

Coopetition as a driver of success for community initiatives in open research

Jennifer Gibson*

Executive Director, Dryad, London, UK

Abstract. “Co-opetition [1]” can help community-driven initiatives to combine strengths, focus resources and thrive in a changing and competitive landscape. Dryad is presented as one such initiative—committed to multi-directional collaboration and co-opetition, which is beginning to nest itself among other complementary services and form a network for open research that increases the collective resilience of aligned organisations.

Keywords: Co-opetition, Dryad, California digital library, collaboration, open science

1. Introduction

The following article is an expansion on remarks offered at the 2022 NISO Plus session on “Open Science: catch phrase, or a better way of doing research?”, organised by Shelley Stall, Senior Director for Data Leadership at the American Geophysical Union. Remarks as presented touched very briefly on the multi-directional collaborations in which Dryad is involved. Here, I offer further context for those relationships, the motivations behind them, and the benefits we hope to achieve for open research through collaboration.

I would like to start with a broad comment on the importance of collaboration for community-driven initiatives in the scholarly communication space. This is a competitive environment, in which programs with limited resources compete with others that are heavily resourced. Myriad organisations such as Dryad, which are community-supported and community-led, but small, compete with some well-staffed, well-oiled machines—for attention and precious investment from institutions, funders, publishing organisations, and other stakeholders.

“Co-opetition [2]”, working with other organisations with complementary strengths, can help less-resourced initiatives to thrive in a changing and competitive landscape. We might collaborate with others in order to:

- achieve common purpose;
- build shared vision;
- limit redundancy;
- maximise strengths;
- leverage shared resources;
- increase impact.

*E-mail: jgibson@datadryad.org.

Below, I use Dryad as an example of an organisation committed to multi-directional collaboration and co-competition, which is beginning to nest itself among other complementary services and form a network for open research that increases our collective resilience in this competitive environment. I will describe our intersections with publishing organisations, different forms of infrastructure, and services—all of which help Dryad to reinforce our own position and strengths while increasing our value for stakeholders. A significant omission from my talk was Dryad’s collaboration with the California Digital Library, which is another excellent example of aligned and complementary organisations working towards a common purpose.

As I joined Dryad as Executive Director only in October 2021, I must credit Dryad’s founders (especially Todd Vision and Michael Whitlock, though there are many people whose efforts are felt at Dryad today), Dryad’s Board over the years, my predecessors (especially Tracy Teal and Elizabeth Hull), and our collaborators at the California Digital Library (especially Daniella Lowenberg and John Chodacki) for developing this vision. I see great potential in it.

For reference, Dryad is a non-profit organisation, open data curation and publishing platform, and multi-stakeholder community dedicated to the open availability and routine reuse of all research data. Dryad started at Duke University in 2007, with funding from the U.S. National Science Foundation and support from the University of North Carolina at Chapel Hill and North Carolina State University, and now operates in collaboration with the California Digital Library.

2. Publishing organisations

Dryad collaborates with publishing organisations to achieve common purpose and build a shared vision for the open availability and full reuse of research data.

Publishing organisations—by which I mean dedicated publishers of different kinds, whether society, non-profit or commercial as well as institution-based publishing programs including but not limited to university presses, plus independent scholar-run programs and others that validate, register, and disseminate new findings—are powerful collaborators in research communication. I don’t need to say that they occupy a point of focus in the researchers’ journey, a milestone on the path that they strive to surmount and pass, to register their claims, gain credit, and advance.

I am excluding here important initiatives to disaggregate traditional publishing functions: such as preprints, which help expedite the communication of new findings before peer review or the publish-review-curate approach, which offers expert peer review as a discrete function after pre-printing. These raise promising questions around where the value of traditional publishing really lies, in my view, and how publishing workflows might flex and change, to become more “liquid [3]”. I am also excluding other types of knowledge custodians with which I’m less familiar, but learning more about. I mean for example rural communities where knowledge of critical domains including health and food production is shared in different ways and excluded by the research communication system that I am describing here [4].

What I want to emphasise is that the organisations currently offering the imprimatur or validation sought by the individual researcher, where they may seek it, have a great deal of influence and are valuable allies. They are able to engage researchers in all kinds of improvements to research communication and use their position to help advance open practice, research integrity, reproducibility, and more.

This form of collaboration is a pillar for Dryad. The researchers who designed Dryad engaged their journal publishers in the first instance, and established policies and processes to capture research data as part of the submission, review, and publication workflow [5]. The journals working with Dryad have varying policies and practices for data sharing, but they do represent some of the most engaged disciplinary

communities outside the specialist realms. One journal secures data publication for every published article. The outcome for every data set submitted to Dryad, regardless of policy or extent of practice, is that a DOI is generated, registered with DataCite along with all of the metadata that we collect, returned to the publisher for inclusion in the paper (and/or reference during peer review), and published as an independently-citable object. Publishing organisations are leveraging Dryad's facility for data curation, registration, publication, and preservation while Dryad taps into the all-important publishing workflow. Together, we're building a shared vision for improving and accelerating research with open and ready access to underlying data.

3. Infrastructure

Through our connections with other forms of infrastructure, Dryad—and the research data that we publish—is nested in the expanding global fabric for open research communication. These connections enable the multidirectional flow of research outcomes and add resilience for Dryad way outside the walls of our small organisation.

I am starting to think of infrastructure in a different way. Until recently I had thought of it as a collection of discrete systems with broad intent, e.g., individual manuscript submission and review system, which hundreds (or thousands) of journals might use; an individual institutional repository for all campus outputs; an individual repository for one discipline's collection; one system supporting the review and publication of one country's research articles; one database of stunning images from one funder.

Increasingly, I am coming to see the power of what colleagues have been building for some time—an open infrastructure that connects systems and traverses their borders in a way that has the potential to unify all the world's knowledge. Infrastructure in 2022 encompasses the unique identifiers that are system-agnostic and can wrap around the thousands of repositories that are out there. By these I mean of course DOIs, but DOIs combined with standardised affiliations (via the Research Organisation Registry), funding information (Funder Registry), research classification (as difficult as global research is to classify), and person identifiers (i.e., ORCID) [6]. Note that here I am omitting efforts that seek to connect communities and knowledge in parallel with this formal, identifier-rich network. For example: the Confederation of Open Access Repositories (COAR) Notify Project, that aims to link objects and resources using a “standard, interoperable, and *decentralised* approach” (emphasis added), in recognition of the fact that some communities may not choose to engage with Persistent Identifier (PID) or Digital Object (DOI) services (for different reasons) [7].

Each of these services is maintained by a different entity: ORCIDs by ORCID, DOIs by DataCite and CrossRef; Funder Registry by CrossRef; and the Research Organization Registry (ROR) by the California Digital Library (CDL) with DataCite and CrossRef (with multiple types of research classification held by multiple entities).

These strike me as the beginning—almost a pilot for proving how information can travel—because it is the more descriptive information about an individual dataset, article, protocol, that will help it reach the right hands. I will say more about that below.

Dryad engages each of these components of the global research infrastructure: ensuring that individual data publications are elevated to the global scene and available for discovery. While keeping researcher interests at the forefront of our priorities and keeping the process as easy as possible, we do require an ORCID for every submitting author plus affiliation, funding, and field of science for every publication. In these cases, our collaborations help us:

- achieve common purpose for connecting open research objects;
- build shared vision for an open and interconnected system for discovery;
- limit redundancy by relying on one another to provide different functions;
- maximise our respective strengths in specific ontologies and indices;
- increase the impact of what we all do.

4. Two infrastructures

At the same time, those discrete systems with broad intent—including the thousands of repositories that have been built to serve specific needs for disciplines, data, or other resources—are embedded in the global fabric of scholarly communication and we must have bridges.

There are challenges to this, of course. One is the capacity for older, deeply-embedded and widely-used systems to update themselves in order to accommodate emerging Persistent Identifiers (PIDs) and standards - a well-known and significant challenge for all large, multi-tenant systems [8]. Another is how to reconcile divergent metadata models that have each been developed with distinct purpose - whether that purpose has been to serve the needs of a distinct discipline or a specific institution.

I am optimistic about the ways in which Dryad has chosen to approach this - in collaboration with academic and research institutions and repositories.

First, my predecessors co-created with members of the academic and research library community a program for mutual support that leverages both Dryad's strengths (in curated research data and traction with the research community) and the institutions' existing investments in repository infrastructure [9]. Through our Institutional Membership Program, libraries and repositories benefit from the flow of data from their affiliated researchers into Dryad by tapping into our Application Programming Interface (API) and replicating that data or the metadata in their own system. Researchers carry on as normal, publishing or sharing data according to their normal habits, and the institution is able to grow its corpus and sense the magnitude of its funded outputs.

Canada's Federal Research Data Repository (FRDR) is one group to take advantage of this invitation to collaborate. Working through our API and using ROR to hone affiliations, they isolated data generated by Canadian researchers and began to ingest it. In this way, we are able to combine our respective strengths and interests for mutual benefit [10]. Our collaborations with existing repositories help us all to:

- achieve common purpose for data sharing and preservation by concerned investors;
- build shared vision for open research and demonstrating institutional impact;
- limit redundant effort by channelling data from one portal to another;
- maximise our respective strengths in repository operations and community engagement;
- leverage shared resources;
- increase the impact of what we do.

Second: we are collaborating with Metadata Game Changers and CEDAR, with support from the U.S. National Science Foundation EAGER program, to devise limited-scope metadata templates that will help us understand more about the data submitted to Dryad [11]. It is important that data pertaining to any specific field of research that holds its own repository be made available there, in order that researchers who rely on that subject repository have access to everything that may benefit them. We're hopeful that by collecting a little more information about data submitted to Dryad we will have the information that we need to create bridges to other repositories and at the same time increase the discoverability and use of the data on our platform.

If we are successful - and there is a lot of work to be done first - the prospective collaboration with specialist repositories would help us both achieve our ambitions for sharing more research data and making it available for reuse, limit redundant effort and cost, and maximise our respective strengths.

5. Service provision

The last thing I would like to highlight is the power of collaborating on service provision. Dryad's focus on publishing data and not other outcomes, and offering high-level rather than specialised curation opens, the potential for collaborating on service provision with other entities to achieve all we hope to achieve. In particular, I will describe our work with Zenodo (see <https://zenodo.org>), the CERN-supported generalist platform, and the U.S.-based Data Curation Network.

Through our work together with Zenodo, we have built a bridge between the European and North American researcher communities, helping them to engage with both our services in tandem for the benefit of their research. Through our partnership, Dryad taps into Zenodo's traction in software sharing and support for software licensing, and Zenodo grows its corpus of curated and FAIR research data. When an author submits to Dryad, we facilitate the diversion of software and supplementary information files to Zenodo—from within the Dryad platform—and publish the data on Dryad along with citable links to the other files.

The ambitions for this partnership are high. It is a powerful example of how complementary open source and community-driven initiatives are collaborating to combine strengths, enhance service to authors, and connect research objects distributed across the open global network.

Dryad's collaborations with data curators have similar power for different reasons. In this case I mean to describe the curation of research data to ensure that necessary metadata is attached to facilitate discovery and enable reuse. Our institutional partners in particular have a vested interest in supporting faculty with this service and collaborate with Dryad and other institutions in a multidirectional network for the exchange and curation of data that is the Data Curation Network (DCN). Through the DCN, individual campuses and Dryad build and leverage the collective competencies in curation of data across subjects and disciplines [12], again expanding each of our individual capacities far beyond the walls of each of our organisations.

6. Conclusion

Through the partnerships that I have described above Dryad and so many other aligned organisations are combining their positions in the space and special skill sets to achieve common purpose, build shared vision, limit redundancy, maximise strengths, leverage shared resources, and increase impact.

Our partnerships with publishing organisations, institutional repositories, and PID overseers are mature and working for mutual benefit.

Our connections with Zenodo and the DCN are operational, address the emerging needs and interests of our stakeholders, and promise to have a significant and enduring impact. As more data flows between us, we will begin to evidence this.

The potential for additional metadata collection through the Metadata Game Changers and CEDAR collaboration is still in front of us, but so exciting.

Looking forward, there are a number of directions we can take, but I am personally really interested in two things, both related to scale. The first is how organisations such as ours might combine operation-level resources, such as marketing, finance, grant management, community engagement, and Information

Technology (IT). Where it is not tied to our core functions and/or value delivery, might we invest in shared resources with the right expertise—in a way that costs less than having dedicated headcount?

The second is how projects such as Dryad are effectively represented to potential investors - the institutions and funders - that so many of us call on for support. It is a recognised challenge [13] both for smaller organisations such as Dryad to reach and engage with the whole potential audience for our service, and for institutions and funders to vet and negotiate agreements with so many different initiatives. The Global Sustainability Coalition for Open Access Services (SCOSS) which vets prospective investments for their members) [14] and Invest in Open Infrastructure (IOI) that is building a catalogue of services that meet specific criteria) [15] will help with this, and with addressing the academic community's concern around the consolidation of "properties" such as Dryad by organisations whose interests are not aligned (op cit). I am encouraged, but also wondering if advocacy and representation to institutions and funders is still a significant gap—and an opportunity—for Dryad and like-minded organisations to further collaborate.

Community-driven initiatives have great power. Buoyed by their support and intellectual investment, we are fundamentally connected to the needs and values of the academic community. Ensuring our healthy operation and longevity in a highly-competitive environment is a further question, though, and requires a different way of thinking. I hope I have successfully presented how co-opetition among organisations that are aligned in values and purpose can help community-driven initiatives to: Achieve common purpose; Build shared vision; Limit redundancy; Maximise strengths; Leverage shared resources; and Increase impact.

About the Author

Jennifer Gibson joined Dryad as Executive Director in October 2021. Since 2005, she has worked with scientists, funders, publishers, libraries, developers, and others to explore fresh paths toward accelerating discovery through open research communication and open-technology innovation. Prior to Dryad, Jennifer was Head of Open Research Communication for eLife, a non-profit and research funder-backed initiative to transform science publishing. She is Chair of the Board of Directors for OASPA (2020–2022) and a former member of the board for FORCE11 (2018–2020). E-mail: jgibson@datadryad.org.

References

- [1] A. Brandenburger and B. Nalebuff, *Co-opetition*. Doubleday, New York, 1996, cited in Phills J.A., *Integrating Mission and Strategy for Non-Profit Organisations*. Oxford University Press, 2005, p. 87.
- [2] A. Brandenburger and B. Nalebuff, *Co-opetition*. Doubleday, New York, 1996, cited in Phills J.A., *Integrating Mission and Strategy for Non-Profit Organisations*. Oxford University Press, 2005, p. 87.
- [3] 'Liquid' is a term I love to borrow from Jean-Claude Burgelman, past advisor on open access to the European Commission. He recently invoked it here: Pascu C., Burgelman J., * (2022). Open data: The building block of 21st century (open) science. *Data & Policy*, 4, E15. doi:[10.1017/dap.2022.7](https://doi.org/10.1017/dap.2022.7).
- [4] I learned a lot from James Tumwine, Professor Emeritus, Makerere University School of Medicine and Professor, Kabale University School of Medicine, Uganda, who presented at the OASPA 2021 conference: <https://copyright.wistia.com/medias/x85bpin5ro>, accessed August 28, 2022.
- [5] Beginning with the Joint Data Archiving Policy (2011): <https://datadryad.org/docs/JointDataArchivingPolicy.pdf>, accessed August 28, 2022.
- [6] My recent thinking is heavily influenced by Ted Habermann, at Metadata Game Changers, through personal communication and his recent talk "Domain Repositories Enriching the Global Research Infrastructure," presented to the EDI meeting and posted online in April 2022: <https://www.youtube.com/watch?app=desktop&v=Dq8ew4Iad60>, accessed August 28, 2022.
- [7] <https://www.coar-repositories.org/notify/>, accessed August 28, 2022.

- [8] Also informed by recent personal communication with Ted Habermann, May 2022.
- [9] <https://blog.datadryad.org/2022/01/27/the-dryad-collaboration-an-invitation-to-libraries/>, accessed August 28, 2022.
- [10] https://www.dal.ca/dept/research-services/opportunities/opportunities-announcements-news/news/2021/07/19/the_federated_research_data_repository_fldr_.html, accessed August 28, 2022.
- [11] <https://blog.datadryad.org/2021/09/29/metadata-game-changers-stanford-university-and-dryad-receive-nsf-funding-to-improve-metadata-quality-and-connect-repositories/>, accessed August 28, 2022.
- [12] <https://datacuriationnetwork.org/>, accessed August 28, 2022.
- [13] <https://sparcopen.org/news/2019/investing-in-open-scholarly-infrastructure-a-community-opportunity/>, accessed August 28, 2022.
- [14] <https://scoss.org/>, accessed August 28, 2022.
- [15] <https://investinopen.org/>, accessed August 28, 2022.