Consider incorporating TK guidelines regarding access to indigenous collections:L

https://localcontexts.org/labels/traditional-knowledge-labels/

In the context of TK, the CARE principles (as an extension to FAIR) might also be relevant. See e.g.

* Global Indigenous Data Alliance. (2019, September). CARE Principles of Indigenous Data Governance. Global Indigenous Data Alliance.<https://www.gida-global.org/care>
* Ruckstuhl, K. (2022). Trust in Scholarly Communications and Infrastructure: Indigenous Data Sovereignty. Frontiers in Research Metrics and Analytics, 6. <https://doi.org/10.3389/frma.2021.752336>



Defining Open Scholarly Infrastructure

Preliminary Investigation

20 April 2022



***Defining Open Scholarly Infrastructure***

***Report by:***

Invest in Open Infrastructure

<https://investinopen.org>

***Contact:***

Richard Dunks

Director of Research and Strategy, Invest in Open Infrastructure

[richard@investinopen.org](mailto:richard@investinopen.org)

***Report authors:***

Saman Goudarzi (Independent Contractor)

<https://orcid.org/0000-0002-4403-8532>

Report DOI: <doi>

Report dated: 20 April 2022

*This report is made available under a* [*Creative Commons Attribution 4.0 International License*](https://creativecommons.org/licenses/by/4.0/)*. Users are free to share, remix, and adapt this work. (Please attribute Invest in Open Infrastructure in any derivative work.)*



Funding for this research was provided by:

[Arcadia Fund](https://www.arcadiafund.org.uk/)

# 

# 

# 

# 0. Executive summary

[To be added, ~200 words]

# 1. Background

Invest in Open Infrastructure (IOI) was founded on the premise that open, community-owned infrastructure is necessary for scholarly research to thrive.

We are certainly not alone in this sentiment. Numerous organisations across the scholarly research ecosystem, including the [Ford Foundation](https://www.fordfoundation.org/campaigns/critical-digital-infrastructure-research/), the [Scholarly Publishing and Academic Resources Coalition (SPARC](https://sparcopen.org/our-work/community-owned-infrastructure/)), the [Global Sustainability Coalition for Open Science Services (SCOSS)](https://scoss.org/help-sustain-open-infra/), and the [Association of Research Libraries (ARL)](https://www.arl.org/principles-and-practices/), as well as many others have recognized the critical role open infrastructures play for the communities they serve — contributing to more [equitable](https://omidyar.com/responsible-technology-2/), [accessible](https://mellon.org/programs/public-knowledge/), and [resilient](https://www.siegelendowment.org/our-interest-areas/infrastructure/) knowledge practices.

IOI strives to build on the efforts of others working to improve funding and resourcing for the open infrastructure on which scholarly research relies. One way we hope to achieve this is by pushing the limits of our current understanding about infrastructure in scientific research and scholarly communication. A deeper understanding will have significant implications for how we collectively engage with and support the organisations providing services that make research and scholarship possible. This report represents the beginning of an iterative process for IOI in deepening its understanding on this topic that we look forward to developing and refining as our work progresses.

# 2. What is the problem?

To ensure its stakeholders understand the essence and scope of IOI’s work, IOI produced in late 2019 a [working definition of both “infrastructure” and “open infrastructure”](https://investinopen.org/about/#defining-open-infrastructure-) in the context of the scholarly research ecosystem. It reads as follows:

*By “infrastructure” we mean the sets of services, protocols, standards and software that the academic ecosystem needs in order to perform its functions throughout the research lifecycle — from the earliest phases of research, collaboration and experimentation through data collection and storage, data organisation, data analysis and computation, authorship, submission, review and annotation, copyediting, publishing, archiving, citation, discovery and more.*

*“Open infrastructure” is the narrower sets of services, protocols, standards and software that can empower communities to collectively build the systems and infrastructures that deliver new improved collective benefits without restrictions, and for a healthy global interrelated infrastructure system.*

Unfortunately, this two-part definition:

1. Is unsystematic (both in breadth and depth) in delineating the kinds of functions and activities scholarly infrastructure needs to support,
2. Is vague in defining the values and goals of “open infrastructure”, and,
3. Does not easily translate to a robust theoretical framework that can provide structure and support to IOI's varying projects.

# 3. What is IOI doing about this problem?

This preliminary report outlines IOI’s initial attempt towards a more sophisticated framework for understanding open infrastructure for research and scholarship. Such a framework will:

1. Facilitate the development of a systematic and standardised *definition* of open infrastructure that mediates the gaps in our current definition while effectively delineating and communicating the essence and scope of our work.
2. Inform the design of our research methods for examining open infrastructure across our varying projects[[1]](#footnote-0) — establishing proven constructs, concepts, and approaches on which we can draw.

For this report, we examined [a body of literature](https://www.zotero.org/groups/4377072/invest_in_open/collections/SDRYGPAI) that includes works across the fields of anthropology, scholarly communications, international development studies, science and technology studies, and infrastructure studies. We aimed to balance foundational understandings of open infrastructure with both recent and peripheral discussions on the topic.

We reviewed, categorised, and annotated this compiled literature in order to develop:

1. An initial assessment of the current state of research on the topic of open infrastructure,
2. Recommendations for the ways in which IOI’s working definition of open infrastructure can be strengthened and,
3. Recommendations for future areas of development and further research for better understanding open infrastructure.

This report is not an exhaustive inventory of literature that attempts to engage with theoretical or practical conceptions of open infrastructure across the social sciences. Instead, this report is a work in progress that will gain breadth, depth, and nuance over time. For more detail on future iterations of this literature review, please refer to Section 6.

This report is divided into three sections:

1. First, we review our curated body of literature in order to outline prevailing conceptualizations of open infrastructure in the context of research and scholarship. We provide categories to better understand the ways in which “infrastructure”, “scholarly infrastructure”, and “open scholarly infrastructure” have each been conceptualised and defined.[[2]](#footnote-1)
2. Second, we synthesise this literature, identifying key takeaways for IOI when conducting its own investigations into open infrastructure for research and scholarship.
3. Lastly, we outline recommendations for future areas of development and further research to better understand and best support open infrastructure.

# 4. Key Elements from the Literature

## 4.1. How have others defined “infrastructure”?

In the digital age, the term “infrastructure” has been used to refer to the “constellations of software technologies and systems usually associated with the Internet” (Karsati et al., 2010, p. 382). With the proliferation of work around this topic, terms such as “information infrastructure”, “cyberinfrastructure”, “e-infrastructure”, and “knowledge infrastructure” have been coined to reflect the distinct frameworks and empirical breadth applied within inquiries into this broad phenomenon.

Our intention in this report is to capture all these varying conceptualizations and we use the all-inclusive term “infrastructure” as a catch-all that encompasses this diversity. However, this preliminary investigation engages to a greater extent with the literature on “information infrastructure” as we wanted to prioritise theoretical approaches that considered both the social and technical dimensions of infrastructure.

In the literature we reviewed, definitions of “infrastructure” (summarised in Table 1, below) often frame the concept as a network (see especially Larkin, 2013). Infrastructure is described as consisting of disparate entities — both technical (hardware and software) and social (practices, norms, and structures) — that as an ensemble, facilitate the linking and/or movement of ideas, signals, objects, and people (Larkin, 2013).

| **Definition** | **Author(s)** | **Research Discipline** |
| --- | --- | --- |
| Cyberinfrastructure refers to a “layer of enabling hardware, algorithms, software, communications, institutions, and personnel. This layer [provides] an effective and efficient platform for the empowerment of specific communities of researchers to innovate and eventually revolutionize what they do, how they do it, and who participates.” | Atkins et al., 2003 (p. 5) | Cyberinfrastructure |
| e-Infrastructure refers to “in the first instance to designate the physical or material components of [a large] technological system, the advanced electronic networks that make use of the Internet and the Web, as well as, secondarily, the organizational networks that are supported by this system.” | Schroeder, 2007 (p. 2) | e-Infrastructure |
| “Superadded to the term ‘information,’ infrastructure refers loosely to digital facilities and services usually associated with the internet: computational services, help desks, and data repositories to name a few.” | Bowker et al., 2010 (p. 98) | Infrastructure Studies; Information Infrastructure |
| Knowledge infrastructure refers to the ‘‘robust networks of people, artefacts, and institutions that generate, share, and maintain specific knowledge about the human and natural worlds.” | Edwards, 2010  (p. 17) | Infrastructure Studies; Knowledge Infrastructure |
| “Infrastructures are built networks that facilitate the flow of goods, people, or ideas and allow for their exchange over space. As physical forms they shape the nature of a network, the speed and direction of its movement, its temporalities, and its vulnerability to breakdown. They comprise the architecture for circulation, literally providing the undergirding of modern societies, and they generate the ambient environment of everyday life.” | Larkin, 2013  (p. 328) | Anthropology |

### **Table 1:** Definitions of Infrastructure Across Varying Research Disciplines

In these frameworks, infrastructure is described as a “supporter” or “enabler”, “sinking into the background” and becoming visible only when it breaks down (Star & Ruhleder, 1996, p. 112). Because of this tendency to fade into the background, infrastructure can appear unremarkable and unexciting in nature (Karasti & Blomberg, 2018: Star & Ruhleder, 1996).[[3]](#footnote-2)

Science and Technology Studies (STS) embraces this boringness — “foregrounding the truly backstage elements” of the mundane, background practices and unnoticed work (of designers, developers, users, managers, and mediators just to name a few) that facilitate the functioning of infrastructure (Star, 2002, p. 16) (refer to Table 2, below).

| **Definition** | **Author(s)** | **Discipline** |
| --- | --- | --- |
| Infrastructure emerges in relation to organised practices. It “occurs when local practices are afforded by a larger-scale technology, which can then be used in a natural, ready-to-hand fashion.” | Star & Ruhleder, 1996  (p. 114) | Ethnography;  Knowledge Management;  Information Systems |
| “Following Star and Ruhleder (1996), an infrastructure emerges when it reaches beyond a single event on a temporal scale or a single site practice on a spatial scale [...occurring] when here-and-now practices are afforded by temporally extended technology that can be used in an everyday, reliable fashion. Infrastructure becomes transparent when it exists as an accessible, ready-to-hand installed base that enables envisioning future usages.” | Karasti et al., 2010  (p. 400) | Information Architecture; Science and Technology Studies |

### **Table 2:** Conceptualization of Infrastructure as Related to Organisational Practice

The field of STS, therefore, understands infrastructure not just in terms of interdependent components of a network but also in terms of “configurations” of practices and activities (Karsati et al., 2010; Star & Ruhleder, 1996). Popularised by Star & Ruhleder (1996), such a framework shifts understanding of infrastructure from being static and definitive (“what infrastructure is”) to dynamic and relationally configured (“we can’t be definitive about what infrastructure is, but rather in the ways infrastructure emerges”) (Bowker et al, 2010; Karsati et al., 2010; Star & Ruhleder, 1996).

This marks a significant shift towards the study of infrastructure’s specific dimensions and characteristics, resulting in the concept most often being “defined by jotting down a laundry list of characteristics” (Bowker et al., 2010, p. 99) (summarised in Table 3, below).

| **Characteristics of Infrastructure** | **Author(s)** |
| --- | --- |
| The configuration of the following nine (9) dimensions form ‘an infrastructure’:  (1) Embeddedness  (2) Transparency  (3) Reach or scope  (4) Learned as part of membership  (5) Links with conventions of practice  (6) Embodiment of standards  (7) Built on an installed base  (8) Becomes visible upon breakdown  (9) Is fixed in modular increments, not all at once or globally | Star & Ruhleder, 1996 |
| Based on a synthesis of characteristics emerging in prominent literature, infrastructures can be characterised by the following five (5) dimensions:  (1) their profoundly relational quality  (2) their intrinsic (at least partial) invisibility  (3) their connectedness, sometimes described as “scaling”  (4) their emerging and accreting quality of infrastructures  (5) the role of intentionality and intervention in delineating infrastructures | Karasati & Bloomberg, 2018 |
| Infrastructures have a “modular, multi-layered, rough-cut character [...]. [They] are not systems, in the sense of fully coherent, deliberately engineered, end-to-end processes. Rather, infrastructures [...] consist of numerous systems, each with unique origins and goals, which are made to interoperate by means of standards, socket layers, social practices, norms, and individual behaviors that smooth out the connections among them. This adaptive process is continuous, as individual elements change and new ones are introduced — and it is not necessarily always successful.” | Edwards et al., 2013  (p. 5) |
| “Information infrastructures are characterised by openness to number and types of users (no fixed notion of “user”), interconnections of numerous modules/systems (i.e. multiplicity of purposes, agendas, strategies), dynamically evolving portfolios of (an ecosystem of) systems and shaped by an installed base of existing systems and practices (thus restricting the scope of design, as traditionally conceived). Information infrastructures are also typically stretched across space and time: they are shaped and used across many different locales and endure over long periods (decades rather than years).” | Monteiro et al, 2013  (p. 576) |

### **Table 3:** Characteristics of Infrastructure Identified by Various Authors

## 4.2. How have others defined “scholarly infrastructure”?

In this report, we use the term “scholarly infrastructure” to refer to infrastructures that are specifically associated with research and scholarly knowledge production. Across the literature reviewed, this phenomenon has also been referred to as “scholarly communication infrastructure”, “scholarly publishing infrastructure”, “e-research infrastructure”, and “knowledge infrastructure” — each reflecting a distinct framework and empirical breadth.[[4]](#footnote-3)

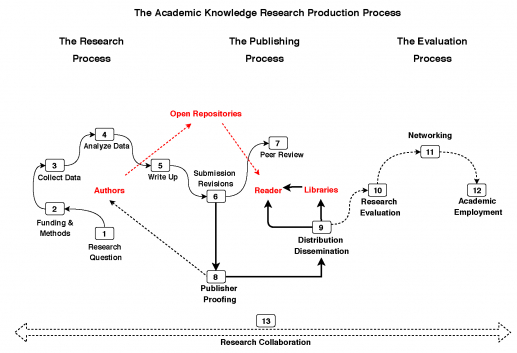
The definitions of scholarly infrastructure that we reviewed (summarised in Table 4, below) utilise the popular metaphors of infrastructure as a “supporter” or “enabler”, describing the phenomenon as the thing upon which the scholarly knowledge production and dissemination — or its particular components — operate. These definitions also often frame scholarly infrastructure as a network, describing it as a system that pulls diverse actors, organisations, and perspectives across domains, disciplines, and geographies together to engage in common practices.

| **Definition** | **Author(s)** |
| --- | --- |
| “e-Research infrastructures are networked systems in which technologies and social institutions are intertwined, [combining] extensive networks of physical artefacts with the organizational capacity to implement and sustain them. [...They are] both: a large technological system insofar as they consist of a number of interdependent social and technical systemic parts (and large because the system covers the globe); and an infrastructure insofar as it supports research.” | Schroeder, 2007 (p. 8) |
| The “fundamental substrate upon which scholarly research operates [...] seamlessly and successfully supporting knowledge work”. | Lagoze et al., 2015  (p. 1054) |
| The “tools and services that underpin the scholarly research life cycle”. | Chen et al., 2019 (p. 1) |
| “Technological infrastructure that runs scholarly communication and publishing.” | Maxwell et al., 2019  (p. 6) |
| “Infrastructure vital to the advancement of the sciences”. | Watkinson & Pitts, 2021 (para. 1) |
| Scholarly communication technologies “includes tools, platforms, and standards that can be locally adopted to support one or more of functions of the lifecycle of scholarly communication, which is conceptualized as including the following activities: creation, evaluation, publication, dissemination, preservation, and reuse.” | SComCaT, n.d. (para. 3) |

### **Table 4:** Definitions of “Scholarly Infrastructure”

Furthermore, we found numerous studies that describe scholarly infrastructure by centering the practices of some individuals and/or organisations within the scholarly knowledge production process (Chen et al., 2019; Kramer & Bosman, 2017; Lewis, 2020). These works mirror understandings of infrastructure common within the field of STS: as emerging in relation to organised practices and connected to particular activities.

For example, Chen et al.’s (2019) investigation into the vertical integration of scholarly infrastructure first outlined the stages of the academic knowledge production process (see Figure 1, below) and then charted varying scholarly tools and services across these stages.



### **Figure 1:** The Academic Research Knowledge Production Lifecycle by Chen et al. (2019)

Furthermore, both Kramer and Bosman (2017) and Lewis’ (2020) works in identifying tools, services, and systems that make up the scholarly infrastructure ecosystem involved a similar approach. These authors first identified a typical workflow for scholarly research and then classified observed tools and services based on their position within the authors’ respective workflows (see Figure 2 and Figure 3 below).

| **Research Phase Number** | **Research Activities (30)** | **Research Phases (7)** |
| --- | --- | --- |
| 0 | project management | preparation |
| 1 | crowdsource/ define research priorities/ ideas/ collaborations |
| 2 | fund get contract |
| 3 | search (lit/data/patents/code) | discovery |
| 4 | get access |
| 5 | get alerts/get (reading) recommendations |
| 6 | reference management |
| 7 | read/view |
| 8 | annotate/tag (during/after reading) |
| 9 | experiment & collect/mine/extract data | analysis |
| 10 | share notebooks/protocols/workflows |
| 11 | analyze |
| 12 | visualize | writing |
| 13 | write (+ code) |
| 14 | cite |
| 15 | translate |
| 16 | archive/share code | publication |
| 17 | archive/share data/video |
| 18 | archive/share publications |
| 19 | archive/share posters |
| 20 | archive/share presentations |
| 21 | present research findings |
| 22 | peer review and comment/recommend (pre-pub) |
| 23 | select journal to submit to |
| 24 | publish |
| 25 | outreach/valorization | outreach |
| 26 | researcher profiling (& social network) |
| 27 | comment | assessment |
| 28 | peer review (post-pub) |
| 29 | measure impact (of output, e.g. article) |
| 30 | assessment (of researcher/research group) |

### **Figure 2:** Research Workflow Phases adapted from Kramer & Bosman (2017)

### **Figure 3:** Research Workflow Developed as Part of the Lewis’ (2020) Bibliographic Scan of Digital Scholarly Communication Infrastructure

## 4.3. How have others defined “open scholarly infrastructure”?

In this report, we use the term “open scholarly infrastructure” to refer to scholarly infrastructure that is owned and/or operated by non-commercial actors, such as academic libraries, consortia, professional associations, communities of practice, independent non-profit organisations, and other research bodies. Across the literature reviewed, this phenomenon has also been referred to as “open infrastructure”, “open science infrastructure”, “open common infrastructure”, “community infrastructure”, and “community-owned infrastructure” — each reflecting a distinct framework and empirical breadth.

| **Definition** | **Author** |
| --- | --- |
| Infrastructure that is “trusted and relied on by the broad community it serves.” | Bilder et al., 2015 (para. 4) |
| “Projects that provide software or services that support open scholarship.” | Lewis et al., 2018 (para. 5) |
| “‘Academy-owned’ and ‘academy governed’ tools, platforms, and services”. | Skinner, 2019 (para. 6) |
| “The structures and services needed for Open Science/Scholarship to operate, e.g. services, protocols, standards and software that the academic ecosystem needs in order to perform its functions during the research lifecycle.” | Ficarra et al, 2020 (p. 10) |
| “Open science infrastructures refer to shared research infrastructures (virtual or physical, including major scientific equipment or sets of instruments, knowledge-based resources such as collections, journals and open access publication platforms, repositories, archives and scientific data, current research information systems, open bibliometrics and scientometrics systems for assessing and analysing scientific domains, open computational and data manipulation service infrastructures that enable collaborative and multidisciplinary data analysis and digital infrastructures) that are needed to support open science and serve the needs of different communities.” | UNESCO, 2021 (p. 12) |
| “In an Open Science context, ‘infrastructure’ — the ‘structures and facilities’ — refers to the scholarly communication resources and services, including software, that we depend upon to enable the scientific and scholarly community to collect, store, organise, access, share, and assess research.” | SCOSS, 2022 (para. 2) |

### **Table 5:** Definitions of “Open Scholarly Infrastructure”

Works that directly explore open scholarly infrastructure or its derivatives often draw from two different frameworks.

Many — often earlier — works examining this phenomenon draw from the extensive literature on “open source”, focusing on the ways in which software, standards, and protocols can promote the accessibility or transparency of infrastructure development, maintenance, and services (Schroeder, 2007; West & O'mahony, 2008). More recently, there has been a shift towards a framework that draws from political economy, instead focusing on the threat of enclosures to community-owned and -operated scholarly infrastructure (see for example, Skinner, 2019; Moore, 2020). As Bilder et al. (2015) note:

*We believe we risk repeating the mistakes of the past, where a lack of community engagement lead[s] to a lack of community control, and the locking up of community resources. In particular our view is that the underlying data that is generated by the actions of the research community should be a community resource – supporting informed decision making for the community as well as providing [a] base for private enterprise to provide value added services.*

While these frameworks draw from different theoretical traditions, both regard infrastructure and its disparate parts beyond commodity production — their valuation existing beyond the logic of the market. In this sense, open scholarly infrastructure functions not just in support of productive practices around scholarship and research, but also in support of social practices and values as well (Helfrich, 2013, as cited in Heinrich Böll Foundation et al., 2013). Some of the explicit values that open scholarly infrastructure has been envisioned to support can be found in Table 6, below.

| **Values for Open Scholarly Infrastructure** | **Author** |
| --- | --- |
| Infrastructure that is characterised by “unrestricted access and use, being free of charge to users, and using non-exclusionary (open) standards.” | Schroeder, 2007 (p. 2) |
| Governance   * Coverage across the research enterprise * Stakeholder-governed * Non-discriminatory membership * Transparent operations * Cannot lobby * Living will * Formal incentives to fulfil mission and wind-down   Sustainability   * Time-limited funds are used only for time-limited activities * Goal to generate surplus * Goal to create contingency fund to support operations for 12 months * Mission-consistent revenue generations * Revenue based on services, not data   Insurance   * Open source * Open data (within constraints of privacy laws) * Available data (within constraints of privacy laws) * Patent non-assertion | Bilder et al., 2015 |
| Infrastructure that “deliberately allow[s] for multiple forms of participation amongst a diverse set of actors, and which purposefully acknowledge[s] and seek[s] to redress power relations within a given context.” | Okune et al., 2019 (p. 2) |

### **Table 6:** Values for Open Scholarly Infrastructure

# 5. Key Takeaways for IOI

To better grasp the fundamental characteristics of infrastructure, we synthesised the characteristics/dimensions defined in the literature we reviewed — paying particular attention to overarching themes and points of contention. From this synthesis, we identified practical considerations IOI should make when conducting its own investigations into open infrastructure for research and scholarship.

## **(1) Infrastructure serves a function**

In the conceptualization of infrastructure by Star and Ruhleder (1996), infrastructure emerges in relation to organised practices. When a cook is preparing dinner, for example — washing ingredients and boiling water — the water system emerges as infrastructure. In this regard, infrastructure’s function is to facilitate (or support) the carrying out of different activities, particularly when: (1) these activities require coordination or agreement, (2) these activities are valued as basic “rights”, and (3) these activities are so widespread that it doesn’t make sense to provision infrastructure individually (Helfrich, 2013, as cited in Heinrich Böll Foundation et al., 2013).

However, many authors have also noted the function of infrastructure beyond supporting productive organisational practices. Numerous case studies across anthropology, politics, and sociology, for example, point to the intrinsic link between “infrastructure” and “modernity” (Larkin, 2013). Most of these case studies — examining traditional “brick and mortar” infrastructures — demonstrate how their function is not just related to organisational activities. Infrastructure, in these instances, also embodies political address — promoting larger ideas and sentiments, related to nationalism (Dalakoglou, 2010), post-colonialism (Harvey & Knox, 2012), and communism (Humphrey, 2005).

**Bottom line for IOI:** When considering open infrastructure, IOI should be mindful of not just the productive functions it serves in supporting/facilitating knowledge practices *but also* its aesthetic and ideological functions as well.

Questions to consider should include:

* What organisational and/or political experiences does open infrastructure produce for its users, designers, builders, and maintainers? Are these experiences different from the experiences produced by commercially-run and -operated infrastructure?
* What emotional values or symbolic meanings are tied to open infrastructure?
* How does the design of open infrastructure’s development, maintenance, and operations impact its productive, aesthetic, and ideological functionalities?
* How do the different communities within the open infrastructure landscape understand and rationalise the various functions of open infrastructure?

## **(2) Infrastructure is a socio-technical system rather than a technical product**

There is a “conceptual unruli[ness]” to infrastructure that can be attributed to the fact that it is simultaneously a “thing” and the “relation between things” (Larkin, 2013, p. 329). Infrastructure is often identified as fixed objects and physical structures, but it is in fact its relational characteristic — connecting disparate entities and facilitating flows — that renders it “an infrastructure” (Larkin, 2013; McArthur, 2019; Star & Ruhleder, 1996).

Developing this idea further, a popular approach taken by STS scholars is to emphasise the practices of people in relation to technical structures. With this approach, understandings of infrastructure broaden from mere technical products to also include dynamic configurations of laborers, communities, and organisations as well (Karasti & Blomberg, 2018). Infrastructure, in this sense, is a complex socio-technical system instead. Bowker et al. (2010) therefore urge “movement beyond seeing the social, organizational, and cognitive sitting somehow on top of or beside the wires and gateways of the physical infrastructure. Each layer is driven through with each of these dimensions” (p. 113).

The category of STS literature following Star and Ruhleder’s (1996) framework emphasises, for example, that infrastructure shapes but is simultaneously shaped by “the conventions of a community of practice” (Karasti & Blomberg, 2018; Star & Ruhleder 1996, p. 113). Infrastructure, Star and Ruhleder (1996) suggest, is intricately linked to social practices in complex ways — pointing to the ways in which the cycles of work throughout a day affect and are affected by electric power rates and needs.

In a more recent attempt to apply STS theories to infrastructure challenges in the global South, Furlong (2014) emphasises on the role the interests and power relations of political-economic actors play — in conjunction with everyday organisational practices — in shaping what he refers to as infrastructures in a state of “malfunction” and “disrepair”. He provides the example of local councillors in India acting as private water vendors, “their conflicting economic and political interests inhibit[ing] the improvement of water supply” (p. 145).

**Bottom line for IOI:** When investigating scholarly infrastructure, it would be a mistake to only consider its technical components. Connected to the design, development, and operation of these technical components are larger social processes and labor (community practices, norms, standards, and values) that infrastructure reinforces and replicates. A successful infrastructure is one able to complement and enhance the organisational and cultural practices of its users, designers, builders, and maintainers, not just provide a technical solution to a particular need.

Questions to consider in our work should include:

* What social forms, practices, and institutions are linked to the development, maintenance, and operation of open infrastructure? In what ways do they facilitate, shape, or accompany these infrastructural processes?
* Which technical or social decisions about infrastructure-design, -development, or -enactment contribute to the “success” or “failure” of an open infrastructure?
* What is the relationship between these technical and social decisions (e.g. are they in conflict or in alignment, are they in a dialectic relationship)?
* Do technical changes in open infrastructure cause social changes? If yes, how does this affect power distribution/dynamics within the open infrastructure landscape?

## **(3) Infrastructure is dynamic**

There is no “one way” for infrastructure to be created or formed. Instead, many authors within STS describe infrastructure formation as being characterised by uncertainty, “including the heterogeneous processes of becoming and the associated temporal complexities full of ups and downs, false starts, disconnects, dead ends and failures” (Karasti & Blomberg, 2018, p. 239).

Infrastructure formation is thus incremental and contiguous, emerging over the long-term and affixing onto an already existing world — of tools, practices, and roles — that both enable and limit its form (Edwards et al., 2007; Karasti & Blomberg, 2018; Star & Ruhleder, 1996). This prolonged and iterative process means that infrastructure is continuously emerging: individual elements change while new ones are introduced — with varying levels of success (Edwards et al., 2013).

The concept of “infrastructuring” was developed by STS scholars struggling with conducting empirical work on “emerging infrastructures” (Karasti & Blomberg, 2018; Star & Bowker, 2002). The fact that infrastructure is continuously emerging often makes it difficult (if not impossible) to foresee all the relevant issues that can occur. Infrastructuring demands the production of stability via ongoing resourcing, labor, monitoring, development, and maintenance — ensuring issues are mediated as they are discovered accordingly.

If infrastructures are not resourced and actively maintained then “the ineluctable pull of decay and decline set in and infrastructures enter the long or short spiral into entropy that – if untended – is their natural fate” (Jackson, 2015, para. 4). Infrastructuring, thus, urges scholars to put an emphasis on the often undervalued and usually-neglected ongoing work of infrastructure monitoring, development, and maintenance (Bowker et al., 2010).

**Bottom line for IOI:** IOI should aim to develop sensitivities towards, and tools and methods for investigating the people, labor, and dynamic processes of infrastructure. It should do this in a way that enables robust and systematic analysis while preserving and recognizing the phenomenon’s ongoing, uncertain, and dynamic qualities.

Questions to consider in our work should include:

* What existing people, labor, systems, and processes is open infrastructure built upon? What are people’s relationships with this “base” (e.g. do they embrace the system or want to replace it)?
* In what ways do these workers, systems, and processes impact open infrastructure’s form and function? How do they accommodate for the integration of new technologies, communities, and practices?
* What resources, workers, relations, and processes facilitate the stability of open infrastructure?
* How can IOI best support the dynamic characteristics of open infrastructure while ensuring scarce resources in this space aren’t wasted or poorly utilised?
* In what ways does the processual (and ongoing) design, development, and operation of open infrastructure contribute to its erasure?
* What existing practical or conceptual methods can we use to make open infrastructure ,its dynamic qualities, and the people who design, build, maintain, and support those qualities visible in our work?

## **(4) Infrastructure operates at different scales, both temporally and spatially**

Infrastructure is large in scale, existing as a constellation of interdependent technical and social phenomena dispersed and distributed not just across space but also across time (Star & Ruhleder, 1996). A city-run water treatment plant emerges as infrastructure when connected to a distribution line that can transport treated water to individuals across the city not just in the present but also the future — considerations having been made for its future repair as well as for potential future changes to the larger water system (Karasati et al., 2010). The key to any infrastructure, as such, “is its ability to permit the distribution of action over space and time” (Bowker et al., 2010, p. 103).

Across the STS literature we reviewed it is generally maintained that disparate technical and social phenomena become connected via the embodiment of standards — “plugging into other infrastructures and tools in a standardised fashion” (Star & Ruhleder 1996, p.113).

It is through these standards that infrastructure exists both in the local and the global contexts, where different local practices “are afforded by larger-scale [global] technology, which can then be used in a natural, ready-to-hand fashion” (Star & Ruhleder, 1996, p. 114). It is also through these standards that infrastructure exists in the short-term and long-term — when “here-and-now practices are afforded by temporally extended technology that can be used in an everyday, reliable fashion”, enabling the envisioning of future usages (Karasati et al., 2010, p. 400). It’s the existence of global standards, for example, that allows for water pumping equipment to be manufactured in Germany and sold on the open market to water utilities in Brazil to meet the particular needs of a local population living in the Amazon.

**Bottom line for IOI:** When identifying open infrastructure in its investigations, IOI should put more emphasis on systems that are or have the ability to become “large” in scale, having the potential to work in both the local and global contexts, as well as in short-term and long-term timeframes through the implementation of recognized standards. Furthermore, IOI’s investigations shouldn’t ignore the role of standards and standard-setting in the development and operations of open infrastructure, but seek to encourage the development, adoption, and regular improvement of standards covering operations, governance, service delivery, community engagement, and other key facets of infrastructure design, development, and operation.

Questions to consider in our work should include:

* What relations and processes allow for the ongoing “growth” of open infrastructure over time and space?
* Do existing standards meet the technical and social needs of the research, scholarly, and maintainer communities?
* What additional standards can and should be created to exemplify the values, norms, and practices of the various research,scholarly, and maintainer communities these services support?

## **(5) Infrastructure is not neutral**

Considering the fact that infrastructure is as much a social as it is a technical system (see Key Takeaway 2, above), there are technical as well as cultural, political, economic, labor, and ethical choices made throughout its design, development, and operation (Clarke and Star, 2008; Star, 1999).

The larger an infrastructure is in scale, the more heterogeneous identities, abilities, needs, experiences, and goals it embodies, often leading to conflict and ambiguity (Star, 1999). Deliberate or accidental design choices can have implications depending both on whose interest and worldviews underpins a particular infrastructure, resulting in particular groups being underserved or not served at all by a particular infrastructure (Karasati, 2014; Okune et al., 2019). Numerous case studies examining water systems infrastructures in the Global South, for example, have underscored the importance of historic colonial and contemporary forms of discriminatory planning in affecting access to essential water infrastructure services for marginalized communities (Kooy & Bakker, 2008; Swyngedouw, 2004).

Similarly, in their exploration of the design and use of knowledge infrastructures for equitable participation in open science practices, Okune et al. (2019), emphasise that not all users of an infrastructure may benefit equally from its installation. An infrastructure may in fact “replicate and reinforce the gendered, raced, and other socio-political imbalances that exist within existing systems of knowledge production” (Okune et al., 2019, p. 6).

**Bottom line for IOI:** When investigating scholarly infrastructure, IOI needs to be mindful of the fact that power and authority are intimately distributed across infrastructural entanglements. The open infrastructure IOI investigates and engages with embodies and legitimises particular social, political, and economic values.

Furthermore, IOI needs to be cognizant of the fact that as an actor within the open infrastructure landscape, our own work can potentially be exclusion-generating and power-distributing. As an organisation, we should acknowledge the diversity and uniqueness of infrastructures and their designers, builders, maintainers, and supporters across the scholarly research ecosystem, and advocate for the development and spread of open infrastructures that account for alternative practices, abilities, and ways of knowing. Most importantly, IOI needs to be open to reflecting on its values and practices within this landscape, mindfully course-correcting when necessary.

Questions to consider in our work should include:

* How can IOI determine if an open infrastructure is not inclusive? What are the markers/indicators of bias?
* Whose practices, abilities, and ways of knowing most influence the design, development and operation of open infrastructure? Whose practices, abilities, and ways of knowing are minimised, ignored, extracted, or silenced?
* In what ways does open infrastructure replicate and/or reinforce existing knowledge, labor, and socio-economic inequities?
* ~~In what ways can IOI be an effective ally to communities working on open infrastructures in niches that aren't well served?~~
* How do we ensure all communities are equitably treated in the development and maintenance of open infrastructures?

# 6. Areas for further work

The following are additional questions or areas of research for IOI to consider in understanding open infrastructure:

* We should expand the breadth of this literature review to better include conceptualisations of infrastructure outside of the more prominent “information infrastructure” literature we reviewed here. Potential areas for a more holistic understanding of open infrastructure include: (1) technical frameworks for infrastructure (e.g. cyberinfrastructure), (2) political economic frameworks for infrastructure (e.g. commons enabling infrastructure), and (3) intersectional frameworks for infrastructure (e.g. inclusive infrastructure, feminist infrastructure, labor infrastructure).
* Somewhat pressing, is our need to develop theoretical and practical approaches to our open infrastructure work. IOI should expand on this literature review in depth by further reviewing existing STS approaches to studying and understanding infrastructure and by adapting them for open infrastructure contexts.
* It’s clear from the literature we reviewed that open infrastructure is intimately related to the scholarly knowledge and research practices of individuals and organisations. In order for us to better understand the broad range of productive functions offered by open infrastructure, IOI should map out these practices and create a comprehensive and robust knowledge production and dissemination workflow. Ultimately, this work can inform the creation of a standardised taxonomy for productive infrastructure functions.
* This review also makes clear that the distinguishing feature between open infrastructure and its commercially-run and -operated counterparts is the fact that it’s valuated not just based on its ability to support productive functions but social practices and values as well. IOI should explicitly identify these values.

# 

# 7. References

Atkins, D. E., Droegemeier, K. K., Feldman, S. I., Garcia-Molina, H., Klein, M. L., Messerschmitt, D. G., Messina, P., Ostriker, J. P., & Wright, M. H. (2003). *Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure*. National Science Foundation. https://repository.arizona.edu/handle/10150/106224

Bilder, G., Lin, J., & Neylon, C. (2015). *Principles for Open Scholarly Infrastructures-v1*. https://doi.org/10.6084/m9.figshare.1314859.v1

Bollier, D. (2013). *The Importance of Infrastructure to Commons*. Bollier. http://www.bollier.org/blog/importance-infrastructure-commons

Bowker, G. C., Baker, K., Millerand, F., & Ribes, D. (2010). Toward Information Infrastructure Studies: Ways of Knowing in a Networked Environment. In *International Handbook of Internet Research*. Springer. https://doi.org/10.1007/978-1-4020-9789-8\_5

Chen, G., Posada, A., & Chan, L. (2019). Vertical Integration in Academic Publishing : Implications for Knowledge Inequality. In P. Mounier (Ed.), *Connecting the Knowledge Commons — From Projects to Sustainable Infrastructure : The 22nd International Conference on Electronic Publishing – Revised Selected Papers*. OpenEdition Press. http://books.openedition.org/oep/9068

Clarke, A. E., & Star, S. L. (2008). The social worlds framework: A theory/methods package. In E. J. Hackett, O. Amsterdamska, M. Lynch, & J. Wajcman (Eds.), *Handbook of science and technology studies* (pp. 113–137). MIT Press. https://www.dhi.ac.uk/san/waysofbeing/data/data-crone-wyatt-2007b.pdf#page=129

Dalakoglou, D. (2010). The road: An ethnography of the Albanian–Greek cross-border motorway. *American Ethnologist*, *37*(1), 132–149. https://doi.org/10.1111/j.1548-1425.2010.01246.x

Edwards, P. N. (2010). *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. MIT Press.

Edwards, P. N., Jackson, S. J., Bowker, G. C., & Knobel, C. P. (2007). *Understanding Infrastructure: Dynamics, Tensions, and Design* [Working Paper]. http://deepblue.lib.umich.edu/handle/2027.42/49353

Edwards, P. N., Jackson, S. J., Chalmers, M. K., Bowker, G. C., Borgman, C. L., Ribes, D., Burton, M., & Calvert, S. (2013). *Knowledge Infrastructures: Intellectual Frameworks and Research Challenges* [Working Paper]. Deep Blue. http://hdl.handle.net/2027.42/97552

Ficarra, V., Fosci, M., Chiarelli, A., Kramer, B., & Proudman, V. (2020). *Scoping the Open Science Infrastructure Landscape in Europe*. Zenodo. https://doi.org/10.5281/zenodo.4159838

Furlong, K. (2014). STS beyond the “modern infrastructure ideal”: Extending theory by engaging with infrastructure challenges in the South. *Technology in Society*, *38*, 139–147. https://doi.org/10.1016/j.techsoc.2014.04.001

Global Sustainability Coalition for Open Science Services. (2022). *Defining open infrastructure*. SCOSS. https://scoss.org/what-is-scoss/defining-open-infrastructure/

Harvey, P., & Knox, H. (2012). The Enchantments of Infrastructure. *Mobilities*, *7*(4), 521–536. https://doi.org/10.1080/17450101.2012.718935

Heinrich Böll Foundation, Commons Strategies Group, Charles Léopold Mayer Foundation, & Remix the Commons. (2013). Economics and the Common(s): From Seed Form to Core Paradigm. https://www.boell.de/sites/default/files/ecc\_report\_final.pdf

Humphrey, C. (2005). Ideology in Infrastructure: Architecture and Soviet Imagination. *Journal of the Royal Anthropological Institute*, *11*(1), 39–58. https://doi.org/10.1111/j.1467-9655.2005.00225.x

Jackson, S. (2015). *Repair*. Society for Cultural Anthropology. https://culanth.org/fieldsights/repair

Karasti, H. (2014). Infrastructuring in participatory design. *Proceedings of the 13th Participatory Design Conference: Research Papers - Volume 1*, 141–150. https://doi.org/10.1145/2661435.2661450

Karasti, H., Baker, K. S., & Millerand, F. (2010). Infrastructure Time: Long-term Matters in Collaborative Development. *Computer Supported Cooperative Work (CSCW)*, *19*(3), 377–415. https://doi.org/10.1007/s10606-010-9113-z

Karasti, H., & Blomberg, J. (2018). Studying Infrastructuring Ethnographically. *Computer Supported Cooperative Work (CSCW)*, *27*(2), 233–265. https://doi.org/10.1007/s10606-017-9296-7

Karasti, H., Millerand, F., Hine, C. M., & Bowker, G. C. (2016). Knowledge infrastructures: Part I. *Science & Technology Studies*, *29*(1), 2–12. https://doi.org/10.23987/sts.55406

Kooy, M., & Bakker, K. (2008). Splintered networks: The colonial and contemporary waters of Jakarta. *Geoforum*, *39*(6), 1843–1858. https://doi.org/10.1016/j.geoforum.2008.07.012

Kramer, B., & Bosman, J. (2017, February 9). *Web-based Tools a Researcher Can Use*. 101 Innovations. https://docs.google.com/spreadsheets/u/1/d/1KUMSeq\_Pzp4KveZ7pb5rddcssk1XBTiLHniD0d3nDqo/htmlview

Lagoze, C., Edwards, P., Sandvig, C., & Plantin, J.-C. (2015). Should I Stay or Should I Go? Alternative Infrastructures in Scholarly Publishing. *International Journal of Communication*, *9*(0), 1052–1071. https://ijoc.org/index.php/ijoc/article/view/2929

Larkin, B. (2013). The Politics and Poetics of Infrastructure. *Annual Review of Anthropology*, *42*, 327–243. https://doi.org/0.1146/annurev-anthro-092412-155522

Lewis, D. W. (2020). *A Bibliographic Scan of Digital Scholarly Communication Infrastructure | Educopia Institute*. Educopia Institute. https://educopia.org/mapping-the-scholarly-communication-landscape-bibliographic-scan/

Lewis, D. W., Goetsch, L., Graves, D., & Roy, M. (2018, February 27). Funding community controlled open infrastructure for scholarly communication: The 2.5% commitment initiative. *College & Research Libraries News*, *79*(3), 133–136. https://crln.acrl.org/index.php/crlnews/article/view/16902

McArthur, J. (2019, June 22). *Infrastructure finance: Historical perspectives and the evolution of financial instruments - UCL Discovery*. Society for the Advancement of Socio-economics Annual Conference, New York. https://discovery.ucl.ac.uk/id/eprint/10117660

Monteiro, E., Pollock, N., Hanseth, O., & Williams, R. (2013). From Artefacts to Infrastructures. *Computer Supported Cooperative Work*, *22*(4–6), 575–607. https://doi.org/10.1007/s10606-012-9167-1

Moore, S. (2020, April 17). Without stronger academic governance, Covid-19 will concentrate the corporate control of academic publishing. *LSE Blogs*. https://blogs.lse.ac.uk/impactofsocialsciences/2020/04/17/without-stronger-academic-governance-covid-19-will-concentrate-the-corporate-control-of-academic-publishing/

Nightingale, A. J. (2019). Commoning for inclusion? commons, exclusion, property and socio-natural becomings. *International Journal of the Commons*, *13*(1), 16–35. https://doi.org/http://doi.org/10.18352/ijc.927

Okune, A., Hillyer, R., Chan, L., Albornoz, D., & Posada, A. (2019). Whose Infrastructure? Towards Inclusive and Collaborative Knowledge Infrastructures in Open Science. In P. Mounier (Ed.), *Connecting the Knowledge Commons — From Projects to Sustainable Infrastructure : The 22nd International Conference on Electronic Publishing – Revised Selected Papers*. OpenEdition Press. http://books.openedition.org/oep/9072

Schroeder, R. (2007). e-Research Infrastructures and Open Science: Towards a New System of Knowledge Production? *Prometheus*, *25*(1), 1–17. https://doi.org/10.1080/08109020601172860

SComCaT. (n.d.). *About SComCaT*. Scholarly Communication Technology Catalogue. Retrieved April 12, 2022, from https://www.scomcat.net/about

Skinner, K. (2019, July 23). *Why Are So Many Scholarly Communication Infrastructure Providers Running a Red Queen’s Race?* [Educopia Institute]. https://educopia.org/red-queens-race/

Star, S. L. (1999). The Ethnography of Infrastructure. *American Behavioral Scientist*, *43*(3), 377–391. https://doi.org/10.1177/00027649921955326

Star, S. L. (2002). Got infrastructure? How standards, categories, and other aspects of infrastructure influence communication. *The 2nd Social Study of IT Workshop at the LSE ICT and Globalization*, 22–23.

Star, S. L., & Bowker, G. C. (2002). How to infrastructure. In L. A. Lievrouw & S. Livingstone (Eds.), *Handbook of New Media - Social Shaping and Consequences of ICTs* (pp. 151–162). SAGE Publications. https://dx.doi.org/10.4135/9781848608245.n12

Star, S. L., & Ruhleder, K. (1996). Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces. *Information Systems Research*, *7*(1), 111–134. https://doi.org/10.1287/isre.7.1.111

Swyngedouw, E. (2004). *Social Power and the Urbanization of Water: Flows of Power*. Oxford University Press.

UNESCO. (2021). *UNESCO Recommendation on Open Science* (SC-2021/SANS COTE). UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000379949.locale=en

Maxwell, J. W., Hanson, E., Desai, L., Tiampo, C., O’Donnell, K., Ketheeswaran, A., Sun, M., Walter, E., & Michelle, E. (2019). *Mind the Gap: A Landscape Analysis of Open Source Publishing Tools and Platforms*. The MIT Press. https://doi.org/10.21428/6bc8b38c.2e2f6c3f

Watkinson, C., & Pitts, M. (2021, February 22). *Re-Envisioning Humanities Infrastructure*. Inside Higher Ed. https://www.insidehighered.com/views/2021/02/22/institutions-and-funders-must-recognize-contributions-university-presses-humanities

West, J., & O’mahony, S. (2008). The Role of Participation Architecture in Growing Sponsored Open Source Communities. *Industry and Innovation*, *15*(2), 145–168. https://doi.org/10.1080/13662710801970142

1. The framework described in this report will directly contribute to our [Costs of Open Infrastructure project](https://investinopen.org/research/costs-of-open-infrastructure/), enabling us to identify the critical providers of open technologies and systems that support research and scholarship. [↑](#footnote-ref-0)
2. These categories emerged organically as we coded our compiled body of literature. We ultimately decided that they offered the most precise classification system while simultaneously facilitating an easy-to-follow narrative for our report. [↑](#footnote-ref-1)
3. While there can be an invisible quality to infrastructure, this invisibility is neither perpetual or constant (Karasti & Blomberg, 2018). Larkin (2013), instead describes the visibility of infrastructures as existing on a spectrum, ranging from “unseen to grand spectacles and everything in between” (p. 336). [↑](#footnote-ref-2)
4. Terms such as big science, data-driven science, networked science, open science, Digital Humanities, science 2.0, e-Science, e-Social Science, and e-Research have also been used by researchers examining knowledge production processes in the digital age (Karasti et al., 2016). These works fall out of the scope of this report. [↑](#footnote-ref-3)