



Welcome!

Emmy Tsang

Why and How to Invest in **Open Research Infrastructure?**













Funded by the Federal Ministry of Education and Research (BMBF) and the state of Berlin under the Excellence Strategy

of the Federal Government and the Länder

Why and how to invest in open infrastructure?

Emmy Tsang | Invest in Open Infrastructure | 13 October 2022

Outline

In the next 30-ish minutes

- What is Invest in Open Infrastructure (IOI)?
- Why change the ways we invest in research infrastructure?
- How to invest in open infrastructure?

About IOI

Increase investment in and adoption of open infrastructure to further equitable access and participation to research

Why change the ways we invest in research infrastructure?

Defining infra



infrastructure (noun):

the systems, protocols, software that research & scholarship *rely on*.

See also: <u>Preliminary investigation: Defining Open Scholarly Infrastructure</u> Dictionary by karina from the Noun Project

The ecosystem is sick.

The current system for scholarly research and knowledge production is geared towards commoditization, is cost-prohibitive, skewed towards profit generation (of higher education and commercial players), and privileging the well resourced.

This is devaluing and underserving the global scholarly research community, and perpetuating a system that rewards prestige, exclusion, and walled gardens.

Why Clarivate Is Buying **ProQuest For \$5.3 Billion**







Roger C. Schonfeld orschon · May 17

Replying to @rschon

"An historic day for both organizations that will create a global leader...The combination will enrich our value chain across the research lifecycle." -- Jerre Stead, CEO of Clarivate

↑7 6



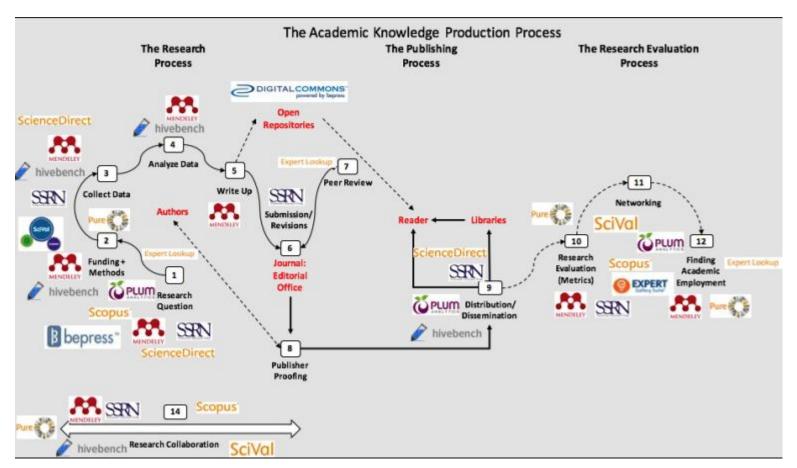
Roger C. Schonfeld @ @rschon · May 17

"Enterprise software is the fastest growing library market." - Stead

Roger C. Schonfeld @ @rschon · May 17

"Significant cross-selling opportunities." -Stead

https://pulse2.com/why-clarivate-is-buying-proquest-for-5-3-billion/ https://scholarlykitchen.sspnet.org/2021/05/18/clarivate-to-acquire-proquest/ https://twitter.com/rschon/status/1394262325411131393



Article: http://knowledgegap.org/index.php/sub-projects/rent-seeking-and-financialization-of-the-academic-publishing-industry/preliminary-findings/Elsevier Acquisitions & Products: https://docs.google.com/spreadsheets/d/1kP22W7kKd6hcuvipzk7xPFk--ZU4kJhAM8CHXhvgZf0/edit#gid=691343153

"Privacy has been getting shortchanged in the education space for years before COVID-19," Fitzgerald said. "What's happening now with everyone having first-hand experiences with software at a level they never have is that areas that have been ignored are getting the attention quickly."

https://iapp.org/news/a/shift-to-online-learning-ignites-student-privacy-concerns/

Software that monitors students during tests perpetuates inequality and violates their privacy

The coronavirus pandemic created a surge in demand for exam proctoring tools. Here's why universities should stop using them.

https://www.technologyreview.com/2020/08/07/1006132/software-algorithms-proctoring-online-tests-ai-ethics/

LIBRARIANSHIP AT THE CROSSROADS OF ICE SURVEILLANCE



These corporations are no longer the publishers that librarians are used to dealing with, the kind that focus on particular data types (academic journals, scientific data, government records, and other staples of academic, public, and specialized libraries). Instead, the companies are data barons, sweeping up broad swaths of data to repackage and sell.



Sarah Lamdan, Professor of Law, CUNY Law School

Infrastructure is not neutral. Our technical choices are not neutral.

We need to reexamine our baseline assumptions that all tech platforms are inherently good and designed to increase access, freedoms, and opportunity.

What are we working to enable?

1

Openness engrained at the operational and design level

Solutions should be open source. Content & data openly licensed. Transparent pricing. Clear governance and guidance on how decisions are made. 2

An enabler for communities to deliver collective benefits

Infrastructure should be maximally accessible for equitable participation, adoption, and usage without restrictions, and governed by community needs.

3

Facilitating the creation & dissemination of open knowledge

For our focus, infrastructure should be contributing to the production and sharing of research and scholarship with minimal restriction.

How to invest in open infrastructure?

Points of friction

- **Time.** Invested vs available, balancing immediacy.
- Prioritization.
 - Near term gain vs. long-term investments.
 - Who's values get applied? (Institution, Dept, budget owner, developer, consortia / society?)
- "Local first" development.
- Staffing & labor to develop, maintain, sustain.
- Favoring usual suspects / "bigger" players.
- Resourcing tradeoffs/influence.
- Perceived resource scarcity (+ the effects of that).

Our approach

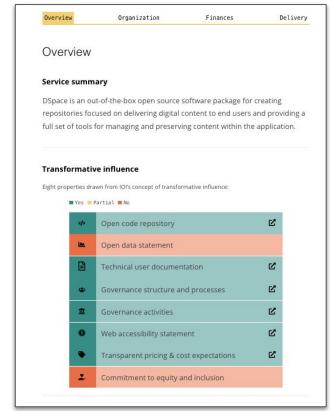
- We employ a research-driven approach to guide strategies and action designed to increase adoption of and investment in open infrastructure.
- We provide resources and analysis to help funders and budget holders assess, evaluate, and make investment decisions about open infrastructure.
- We pilot solutions and coordinate stakeholders to increase the sustainability of the sector, to further a shared agenda for making open infrastructure the default in research.

Costs & characteristics of Ol

- Address information asymmetries
- Foster greater understanding of open infrastructure services
- Cultivate a deeper awareness of how the services are provided
- **Prototype** a means of standardizing key pieces of information
- Meet the needs of various stakeholders (funders, providers, and users)

Costs & characteristics of Ol

Catalog of Open Infrastructure Services (COIs) #Research #Projects The **Catalog of Open Infrastructure Services (COIs)** is a step towards addressing the information asymmetries that exist in understanding and assessing open infrastructure projects. This effort is designed to model a means of standardizing information about core open infrastructure services for decision makers and members of the community.



Data collection:

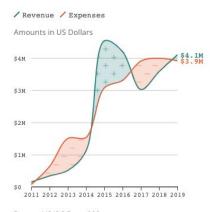
- Provider and funder websites
- Annual Reports
- US Internal Revenue Service
 Form 990 data
- Surveys and interviews with service providers
- Evidence from documented practices

Dimension	Indicator	Evidence
Reliable technologies	Open code repository	One or multiple open code repositories must be available and accessible to the public.
Reliable technologies	Open data statement	If applicable, a statement indicating licensing and usage rights of serviced data in addition to clarifications of how user data is handled.
Reliable technologies	Technical user documentation	Technical documentation outlining not only how to use the service but also providing enough insights to fully reproduce the product.
Trustworthy organizations	Governance structure and processes	Examples include pages or content on project websites describing board structure or posted bylaws.
Trustworthy organizations	Governance activities	Examples include public summaries or minutes of board meetings (or other governance activities) in any form.
Equitable & inclusive services	Web accessibility statement	Indicators of a commitment to web accessibility standards such as a public statement or dedicated working groups.
Equitable & inclusive services	Transparent pricing & cost expectations	In addition to a transparent pricing schema potential users should be able to determine estimated costs.
Equitable & inclusive services	Commitment to equity and inclusion	Indicators of a critical and self-reflected commitment to equity and inclusion in the form of public statements or working groups. This also applies to projects with inherently equitable missions as social justice concerns many modes of discrimation and privilege as well as internal organizational practices.

Questions:

- What key information do decision makers look for to guide investment? What's missing?
- What surfaced in this work that may not be obvious to the broader community that feels important to share?
- What trends & observations do we notice? Have these been reported before? Why/why not?

Total revenue and expenses



Source: US IRS Form 990

Amounts in US Dollars \$3M \$2M

2011 2012 2013 2014 2015 2016 2017 2018 2019

Investment income vs. other

/ Investment income / Other revenue

Total assets and liabilities

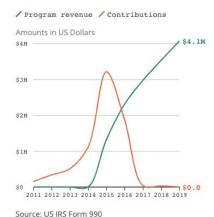
/ Assets / Liabilities / Net assets

Source: US IRS Form 990

revenue

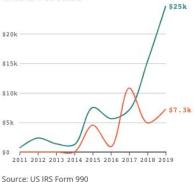
-\$1M

Contributions, gifts and grants vs. program revenue



\$20k

Amounts in US Dollars



What is good community governance?

Governance in Nonprofit Organizations: A Literature Review

Authors

Ravin Cline, University of Texas, Dallas, <u>0000-0001-8216-6861</u>
Tania Hernandez Ortiz, Invest in Open Infrastructure, <u>0000-0002-8046-6024</u>
Richard Dunks, Invest in Open Infrastructure, <u>0000-0001-6999-7935</u>

Introduction

The aim of this work is to explore the importance of governance within a nonprofit organization, the key features commonly ascribed to effective governance, and the impact of governance on the operations of a nonprofit organization as represented in a selection of the available scholarly literature from the nonprofit management and evaluation discipline.

Community Governance In Scholarly Communication

Author

Dr. Samuel A. Moore, for Invest in Open Infrastructure, 0000-0002-0504-2897

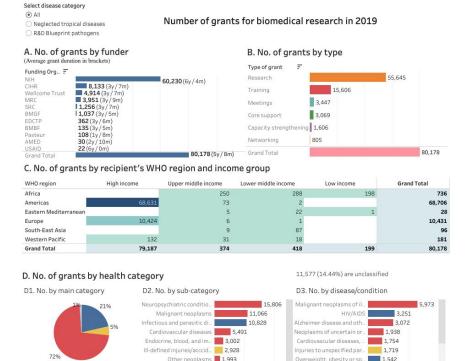
Summary

Part 1 of this document provides an overview of governance in the context of scholarly communication and open research infrastructures. With reference to the theoretical literature on governance in scholarly communication and open infrastructure, the piece illustrates some of the key considerations that relate to governance in this space, the importance of good governance and a discussion of some of the work that IOI is doing to improve our understanding of governance of open infrastructures. Part 2 of this document proposes a framework for community governance based on the elements discussed in Part 1, with relevant examples to demonstrate important pieces

Funding trends

What we're building towards

- Where is investment currently flowing?
- What gaps and concentrations are there in support that may inform our recommendations?
- What trends & observations do we notice? Have these been reported before? Why/why not?



WHO Global Observatory on Health Research & Development

Skin diseases 697

Nutritional deficiency 1,898

Sense organ diseases 1,689

Musculoskeletal diseases 1.630

Respiratory diseases 1.621

Genitourinary diseases 1,607

Congenital anomalies 1,394
Digestive diseases 1,354

Respiratory infections | 944

Health related- medical br.. | 727

Diabetes mellitus 1,732

Drug use disorders 1,409

Tuberculosis 846

Neurodevelopmental diso.. 1.180

Diabetes mellitus, unspeci.. 1.154

Cerebrovascular diseases 841

Unipolar depressive disor.. 786

Certain disorders of infect.

Disorders of the nervous s.

Noncommunicable

maternal, perinata

diseases

Others

Funding models exploration

What we're exploring and building towards

- Set values-aligned conditions for ongoing investment
- Explore investment that provides additional flexibility for projects
- Explore mechanisms that defy traditional timelines for investment of 3-5 years
- Catalyze and incentivize collective investment from a variety of supporters
- Ways to augment and build on / with others in the space

Convening stakeholders

Invest in Open Infrastructure (IOI) Funders Summit Hands on Deep dives 31 October - 4 November 2022 | Online Day 1 (Oct 31): Day 2 (Nov 1): Day 3-4 (Nov 2-3): Day 5 (Nov 4): Why collective, transparent, How to fund? What to fund? Next steps and a roadmap and evidence-based funding forward mechanisms for open infrastructure? Opening session: Funding Funding models for open Hands on: Designing the Closing session: Towards Hands on: Setting up the A funding framework for collective fund pilot trends in the open infrastructure collective, evidence-based open infrastructure collective fund pilot Nov 3, 21:00-22:00 GMT OR infrastructure ecosystem Nov 1, 15:00-16:30 GMT Nov 1, 17:30-18:30 GMT OR Nov 2, 15:00-16:30 GMT investments in open Oct 31, 15:00-16:30 GMT Nov 4, 08:00-09:00 GMT infrastructure Nov 2, 08:00-09:00 GMT Nov 4, 15:00-16:30 GMT Mapping funding for open Supporting viable and Defining criticality and Collective funding infrastructure: Challenges sustainable open at-risk-ness of mechanisms: Next steps and opportunities infrastructure for preprints infrastructure Nov 4, 18:00-19:30 GMT Oct 31, 17:00-18:00 GMT Nov 2, 17:00-18:00 GMT Nov 3, 15:00-16:00 GMT Catalog of Open Technical reliability and Good community Infrastructure Services governance: insights from security of open (COIs): Overview, recent the nonprofit and scholarly infrastructure work, and next steps communication sectors Nov 3, 17:00-18:00 GMT Oct 31, 21:00-22:00 GMT Nov 2, 18:30-19:30 GMT Financial health and risks Furthering the adoption for nonprofits in research of open infrastructure Nov 2, 21:00-22:00 GMT Nov 3, 19:30-20:30 GMT

Equitable access ... for whom?

When we speak about the value of "shared infrastructure" the "for whom?" often favors the well-resourced, Western institutions and scholars.

Systemic inequities create multi-level barriers to participation, knowledge exchange, resourcing, and equitable access, countering our shared community values. What voices are missing from the discussion?

Open research & access to knowledge requires open infrastructure.

Equitable and accessible participation in knowledge production and dissemination requires that our infrastructure is similarly designed, and anchored in community values and governance.

That means shifting reliance from players misaligned with core values of the community, transparency, and collective responsibility.

Thank you!

