

## Knowledge Infrastructures: Part IV

*Helena Karasti*

*Information Systems, Luleå University of Technology, Sweden / [helena.karasti@ltu.se](mailto:helena.karasti@ltu.se)  
INTERACT, University of Oulu, Finland / [helena.karasti@oulu.fi](mailto:helena.karasti@oulu.fi)*

*Florence Millerand*

*Department of Public and Social Communication, University of Quebec at Montreal, Canada /  
[millerand.florence@uqam.ca](mailto:millerand.florence@uqam.ca)*

*Christine M. Hine*

*Department of Sociology, University of Surrey, UK / [c.hine@surrey.ac.uk](mailto:c.hine@surrey.ac.uk)*

*Geoffrey C. Bowker*

*Department of Informatics, University of Irvine, CA, USA / [gbowker@uci.edu](mailto:gbowker@uci.edu)*

This issue of Science and Technology Studies is the final one of four in total published this year focusing on the topic of Knowledge Infrastructures. Across the four issues we have presented fourteen papers (thirteen research articles and one discussion paper) and four book reviews. In this final editorial we first take a look at the issues raised by the final batch of articles, then take a step back to review the collection as a whole, considering what it tells us about the state of the art in Science and Technology Studies' understanding of knowledge infrastructures and looking forward to the challenges still on the horizon.

### **Articles in This Fourth and Last Part of the Special Issue**

The first article 'The Daily Shaping of State Transparency: Standards, Machine-Readability and the

Configuration of Open Government Data Policies' addresses the issue of open standards for diffusing online data in the context of government bureaucracies. In common with open data initiatives in other substantive fields, such as science (Borgman, 2007) and cultural heritage (Stuedahl et al., in this issue), many governments are now committed to the release of open data. Open Government Data (OGD) initiatives are constructing ways to store and share data, forming a new layer of 'open data infrastructure' shaped by the development and deployment of data standards (Lampland and Star, 2009). While OGD movements to sharing data under non-proprietary standardized formats have been highly visible, Samuel Goëta and Tim Davies point out that considerably less attention has been given to what is happening on the ground around the production of standards and the actual consequences of

standards for knowledge workers, the issues that form the authors' focus in this article.

Goëta and Davies study three very different open data standards, namely Comma Separated Value (CSV), General Transit Format Specification (GTFS) and the International Aid Transparency Initiative (IATI). They operate an 'infrastructural inversion' by looking at the historical development of the named standards and by studying ethnographically the 'back rooms' of government bureaucracies with a focus on the invisible work necessary to open the data by using these standards. The authors pay particular attention to the concrete work practices that go along with aligning the standards, the organizational arrangements they create, and the way they shape the data for others to access and use.

Through the empirical work the authors discuss how transparency is or is not achieved by the demands for openness and standardization. The authors show that the standards substantively shape the production of open data. They describe how the use of open standards requires intensive work in order to transform and adjust datasets to the standards; thus, the making of datasets machine-readable may increase the complexity of releasing data. The authors further show how enacting open standards operates "a quiet and localised transformation of bureaucracies", with consequences for how open government data and transparency agendas are performed. The use of open standards has become interpreted not only as a sign of a quality dataset, but also used to evaluate the progress of the open data program itself. The adoption of open standards is increasingly becoming (used as) an indicator of the advancement of open data programmes. Furthermore, the authors discuss the particular kind of transparency delivered by OGD which reveals a rationalisation and representation of the information held inside the state, focussing on machine-mediated transparency rather than transparency as a relationship between citizen and account-giving state.

In addition to the above 'producer' side inside the 'back rooms' of government bureaucracies, the authors also discuss the 'user' side of OGD. They see that the emphasis on machine-readability in OGD projects configures the primary users as

'advanced users' with a need for technical skills, financing and capability to create services to make desired re-use of the published data. These set-ups (of professional developers and ecosystems) introduce other layers of infrastructure and eventually intermediation between citizens and the state.

In the second article, Ayelet Shavit and Yael Silver discuss the development of long term biodiversity surveys and specifically focus in on tensions inherent in recording locality within such surveys. The first case study in the article discusses the evolving treatment of locality information within the specimen collections of the Museum of Vertebrate Zoology at the University of California, Berkeley. A formalized approach to recording was established early on in the museum's history, requiring both a standardized set of information including a record of locality and a narrative account of the circumstances surrounding collection of the specimen in a field journal. This system of recording thus combined what Shavit and Silver term 'exogenous' and 'interactionist' approaches to locality. The two approaches are associated with contrasting epistemic values: an exogenous approach to 'location' focuses on production of representative and reliable data whilst the interactionist approach attends to the need for comprehensive and accurate data for the location in question. Both systems co-existed in the pre-computerised system of journals, index cards and tags, but the advent of computerized records in the 1970s began a push towards inclusion of a searchable and generalizable version of specimen locality in specimen databases and prompted the development of a system to map historical localities to estimated longitude and latitude using a standard georeferencing protocol. Subsequently, new challenges for the recording of locality emerged, as new devices used by researchers in the field occasioned a more precise georeferencing, producing new forms of data and shifting away from narrative field journals to numerical data. A separation emerged between the requirement for a globally interoperable and easily searchable form of locality information and the historical collections of narrative data on circumstances of collection that were locally held at the museum and mined by relatively few researchers.

A subsequent workaround involved digitization of field journals, allowing this information to be linked to specimen records and hence made available albeit not in an equivalent searchable form to the exogenous locality information.

The second case study in Shavit and Silver's paper focuses on a biological monitoring project 'Hamaarag' initially associated with Long Term Ecological Research (LTER) stations funded by Israel's Science Foundation. Shavit and Silver track the changing political, financial and scientific focus of the project over time, and also the tensions over the version of locality embedded within the project. As with the Museum of Vertebrate Zoology, tensions focused on a clash between the possibility of developing an interoperable infrastructure across the various LTERs involved and the very different demands imposed by the different species each were monitoring and the practices of the groups of scientists involved. Shavit and Silver track the diverse and shifting pressures that beset the project over time and challenge attempts to produce a single overarching infrastructure for the project, leading ultimately to an approach that favours an interactionist approach to location and includes citizen science initiatives alongside research team efforts. Across the two case studies, Shavit and Silver identify a tension between different notions of locality and an emergent recognition that to focus only on a globally interoperable exogenous version of locality may entail a loss of a significant flexibility. They conclude that developing an infrastructure to sustain local memories of a locality and alternating between local and global memory practices (Bowker, 2005) may be better justified, both rationally and sometimes morally. Tracking the movement from a technical thing (the technical category of 'location') becoming a problematic epistemic thing, the article demonstrates a recurring issue in knowledge infrastructure work more broadly i.e. the weight that may be carried by technical decisions on the representation of key concepts.

The third article in the special issue, by Dagny Stuedahl, Mari Runardotter and Christina Mörtberg, focuses on the substantive field of the cultural heritage sector. The authors develop two case studies of digital infrastructure projects that

are involved in opening up cultural heritage institutions to engagement with the public. Whilst both projects are working within an environment that encourages openness and public involvement, the two case studies contrast significantly in their institutional form and in the approach they take to defining what will count as an acceptable open engagement with the public. The first study focuses on a "top-down" initiative in the design phase: a new infrastructure intended to facilitate public access to archival materials. By studying discussions in the design phase Stuedahl et al. are able to identify tensions and controversies around the implementation of the high-level policy imperative to open data and engage with citizens. When these imperatives meet with local practices they encounter considerable concerns that revolve around the extent of openness deemed desirable and the quality of content acquired through crowd-sourcing, leading ultimately to adoption of an approach focused on providing access to existing archival data rather than acquiring new data. The second case study explores a 'bottom-up' initiative: a local history wiki used by professional and amateur local historians. Here Stuedahl et al. encounter the project when it is already up and running, and analyse threads from the discussion forum that demonstrate ongoing negotiations over the categories to be used to structure contributions to the wiki and tensions between wiki administrators and local historians over the extent to which diverse understandings can be accommodated within the wiki.

To draw together the comparison between these two substantively similar yet contrasting initiatives Stuedahl et al. rely on the concept of 'attachments' used within STS variously by Gomart and Hennion (1999), Latour (1999), Marres (2007) and Hennion (2012) to denote an array of resources that are drawn on to inform and make sense of engagements and actions. Attachments are potentially more diffuse than motivations and more emotionally charged than influences, offering a means to identify what matters to people as they decide on a course of action or design an intervention. In the participatory knowledge infrastructures that they study Stuedahl et al. identify attachments used by actors to outline what matters to them and position

themselves in relation to past, present, and future. The authors argue that attachments offer a useful alternative way to explore the temporality of knowledge infrastructuring, stressing that sustainable infrastructures may need not only to work with the long now (Ribes & Finholt, 2009) of an anticipated future but also to display an appropriate attachment to relevant values and practices of the past as well as attachments to other pressures and policies in the present. By highlighting the various attachments that actors bring to the two case studies they outline, Stuedahl et al. bring out the process through which the contrasting (and sometimes internally conflicting) notions of openness and engagement that the two projects arrive at come into being.

### **An Overview and Emerging Themes**

The fourteen articles published in this special issue, while all viewing their material through the lens of the knowledge infrastructure, have covered a range of substantive fields: biodiversity (Taber, 2016); cultural heritage (Stuedahl et al., in this issue); disease genetics (Dagiral & Peerbaye, 2016); drug discovery (Fukushima, 2016); e-health (Aspria et al., 2016); ecological science (Stuedeahl et al., 2016; Shavit & Silver, in this issue); environmental monitoring (Jalbert, 2016; Parmiggiani & Monteiro, 2016); open government (Goëta & Davies, in this issue); public health (Boyce, 2016); social science data archiving (Shankar et al., 2016); weather recording (Goëta & Davies, 2016); wikipedia content (Wyatt et al., 2016). While many have at their heart a database or other form of digital technology, this has not been universally the case: Taber (2016) views the herbarium as the focus of a knowledge infrastructure. The articles exemplify the interdisciplinary trend within Science and Technology Studies more broadly. While we have not conducted a systematic census of the disciplinary origins of the scholars represented here, it is clear from their institutional addresses as much as their substantive foci that the authors come from an array of backgrounds including anthropology, informatics and information science, media and communications, public health and social science in addition to science and technology studies departments. The geographical spread

is also broad, including authors from Australia, France, Ireland, Israel, Japan, Netherlands, Norway, Sweden, United States of America and United Kingdom.

In the three previous editorials (Karasti et al., 2016a, 2016b, 2016c) we have identified some emerging themes that tie together the contributions made by individual articles and suggest areas of common significance across quite diverse manifestations of knowledge infrastructures. In the first issue we discussed themes of scale, invisibility, tensions, uncertainty, and accountability. We also explored methodological issues, focusing on the infrastructural inversion and the challenges inherent for the researcher in choosing levels, locations, and scales to examine. In the second issue we explored the performativity of knowledge infrastructures and the struggles over power, values, and voice that prevail at the very heart of infrastructural work. The third issue highlighted temporality and labour as key areas of connection across infrastructural studies.

These themes continue to resonate across the three articles presented in this fourth issue to focus on knowledge infrastructures. All three articles deploy a methodological focus that encompasses the diverse scales of infrastructural work and each in its own way highlights an otherwise invisible or neglected aspect of that work and brings it into the foreground as consequential site for the enactment of values and the experience of tensions between different practices and sets of accountability. Temporality arises with particular significance in Stuedahl et al.'s exploration of the notion of attachments, as they argue that an attachment to aspects of the past can give meaning to infrastructural work as much as visions of an anticipated future.

Beyond the themes already identified, a further theme deserves exploration in this editorial: the notion of openness. As a value and a set of practices the notion of openness has a considerable contemporary significance and yet, as studied here, it emerges as a problematic concept not necessarily easy to achieve. Openness appears repeatedly across the papers collected here: in the first issue, Parmiggiani and Monteiro (2016) explore the development of an infrastructure for monitoring subsea ecosystems and evaluating

environmental risk and here achieving a portrayal of the openness of data in a public portal plays a part in building a new sense of trust; in the second issue, Shankar et al. (2016) propose a study of social science data archives that pays attention to the specificity of circumstances under which open sharing of data arises; in the third issue Aspria et al. (2016) explore the metaphors that underpin operationalization of a patient information portal that aspires to be seen as open and inclusive. In this fourth issue, openness receives further significant attention: Goëta and Davies place the standards that underpin open data sharing under the spotlight, and find that these standards are a site of considerable labour both in development and use and far from a smooth route to automatic transparency; Stuedahl et al. focus on the movement towards open data sharing in cultural heritage contexts and find that whilst aspiring to openness may be dictated by policy, it still requires considerable negotiation to make manageable in practice. When we study contemporary knowledge infrastructures we find values of openness often embedded there, but translating the values of openness into the design of infrastructures and the practices of infrastructuring is a complex and contingent process.

In putting together the special issue we aimed to assess the current state of Science and Technology Studies' contribution to the understanding of knowledge infrastructures. This set of emergent themes, connecting across together, exemplify the contribution that a set of sensibilities drawn from Science and Technology Studies can make in this area: by a detailed attention to technology as it is enacted in situ and as it is embedded in and embeds policies and practices, we can see the knowledge infrastructure as a very particular kind of achievement with far-reaching yet often overlooked consequences. We learn in detail about the modes of governance that depend upon and are enabled by knowledge infrastructures and we find out how great the gulf may be between an aspiration in the domain of policy and its realisation on the ground. STS scholars are studying the processes of infrastructuring in detail but also considering the consequences: what kind of ways of being in the world do knowledge infrastructures enable, to whom do they give voice and who

do they silence, what do they prioritise and what do they neglect or negate?

Viewed as a whole, this collection of papers suggests that the STS-enabled study of knowledge infrastructures is on increasingly solid theoretical and methodological ground. Across the papers we see a confidence in identifying diverse sets of technological developments as knowledge infrastructures and applying to them a relatively stable set of theoretical resources. Among the papers we find also theoretical innovations, such as Fukushima's (2016) recourse to a Marxist-inflected notion of infrastructure alongside the resources of STS or Stuedahl et al.'s (in this issue) deployment of attachments as a means to uncover the meanings that pervade infrastructural work. On the whole, however, the articles wear their theoretical development relatively lightly and concentrate on illuminating what is being achieved through the medium of knowledge infrastructural work and how this is being brought about.

Methodologically speaking, also, this collection of papers speaks to a relatively confident set of resources being deployed to good effect. Most of the papers make a broad claim to ethnographic approaches, with the notable exceptions of Wyatt et al. (2016) in their study of data from editorial discussions on Wikipedia and Taber (2016) and Shankar et al. (2016) with historical approaches founded on archival data. Ethnography, in the knowledge infrastructure context, often means a foundation of participant observation within a key location, taking part in ongoing discussions and attending meetings. The temporal and spatial complexity of infrastructural work is handled through a combination of mobility from the research and recourse to programmes of interviewing and documentary analysis. Online discussions appear as sources of data that give a useful insight into day-to-day negotiations into the meaning of data, capturing as they do a level of detail often otherwise ephemeral and hard to capture when work goes on in face-to-face settings, even for an ethnographer on the spot. The increasing recourse to online discussion forums for getting infrastructural work done has, as a by-product, provided a useful set of data for STS scholars interested in how this work is done.

Studying the otherwise invisible becomes easier when this work is captured in a persistent form.

The notion of the infrastructural inversion has clearly become one of the established resources of an STS approach to knowledge infrastructures. Responding to Geof Bowker's call to make material infrastructures the central object of study (Bowker, 1994), many of the papers in this collection used the infrastructural inversion in the standard sense of a methodological sensitivity associated with making otherwise neglected things visible, as exemplified by Bowker and Star (1999). In doing so, these papers confirmed the pertinence of this methodological lens to scrutinize the inter-dependences between technical components and the politics of knowledge production. Three articles elaborated on the infrastructural inversion to a significant extent: Fukushima (2016) drawing out an isomorphism with the Marxist inversion of the infrastructure/superstructure relation; and both Parmiggiani and Monteiro (2016) and Dagiral and Peerbaye (2016) drawing out the use of the inversion as a resource by actors themselves.

There are, thus, promising signs for future knowledge infrastructure studies in STS, confidently adopting and developing a mature set of methodological and theoretical resources. Promising future prospects include possible pay-offs from making further use of online data and myriad digital traces left by digital work, taking on board Edwards et al.'s (2013) challenge to infrastructural studies to take more account of big data. Future studies may also do more to engage in depth with the reflexive work done by the actors in infrastructural projects, building on the recognition that concepts such as the infrastructural inversion resonate strongly with what actors themselves do. New methodological forms may yet emerge. The majority of the articles collected here represent either the work of one

scholar, or a small group of scholars pooling or contrasting a small number of case studies. We see little as yet of the larger team-based and multi-sited studies that may be necessary in order to scale up knowledge infrastructure studies and more extensively explore their ramifications across time and space as Edwards et al. (2013) exhort. Similarly, while historical and archival studies promise to allow us to extend our interest in the evolution of knowledge infrastructures across greater time spans, as yet our analytic resources for conducting archival studies are relatively under-developed (Bowker, 2015). The collection of articles presented here demonstrate a healthy and vibrant field, with a clearly significant pay-off in terms of illuminating some very powerful aspects of contemporary world, yet there is clearly still further to go in developing the STS contribution in this area.

## Acknowledgements

Beyond the team of four guest editors responsible for putting these special issues together there has been a considerable input from the journal *Science and Technology Studies* and from an array of anonymous reviewers. We are very grateful to the reviewers who have read articles so thoroughly and offered their wise and constructive advice. For the journal, Antti Silvast has taken editorial responsibility for the Knowledge Infrastructures special issue. Antti has been hugely generous with his time and knowledge of the field as he has guided and advised us across what has been, for him, a much greater commitment than initially anticipated when the response to the call for papers produced four special issues rather than a single one. Louna Hakkarainen has acted for the journal to see the issues efficiently through the production process.

## References

- Aspria M, de Mul M, Adams S & Bal R (2016) Of Blooming Flowers and Multiple Sockets: The Role of Metaphors in the Politics of Infrastructural Work. *Science & Technology Studies* 29(3): 68-87.
- Borgman, CL (2007) *Scholarship in the digital age: Information, infrastructure, and the Internet*. Cambridge: MIT.
- Bowker, GC (1994) *Science on the run: Information management and industrial geophysics at Schlumberger, 1920–1940*. Cambridge: MIT Press.
- Bowker, GC (2005) *Memory Practices in the Sciences*. Cambridge: MIT Press.
- Bowker GC (2015) Temporality: Theorizing the Contemporary. *Cultural Anthropology*. September 24, 2015. Available at: <https://culanth.org/fieldsights/723-temporality> (accessed: 22.8.2016)
- Bowker GC & Star SL (1999) *Sorting Things Out: Classification and Its Consequences*. Cambridge: MIT Press.
- Boyce AM (2016) Outbreaks and the Management of ‘Second-Order Friction’: Repurposing Materials and Data From the Health Care and Food Systems for Public Health Surveillance. *Science & Technology Studies* 29(1): 52-69.
- Dagiral É & Peerbaye A (2016) Making Knowledge in Boundary Infrastructures: Inside and Beyond a Database for Rare Diseases. *Science & Technology Studies* 29(2): 44-61.
- Edwards PN, Jackson SJ, Chalmers MK, Bowker GC, Borgman CL, Ribes D, Burton M & Calvert S (2013) *Knowledge Infrastructures: Intellectual Frameworks and Research Challenges*. Ann Arbor: Deep Blue. <http://knowledgeinfrastructures.org/> (accessed: 5.11.2016)
- Fukushima M (2016) Value Oscillation in Knowledge Infrastructure: Observing its Dynamic in Japan’s Drug Discovery Pipeline. *Science & Technology Studies* 29(2): 7-25.
- Gomart E& Hennion A (1999) A Sociology of Attachment: Music Amateurs, Drug Users. In: Law J & Hassard J (eds) *Actor Network Theory and After*. Oxford: Blackwell, 220-47.
- Granjou C & Walker J (2016) Promises that Matter: Reconfiguring Ecology in the Ecotrons. *Science & Technology Studies* 29(3): 49- 67.
- Hennion A (2012) Attachments: A Pragmatist View Of What Holds Us In: *The First European Pragmatist Conference*, Roma 19-21 September 2012.
- Jalbert K (2016) Building Knowledge Infrastructures for Empowerment: A Study of Grassroots Water Monitoring Networks in the Marcellus Shale. *Science & Technology Studies* 29(2): 26-43.
- Karasti H, Millerand F, Hine CM, & Bowker GC (2016a) Knowledge infrastructures: Part I. *Science & Technology Studies* 29(1): 2-12.
- Karasti H, Millerand F, Hine CM, & Bowker GC (2016b) Knowledge infrastructures: Part II. *Science & Technology Studies* 29(2): 2-6.
- Karasti H, Millerand F, Hine CM, & Bowker GC (2016c) Knowledge infrastructures: Part III. *Science & Technology Studies* 29(3): 2-9.
- Lampland M & Star SL (eds) (2009) *Standards and Their Stories: How Quantifying, Classifying, and Formalizing Practices Shape Everyday Life*. Ithaca, NY: Cornell University Press.
- Latour B (1999 [2004]) *Politics of Nature: How to Bring the Sciences into Democracy*. Cambridge MA: Harvard University Press.
- Lin Y-W, Bates J & Goodale P (2016) Co-Observing the Weather, Co-Predicting the Climate: Human Factors in Building Infrastructures for Crowdsourced Data. *Science & Technology Studies* 29(3): 10-27.

- Marres N (2007) The Issues Deserve More Credit: Pragmatist Contributions to the Study of Public Involvement in Controversy. *Social Studies of Science* 37:759.
- Parmiggiani E & Monteiro E (2016) A Measure of 'Environmental Happiness': Infrastructuring Environmental Risk in Oil and Gas Offshore Operations. *Science & Technology Studies* 29(1): 30-51.
- Ribes D & Finholt TA (2009) The Long Now of Technology Infrastructure: Articulating Tensions in Development. *Journal of the Association for Information Systems* 10(5): Article 5. Available at: <http://aisel.aisnet.org/jais/vol10/iss5/5> (accessed: 4.11.2016)
- Shankar K, Eschenfelder KR & Downey G (2016) Studying the History of Social Science Data Archives as Knowledge Infrastructure. *Science & Technology Studies* 29(2): 62-73.
- Taber P (2016) Taxonomic Government: Ecuador's National Herbarium and the Institution of Biodiversity, 1986-1996. *Science & Technology Studies* 29(3): 28-48.
- Wyatt S, Harris A & Kelly SE (2016) Controversy goes online: Schizophrenia genetics on Wikipedia. *Science & Technology Studies* 29(1): 13-29.