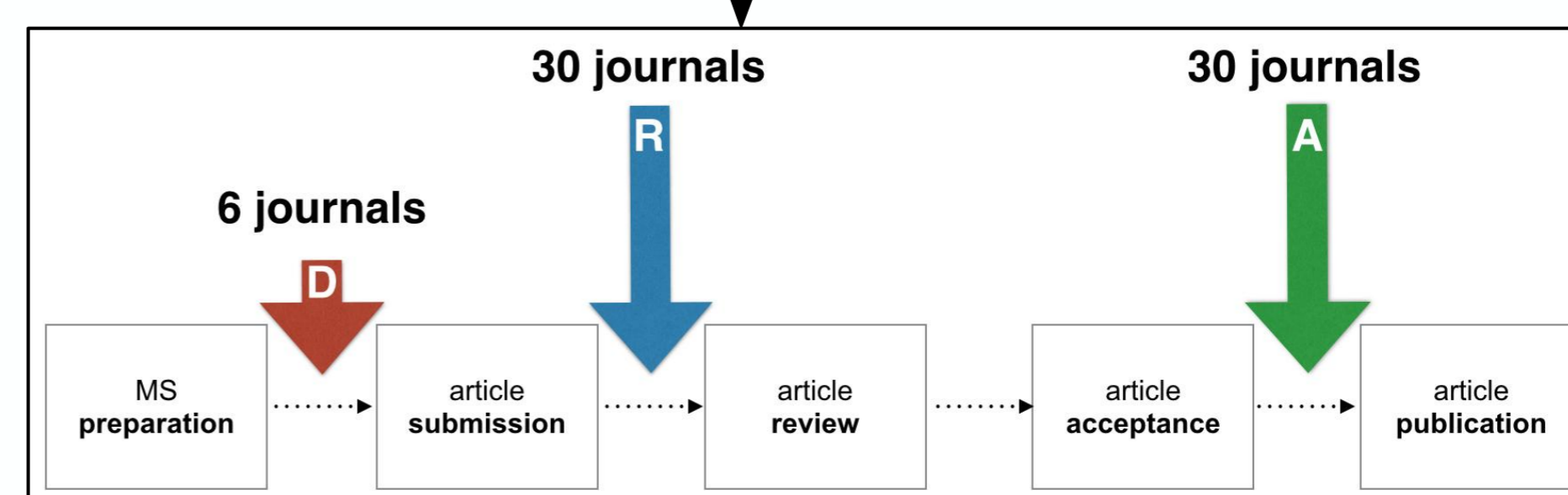
Journals and publishers tell Dryad how they wish to configure integration to meet their needs. Key questions include:

- whether data should be available **before** MS review
- whether secure **data review** is needed
- whether **data embargoes** are allowed

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

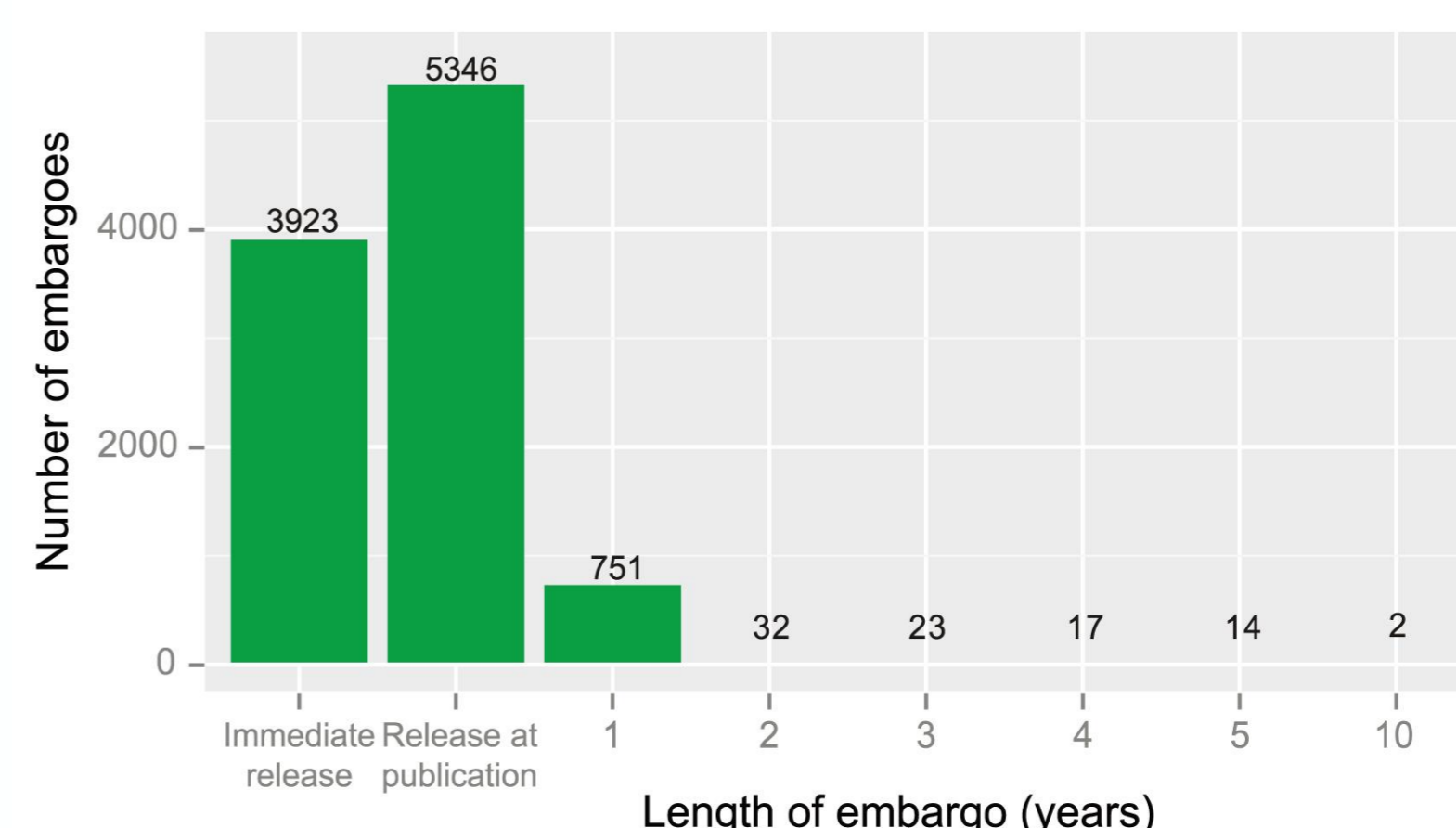
*author submits via custom submission link

**also allows for secure data review



When is data released?

70% of integrated journals allow data embargoes extending beyond article publication, but most authors don't choose that option:



Embargo selections of Dryad data authors for the 10,108 files in Dryad deposited from inception to September 20, 2013. Data include only datasets related to articles published in journals for which the authors had the option of selecting an embargo. Roche DG, Lanfear R, Binning SA, Haff TM, Schwanz LE, et al. (2014) Troubleshooting Public Data Archiving: Suggestions to Increase Participation. PLoS Biol 12(1): e1001779. doi:10.1371/journal.pbio.1001779

Journal case study: Molecular Ecology



Integration workflow: **Review**
Data embargo allowed?: **No**

ME data packages in Dryad:

1,147 (15% of total)

ME data files in Dryad:

4,313 (avg 3.7/package)

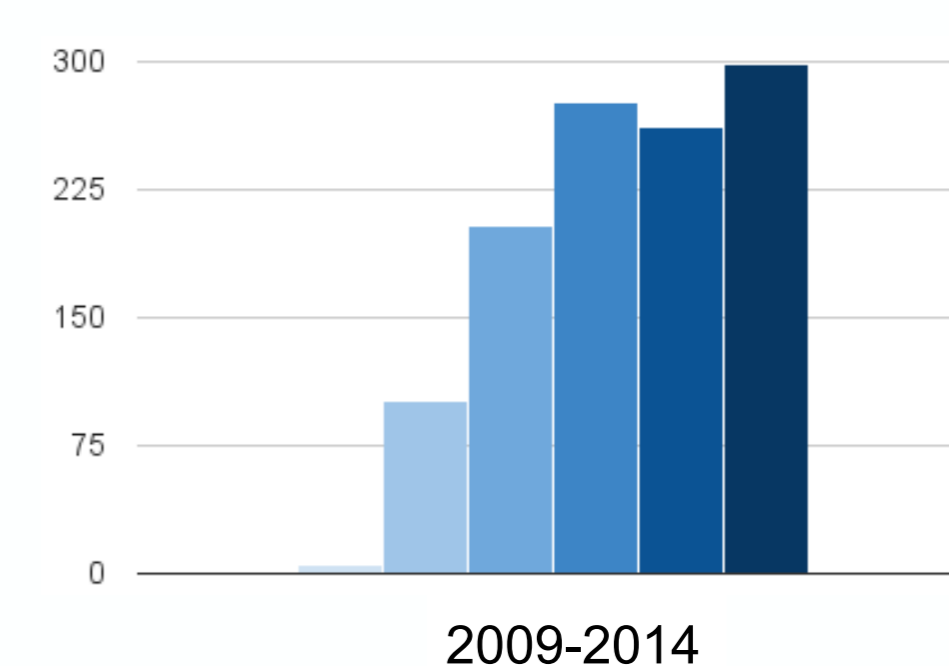
Total size of data packages:

162 Gb

Total ME file downloads:

56,749

ME data packages released by year



Dryad is funded in part by a "Sustainable and Scalable Infrastructure for the Publication of Data" grant (award 1147166) from the U.S. National Science Foundation.

1. Dryad, 2. Dryad & National Evolutionary Synthesis Center, 3. National Evolutionary Synthesis Center & University of North Carolina at Chapel Hill.

Learn more: <http://datadryad.org> @datadryad /DataDryad