



PROMOTING SCHOLARSHIP THROUGH OPEN DATA

## 2016 Organizational and Membership Report

### **Table of contents**

[Message from the Executive Director](#)

[Partner journals](#)

[Data packages published in 2016](#)

[Membership](#)

[Data publishing charges](#)

[Downloads and web traffic](#)

[Outreach](#)



## Message from the Executive Director

In 2016, Dryad celebrated many accomplishments surrounding sustainability and growth.

The Dryad business model enables a variety of stakeholders with different needs, constituents, and budgets to contribute to the sustainability of data preservation, curation, and publication. Our goal has always been to grow beyond an initial reliance on grants for basic operational costs and to limit the cost burden on individual researchers. We aim to accomplish this by developing partnerships with a wide variety of sponsors including institutions, libraries, publishers, societies, and funders. In 2016, we welcomed 18 new partner journals that integrated their manuscript submission system with data submission to Dryad.

We seek to identify funding opportunities that are designed to support growth. We received a [new three-year Sustaining award](#) from the U.S. National Science Foundation (NSF) to support “cyberinfrastructure that is critical for the continued advance of priority biological research.” We also received [pilot funding from NSF](#) to study how a funder might directly pay for Data Publishing Charges (DPC). This pilot allowed us to learn more about [the ways researchers pay for data](#) when there is no sponsorship in place and the challenges they face.

For a growing non-profit organization, financial security is a significant goal. We revised our pricing structure in 2016 to reflect what we’ve learned about curating data and preserving it over the long term. We made pricing consistent for all journals, regardless of integration; increased the discounts Dryad members receive; doubled the maximum size of datasets before overage fees kicked in; and increased DPCs.

In spite of the current threat to the accessibility of research data, more is being published than ever before. At the end of 2016, Dryad contained 15,325 data packages linked to articles. 4,307 of those were from 2016 alone. The most popular data packages for 2016 represented a broader variety of disciplines, including social media and medical imaging.

As an independent, non-profit organization, we are committed to good stewardship of data and serving our stakeholder community of researchers and research organizations. We are excited for the year ahead as we continue to grow, broaden our support and disciplinary variety, and open up data for the advancement of science and medicine.

A handwritten signature in black ink that reads "Meredith Morovati". The signature is fluid and cursive.

*Meredith Morovati, Executive Director*

## Partner journals

Submission integration is a free service that allows journal publishers to coordinate the submission of manuscripts with submission of data to Dryad, used by over 100 partner journals. Dryad supports submission integration for any organization that publishes a scientific journal irrespective of whether the organization sponsors submissions through a payment plan or is a member of Dryad. The submission integration process works with many different online manuscript processing systems. It is lightweight, customizable to each journal's needs, and brings a number of benefits:

- Simplify the process of data submission for authors.
- Allow authors to deposit, to a single repository, gigabytes of data files in their original formats.
- Reduce the rate of noncompliance with journal data policy.
- Have the option of making data available for editorial or peer review, via secure access for editors and reviewers.
- Ensure bidirectional links between the article and the data and increased visibility for both.
- Ensure that the data is accessible once the article becomes available online.
- Give authors the option to embargo public access to data for a limited time after publication, if permitted by the journal's data policy.

## New partner journals in 2016

Title	Workflow	Sponsorship
<i>Biotropica</i>	basic	Yes
<i>Behavioral Ecology</i>	review	Yes
<i>Work, Aging and Retirement</i>	review	No
<i>Virus Evolution</i>	review	No
<i>Urban Ecology</i>	review	No
<i>Environmental Epigenetics</i>	review	No
<i>BioScience</i>	review	No
<i>Toxicological Sciences</i>	review	Yes
<i>Journal of Systematics and Evolution</i>	review	Yes

Title	Workflow	Sponsorship
<i>Applications in Plant Sciences</i>	review	No
<i>European Respiratory Journal</i>	basic	Yes
<i>ERJ Open Research</i>	basic	Yes
<i>Journal of Experimental Botany</i>	review	Yes
<i>Annals of the Entomological Society of America</i>	review	No
<i>Environmental Entomology</i>	review	No
<i>Journal of Economic Entomology</i>	review	No
<i>Journal of Medical Entomology</i>	review	No
<i>Journal of Insect Science</i>	review	No

## Data packages published in 2016

Dryad published 4,307 data packages in 2016, a 10% increase from 2015. They were associated with articles from over 268 different journals. Dryad also continued to receive a small number of data packages associated with books and theses.

There were an average of 3.15 files per data package, about the same as the prior year, while the average size per data package was 573 MB, an increase of 20%. Only 4% of data packages contained files embargoed for one year or longer.

### ***The ten journals with the most data packages published in 2016***

Journal	Data packages
<i>PLOS ONE</i>	421
<i>Proceedings of the Royal Society B</i>	295
<i>Molecular Ecology</i>	269
<i>Evolution</i>	199
<i>Royal Society Open Science</i>	186

Journal	Data packages
<i>Ecology and Evolution</i>	150
<i>Biology Letters</i>	132
<i>Journal of Evolutionary Biology</i>	131
<i>Journal of Applied Ecology</i>	129
<i>Journal of Ecology</i>	115

## Membership

Membership is open to any stakeholder organization, including but not limited to journals, scientific societies, publishers, research institutions, libraries, and funding organizations. Members help advance Dryad's mission and have a say in the governance of the organization through nomination and election of its Board of Directors, and voting on amendments to its Bylaws at the Annual Meeting. Members benefit from discounts on submission fees and participation in an active knowledge sharing network. Memberships run from January through December. Annual dues are \$1,000 for organizations with gross income below 10 million USD, and \$5,000 if above. Members can also join for \$500, irrespective of budget size, but will not benefit from DPC savings. Current Dryad members:

- [American Association for the Advancement of Science](#)
- [American Society of Naturalists](#)
- [The American Genetic Association](#)
- [Association for Tropical Biology and Conservation](#)
- [Botanical Society of America](#)
- [British Ecological Society](#)
- [BMJ Publishing Group, Ltd.](#)
- [Cambridge University Press](#)
- [European Society for Evolutionary Biology](#)
- [German National Library of Medicine](#)
- [HighWire](#)
- [Oikos](#)
- [Oxford University Press](#)
- [The Royal Society](#)
- [Society for the Study of Evolution](#)
- [Society of Systematic Biologists](#)
- [United States Fish and Wildlife Service](#)
- [Wiley](#)

# Data publishing charges

## Payment plans

Independent of membership, organizations may contract with Dryad to provide a valuable service to researchers: allowing them to deposit data without paying a deposit fee at the time of submission. To do this, Dryad offers flexible payment plans, all of which provide a volume discount. If no payment plan is in place, authors pay a \$120USD fee at the time of archiving, with the exception of researchers from developing countries, for whom the fee is waived.

At the beginning of 2016, Dryad implemented a revised pricing structure that simplifies payment plans and individual charges, increases the base Data Publishing Charge (DPC), and raises the maximum package size before overage fees kick in (to 20GB). Payment plan details were updated in 2016 to the following:

**1. Subscription Plan.** A journal (or journals) with an active Subscription Plan may make an unlimited number of submissions for a fixed cost. This cost is calculated based on the total number of research articles published by the journal(s) in the prior year, or a reasonable estimate of anticipated volume for new titles.

**2. Usage Plan.** Payment is based on the actual number of data submissions Dryad receives, with prepaid or deferred options:

- **Prepaid voucher option:** Organizations can purchase a block of vouchers for future submissions at an additional discount. The minimum purchase to receive a volume discount is 25 vouchers. Vouchers can be used on any submission for any publication.
- **Deferred payment option:** Organizations will be invoiced quarterly for the number of packages submitted by their authors in the previous quarter. Participants must enter into a contract of at least one year in duration and may cap the number of submissions for which they are charged.

While base DPCs cover the entire cost of most submissions, additional charges will apply when data packages are in excess of the new 20GB size limit. Submitters will be charged \$50 for each additional 10GB, or part thereof. (Packages between 20 and 30GB = \$50, between 30 and 40GB = \$100, and so on).

## Summary of payment plan options

Payment plan	For Sustainer/ Supporter members	For all others
Subscription	annual fee based on \$40 per published research article	annual fee based on \$50 per published research article
Usage: Prepaid	\$95 per data package	\$105 per data package
Usage: Deferred	\$105 per data package	\$115 per data package

## Sponsoring organizations

As of 2016, more than 30 organizations sponsor DPCs on behalf of researchers and authors. The publishing organizations in this list collectively sponsor DPCs for 67 titles, while the research organizations sponsor DPCs on behalf of their affiliated researchers wherever they choose to publish. An up-to-date list of journals with sponsored DPCs is available at <http://datadryad.org/pages/journalLookup>. Sponsoring organizations:

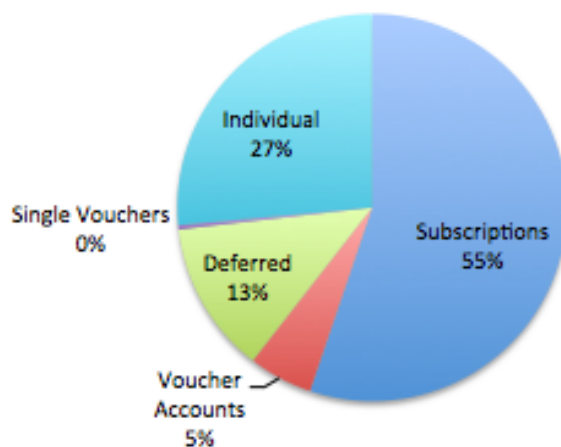
- American Genetic Association
- American Society of Naturalists
- American Society of Plant Biologists
- American Society of Plant Taxonomists
- Behavioral Ecology
- Biotropica
- British Ecological Society
- British Editorial Society of Bone and Joint Surgery
- Canadian Healthy Oceans Network CHONE
- Ecological Society of America
- Ecology and Evolution
- eLife
- European Respiratory Society
- European Society for Evolutionary Biology
- Evidence-Based Preclinical Medicine
- Evolutionary Applications
- German National Library of Medicine
- Journal of Consumer Research
- Journal of Experimental Botany
- Journal of Systematics and Evolution
- Molecular Ecology
- Molecular Ecology Resources
- NIOO-KNAW
- Nordic Society Oikos

- Pensoft Publishers
- Society for the Study of Evolution
- Society of Systematic Biologists
- The Genetics Society
- The Palaeontological Association
- The Paleontological Society
- The Royal Society
- Toxicological Sciences
- University of Chicago Press
- University of Rochester
- US Fish and Wildlife Service

### Revenue, by payment plan, 2016 (in USD)

	Q1	Q2	Q3	Q4	Total
Subscription Plan	\$27,576	\$17,500	\$13,228	\$2,173	\$60477
Voucher Plan	7,135	2,625	28,900	5,670	44330
Deferred Payment Plan	12,707	10,710	-	10,605	34022
Single vouchers	-	240	-	-	240
Individuals	30,035	35,825	32,130	30,190	128180
<b>Total</b>	<b>\$77,453</b>	<b>\$66,900</b>	<b>\$74,258</b>	<b>\$48,638</b>	<b>\$267249</b>

### Percentage of data packages published by payment type



### Waivers

Waivers are granted upon request for submissions originating from researchers based in countries classified by the World Bank as low-income or lower-middle-income economies. Waivers were granted for 82 data packages in 2016, compared with 74 in 2015.



## Financials

In addition to the budget for the nonprofit, a three-year \$762K sustaining award from the US National Science Foundation started in September 2016. An additional \$140K award from the US NSF started in May of 2016 to run for 10 months. This last award was a pilot project to test the feasibility of funders paying directly for data publication charges. The University of North Carolina at Chapel Hill partnered with Dryad on these awards.

## Downloads and web traffic

There were approximately 475,000 data file downloads in 2016. Although most activity from automated “bots” has been filtered out of these statistics, they may still include counts from individual users downloading files multiple times.

### *The five most frequently downloaded data packages published in 2016*

Publication	Data package	Downloads
Wagner MR, et al (2016) Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. <i>Nature Communications</i> 7: 12151. <a href="http://dx.doi.org/10.1038/ncomms12151">http://dx.doi.org/10.1038/ncomms12151</a>	Wagner MR, et al (2016) Data from: Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. Dryad Digital Repository. <a href="http://dx.doi.org/10.5061/dryad.g60r3">http://dx.doi.org/10.5061/dryad.g60r3</a>	3,123
Bohannon J (2016) Who's downloading pirated papers? <i>Everyone</i> . <i>Science</i> 352(6285): 508-512. <a href="http://dx.doi.org/10.1126/science.352.6285.508">http://dx.doi.org/10.1126/science.352.6285.508</a>	Bohannon J, Elbakyan A (2016) Data from: Who's downloading pirated papers? <i>Everyone</i> . Dryad Digital Repository. <a href="http://dx.doi.org/10.5061/dryad.q447c">http://dx.doi.org/10.5061/dryad.q447c</a>	2,969
D'Orangeville L, et al (2016) Northeastern North America as a potential refugium for boreal forests in a warming climate. <i>Science</i> 352(6292): 1452-1455. <a href="http://dx.doi.org/10.1126/science.aaf4951">http://dx.doi.org/10.1126/science.aaf4951</a>	D'Orangeville L, et al (2016) Data from: Northeastern North America as a potential refugium for boreal forests in a warming climate. Dryad Digital Repository. <a href="http://dx.doi.org/10.5061/dryad.785cv">http://dx.doi.org/10.5061/dryad.785cv</a>	741
Johnson NP, et al (2016) Continuum of vasodilator stress from rest to contrast medium to adenosine hyperemia for fractional flow reserve assessment. <i>JACC. Cardiovascular Interventions</i> 9(8): 757-767. <a href="http://dx.doi.org/10.1016/j.jcin.2015.12.273">http://dx.doi.org/10.1016/j.jcin.2015.12.273</a>	Johnson NP, et al (2016) Data from: Continuum of vasodilator stress from rest to contrast medium to adenosine hyperemia for fractional flow reserve assessment. Dryad Digital Repository. <a href="http://dx.doi.org/10.5061/dryad.f76nv">http://dx.doi.org/10.5061/dryad.f76nv</a>	453
Lu J, et al (2016) The oldest actinopterygian highlights the cryptic early history of the hyperdiverse ray-finned fishes. <i>Current Biology</i> 26(12): 1602–1608. <a href="http://dx.doi.org/10.1016/j.cub.2016.04.045">http://dx.doi.org/10.1016/j.cub.2016.04.045</a>	Lu J, et al (2016) Data from: The oldest actinopterygian highlights the cryptic early history of the hyperdiverse ray-finned fishes. Dryad Digital Repository. <a href="http://dx.doi.org/10.5061/dryad.t6j72">http://dx.doi.org/10.5061/dryad.t6j72</a>	423

## Web traffic

Dryad collects a variety of web analytics for <http://datadryad.org>, but pays particular attention to the number of “sessions,” or visits from the same user with gaps of no more than 30 minutes between subsequent hits. The number of sessions increased 23%, from 247,185 in 2015 to 304,225 in 2016.

### *Sessions by language (top 5)*

Language	Percentage
English	72%
Chinese	4%
French	3%
Spanish	2.4%
German	2%

### *Sessions by country (top 10)*

Country	Percentage
United States	37.7%
United Kingdom	8.8%
Canada	5%
Germany	4.6%
China	3.5%
Australia	3.3%
France	3.1%
Brazil	2.6%
Japan	2.5%
Spain	2.2%

## Outreach

### Blog posts

There were 11 posts to <http://blog.datadryad.org> in 2016, which collectively received 13,579 pageviews from 9,947 visitors, nearly double the traffic received in 2015. The most popular by far was a guest post by science journalist John Bohannon on usage of the Sci-Hub repository. The second most popular post, “Piloting new ways for funders to support data stewardship”, described Dryad’s NSF pilot study of funding models for data archiving.

#### ***Blog posts from 2016, in reverse chronological order***

- Researcher Profile: Zach Gompert
- A month of open
- NSF sustaining their support for open data
- Introducing Dryad’s new board members
- Piloting new ways for funders to support data stewardship
- Dryad’s first virtual community meeting: members share their expertise
- Making open data useful: A drug safety case study
- Sci-Hub stories: Digging into the downloads
- 2015 stats roundup
- Newly integrated titles from Oxford University Press Journals
- New partnership with The Company of Biologists

#### ***Blog readership, by country (top 10)***

Country	Number of visits
USA	5,517
UK	1,352
Germany	575
Canada	566
India	551
France	372
Brazil	313
Australia	308
Netherlands	268
China	263

## Twitter

The @datadryad Twitter account had roughly 11,000 followers by the end of 2016 (up approximately 17% from the year before). We also continue to post all new data packages, and Featured Data, to @datadryadnew.

### A sample of tweets from 2016

 **Micaela E Martinez** @ME\_\_Martinez · 19 Apr 2016  
Replying to @ME\_\_Martinez  
@MartinezBakker by @EvolveDotZoo and many colleagues, they have posted their data on @datadryad [datadryad.org/resource/doi:1...](http://datadryad.org/resource/doi:1...)

 **Sciencegurrl**  
@sciencegurrlz0

The @datadryad help desk is amazing.

 **Erin E. McCabe** @e\_erinm · 29 Apr 2016  
@sciencemagazine and @datadryad released a dataset of 28mil download request records from Sci-Hub. Soooo I know what I'll be doing all wknd

 **Thomas E. Keller** @tek\_keller · 29 Apr 2016  
Kudos to @datadryad for hosting the #scihub dataset, really looking forward to digging into that. They are a good repo for scientific data.

 **Emanuel A. Fronhofer** @efronhofer · 25 Apr 2016  
Adding code to @github and datasets to @datadryad for all my published work. Making my research more reproducible: [emanueifronhofer.net/publications.h...](http://emanueifronhofer.net/publications.h...)

 **Alex Hall** @allopatry · 17 May 2016  
Replying to @allopatry  
These data are available on @datadryad. Also my first experience with this. Very professional staff! [datadryad.org/resource/doi:1...](http://datadryad.org/resource/doi:1...)

 **John Chodera** @jchodera · 25 Jun 2016  
How do you all share GB-scale datasets? We made the Ensembler human tyrosine kinase dataset available on Dryad: [datadryad.org/resource/doi:1...](http://datadryad.org/resource/doi:1...)

 **Sebastian Eves-van d** @Seb\_EvdA · 13 Jul 2016  
Now I found @datadryad, Ive made more useful data available for N. aberrans transcriptome GBE 2014 [datadryad.org/resource/doi:1...](http://datadryad.org/resource/doi:1...) Sorry its late!

 **Biotropica** @Biotropica · 12 May 2016  
I'm excited our first dataset has been archived as part of the @datadryad + @Biotropica partnership. Thanks to E. Padyšáková & Š. Janeček!

 **Kristina Riemer** @KristinaRiemer · 6 May 2016  
I spent a good 20 minutes at a party last night convincing someone to put their master's mussels data on @datadryad. Who am I?

 **Craig Primmer** @FishConGen · 19 Jul 2016  
Graham: having data on @datadryad helped solve a weird result (other scientists worked it out): salmon DNA replaced by bacteria DNA #fsbi16

## Conferences and seminars

Dryad had a presence through either a presentation, conference table or sponsorship at the following events in 2016:

- Data & Publication Linking Workshop, 5 January, Washington DC
- CurateGear, 14 January, Chapel Hill, NC, USA
- Academic Publishing in Europe, 19-20 January, Berlin, Germany
- International Digital Curation Conference, 22-25 February, Amsterdam
- National Academy of Sciences Journal Summit, 17 March, Washington, DC
- 3rd Life Discovery – Doing Science Education, 18-19 March, Baltimore, MD, USA
- National Data Service Consortium Workshop, 4-6 April, Chapel Hill, NC, USA
- FORCE, 17-19 April, Portland, OR, USA
- NSF AVATOL PI meeting, 26-27 April, Washington DC
- Council of Science Editors, 14-17 May, Denver, CO, USA
- NSF Workshop on Data and Software Citation - Connecting Journals to Data Repositories Workshop, 7-10 June, Cambridge, MA, USA
- Open Repositories, 13-17 June, Dublin, Ireland
- American Association of University Presses, 16-18 June, Philadelphia, PA, USA
- Summer Institute on Quantitative Biology, 19-24 June, Raleigh, NC, USA

- Force11 Data Citation Early Adopters Workshop, 22 June, San Diego, CA, USA
- BioCaddie Repository Workshop, 23 June, San Diego, CA, USA
- SHARE Community Meeting and Hackathon, 11-14 July, Charlottesville, VA, USA
- CODATA Data Citation workshop, 12 July, Washington DC
- DataONE Users Group Meeting, 17-18 July, Durham, NC, USA
- SciDataCon, 11-13 September, Denver, CO, USA
- Research Data Alliance 8th Plenary, 13-15 September, Denver, CO, USA
- Open Access Scholarly Publishers Association, 21-22 Sept, Arlington, VA, USA
- Society for Scholarly Publishing Regional Event, 8 December, Durham, NC, USA