

Principles of Open Scholarly Infrastructure (POSI)

김기홍 (과편협, 아주대)

Background and motivation

- 2015년에 Geoffrey Bilder, Jennifer Lin, Cameron Neylon이 제안
“Everything we have gained by opening content and data will be under threat if we allow the enclosure of scholarly infrastructures. We propose a set of principles by which Open Infrastructures to support the research community could be run and sustained.”
- Bilder G, Lin J, Neylon C (2020), The Principles of Open Scholarly Infrastructure <https://doi.org/10.24343/C34W2H>
- 대상: Infrastructure that supports any kind of scholarly activities including infrastructure services that support the management and publication of research data and metadata, scholarly research archives, identifier schemes, etc.
- 16 principles on governance, sustainability, and insurance
- Crossref를 포함한 11개 국제 기관들이 참여 선언

Governance

7 principles on how the infrastructure is governed and managed

Stakeholder led governance

Transparency

Plan across the entire lifecycle of the infrastructure, including its wind-down

- **Coverage across the research enterprise**

Research transcends disciplines, geography, institutions and stakeholders. The infrastructure that supports it needs to do the same.

- **Stakeholder Governed**

Board-governed organization drawn from the stakeholder community

Decisions driven by community consensus and consideration of different interests

Governance

- **Non-discriminatory membership**

Any stakeholder group may express an interest and should be welcome. The process of representation in day to day governance must be inclusive with governance that reflects the demographics of the membership.

- **Transparent operations**

Achieving trust in the selection of representatives to governance groups will be best achieved through transparent processes and operations.

- **Cannot lobby**

The community, not infrastructure organizations, should collectively drive regulatory change. An infrastructure organization's role is to provide a base for others to work on and should depend on its community to support the creation of a legislative environment that affects it.

Governance

- **Living will**

Publicly describe a plan addressing the condition under which an organization would be wound down, how this would happen, and how any ongoing assets could be archived and preserved when passed to a successor organization. Any such organization would need to honor this same set of principles.

- **Formal incentives to fulfil mission & wind-down**

Infrastructures exist for a specific purpose and that purpose can be simplified or rendered unnecessary by technological or social change. If it is possible the organization (and staff) should have direct incentives to deliver on the mission and wind down.

Sustainability

5 principles on

Revenue generation to align with the mission of the organization

Generate revenue from services rather than data

Generate a surplus and find long-term sources of revenue

- **Time-limited funds are used only for time-limited activities**

Day to day operations should be supported by day to day sustainable revenue sources. Grant dependency makes them fragile and more easily distracted from building core infrastructure.

- **Goal to generate surplus**

To weather economic, social and technological volatility, organizations need financial resources beyond immediate operating costs.

Sustainability

- **Goal to create contingency fund to support operations for 12 months**

Generate a contingency fund that can support a complete, orderly wind down (12 months in most cases). This fund should be separate from those allocated to covering operating risk and investment in development.

- **Mission-consistent revenue generation**

Potential revenue sources should be considered for consistency with the organizational mission and not run counter to the aims of the organization.

- **Revenue based on services, not data**

Data related to the running of the research enterprise should be a **community property**. Appropriate revenue sources might include value-added services, consulting, API Service Level Agreements or membership fees.

Insurance

4 principles on openness

Open source, open data

IP issues

Ensure that if the infrastructure fails, its assets can be reused.

- **Open source**

All software required to run the infrastructure should be available under an open source license. This does not include other software that may be involved with running the organization.

- **Open data (within constraints of privacy laws)**

For an infrastructure to be forked, it will be necessary to replicate all relevant data. The CC0 waiver is best practice in making data legally available. Privacy and data protection laws will limit the extent to which this is possible

Insurance

- **Available data (within constraints of privacy laws)**

Underlying data should be made easily available via periodic data dumps.

- **Patent non-assertion**

The organization should commit to a patent non-assertion covenant. The organization may obtain patents to protect its own operations, but not use them to prevent the community from replicating the infrastructure.

참여 기관

- Crossref
- DataCite
- Dryad
- Europe PMC
- JOSS (Journal of Open Source Software)
- OA Switchboard
- OpenAIRE (Open access infrastructure for research in Europe)
- OpenCitations
- OurResearch
- ROR (Research Organization Registry)
- Society

자체 평가

		Crossref	Dryad	ROR	JOSS	OurResearch	OpenCitations	DataCite	OA Switchboard	Society	Europe PMC
	Date	2020-12-02	2020-12-08	2020-12-16	2021-02-14	2021-06-10	2021-08-09	2021-08-30	2021-10-07	2021-12-22	2022-02-21
	Source	Assessment									
Category	Principle										
Governance	Coverage across the research enterprise	G	G	G	A	G	G	G	G	A	G
	Stakeholder Governed	R	G	G	A	A	R	G	A	A	G
	Non-discriminatory membership	G	G	G	G	A	G	G	G	A	G
	Transparent operations	G	G	A	G	G	G	G	G	G	G
	Cannot lobby	G	G	G	G	G	G	G	G	G	G
	Living will	G	G	G	A	G	G	A	G	A	A
	Formal incentives to fulfil mission & wind-down	G	G	G	G	G	A	G	G	G	G
Sustainability	Time-limited funds are used only for time-limited activities	G	G	G	G	G	G	G	A	G	A
	Goal to generate surplus	G	A	A	A	G	A	G	G	A	R
	Goal to create contingency fund to support operations for 12 months	A	A	A	A	G	A	A	G	G	R
	Mission-consistent revenue generation	G	G	G	G	G	G	G	G	G	G
	Revenue based on services, not data	G	G	G	G	G	G	G	G	G	G
Insurance	Open source	A	G	G	G	G	G	G	A	G	A
	Open data (within constraints of privacy laws)	A	G	G	G	G	G	G	?	G	G
	Available data (within constraints of privacy laws)	G	G	G	G	G	G	G	?	G	G
	Patent non-assertion	A	G	G	G	G	G	G		G	A

Some comments

- A lack of community engagement lead to a lack of community control, and the locking up of community resources.
- Underlying data that is generated by the actions of the research community should be a community resource.
- The need for forkability implies centralization of control.