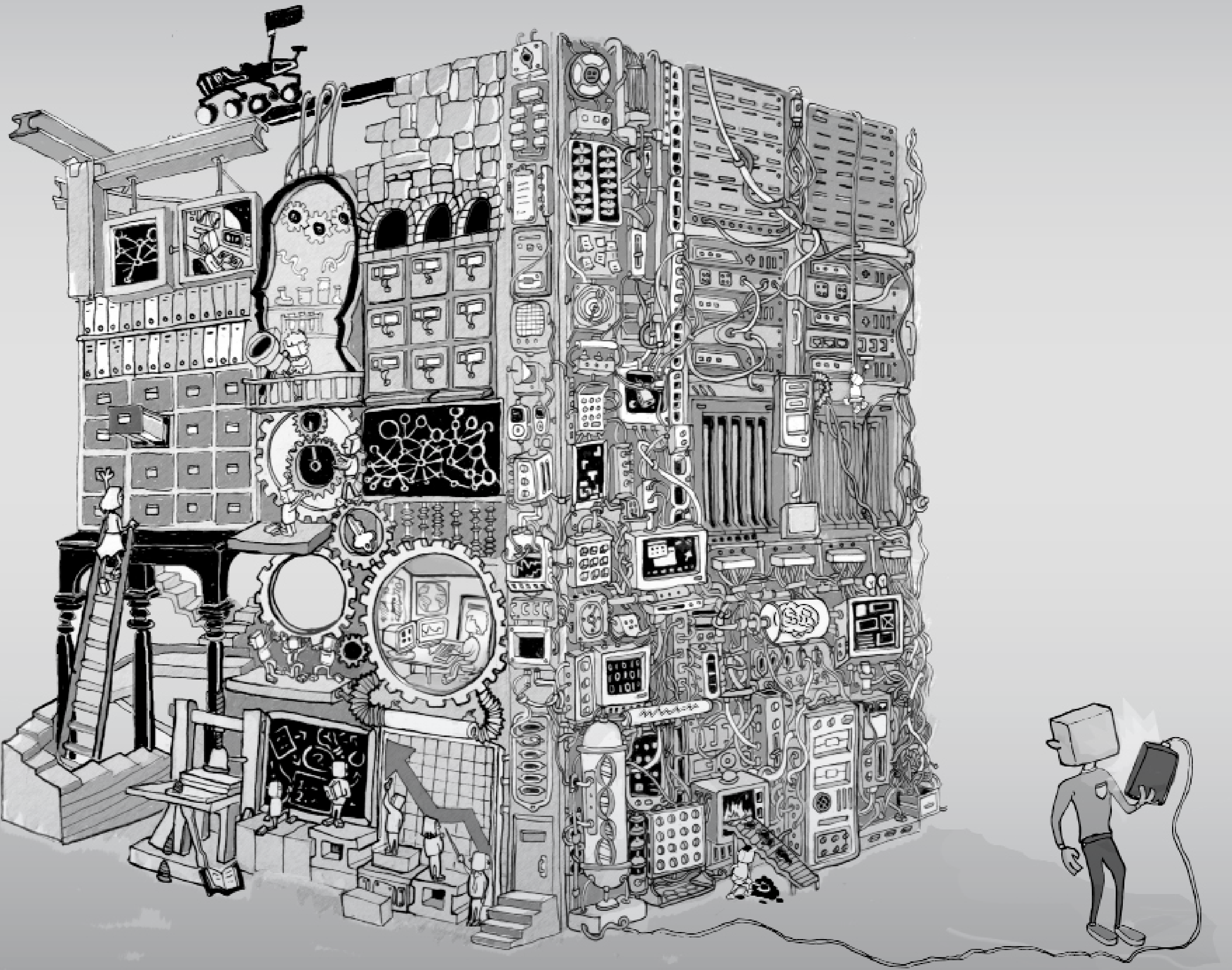
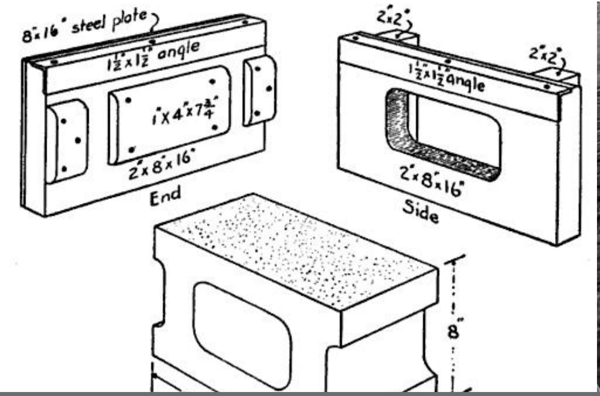
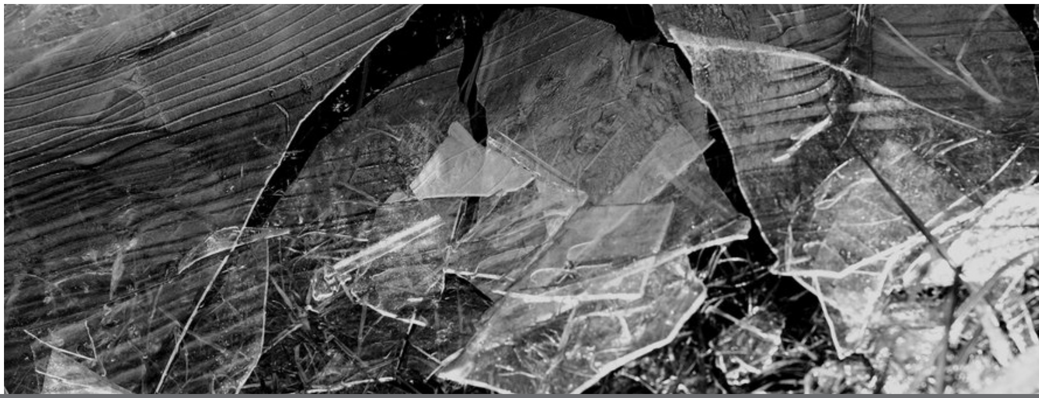


Knowledge Infrastructures: Intellectual Frameworks and Research Challenges

Report of a workshop sponsored by the National Science Foundation and the Sloan Foundation

University of Michigan School of Information, 25-28 May 2012





19th century knowledge mechanisms homemade: concrete block mold parts, diagram by Henry Colin Campbell

Ice texture, photograph by Aimelle-Stock on Deviant Art.

Knowledge Infrastructures: Intellectual Frameworks and Research Challenges

Report of a workshop sponsored by the National Science Foundation and the Sloan Foundation
University of Michigan School of Information, 25-28 May 2012

Report Authors:

Paul N. Edwards, *University of Michigan and SciencesPo, Paris*
Steven J. Jackson, *Cornell University*
Melissa K. Chalmers, *University of Michigan*
Geoffrey C. Bowker, *University of California, Irvine*
Christine L. Borgman, *University of California, Los Angeles and Oxford University, UK*
David Ribes, *Georgetown University*
Matt Burton, *University of Michigan*
Scout Calvert, *University of California, Irvine*

Workshop participants:

Margy Avery, *MIT Press*
Jean-Francois Blanchette, *University of California, Los Angeles*
Finn Brunton, *University of Michigan*
David DeRoure, *University of Oxford*
Greg Downey, *University of Wisconsin, Madison*
Carole Goble, *University of Manchester*
Josh Greenberg, *Sloan Foundation*
James Howison, *University of Texas, Austin*
Hrönn Brynjarsdóttir Holmer, *Cornell University*
John King, *University of Michigan*
Carl Lagoze, *University of Michigan*
Sandy Payette, *Cornell University*
Daniela Rosner, *University of Washington*
Christian Sandvig, *University of Michigan*
Phoebe Sengers, *Cornell University*
Katie Shilton, *University of Maryland*
Victoria Stodden, *Columbia University*
Janet Vertesi, *Princeton University*
Robin Williams, *University of Edinburgh*
Alex Wade, *Microsoft Research*

Release date: May 2013

Suggested citation:

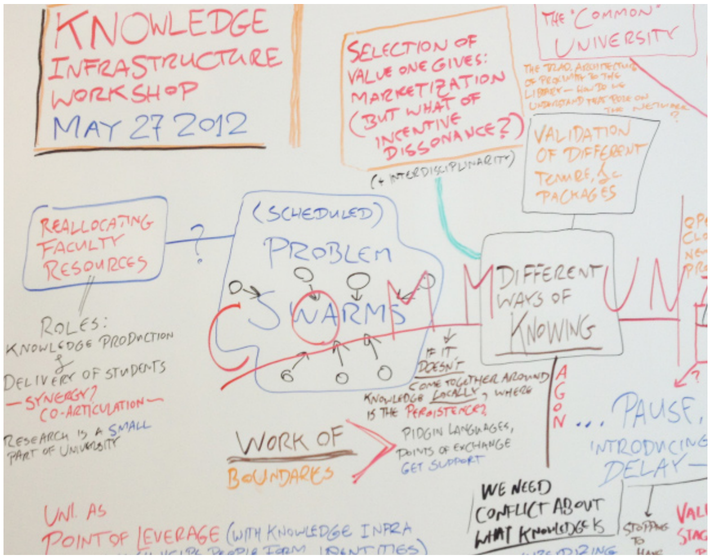
Edwards, P. N., Jackson, S. J., Chalmers, M. K., Bowker, G. C., Borgman, C. L., Ribes, D., Burton, M., & Calvert, S. (2013) *Knowledge Infrastructures: Intellectual Frameworks and Research Challenges*. Ann Arbor: Deep Blue. http://hdl.handle.net/2027.42/97552.

Cover art and website design by Jake Fagan. Report design and layout by Jillian C. Wallis. The report was set in Myriad Pro using Adobe InDesign CS3, and printed on 70# Neenah digital text weight stock. All images besides those from the workshop are under CC-BY and CC-BY-SA licenses. Image titles and authors are acknowledged in footnotes on the bottom of the page the image appears.

The Report at a Glance

The Knowledge Infrastructures research group convened a workshop in May 2012, sponsored by the U.S. National Science Foundation and the Sloan Foundation. Some 25 international scholars from many domains, including sociology, science and technology studies, computer science, human-computer interaction, and the digital humanities, participated in three days of intensive discussions and breakout groups. This document reports the outcomes, organized around three central questions: How are knowledge infrastructures changing? How do changes in knowledge infrastructures reinforce or redistribute authority, influence, and power? And how can we best study, know, and imagine knowledge infrastructures moving forward?

Our report offers key examples of change, considers the consequences (good and bad) of emergent practices, and offers some rough tools and approaches that might support new ways of thinking and acting on the changing knowledge infrastructures around us. We conclude with recommendations for a more effective program of research and action in this space. As always, the real-world terrain is more vast and complex than any single representation can capture. The report that follows is meant to open conversations rather than close them. Our goal is to gather and connect existing threads in a way that supports learning, insight, and more effective modes of infrastructure development moving forward.



Knowledge infrastructures and our understanding of them are changing rapidly. Therefore, in addition to this report we have developed a website for continuing the conversation begun at this workshop. Please visit www.knowledgeinfrastructures.org and lend your voice to the discussion.

Acknowledgements

The authors gratefully acknowledge the support of the Sloan Foundation and the National Science Foundation (grant BCS-0827316). We also acknowledge the vital assistance of Melissa Chalmers, Matt Burton, and Todd Stuart in organizing workshop logistics.

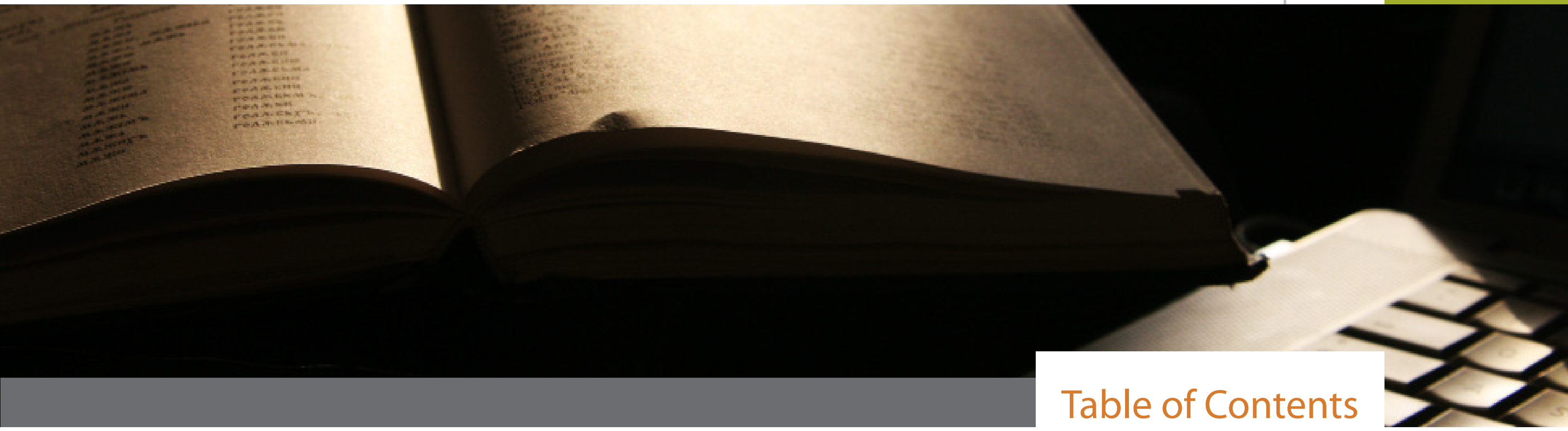
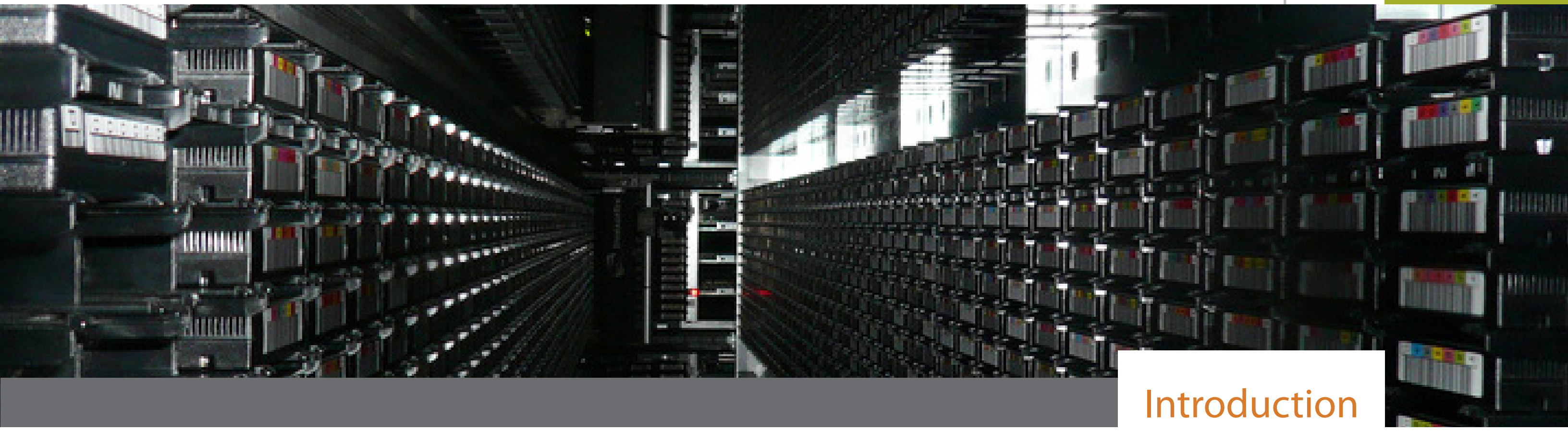


Table of Contents

"Day 158: Diffusion of Knowledge", photograph by Quinn Dombrowski



Introduction

This report lays groundwork for a new approach to understanding the massive transformations currently underway in how people create, share, and dispute knowledge. We explore some of the major questions that need to be addressed if these changes are to reach their full potential, and the types of inquiries they will require. We seek to inspire new ways of thinking around issues that have been obscured by older approaches and assumptions – some of them in the process of being undermined and remade by the very forces described here. Our report is at the same time a manifesto and an unfinished agenda, a statement and a provocation we hope will inspire others to further investigation.

Enormous transformations have occurred over the last 20 years in our systems for generating, sharing, and disputing human knowledge. Changes associated with Internet technologies — such as social media, "big data," open source software, ubiquitous computing, and Wikipedia — have altered the basic mechanics by which knowledge is produced and circulated. Remarkable new knowledge practices have emerged, captured under the language of crowdsourcing, cyberinfrastructure, personal informatics, citizen science, open access, MOOCs, and dozens of other terms that wouldn't have shown up in the Wikipedia pages of a decade ago; academic studies of some of these phenomena have become virtual scholarly fields unto themselves. Knowledge institutions like universities, libraries, and government agencies (and increasingly private entities like Facebook, Google, and Twitter) have begun to adjust, opening up vast stores of anonymized data to analysis and exploitation, engaging users and publics in new ways, and in some cases rethinking logics and practices that have been decades if not centuries in the making.

These developments have emerged in part from deliberate strategies on the part of funders and policymakers. For example, National Science Foundation programs including Knowledge and Distributed Intelligence (late 1990s), the Digital Libraries Initiative (late 1990s to early 2000s), Information Technology Research (early 2000s), the Office of Cyberinfrastructure (mid 2000s),

---

Tape Library at CERN, photographed by Cory Doctorow

University of Michigan Library Card Catalog, photographed by David Fulmer and Human and Social Dynamics (late 2000s) encouraged researchers to experiment with new modes of knowledge production and dissemination, as well as to study how such forms emerge. The Sloan Foundation (and others) funded important scientific initiatives, such as the Sloan Digital Sky Survey, that exploited these new modes. Finally, the Obama Administration's data.gov initiative represents the latest in a series of experiments in opening government databases to use by non-governmental entities.

Such has been the power of the Internet, both as a new medium and as a metaphor for knowledge, that much of the research surrounding these phenomena has attended mainly to two principal axes of change: first, technical systems and standards (computers and networks, of course, but also metadata, federated data systems, and middleware), and second, new modes of analyzing social (re)organization that exploit the extensive traces left behind by users of information technology. Major, related social and institutional changes in knowledge infrastructure include at least the following:

- Education: the rise of for-profit and online universities; open courseware; massively open online courses; a generalized crisis of traditional pedagogies
- Libraries: changing structures, services, and physical spaces
- The publishing industry: e-books vs. paper; prohibitive pricing of scientific journals; the collapse of university presses
- Intellectual property: distortions of copyright and patent law; creative commons practices; stark and growing differences between legal frameworks and actual use practices
- Global flows: increasingly rapid and supple transborder movement of researchers, students, professional expertise, and knowledge-based industries
- Knowledge politics: the "filter bubble" ; counter-expertise; challenges to expert knowledge organizations

This list — which could easily be far longer — makes clear that we are living through a period of fundamental transformations that profoundly challenge our understanding of the basic processes by which knowledge is created, debated, and spread.

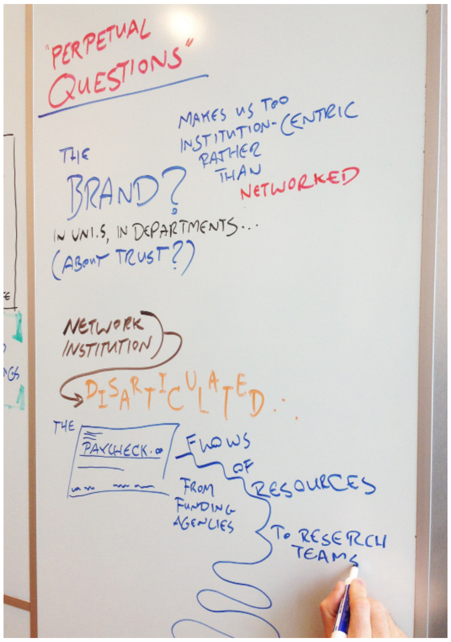
This challenge is of more than intellectual concern. The institutions in which most knowledge workers live and labor have not kept pace, or have done so piecemeal, without a long-term vision or a strategy. For example, the widespread excitement about crowdsourced knowledge, assembled by unpaid individuals who volunteer their time out of personal interest, ignores the fact that most knowledge workers' salaries are still paid by bricks-and-mortar organizations with hierarchical structures, established institutional cultures, systems of credit and compensation, and other "sticky" processes and routines. Similarly, our educational systems, libraries, publishers, news organizations, intellectual property structures, and political mechanisms have struggled to match or adapt to the changing information environment (Borgman 2007). The result is a patchwork of unsatisfactory kludges, contradictions, and inconsistencies that may undermine the prospects for change.

Popular attention and academic research on changing knowledge systems has tended to follow the new, fast-moving, and dramatic parts of the current transition. For example, in *Reinventing Discovery*, Nielsen (2012) extrapolates from current events to the eventual rise of a scientific culture of "extreme openness" where "all information of scientific value, from raw experimental data and computer code to all the questions, ideas, folk knowledge, and speculations that are currently locked up inside the heads of individual scientists" is moved onto the network, "in forms that are not just human-readable, but also machine-readable, as part of a data web." Shirky (2010) argues that a "cognitive surplus" will permit massively distributed contributions to the analysis of information and the production of new knowledge. While surely partially correct, these breathless assessments too often lose track of crucial questions about the complex processes of mutual adjustment by which older knowledge institutions adapt to emergent ones, and vice versa. Charmed by the novelty of the first date, they miss the complexity of the marriage that ensues: the dynamics of scale, time, and adjustment by which new practices emerge.

We think the time has come to reconceive our object(s) of interest around the idea of knowledge infrastructures.* To help configure this interest, we posed three themes for workshop participants to deliberate:

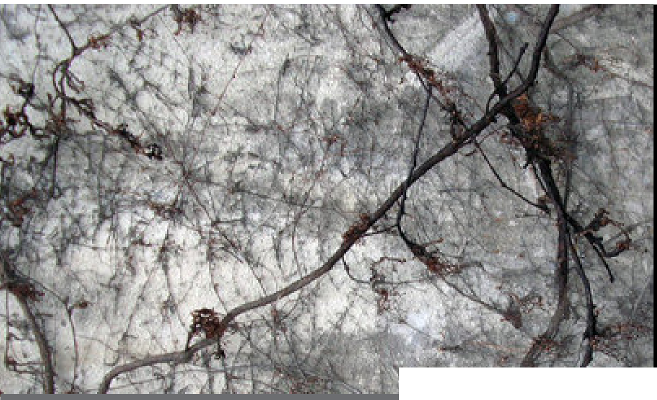
1. How are knowledge infrastructures changing?
2. How do knowledge infrastructures reinforce or redistribute authority, influence, power and control?
3. How can we best study, know, and imagine today's (and tomorrow's) knowledge infrastructures?

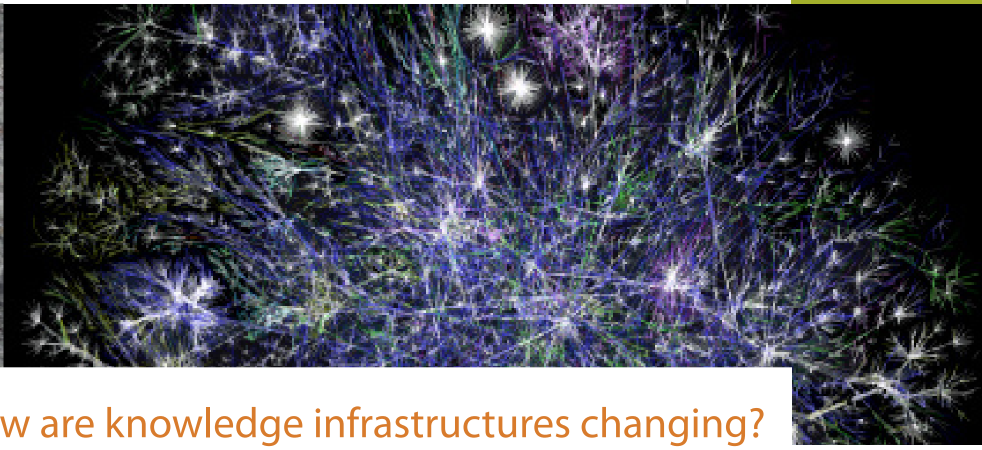
The remainder of this report summarizes the intense discussions that ensued. It highlights the issues we believe will be most salient for at least the next ten years — i.e., a set of research questions urgently in need of study — and discusses the tools and methods we will need.



---

* The workshop organizers deliberately decided to bypass the problem of definition, instead allowing the phrase "knowledge infrastructure" to serve as a suggestive provocation. This report adopts a similar strategy, though it provides numerous pointers to the burgeoning literature in infrastructure studies.





Theme 1: How are knowledge infrastructures changing?

What does it mean to "know" in an age of social networks, big data, interdisciplinary research, and new modes of access to "bigger," "wider," "longer," and "faster" information? How is knowledge now being generated, maintained, revised, and spread? How are open data, web publication, and commodity tools affecting concepts of expertise, processes of peer review, and the quality of knowledge?

Building on extensive literatures in science & technology studies, including previous work by members of this group (Edwards et al. 2007), Edwards (2010) defined knowledge infrastructures as "robust networks of people, artifacts, and institutions that generate, share, and maintain specific knowledge about the human and natural worlds." This framing aimed to capture routine, well-functioning knowledge systems such as the world weather forecast infrastructure, the Centers for Disease Control, or the Intergovernmental Panel on Climate Change. Under this definition, knowledge infrastructures include individuals, organizations, routines, shared norms, and practices.

Key to the infrastructure perspective is their modular, multi-layered, rough-cut character. Infrastructures are not systems, in the sense of fully coherent, deliberately engineered, end-to-end processes. Rather, infrastructures are ecologies or complex adaptive systems; they 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. The current situation for knowledge infrastructures is characterized by rapid change in existing systems and introduction of new ones, resulting in severe strains on those elements with the greatest inertia.

> *<< Knowledge is perpetually in motion. Today, what we call "knowledge" is constantly being questioned, challenged, rethought, and rewritten. >>*

The workshop concluded that at least the following phenomena require sustained attention:

KNOWLEDGE IN PERPETUAL MOTION. A transition is underway from what Weinberger (2012) calls "knowledge as a series of stopping points" — printed journal articles, books, textbooks, and

---

Climber plants, photograph by brunilde-stock on Deviant Art

Classic OPTE Map of the Internet, source www.opte.org/maps/

other fixed products — to a world where *knowledge is perpetually in motion. Today, what we call "knowledge" is constantly being questioned, challenged, rethought, and rewritten*. As Weinberger describes the current situation, we face a world of abundant information, hyperlinked ideas, permission-free resources, highly public interaction, and massive, unresolved disagreement. Individual expertise is (many argue) being replaced by the wisdom of crowds: noisy and endlessly contentious, but also rich, diverse, and multi-skilled. In part, this means that the divide between knowledge producers and knowledge consumers is increasingly and radically blurred. In such a world, the missions of educational institutions such as schools and colleges, research institutions such as laboratories and universities, and memory institutions such as libraries, archives, and museums are bleeding into each other more than ever before. New forms of collective discovery and knowledge production, such as crowdsourced encyclopedias, wikis of all sorts, shared scientific workflows, and citizen science are springing up within and across many academic disciplines (De Roure et al. 2011; De Roure et al. 2010; Goble & De Roure 2007; Shilton 2009; Shirky 2009, 2010; Takeda et al. 2013; Wade & Dirks 2009). The quality and durability of knowledge produced by such efforts remains uncertain, but their tremendous vigor and growing utility cannot be questioned.

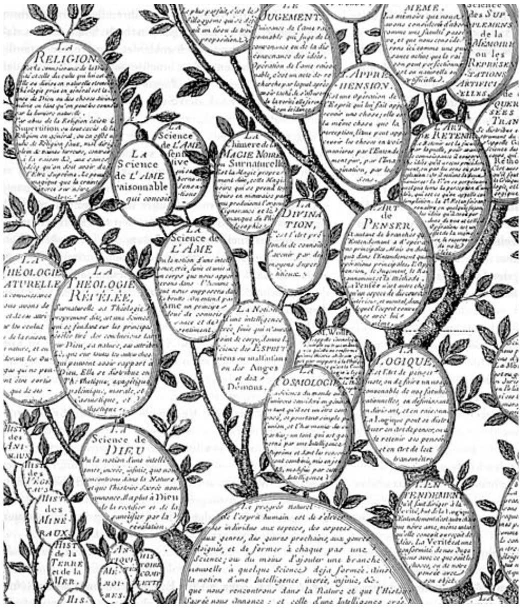
<< In a world of Skype, Google Hangouts, Twitter, YouTube videos, and highly developed visualization techniques, the roles of tacit knowledge and common ground are changing, and a renewal of our understanding is required. >>

SHIFTING BORDERS OF TACIT KNOWLEDGE AND COMMON GROUND. Consider that the study of knowledge by the social sciences and the humanities has been based on the same premises now being challenged by emerging forms. For example, several decades of scholarship in sociology and anthropology of knowledge established the difficulty of communicating local practices and understandings without face-to-face contact (Collins 1985; Collins & Pinch 1993). The phrase "distance matters" — because technology-mediated communication makes it more difficult to establish common ground — became a watchword in computer-supported cooperative work. Tacit knowledge and common ground were, and still are, regarded as major stumbling blocks to long-distance collaboration (Olson & Olson 2000; Olson et al. 2009). Yet an increasing amount of important knowledge work occurs under precisely these conditions; both technology and human skills are evolving to meet the challenge (Rosner 2012; Rosner et al. 2008; Vertesi 2012; Wiberg et al. 2012). *In a world of Skype, Google Hangouts, Twitter, YouTube videos, and highly developed visualization techniques, the roles of tacit knowledge and common ground are changing, and a renewal of our understanding is required* (Cummings et al. 2008).

COMPLEXITIES OF SHARING DATA ACROSS DISCIPLINES AND DOMAINS. Excitement continues to mount over new possibilities for sharing and "mining" data across scientific disciplines. Vast data repositories are already available to anyone who cares to use them, and many more are on the way. Yet data sharing begs urgent questions (Borgman 2012). In science, at least, the meaning of data is tightly dependent on a precise understanding of how, where, and when they were created (Bechhofer et al. 2010; Burton & Jackson 2012; Gitelman 2013; Ribes & Jackson 2013;

Vertesi & Dourish 2011). But *the rapid "commodification" of data — the presentation of datasets as complete, interchangeable products in readily exchanged formats — may encourage misinterpretation, over reliance on weak or suspect data sources, and "data arbitrage" based more on availability than on quality. Will commodified data lead to dangerous misunderstandings, in which scientists from one discipline misinterpret or misuse data produced by another?* How far can the standardization of data and metadata be carried, and at what scale? What new kinds of knowledge workers are needed to bridge the gaps, both technical and social, among the many disciplines called on to address major scientific and social issues such as climate change, biodiversity triage, or health care for an aging population? Can the reputation systems of science be re-tuned to recognize and compensate these vital, but too often invisible and unrewarded, workers?

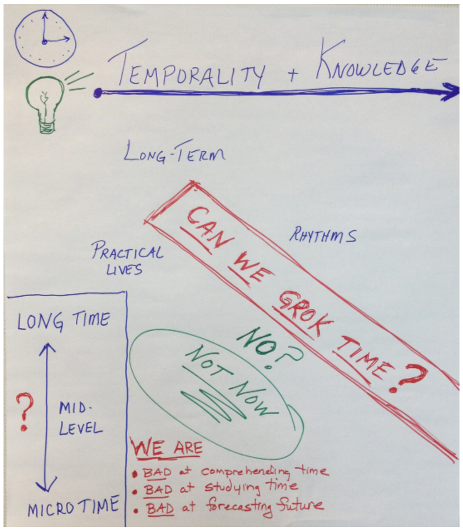
NEW NORMS FOR WHAT COUNTS AS KNOWLEDGE. Scientific data analysis increasingly uncovers significant and useful patterns we cannot explain, while simulation models too complex for any individual to grasp make robust predictions (e.g., of weather and climate change). *Will these phenomena add up, as some predict, to an "end of theory"* (Anderson 2008)? The question of how to evaluate simulation models — of whether they can be "validated" or "verified," and whether they require a fundamentally different epistemology than theory and experiment — had already been puzzling both scientists and philosophers for several decades (Giere 1999; Heymann 2010; Jackson 2006; Morgan & Morrison 1999; Norton & Suppe 2001; Oreskes et al. 1994; Petersen 2007; Sismondo 1999; Sundberg 2009, 2010a, 2010b; Suppe 2000). Data-driven science poses a similar, even harder problem of evaluation. *Do we "know" things if we cannot explain why they are true?* Whatever the case, norms for what can count as "knowledge" are clearly changing (Anderson 2008; Hey et al. 2009).*



MASSIVE SHIFTS IN PUBLISHING PRACTICES, LINKED TO NEW MODES OF KNOWLEDGE ASSESSMENT. Historically, knowledge institutions depended on costly, hierarchically organized forms of credentialing, certification, and publishing. These set severe limits not only on outputs (in the form of published articles, books, etc.), but also on who could count as a valid participant in knowledge assessment practices such as peer review. Today, these mechanisms are challenged on all fronts. Much less costly modes of publication permit the early release and broad dissemination of virtually all data and models used in science; one result is a broad-based movement toward publication practices that permit results to be readily reproduced, at least in the computational sciences (Stodden 2010a, 2010b, 2011). *Commodified data analysis tools and widely available software skills permit a much larger number of participants to analyze data and run models. Networked social forms permit many more participants to comment publicly on knowledge products, bypassing traditional credentialing and certification mechanisms* (De Roure et al. 2011; De Roure et al. 2010; Kolata 2013).

Fold-out frontspiece in volume 1 of Pierre Mouchon, Table analytique et raisonnée (Paris, Panckoucke, 1780)

CHALLENGES TO TRADITIONAL EDUCATIONAL INSTITUTIONS. Both research universities and teaching colleges face extraordinary challenges. For decades, costs to students have risen faster than inflation, while Coursera, open courseware, and online universities offer new, lower-cost alternatives. The majority of university students no longer attend 4-year residence programs. Many of those who do appear more motivated by the university as a rite of passage and a lifestyle than by learning itself, as reflected in numerous measures of student learning and the amount of time spent studying (Babcock & Marks 2010, 2011; Mokhtari et al. 2009). Classroom teaching competes directly with online offerings; professors are no longer seen as infallible experts, but as resources whose facts can be checked in real time. As institutions, research universities display patent-seeking behavior that makes them increasingly difficult to distinguish from corporations, and indeed corporate sponsorship and values have penetrated deeply into most universities. Some have been more effective than others at building firewalls between sponsors' interests and researchers to protect their objectivity, but no institution is immune to these challenges (Borgman et al. 2008). K-12 education faces related, but different challenges, as schools struggle to adapt teacher training, equipment, and teaching methods to the screen-driven world most children now inhabit. *Major benefits will accrue to institutions and students that find effective ways to meet these challenges — and doing so will require new visions of their place in larger infrastructures of knowledge, from national science foundations to corporate laboratories to educating new generations of researchers.*



NAVIGATING ACROSS SCALES OF SPACE AND TIME, AND RATES OF CHANGE. Given the layered nature of infrastructure, navigating among different scales — whether of time and space, of human collectivities, or of data — represents a critical challenge for the design, use, and maintenance of robust knowledge infrastructures. A single knowledge infrastructure must often track and support fluid and potentially competing or contradictory notions of knowledge. Often invisible, these notions are embodied in the practices, policies, and values embraced by individuals, technical systems, and institutions. For example, sustainable knowledge infrastructures must somehow provide for the long-term preservation and conservation of data, of knowledge, and of practices (Borgman 2007; Bowker 2000, 2005; Ribes & Finholt 2009). In the current transformation, sustaining knowledge requires not only resource streams, but also conceptual innovation and practical implementation. *Both historical and contemporary studies are needed to investigate how knowledge infrastructures form and change, how they break or obsolesce, and what factors help them flourish and endure.*

STANDARDS AND ONTOLOGIES. A quintessential tension surrounds the deployment of standards and ontologies in knowledge infrastructures. Fundamentally, it consists in *the opposition between the desire for universality and the need for change.*

Robust hypotheses require information in standardized formats. Thus the spread of a particular disease around the world cannot be tracked unless everyone is calling it the same thing. At the same time, medical researchers frequently designate new diseases, thus unsettling the existing order. For example, epidemiologists have sought to track the phenomenon of AIDS to periods predating its formal naming in the 1980s (Grmek 1990; Harden 2012). However, using historical medical records to do so has proven difficult because prior record-keeping standards required the specification of a single cause of death, precluding recognition of the more complex constellation of conditions that characterize diseases such as AIDS.

How might one solve this problem (if it is solvable at all)? One could review the old records and try to conjure them into modern forms. This could work to an extent; some fields, such as climate science, routinely investigate historical data before adjusting and re-standardizing them in modern forms to deepen knowledge of past climates (Edwards 2010). Yet this is possible largely because the number of records and their variety is relatively limited. In many other fields such a procedure would be extremely difficult and prohibitively expensive. Alternatively, one could introduce a new classificatory principle, such as the Read Clinical Classification, which would not permit that kind of error to propagate. Here too, due to the massive inertia of the installed base, it would cost billions of dollars to make the changeover. On top of that, it would complicate backward compatibility: every new archival form challenges the old (Derrida 1996). *In practice, this adds up to very slow updating of classification standards and ontologies, marked by occasional tectonic shifts.*

<< In practice, this adds up to very slow updating of classification standards and ontologies, marked by occasional tectonic shifts. >>

Today, hopes for massively distributed knowledge infrastructures operating across multiple disciplines consistently run headlong into this problem. Such infrastructures are vital to solving key issues of our day: effective action on biodiversity loss or climate change depends on sharing databases among disciplines with different, often incompatible ontologies. If the world actually corresponded to the hopeful vision of data-sharing proponents, one could simply treat each discipline's outputs as an "object" in an object-oriented database (to use a computing analogy). Discipline X could simply plug discipline Y's outputs into its own inputs. One could thus capitalize on the virtues of object-orientation: it would not matter what changed within the discipline, since the outputs would always be the same. Unfortunately, this is unlikely — perhaps even impossible — for both theoretical and practical reasons (Borgman et al. 2012).

An "object-oriented" solution to these incompatibilities is theoretically improbable because the fundamental ontologies of disciplines often change as those disciplines evolve. This is among the oldest results in the history of science: Kuhn's term "incommensurability" marks the fact that "mass" in Newtonian physics means something fundamentally different from "mass" in Einsteinian physics (Kuhn 1962). If Kuhnian incommensurability complicates individual disciplines, it has even larger impacts across disciplines. A crisis shook virology, for example, in the 1960s when it was discovered that "plant virus" and "animal virus" were not mutually exclusive categories. Evolutionary biology suffered a similar, and related, crisis when it was learned that some genes could jump between species within a given genus, and even between species of different genera

(Bowker 2005). Suddenly, disciplines that previously had no need to communicate with each other found that they had to do so, which then required them to adjust both their classification standards and their underlying ontologies.

In practice, an object-oriented solution to ontological incompatibilities is unlikely because *we have not yet developed a cadre of metadata workers who could effectively address the issues, and we have not yet fully faced the implications of the basic infrastructural problem of maintenance.* We do know that it takes enormous work to shift a database from one medium to another, let alone to adjust its outputs and algorithms so that it can remain useful both to its home discipline and to neighboring ones. Thus three results of today's scramble to post every available scrap of data online are, first, a plethora of "dirty" data, whose quality may be impossible for other investigators to evaluate; second, weak or nonexistent guarantees of long-term persistence for many data sources; and finally, inconsistent metadata practices that may render reuse of data impossible — despite their intent to do the opposite.

We expect our knowledge infrastructures to permit effective action in the world; this is the whole impulse behind Pasteur's Quadrant or Mode II science (Gibbons et al. 1994; Jackson et al. 2013; Stokes 1997). And yet, in general, scientific knowledge infrastructures have not been crafted in such a way as to make this easy. What policymakers need and what scientists find interesting are often too different — or, to put it another way, a yawning gap of ontology and standards separates the two. Consider biodiversity knowledge. In a complex series of overlapping and contradictory efforts, taxonomists have been trying to produce accounts of how species are distributed over the Earth. However, the species database of the Global Biodiversity Information Facility, which attempts to federate the various efforts and is explicitly intended for policy use, does not produce policy-relevant outputs (Slota & Bowker forthcoming).The maps of distribution are not tied to topography (necessary to consider alternative proposals such as protecting hotspots or creating corridors), they give single observations (where what is needed is multiple observations over time, so one can see trends), and for political reasons, they do not cover many parts of the planet (which one needs in order to make effective global decisions). Similarly, in the case of climate change, for decades the focus on "global climate" — an abstraction relevant for science, but not for everyday life — has shaped political discourse in ways that conflicted with the local, regional, and national knowledge and concerns that matter most for virtually all social and political units. Climate knowledge infrastructures have been built to produce global knowledge, whereas the climate knowledge most needed for policymaking is regional, culturally specific, and focused on adaptation (Hulme 2009).

<< We have not yet developed a cadre of metadata workers who could effectively address the issues, and we have not yet fully faced the implications of the basic infrastructural problem of maintenance. >>

Theme 2: How do knowledge infrastructures reinforce or redistribute authority, influence, and power?

What new forms of organization and community are emerging? What power relations do they rely on, create, or destroy? Who wins and who loses as knowledge infrastructures change?

New knowledge infrastructures hold great promise, and they may help address key issues of public import. But *knowledge infrastructures also face limits, create tensions, and raise concerns. These must be addressed early and often.* Many of these limits and concerns are simply the flip side of the potentials sketched above. Systems that develop and elevate new forms of knowledge may demote or undermine others. The growth of new ways of knowing may come at the expense of old or alternative ones. Efforts to expand access to knowledge for some groups may curtail or limit the effective access of others.

A classic example of this dynamic can be found in the "two systems" problem noted by early digital government researchers. In the 1990s and early 2000s, as governments moved to implement new digital record systems and online interfaces with the public, they were faced with the need to maintain and staff a parallel world of paper forms and public kiosks for the sizeable component of the population that remained offline — who, in many social service contexts, were in fact the chief clients of the records and services in question. As a result, early digital government investments usually functioned as add-ons rather than substitutes for existing services, and usually ended up costing more than older paper-based systems — even though justified initially on cost reduction grounds (Chongthammakun & Jackson 2012). Similar phenomena have been observed in the context of electronic medical records (Monteiro et al. 2012). More recently, governments have begun doing away with these parallel paper-based systems, leaving those without meaningful access to digital media even worse off than before.

<< knowledge infrastructures also face limits, create tensions, and raise concerns. These must be addressed early and often. >>

In the world of science, new forms of knowledge infrastructure may disadvantage and devalue older forms of knowledge production, even while producing genuinely exciting advances. For example, new sensor networks are replacing ecological fieldwork, while visions of "instrumenting the ocean" are supplanting traditions of the oceanographic cruise (Borgman et al. 2006; Borgman et al. 2012; Borgman et al. 2007; Jackson & Barbrow 2013; Mayernik et al. 2013; Wallis et al. 2007).

"Reinforce Forest", photograph by b0cah on flickr

Along the way, these new infrastructures are disrupting social and material relationships that formerly sustained human interest and commitment to particular modes of scientific life.

As these two quick examples suggest, knowledge infrastructures – past, present, and future – must be grasped in their full range of effects and dimensions. Analysts must recognize that valuable gains in some registers often entail losses, costs, and adjustments in others. This requires thinking beyond the instrumental languages of utility and function, which tend to cast knowledge infrastructures as neutral instruments whose positive and negative effects lie solely in the way they are operationalized and used. Like technology in general, knowledge infrastructures "are neither good nor bad, nor are they neutral" (Kranzberg 1986). As Palfrey and Gasser (2012) put it, even as societies have constructed vast information and communications infrastructures to "enhance connectivity and enable the flow of information," they have failed to build a corresponding normative theory of interconnectivity that would help define the social purposes and values we intend our infrastructures to serve.

<< Knowledge infrastructures must be understood in their entirety, as hybrids that join and rely on elements too often separated under the (bogus) headings of "technical" and "social." >>

The following point united workshop participants across multiple areas and backgrounds: despite their frequently significant technological components, *knowledge infrastructures must be understood in their entirety, as hybrids that join and rely on elements too often separated under the (bogus) headings of "technical" and "social."* Programmatic efforts to improve science and other knowledge infrastructures have frequently prioritized investments in technical systems over research on how to effectuate equally crucial cultural, social, and organizational transformations. The imbalance has been great enough to require periodic efforts to "bring the people back in" — for example, work by Lee et al. (2006) emphasizing the "human infrastructure of cyberinfrastructure," and this group's 2007 workshop emphasizing the irreducibly "sociotechnical" character of scientific cyberinfrastructure (Edwards et al. 2007; Jackson et al. 2007; Lee et al. 2006). Similarly, Pipek and Wulf (2009) developed the concept of "infrastructuring" in reference to the ongoing co-design of infrastructures-in-use that takes place as new systems are adapted to interface with existing ones through combinations of improvisation, work practices, and continuing innovation by both designers and users. Monteiro et al., following Pollock & Williams, Karasti, and others, have argued that the very concept of "design" as a local, punctual activity of system developers needs to be rethought, at least in the context of large-scale enterprise software infrastructures (Monteiro et al. 2012; Pollock & Williams 2010; Pollock et al. 2009).

Thinking the social and technical together helps in the analysis of costs and benefits attending new infrastructure development. It may also open up the effective range of action available to infrastructure designers and developers, in particular in cases where solely technical programs of development run up against immediate and insuperable 'social' difficulties, and vice versa. It may also help to define and characterize the success of many of those forms of infrastructure that we find most compelling and sometimes surprising, and which we sometimes hold out as models or examples of change. Transformative infrastructures cannot be merely technical; they must engage fundamental changes in our social institutions, practices, norms and beliefs as well. For that reason, many scholars have dropped the dualistic vocabulary of "technical" and "social" altogether as anything other than a first order approximation, replacing those terms with concepts such as collectives (Latour 2005), assemblages (Ong & Collier 2005), or configurations (Suchman 2007) in which "technical" and "social" figure as no more than moments or elements within what are at their core heterogeneous and uneven but ultimately unified wholes. The empirical veracity of this approach becomes apparent when we apply it to any of the examples cited above: try to tell yourself a "technology only" or "social only" story of knowledge infrastructures, and the real-world impossibility of the old conceptual split becomes immediately apparent. Brunton's *Spam: A Shadow History of the Internet* (2013) provides a fascinating case study in the co-evolution of technical systems, communities, and social norms.

Once the world of knowledge infrastructures is reconnected in this way, it becomes possible to tease out some of the limits, tensions, and problems associated with infrastructural development in a more direct and fruitful way. To begin with the most obvious point, *knowledge infrastructures often carry significant distributional consequences, advancing the interests of some and actively damaging the prospects of others.* For example, moves toward expanded data sharing may simultaneously devalue or commodify certain kinds of data production – for example, the labor-intensive collecting practices and site-specific expertise required to produce 'raw' data in field ecology (Jackson & Barbrow 2013; Jackson et al. 2013; Ribes & Jackson 2013) or the intimate familiarity with specific sensors and robots needed to gather reliable data in interplanetary space exploration (Vertesi & Dourish 2011). Similarly, attempts to open access to education by posting teaching materials and recorded lectures may devalue the work and expertise that goes into the thoughtful production of lectures, syllabi, and other teaching materials. The celebration or move towards digital modes of production may undermine the practice and recognition of long-standing craft traditions, from artistic production to fishing to biological research (Pierce et al. 2011; Rosner 2012; Rosner et al. 2008). Finally, more efficient forms of information exchange have redistributed labor and eliminated whole categories of workers, from telegraph messenger boys (Downey 2001, 2002, 2003, 2008) to the once-vast secretarial ranks of large organizations (Hedstrom 1991).

<< Knowledge infrastructures often carry significant distributional consequences, advancing the interests of some and actively damaging the prospects of others. >>

Our point is not that these are always bad things, nor that we should abandon the language of sharing, access, or any of the other real and positive potentials of new knowledge infrastructure development. Rather, we argue that the consequences of change are rarely socially, culturally, or economically neutral. Neglecting this point has obvious and troubling normative implications, including the callousness towards distributional outcomes sometimes associated with revolutionary or reformist programs. Yet lack of attention to the redistribution of labor has immediate practical implications, in the form of opposition, resistance, work-arounds, non-adoption and a thousand other effects that show up as brakes or challenges to new infrastructure development when such consequences are neglected or ignored. The study and practice of knowledge infrastructures therefore require new languages of distributive justice that can map change to consequence in more ethical and effective ways.

Beyond their effects on individual actors and interests, knowledge infrastructures exert effects on the shape and possibility of knowledge in general – and once again, these effects are neither neutral nor uniformly positive. *One of the most powerful and exciting effects of new knowledge infrastructures is their ability not only to answer existing questions, but to make new questions thinkable.* Indeed, modern experimental science owes its very existence to the infrastructure of laboratories, scientific societies, and journals developed in the 17th century, described by Shapin and Schaffer as a set of "technologies" for "virtual witnessing" (Shapin & Schaffer 1985). Other examples include the subdisciplines of cognitive psychology and artificial intelligence, which developed directly from encounters with computers in the 1950s (Edwards 1996), and the transformations that occurred in geography, cartography, meteorology, and many other fields when global satellite observing systems became available to them. It is this shift in imagination and possibility, more than anything else, that marks the great transitions in the history of science and human knowledge, whether conceptualized after the manner of Kuhn's (1962) "paradigms," Lakatos' (1970) "research programmes," or Foucault's (1971) "epistemes". At such moments, we see the relationship between knowledge and the infrastructures that support it as both intimate and co-productive (Edwards 2010; Jasanoff 2004). Knowledge infrastructures do not only provide new maps to known territories – they reshape the geography itself.

<< One of the most powerful and exciting effects of new knowledge infrastructures is their ability not only to answer existing questions, but to make new questions thinkable. >>

This too, however, is not a neutral feature. As knowledge infrastructures shape, generate and distribute knowledge, they do so differentially, often in ways that encode and reinforce existing interests and relations of power. Evidence for this may be found in the long-standing differential attention of medical research to women's vs. men's cancers, or diseases of the rich vs. diseases of the poor (Epstein 2008). As Bowker (2000) points out, it also accounts for the relative overrepresentation of research on "charismatic megafauna" (e.g. cuddly pandas or expressive, human-like chimps) vs. other biota (e.g. blue green algae) that are arguably more central to ecological process, but fare poorly as mascots for conservationism. At scale, the effect of these choices may be an aggregate imbalance in the structure and distribution of our knowledge.

This effect goes beyond the lack of development of some areas relative to others — a point which, taken on its own, fits comfortably within what we might call the "dark continent fallacy" of knowledge and ignorance. Under this fallacy, ignorance (or non-knowledge) is simply the absence of knowledge: a site, phenomenon, or set of questions that we haven't yet been able or thought to investigate, as darkness is the absence of light. But as recent history and sociology of science has begun to explore, the relationship between knowledge and non-knowledge may not be as simple or innocent as that. New ways of knowing and new forms of knowledge infrastructure may do more than add to existing stocks of knowledge: they may rework them, reorder our sense of value and structure in the world, write new ontologies over old ones. In the process, whole classes of questions, phenomena and forms of knowledge may be lost or rendered unthinkable. Any effective sociology of knowledge (or knowledge infrastructure) ought therefore to provide some account of its opposite: the accidental and systematic, means by which non-knowledge is produced and maintained. A group of science studies scholars led by Robert Proctor has christened this effort "agnotology," i.e., "the systematic study of ignorance" (Proctor & Schiebinger 2008). From a different direction, neurobiologist Stuart Firestein has called on science to reorient its conversations and public presentation from knowledge to ignorance (Firestein 2012).

Two examples will serve to flesh out these broad claims in specific cases. First, new modes of Internet-supported citizen science — headlined by leading initiatives such as GalaxyZoo, Zooniverse, FoldIt, Project Budburst, and the Cornell Laboratory of Ornithology's eBird program — are often cited as harbingers of new strategies for practicing science at scale, while engaging, educating, and energizing a science-friendly public (Kelling et al. 2012). But they also raise questions and tensions around the nature of participation, expertise, and the relative authority of credentialed experts vs. lay publics in the production of scientific data and knowledge. While new Internet-supported collection, reporting, and analysis activities may help with sourcing and processing more scientific data, they may raise problems of quality control and "bad" data. For their part, lay participants may find their opportunities for input, learning, and progression beyond simple manual processing too limited, and long for opportunities to engage the research process in a more substantive way. To keep the new fountain of citizen assistance flowing, scientists will need to find effective ways of responding to these aspirations. This may mean reworking long-standing traditions of public engagement in their fields — some of which began decades or centuries ago with professional bids to separate the expertise of credentialed science from the contributions of amateur publics.

Second, the recent phenomenon of Massively Open Online Courses (MOOCS) is extending the reach of university teaching to remote participants numbering from the hundreds to the tens of thousands. In some cases, these activities are promoted by established institutions of higher education, sometimes building on other open education initiatives, such as MIT's Open Courseware). But in others, MOOCs are offered by independent businesses, which piggyback on the prior experience, institutional reputation and/or day jobs of university professors (as in the case of the Coursera initiative, started as a business by former Stanford professors in April 2012). These

---

Suzallo Library, Seattle, Washington, photograph by wonderlane on flickr efforts have attracted well-deserved attention: at their best, MOOCs might help to mitigate the exclusivity and expense of high-quality post-secondary education. At the same time, it is at best unclear how well MOOCs or similar forms can communicate knowledge to students in the absence of the face-to-face group experiences and systematic educational approaches universities have traditionally supplied. In the meantime, researchers, educators, and university administrators are scrambling to figure out their positions with respect to this emerging element of knowledge infrastructure, which raises major new questions and tensions. How should faculty-produced teaching tools be credited reputationally and rewarded financially? How should adjunct and part-time instructors working in MOOCs be compensated? What happens when universities and for-profit businesses compete for the same pool of remote students, untethered from classrooms located in a particular place? How can we evaluate student work effectively and accurately at enormous scales? Is it possible to safeguard against cheating and misappropriation of others' work? All of these questions tie into future judgments about the value and limits of distinct institutional brands, including how official credit and degrees can be meaningfully conferred.

Facing the redistributions of authority, influence, and power that come with changing knowledge infrastructures requires new approaches on many fronts and scales. We close this section with three possibilities that seemed particularly promising to workshop participants:

APPROACH THIS PROBLEM AS A DESIGN OPPORTUNITY. The issue of sharing knowledge across different social worlds and conflicting conceptual frameworks has been extensively treated in the history and sociology of science and technology, leading to such widely used ideas such as "boundary objects," "trading zones," and "actor networks," as well as to a well developed understanding of how such sharing works in practice (Callon & Latour 1981; Galison 1996, 1997; Latour 1983, 2004, 2005; Latour & Woolgar 1979; Star & Griesemer 1989; Star & Ruhleder 1996). Similarly, the history and sociology of standards has produced considerable insight into how large communities of heterogeneous stakeholders come to (often rough) agreement on shared norms, practices, and technical systems (Bartky 1989; Blanchette 2012; Busch 2011; Egyedi 2001; Fujimura 1992; Hanseth et al. 2006; Russell 2006; Sundberg 2011). A design community versed in these literatures might transform such concepts into design principles and practices — but doing so would require deliberate efforts to "scale up" the generally lower-level focus of design thinking (Le Dantec & DiSalvo 2013; Monteiro et al. 2012; Pipek & Wulf 2009).

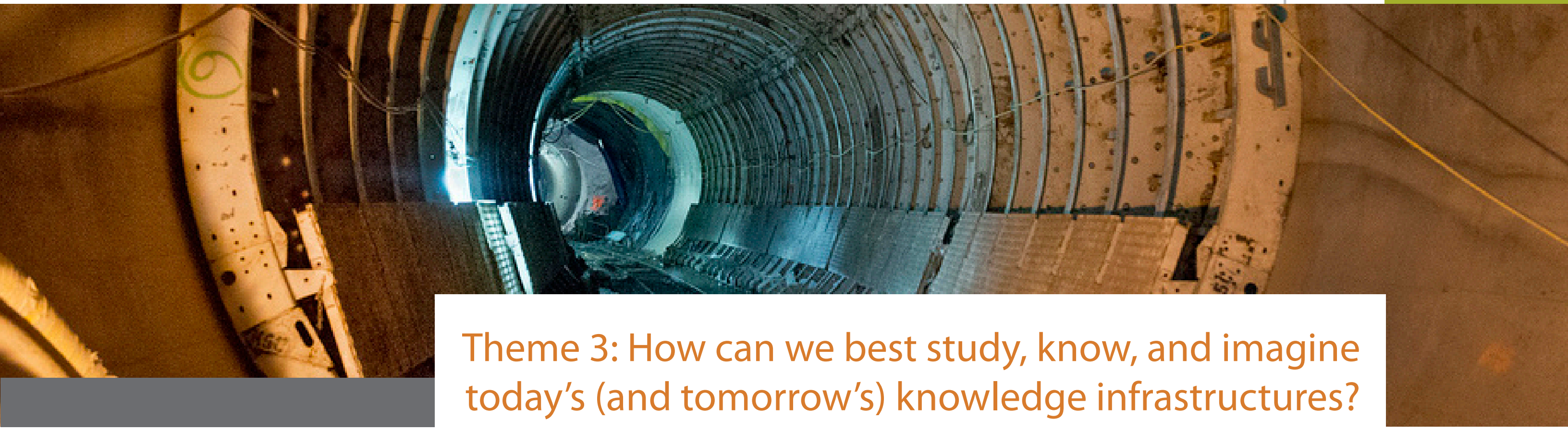
CREATE A PROFESSIONAL CADRE AT THE INTERFACE BETWEEN SCIENTISTS AND SOFTWARE DEVELOPERS. So often neither side understands the vocabulary, the work practices, and/or the products of the other well enough to communicate effectively. Scientific software is too frequently

---

Photograph by Mark Brannon

a byproduct of "real" scientific work, with code written by scientists who often do not even know about, let alone apply, best practices for software work (Clune & Rood 2011; Easterbrook & Johns 2009; Howison & Herbsleb 2011; Ribes et al. 2012). Meanwhile, software developers are too quick to construct "ontologies" — a term of art in their social world — that can rapidly diverge from scientific usage. Professionals who understand domain science, software best practices, and the issues of social interfacing would be enormously beneficial in almost any area of knowledge. One possible source of such a cadre is the growing iSchool movement, where curricula and research on "data science" are rapidly emerging.

USE THE PARTICIPATORY DESIGN MOVEMENT AS A MODEL FOR WORK PRACTICES. Participatory design started when Scandinavian countries developed a requirement that new IT introduced into the workplace must be developed in consultation with workers who would use it. Workers could not effectively express their IT needs and concerns, while software developers often failed to fully understand their work processes. The flourishing design community that has developed sits between both (Kensing & Blomberg 1998; Muller 2007). We need the same function for the development of knowledge infrastructures (Shilton et al. 2008).



Theme 3: How can we best study, know, and imagine today's (and tomorrow's) knowledge infrastructures?

What methods and techniques still work, and which new methods and techniques show promise — especially at scale? How should we organize future studies? What assumptions and practices should we abandon?

Considering knowledge infrastructures historically reveals what might otherwise appear as necessary features to be, instead, historical creations which could have followed other paths. For example, the nineteenth century witnessed a massive and global shift in modes of classification (to the genetic form, favored by Darwin) — and yet this occurred without generalized academic conferences on classification. Instead, each discipline painfully learned the same lessons in isolation. In retrospect, the degree of redundant effort seems staggering. Clearly the same pattern is presently underway, with many disciplines each struggling to find its own path through the maze of related issues, including appropriate cyberinfrastructure, trustworthy and enduring institutions, data management practices, and handoffs with other disciplines. One role for scholars of knowledge infrastructures, then, might be to help decrease the amount of effort this struggle demands, for example by organizing and synthesizing collective conversations about how to shape infrastructures for knowledge work in the 21st century.

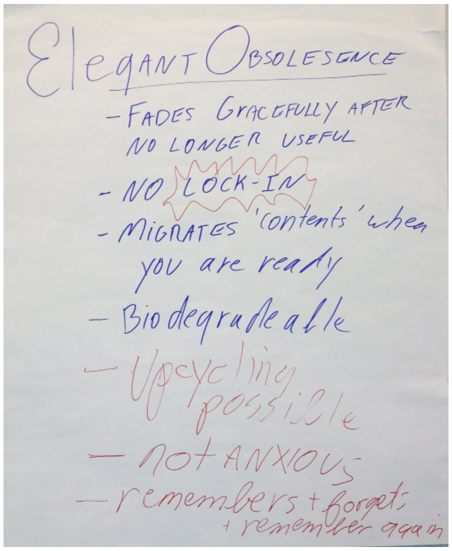
Workshop participants emphasized that we cannot remain simple bystanders to the current transformations. Nor should we be mere critics of the emerging inequalities and tensions described in Theme 2 of this report. Instead, our task is to co-design new infrastructures, and refashion old ones, with scientists, knowledge institutions, and policymakers as our partners. Put another way, we can play a key part in today's grand challenge: to debalkanize scholarship by assembling a methodological repertoire that can match the geographic and temporal scale of emerging knowledge infrastructures. This is an exhilarating possibility. Imagine what might have happened if scholars of the 15th and 16th centuries could have experimented directly with the sociotechnical reconfigurations that accompanied the advent of the printing press — as we can do today.

Our call for methodological and collaborative innovation is best explained via an analogy in the natural sciences. Twenty years ago, the average ecologist worked on a patch of land no larger than a hectare, typically for a few months or a year, gathered data over a thirty-year career, published results, and then gradually lost the data. With the creation of the Long Term Ecological Research Network (LTER), the National Science Foundation began to change the nature of research. Today,

---

East Side Access Tunnel, Metropolitan Transit Authority, photograph by Patrick Cashin at a number of sites nationally and in consonance with international projects, ecologists are able to look beyond the scale of a field and timeframe of a career: they now have the prospect of studying ecology and climate locally, nationally, globally, and over spans of time that more closely match those of ecological change.

How did this happen? In the last twenty years, new sensor grids have come to cover the oceans, land, sky and space. These technologies did not solve the question of scaling by themselves; instead, they posed new problems, as streams of data from extremely heterogeneous sources poured into the hands of scientists (Courain 1991; Edwards 2010; Hey et al. 2009). Standardizing data has proven to be a crucial activity in scaling up the sciences, but it is never easy and rarely, if ever, complete (Bowker 2000; Edwards et al. 2011; Gitelman 2013). While preservation has been recognized as an issue (Blue Ribbon Task Force on Sustainable Digital Preservation and Access 2010), no general response to long-term preservation of datasets exists in any branch of the sciences; instead, we find a conflicted field of partial solutions ranging from supercomputer centers to university libraries (Borgman 2007; Bowker 2005). Preserving the meaning of data is a human affair, requiring continuous curation. For these reasons, managing and preserving ecological data for the long term ultimately required new organizational forms. LTER represents the beginning, not the end, of that transformation.



Elegant Obsolesence

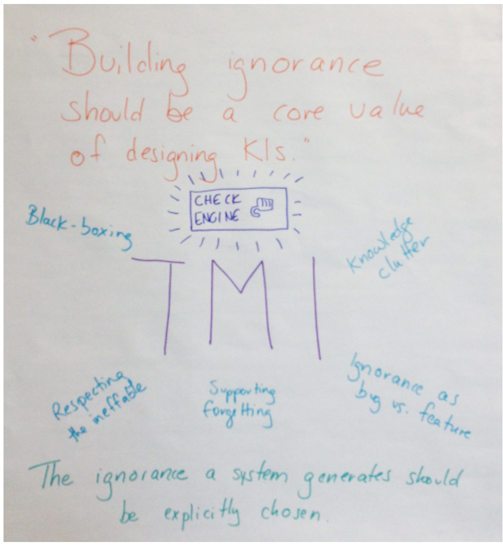
- Fades gracefully after no longer useful
- No lock-in
- Migrates 'contents' when you are ready
- Biodegradeable
- Upcycling possible
- not anxious
- remembers + forgets + remember again

We advocate a similar revolution in the study of knowledge infrastructures, using the lens of what Stewart Brand has called "the long now" (Bowker et al. 2010; Brand 1999; Ribes & Finholt 2009). The need for thinking in stretches of years to decades is quite apparent. Paul David's classic study on the "productivity paradox" of computing showed that introducing computers into the workplace did not immediately yield the productivity gains promised. In fact, productivity declined for twenty years before moving upwards. The cause, he argued, was that it took about 20 years to "think" the new technology: to move from using the computer as a bad, very expensive typewriter to realizing the potentials of new ways of working, which could happen only after a substantial period of social, cultural, organizational and institutional adjustments (David 1990; Landauer 1995).

With the advent of the Internet, we changing our knowledge generation and expression procedures root and branch. Yet currently we remain bound to the book and article format and to the classic nineteenth century technology of files and folders. It took well over 200 years for printed books to acquire the intellectual armature we now consider intuitive (such as the index, table of contents, bibliography, footnotes, and generally agreed rules on plagiarism). Even page numbers were once an innovation. Infrastructure researchers need a form of analysis that is actually responsive to the scale, scope and rhythms of the changes we are studying. Yet we are caught in the same cycle as the early ecologists mentioned above: our projects for studying social change come in three to five year chunks, in projects usually limited to three to five sites. Unlike the quantitative social sciences, which have benefitted enormously from now-vast stores of accumulated demographic, economic, and polling data, the qualitative social sciences have accumulated relatively little data across the years — and particularly across sites of research or across researchers. We reinvent the wheel with each investigation.

How can the qualitative social sciences accumulate, compare, and share data? Potential solutions exist. We present seven interlocking steps to meet the challenges for the future of sociotechnical studies. Together these make up our vision for an institution supporting long-term and large-scale qualitative research:

CREATE AND NOURISH MECHANISMS FOR LARGE-SCALE, LONG-TERM RESEARCH. We need to go beyond one-off projects to develop systems and standards for collecting, curating and using similar kinds of data, while simultaneously protecting subjects' identities and interests. Similarly, we need to build mechanisms to build and nourish larger, far more persistent research teams than the short-term, project-by-project work currently (and historically) typical of qualitative research. Organized research efforts at the scale of NSF science & technology centers would be a start — and funder investments on that scale could provide a powerful signal of need and reward —but innovation at all institutional levels and across disciplines will also be critical.



"Building ignorance should be a core value of designing KIs."

The ignorance a system generates should be explicitly chosen.

BUILD INTERDISCIPLINARY COLLABORATIONS ACROSS NATURAL AND SOCIAL SCIENCES. Sociotechnical phenomena do not rest within the domain of a single discipline or research approach. For example, climate change is simultaneously a matter of individual action and state policy, of technological innovation and economic reorganization. It demands the participation of social science but stretches well beyond it, requiring collaboration with ecological, hydrological, and biological scientists. Integrated assessment modeling — a popular and powerful tool for studying climate change impacts and adaptation —desperately needs better, more constructive contributions from the qualitative disciplines (Beck 2010; Hulme 2009, 2010; Lahsen 2010; van der Sluijs et al. 2008). This insight is far from new, but the fruits of previous integrative efforts have been modest; real innovation in knowledge infrastructures is needed.

DEVELOP COMPARATIVE ANALYSIS TECHNIQUES FOR STUDYING LARGE-SCALE, LONG-TERM DATA. Comparison across cases is among the most revealing qualitative research methods, encouraging the identification of crucial similarities and differences as well as enabling generalization. The key to comparison is sharing data across teams of investigators. This means investing in the creation of comparable data, i.e. data that are properly documented to facilitate sharing.

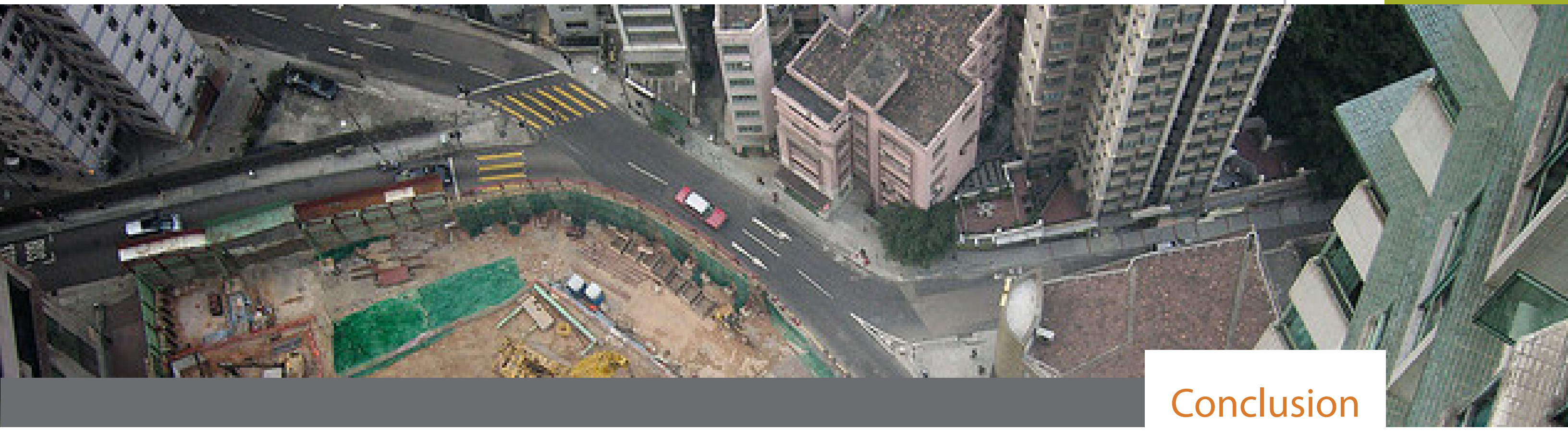
CREATE SUSTAINABLE, SHAREABLE DATA ARCHIVES. We must explore ways to federate the data collected over multiple investigative projects. Researchers need to publish their data alongside their articles, as they are in the natural sciences and economics today. Only in this way can researchers discern trends happening beyond their noses, long as these may be. Significant confidentiality issues exist, and should be addressed through the creation of new kinds of consent form and anonymization procedures). The Human Relations Area File and the Pew Research Center's Internet and American Life Project are two of the few extant examples of shared qualitative, long-term data; these models should be emulated and extended.

BUILD BETTER SOFTWARE FOR QUALITATIVE WORK. The infrastructure of knowledge infrastructures research has not kept up with the ambitions of this emerging area. Tools for collecting and organizing qualitative data remain tedious, fragile, and intended for small-scale efforts. As examples, consider NVivo and AtlasTI, the best-developed such tools. Although their current incarnations claim to support teams of researchers, anyone who has worked with them will know that the single-investigator paradigm continues to dominate their function; at best, each can support a handful of investigators working simultaneously. Each claims to support "large-scale" analysis, but quickly becomes unusable when handling more than a few hundred documents. The result is that even after decades of development, project after project continues to confront, and often to fail at, this challenge.

INTEGRATE QUALITATIVE WORK WITH STATISTICAL TECHNIQUES AND SOCIAL NETWORK ANALYSIS. The strengths of qualitative research (detailed, in-depth, meaning-oriented investigations) must be combined with those of quantitative and semi-quantitative approaches, such as social network analysis, whose strengths are scope and summation. This kind of integration has proven very powerful in the field of history (through the work of the Annales school, such as Fernand Braudel and Emmanuel le Roy Ladurie), yet it remains unusual. No software of which we are aware has effectively surmounted this challenge, though the "controversy mapping" tools under development at Bruno Latour's Médialab (Sciences Po, Paris) show promise. *

IMAGINE NEW FORMS OF CYBERSCHOLARSHIP. In the social sciences, we continue to use the computer as a glorified typewriter. Some remarkable experiments, often in conjunction with new media artists, have demonstrated new possibilities (see, for example, the multi-modal journal Vectors). However, these remain one-off ventures and generally suffer from the marked absence of funding for new forms of expression. When we begin to actively scale up qualitative social science, we will have to deploy the data storage, visualization, hypertext, and collective-creation possibilities of the web and social media. Further, we must, as a community, develop new tools for textual analysis that match the availability of electronic data. The digital humanities are already making remarkable strides in this area.

Body of Knowledge, by Jaume Plensa, photograph by Dontworry on wikimedia



Conclusion

Knowledge infrastructures are robust internetworks of people, artifacts, and institutions which generate, share, and maintain specific knowledge about the human and natural worlds. Like all infrastructures, they are composed of many systems and networks, each with its own unique dynamics. Because shared, reliable knowledge is among human society's most precious resources, the institutional elements of knowledge infrastructures – such as universities, libraries, and scientific societies – have typically adopted conservative, slow-changing forms. Yet recently key elements of knowledge infrastructures, especially information technologies and communication practices, have changed very rapidly, creating a growing sense of disarray and disjuncture between established forms and new and exciting, but unproven, possibilities. This report argues for the need to consider knowledge infrastructures as wholes, rather than focusing only on their most rapidly evolving elements. It poses a series of challenges and unresolved questions as the basis for a new area of research, practice, and design. These include the changing status of expertise as knowledge becomes more open to contestation from all quarters, the shifting borders of tacit knowledge and common ground, the unrecognized complexities of sharing data across disciplines and domains, and massive shifts in publishing practices linked to new modes of knowledge assessment. The new knowledge ecologies will necessarily involve transformations of the research process: traditional institutions will adapt or die; new forms will come into being.[*]

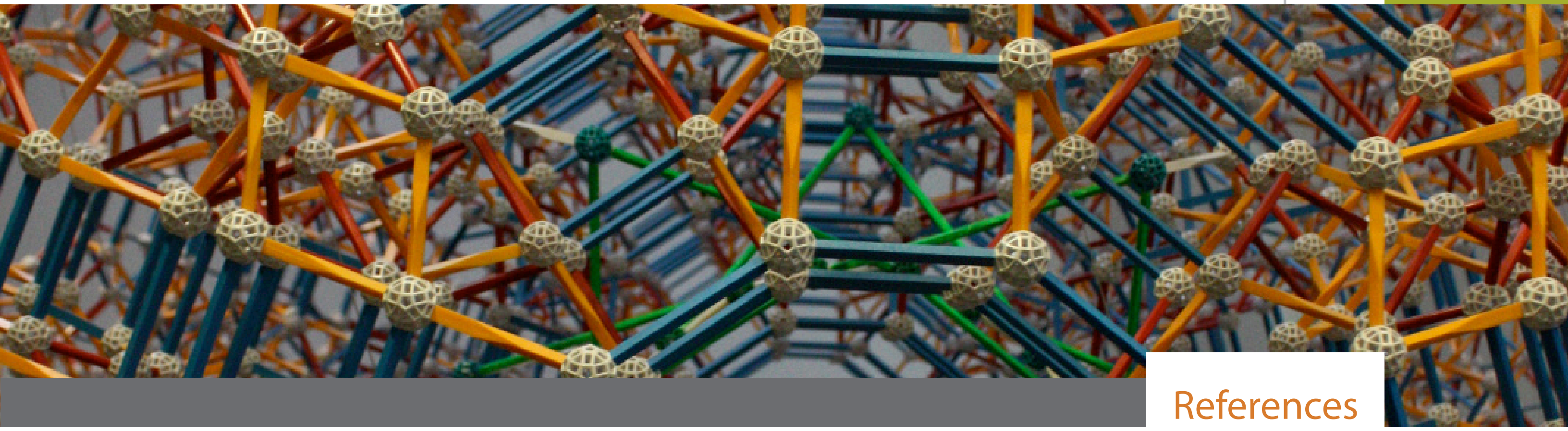
All infrastructures embed social norms, relationships, and ways of thinking, acting, and working. As a corollary, when they change, authority, influence, and power are redistributed. Knowledge infrastructures are no different; they create tensions and raise concerns that are best addressed early and often. New kinds of knowledge work and workers displace old ones; increased access for some may mean reduced access for others. As knowledge infrastructures evolve, attending to the social relations both created and broken by new modes may help societies reduce the negative distributional consequences of change. For example, citizen science projects can be designed in ways that maximize labor exploitation, on the one hand, or co-production and engagement, on the other. Approaching these tensions and redistributive consequences as a design opportunity — perhaps using the Scandinavian participatory design movement as a model — could help to energize a new kind of thinking about scale and structure in design.

The final section of this report reflects on what kinds of research might best engage the question of knowledge infrastructures. Participants emphasized that social scientists cannot remain simple

---

[*] Goldwin Heights, photograph by zellerludwig81 on flickr

bystanders or critics of the current transformations, which will not be reversed; instead, we need research practices that can help innovate, rethink, and rebuild. For example, a long-time-scale, historically informed framework can help situate our thinking by reminding us that infrastructural change normally takes decades rather than years, and that very substantial social learning must take place before the full benefits of new sociotechnical systems can manifest. Creating and nourishing standards and mechanisms for large-scale, long-term research in the qualitative social sciences, such as sustainable, accumulative, and shareable qualitative databases, could contribute to this goal. Improvements in qualitative data analysis software are urgently needed. New forms of cyberscholarship, such as new modes of writing or what one participant called a "knowledge zoom lens" for presenting qualitative evidence at any desired level of detail, need support and creative thought. Building better interdisciplinary collaborations across the natural and social sciences is an old goal, rarely realized — but more crucial now than ever in the face of such problems as climate change and biodiversity loss. A knowledge infrastructures perspective on the study of scholarship will promote more sustained, collective progress in research, design, and policy for 21st century scholarship.



References

Anderson, C. (2008). The End of Theory. *Wired Magazine, 16.*

Babcock, P., & Marks, M. (2010). Leisure College, USA: The Decline in Student Study Time. *Education Outlook, 7*, 1-7.

Babcock, P., & Marks, M. (2011). The Falling Time Cost of College: Evidence from Half a Century of Time Use Data. *Review of Economics and Statistics, 93*(2), 468-478.

Bartky, I. R. (1989). The Adoption of Standard Time. *Technology and Culture, 30*(1), 25-56.

Bechhofer, S., Ainsworth, J., Bhagat, J., Buchan, I., Couch, P., Cruickshank, D., Roure, D. D., Delderfield, M., Dunlop, I., & Gamble, M. (2010). Why Linked Data Is Not Enough for Scientists. *2010 IEEE Sixth International Conference on e-Science*, 300-307.

Beck, U. (2010). Climate for Change, or How to Create a Green Modernity? *Theory, Culture & Society, 27*(2-3), 254-266.

Blanchette, J.-F. (2012). *Burdens of Proof: Cryptographic Culture and Evidence Law in the Age of Electronic Documents*. Cambridge: MIT Press.

Blue Ribbon Task Force on Sustainable Digital Preservation and Access. (2010). Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information. *Final Report of the Blue Ribbon Task Force on Sustainable Digital Preservation and Access.*

Borgman, C. L. (2007). *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. Cambridge, MA: MIT Press.

Borgman, C. L. (2012). The Conundrum of Sharing Research Data. *Journal of the American Society for Information Science and Technology, 63*(6), 1059-1078.

Borgman, C. L., Abelson, H., Dirks, L., Johnson, R., Koedinger, K., Linn, M., Lynch, C., Oblinger, D., Pea, R., Salen, K., Smith, M., & Szalay, A. (2008). Fostering Learning in the Networked World: The Cyberlearning Opportunity and Challenge. A 21st Century Agenda for the National Science Foundation. *Report of the NSF Task Force on Cyberlearning*. Washington, DC: Office of Cyberinfrastructure and Directorate for Education and Human Resources.

Borgman, C. L., Wallis, J. C., Enyedy, N., & Mayernik, M. S. (2006). *Capturing Habitat Ecology in Reusable Forms: A Case Study with Embedded Networked Sensor Technology*. Paper presented at the Annual Meeting of the Society for the Social Studies of Science, Vancouver, BC.

Borgman, C. L., Wallis, J. C., & Mayernik, M. S. (2012). Who's Got the Data? Interdependencies in Science and Technology Collaborations. *Computer Supported Cooperative Work (CSCW), 21*(6), 485-523.

Borgman, C. L., Wallis, J. C., Mayernik, M. S., & Pepe, A. (2007). Drowning in Data: Digital Library Architecture to Support Scientific Use of Embedded Sensor Networks. *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries*, 269-277.

Bowker, G. C. (2000). Biodiversity Datadiversity. *Social Studies of Science, 30*(5), 643-683.

Bowker, G. C. (2005). *Memory Practices in the Sciences*. Cambridge, MA: MIT Press.

Bowker, G. C., Edwards, P. N., & Jackson, S. J. (2010). The Long Now of Cyberinfrastructure. In W. Dutton & P. Jeffreys (Eds.), *World Wide Science*. Cambridge, MA: MIT Press.

Zome Ball at the London Knowledge Lab, photograph by Juliette Culver

Brand, S. (1999). *The Clock of the Long Now: Time and Responsibility.* New York: Basic Books.

Brunton, F. (2013). *Spam: A Shadow History of the Internet*. Cambridge, MA: MIT Press.

Burton, M., & Jackson, S. J. (2012). Constancy and Change in Scientific Collaboration: Coherence and Integrity in Long-Term Ecological Data Production. *Proceedings of the 2012 Hawaii International Conference on Systems Science*, 353-362.

Busch, L. (2011). *Standards: Recipes for Reality.* Cambridge, MA: MIT Press.

Callon, M., & Latour, B. (1981). Unscrewing the Big Leviathan: How Actors Macro-Structure Reality and How Sociologists Help Them to Do So. In K. D. Knorr-Cetina & A. V. Cicourel (Eds.), *Advances in Social Theory and Methodology: Towards an Integration of Micro- and Macro-Sociologies* (pp. 277-303). Boston, MA: Routledge and Kegan Paul.

Chongthammakun, R., & Jackson, S. J. (2012). Boundary Objects, Agents, and Organizations: Lessons from E-Document System Development in Thailand. *Proceedings of the 2012 Hawaii International Conference on System Science*, 2249-2258.

Clune, T. L., & Rood, R. B. (2011). Software Testing and Verification in Climate Model Development. *IEEE Software, 28*(6), 49-55.

Collins, H. M. (1985). *Changing Order: Replication and Induction in Scientific Practice.* London: Sage Publications.

Collins, H. M., & Pinch, T. (1993). *The Golem: What Everyone Should Know About Science.* Cambridge: Cambridge University Press.

Courain, M. E. (1991). *Technology Reconciliation in the Remote-Sensing Era of United States Civilian Weather Forecasting: 1957-1987*. Unpublished doctoral dissertation, Rutgers University, New Brunswick.

Cummings, J., Finholt, T. A., Foster, I., Kesselman, C., & Lawrence, K. A. (2008). Beyond Being There: A Blueprint for Advancing the Design, Development, and Evaluation of Virtual Organizations: Report from an NSF Workshop on Developing VIrtual Organizations.

David, P. (1990). The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox. *The American Economic Review, 80*(2), 355-361.

De Roure, D., Bechhofer, S., Goble, C., & Newman, D. (2011). Scientific Social Objects: The Social Objects and Multidimensional Network of the Myexperiment Website. *Proceedings of the 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, 1398-1402.

De Roure, D., Goble, C., Aleksejevs, S., Bechhofer, S., Bhagat, J., Cruickshank, D., Fisher, P., Hull, D., Michaelides, D., & Newman, D. (2010). Towards Open Science: The Myexperiment Approach. *Concurrency and Computation: Practice and Experience, 22*(17), 2335-2353.

Derrida, J. (1996). *Archive Fever : A Freudian Impression*. Chicago: University of Chicago Press.

Downey, G. (2001). Virtual Webs, Physical Technologies, and Hidden Workers: The Spaces of Labor in Information Internetworks. *Technology and Culture, 42*, 209-235.

Downey, G. (2002). *Telegraph Messenger Boys: Labor, Technology, and Geography, 1850-1950*. New York: Routledge.

Downey, G. (2003). Commentary: The Place of Labor in the History of Information-Technology Revolutions. *International Review of Social History, 48*(S11), 225-261.

Downey, G. (2008). Human Geography and Information Studies. *Annual Review of Information Science and Technology, 41*(1), 683-727.

Easterbrook, S. M., & Johns, T. C. (2009). Engineering the Software for Understanding Climate Change. *Computing in Science & Engineering, 11*(6), 65-74.

Edwards, P. N. (1996). *The Closed World: Computers and the Politics of Discourse in Cold War America*. Cambridge, MA: MIT Press.

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

Edwards, P. N., Jackson, S. J., Bowker, G. C., & Knobel, C. P. (2007). *Understanding Infrastructure: Dynamics, Tensions, and Design*. Ann Arbor: Deep Blue. http://hdl.handle.net/2027.42/49353.

Edwards, P. N., Mayernik, M. S., Batcheller, A. L., Bowker, G. C., & Borgman, C. L. (2011). Science Friction: Data, Metadata, and Collaboration. *Social Studies of Science, 41*(5), 667-690.

Egyedi, T. (2001). Infrastructure Flexibility Created by Standardized Gateways: The Cases of Xml and the Iso Container. *Knowledge, Technology & Policy, 14*(3), 41-54.

Epstein, S. (2008). *Inclusion: The Politics of Difference in Medical Research*. Chicago: University of Chicago Press.

Firestein, S. (2012). *Ignorance: How It Drives Science*. New York: Oxford University Press.

Foucault, M. (1971). *The Order of Things: An Archaeology of the Human Sciences* (1st American ed.). New York: Pantheon Books.

Fujimura, J. (1992). Crafting Science: Standardized Packages, Boundary Objects, and 'Translation'. In A. Pickering (Ed.), *Science as Practice and Culture* (pp. 168-214). Chicago: University of Chicago Press.

Galison, P. L. (1996). Computer Simulations and the Trading Zone. In P. L. Galison & D. J. Stump (Eds.), *The Disunity of Science: Boundaries, Contexts, and Power* (pp. 118-157). Stanford: Stanford University Press.

Galison, P. L. (1997). *Image and Logic: A Material Culture of Microphysics*. Chicago: University of Chicago Press.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The New Production of Knowledge : The Dynamics of Science and Research in Contemporary Societies*. London: Sage Publications.

Giere, R. N. (1999). Using Models to Represent Reality. In L. Magnani, N. J. Nersessian & P. Thagard (Eds.), *Model-Based Reasoning in Scientific Discovery* (pp. 41-58). New York: Springer.

Gitelman, L. (Ed.). (2013). *Raw Data Is an Oxymoron*. Cambridge, MA: MIT Press.

Goble, C., & De Roure, D. (2007). Grid 3.0: Services, Semantics and Society. *2007 8th IEEE/ACM International Conference on Grid Computing*.

Grmek, M. (1990). *History of Aids: Emergence and Origin of a Modern Pandemic*. Princeton: Princeton University Press.

Hanseth, O., Jacucci, E., Grisot, M., & Aanestad, M. (2006). Reflexive Standardization: Side Effects and Complexity in Standard Making. *Management Information Science Quarterly, 30*, 563-581.

Harden, V. (2012). *Aids at 30: A History*. Dulles, VA: Potomac Books Inc.

Hedstrom, M. (1991). Understanding Electronic Incunabula: A Framework for Research on Electronic Records. *The American Archivist, 54*(3), 334-354.

Hey, T., Tansley, S., & Tolle, K. (Eds.). (2009). *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Redmond, WA: Microsoft Research.

Heymann, M. (2010). Understanding and Misunderstanding Computer Simulation: The Case of Atmospheric and Climate Science--an Introduction. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics, 41*(3), 193-200.

Howison, J., & Herbsleb, J. D. (2011). Scientific Software Production: Incentives and Collaboration. *Proceedings of the ACM 2011 conference on Computer Supported Cooperative Work*, 513-522.

Hulme, M. (2009). *Why We Disagree About Climate Change: Understanding Controversy, Inaction and Opportunity*. Cambridge: Cambridge University Press.

Hulme, M. (2010). Problems with Making and Governing Global Kinds of Knowledge? *Global Environmental Change, 20*(4), 558-564.

Jackson, S. J. (2006). Water Models and Water Politics: Deliberative Design and Virtual Accountability. *Proceedings of the 2006 Digital Government Conference*.

Jackson, S. J., & Barbrow, S. K. (2013). Infrastructure and Vocation: Field, Calling and Computation in Ecology. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*.

Jackson, S. J., Edwards, P. N., Bowker, G. C., & Knobel, C. P. (2007). Understanding Infrastructure: History, Heuristics, and Cyberinfrastructure Policy. *First Monday, 12*(6).

Jackson, S. J., Steinhardt, S., & Buyuktur, A. (2013). Why Cscw Needs Science Policy (and Vice-Versa). *Proceedings of the ACM 2013 conference on Computer Supported Cooperative Work*, 1113-1124.

Jasanoff, S. (2004). *States of Knowledge : The Co-Production of Science and Social Order*. New York: Routledge.

Kelling, S., Gerbracht, J., Fink, D., Lagoze, C., Wong, W., Yu, J., Damoulas, T., & Gomes, C. (2012). Ebird: A Human/Computer Learning Network for Biodiversity Conservation and Research. *Proceedings of the 24th Conference on Innovative Applications of Artificial Intelligence.*

Kensing, F., & Blomberg, J. (1998). Participatory Design: Issues and Concerns. *Computer Supported Cooperative Work (CSCW), 7*(3-4), 167-185.

Kolata, G. (2013, April 7). Scientific Articles Accepted (Personal Checks, Too), *New York Times*. Retrieved from http://www.nytimes.com/2013/04/08/health/for-scientists-an-exploding-world-of-pseudo-academia.html?pagewanted=all&_r=1&

Kranzberg, M. (1986). Technology and History: "Kranzberg's Laws". *Technology and Culture, 27*(3), 544-560.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago.

Lahsen, M. (2010). The Social Status of Climate Change Knowledge: An Editorial Essay. *Wiley Interdisciplinary Reviews: Climate Change, 1*(2), 162-171.

Lakatos, I., & Musgrave, A. (1970). *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press.

Landauer, T. K. (1995). *The Trouble with Computers: Usefulness, Usability, and Productivity*. Cambridge, MA: MIT Press.

Latour, B. (1983). Give Me a Laboratory and I Will Raise the Earth. In K. Knorr-Cetina & M. Mulkay (Eds.), *Science Observed* (pp. 141-170). London: Sage Publications.

Latour, B. (2004). *Politics of Nature: How to Bring the Sciences into Democracy*. Cambridge, MA: Harvard University Press.

Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford University Press.

Latour, B., & Woolgar, S. (1979). *Laboratory Life: The Social Construction of Scientific Facts*. London: Sage Publications.

Le Dantec, C. A., & DiSalvo, C. (2013). Infrastructuring and the Formation of Publics in Participatory Design. *Social Studies of Science, 43*(2), 241-264.

Lee, C. P., Dourish, P., & Mark, G. (2006). The Human Infrastructure of Cyberinfrastructure. *Proceedings of the ACM 2006 20th anniversary conference on Computer Supported Cooperative Work*, 483-492.

Mayernik, M. S., Wallis, J. C., & Borgman, C. L. (2013). Unearthing the Infrastructure: Humans and Sensors in Field-Based Scientific Research. *Computer Supported Cooperative Work (CSCW), 22*(1), 65-101.

Mokhtari, K., Reichard, C. A., & Gardner, A. (2009). The Impact of Internet and Television Use on the Reading Habits and Practices of College Students. *Journal of Adolescent & Adult Literacy, 52*(7), 609-619.

Monteiro, E., Pollock, N., Hanseth, O., & Williams, R. (2012). From Artefacts to Infrastructures. *Computer Supported Cooperative Work (CSCW)*, 1-33.

Morgan, M. S., & Morrison, M. (1999). *Models as Mediators: Perspectives on Natural and Social Sciences*. Cambridge: Cambridge University Press.

Muller, M. J. (2007). Participatory Design: The Third Space in Hci (Revised). In J. Jacko & S. Sears (Eds.), *Handbook of Hci* (2nd Edition ed.). Mahway, NJ: Erlbaum.

Nielsen, M. A. (2012). *Reinventing Discovery: The New Era of Networked Science*. Princeton, NJ: Princeton University Press.

Norton, S. D., & Suppe, F. (2001). Why Atmospheric Modeling Is Good Science. In C. A. Miller & P. N. Edwards (Eds.), *Changing the Atmosphere: Expert Knowledge and Environmental Governance* (pp. 67-106). Cambridge, MA: MIT Press.

Olson, G. M., & Olson, J. S. (2000). Distance Matters. *Human-Computer Interaction, 15*(2), 139-178.

Olson, G. M., Olson, J. S., & Venolia, G. (2009). What Still Matters About Distance. *Proceedings of the Human Computer Interaction Consortium.*

Ong, A., & Collier, S. J. (Eds.). (2005). *Global Assemblages: Technology, Politics, and Ethics as Anthropological Problems*. Malden, MA: Blackwell Publishing.

Oreskes, N., Shrader-Frechette, K., & Belitz, K. (1994). Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences. *Science, 263*(5147), 641-646.

Palfrey, J., & Gasser, U. (2012). *Interop: The Promise and Perils of Highly Interconnected Systems*. New York: Basic Books.

Petersen, A. (2007). *Simulating Nature: A Philosophical Study of Computer-Simulation Uncertainties and Their Role in Climate Science and Policy Advice*. Antwerp: Het Spinhuis.

Pierce, J., Brynjarsdottir, H., Sengers, P., & Strengers, Y. (2011). Everyday Practice and Sustainable Hci: Understanding and Learning from Cultures of (Un) Sustainability. *CHI '11 Extended Abstracts on Human Factors in Computing Systems, Proceedings of the 2011 ACM Annual Conference on Human Factors in Computing Systems Extended Abstracts*, 9-12.

Pipek, V., & Wulf, V. (2009). Infrastructuring: Towards an Integrated Perspective on the Design and Use of Information Technology. *Journal of the Association for Information Systems, 10*(5), 447-473.

Pollock, N., & Williams, R. (2010). E-Infrastructures: How Do We Know and Understand Them? Strategic Ethnography and the Biography of Artefacts. *Computer Supported Cooperative Work (CSCW), 19*(6), 521-556.

Pollock, N., Williams, R., D'Adderio, L., & Grimm, C. (2009). Post Local Forms of Repair: The (Extended) Situation of Virtualised Technical Support. *Information and Organization, 19*(4), 253-276.

Proctor, R., & Schiebinger, L. L. (Eds.). (2008). *Agnotology: The Making and Unmaking of Ignorance*. Stanford, CA: Stanford University Press.

Ribes, D., & Finholt, T. A. (2009). The Long Now of Technology Infrastructure: Articulating Tensions in Development. *Journal of the Association for Information Systems, 10*(5), 375-398.

Ribes, D., & Jackson, S. J. (2013). Data Bite Man: The Work of Sustaining Long-Term Data Collection. In L. Gitelman (Ed.), *Raw Data Is an Oxymoron*. Cambridge, MA: MIT Press.

Ribes, D., Jackson, S. J., Geiger, R. S., Burton, M., & Finholt, T. A. (2012). Artifacts That Organize: Delegation in the Distributed Organization. *Information and Organization, 23*(1), 1-14.

Rosner, D. K. (2012). The Material Practices of Collaboration. *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, 1155-1164.

Rosner, D. K., Oehlberg, L., & Ryokai, K. (2008). Studying Paper Use to Inform the Design of Personal and Portable Technology. *CHI '08 Extended Abstracts on Human Factors in Computing Systems*, 3405-3410.

Russell, A. L. (2006). 'Rough Consensus and Running Code' and the Internet-Osi Standards War. *IEEE Annals of the History of Computing, 28*(3), 48-61.

Shapin, S., & Schaffer, S. (1985). *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*. Princeton, NJ: Princeton University Press.

Shilton, K. (2009). Four Billion Little Brothers? Privacy, Mobile Phones, and Ubiquitous Data Collection. *Communications of the ACM, 52*(11), 48-53.

Shilton, K., Ramanathan, N., Reddy, S., Samanta, V., Burke, J., Estrin, D., Hansen, M., & Srivastava, M. (2008). Participatory Design of Sensing Networks: Strengths and Challenges. *Proceedings of the Tenth Anniversary Conference on Participatory Design 2008*, 282-285.

Shirky, C. (2009). *Here Comes Everybody: The Power of Organizing without Organizations*. New York: Penguin Group USA.

Shirky, C. (2010). *Cognitive Surplus: Creativity and Generosity in a Connected Age*. New York: Penguin Press.

Sismondo, S. (1999). Models, Simulations, and Their Objects. *Science in Context, 12*(2), 247-260.

Slota, S., & Bowker, G. C. (forthcoming). *The Curious Case of the Aggregated Database: The Global Biodiversity Information Facility*.

Star, S. L., & Griesemer, J. (1989). Institutional Ecology, 'Translations,' and Coherence: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-1939. *Social Studies of Science, 19*(3), 387-420.

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.

Stodden, V. C. (2010a). Open Science: Policy Implications for the Evolving Phenomenon of User-Led Scientific Innovation. *Journal of Science Communication, 9*(1).

Stodden, V. C. (2010b). The Scientific Method in Practice: Reproducibility in the Computational Sciences *MIT Sloan Research Paper No. 4773-10*.

Stodden, V. C. (2011). The Credibility Crisis in Computational Science: A Call to Action. Plenary Keynote, Cyberinfrastructure Days, University of Michigan

Stokes, D. (1997). *Pasteur's Quadrant: Basic Science and Technological Innovation*. Washington: Brookings Institution.

Suchman, L. A. (2007). *Human-Machine Reconfigurations : Plans and Situated Actions* (2nd ed.). Cambridge: Cambridge University Press.

Sundberg, M. (2009). The Everyday World of Simulation Modeling: The Development of Parameterizations in Meteorology. *Science, Technology & Human Values, 34*(2), 162-181.

Sundberg, M. (2010a). Creating Convincing Simulations in Astrophysics. *Science, Technology & Human Values, 37*(1), 64-87.

Sundberg, M. (2010b). Cultures of Simulations Vs. Cultures of Calculations? The Development of Simulation Practices in Meteorology and Astrophysics. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics, 41*(3), 273-281.

Sundberg, M. (2011). The Dynamics of Coordinated Comparisons: How Simulationists in Astrophysics, Oceanography and Meteorology Create Standards for Results. *Social Studies of Science, 41*(1), 107-125.

Suppe, F. (2000). Understanding Scientific Theories: An Assessment of Developments, 1969-1998. *Philosophy of Science, 67*, 102-115.

Takeda, K., Earl, G., Frey, J., Keay, S., & Wade, A. (2013). Enhancing Research Publications Using Rich Interactive Narratives. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 371*(1983).

van der Sluijs, J. P., Petersen, A. C., Janssen, P. H. M., Risbey, J. S., & Ravetz, J. R. (2008). Exploring the Quality of Evidence for Complex and Contested Policy Decisions. *Environmental Research Letters, 3*(2).

Vertesi, J. (2012). Seeing Like a Rover: Visualization, Embodiment, and Interaction on the Mars Exploration Rover Mission. *Social Studies of Science, 42*(3), 393-414.

Vertesi, J., & Dourish, P. (2011). The Value of Data: Considering the Context of Production in Data Economies. *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*, 533-542.

Wade, A., & Dirks, L. (2009). Tools for Researchers: Microsoft Research and the Scholarly Information Ecosystem.

Wallis, J. C., Borgman, C. L., Mayernik, M. S., Pepe, A., Ramanathan, N., & Hansen, M. (2007). Know Thy Sensor: Trust, Data Quality, and Data Integrity in Scientific Digital Libraries. In L. Kovacs, N. Fuhr & C. Meghini (Eds.), Research and Advanced Technology for Digital Libraries. *Proceedings of the 11th European Conference on Digital Libraries* (pp. 380-391). Berlin: Springer.

Weinberger, D. (2012). *Too Big to Know: Rethinking Knowledge Now That the Facts Aren't the Facts, Experts Are Everywhere, and the Smartest Person in the Room Is the Room*. New York: Basic Books.

Wiberg, M., Ishii, H., Dourish, P., Rosner, D. K., Vallgårda, A., Sundström, P., Kerridge, T., & Rolston, M. (2012). "Material Interactions": From Atoms & Bits to Entangled Practices. *CHI '12 Extended Abstracts on Human Factors in Computing Systems*, 1147-1150.

Back cover:

"Stripes of Change", photograph by SchallaArt on Deviant Art

Engines, photograph by Jjb@nalog on flickr

WWF Climate Change Program, photograph by @NonprofitOrgs on flickr

"Rusty Manhole Cover", photograph by GrungeTextures on Deviant Art

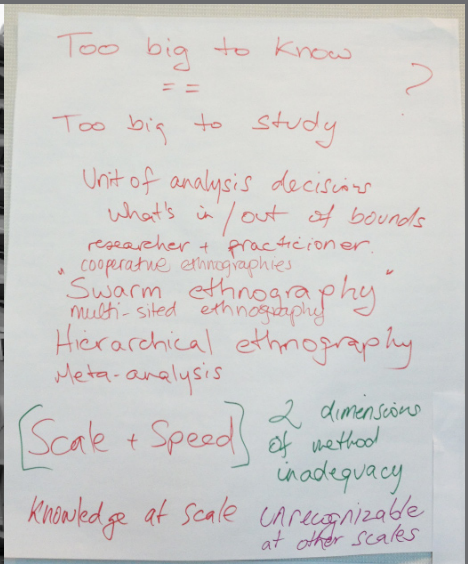
Salford's Tree of Knowledge, photograph by Edward Smith

Kegg Apoptosis network, visualised in Ondex by Simon Cockell

"Autopilot engaged", photograph by Mike Miley

Civil Aviation Authority Head Office, London, photograph by Dan Taylor

"Classroom", photograph by Lauren Manning

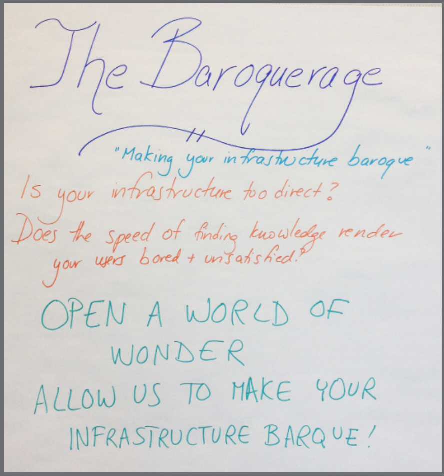


Too big to know
==
Too big to study
?

Unit of analysis decisions
what's in / out of bounds
researcher + practicioner.
cooperative ethnographies
"Swarm ethnography"
multi-sited ethnography
Hierarchical ethnography
Meta-analysis

[Scale + Speed] 2 dimensions of method inadequacy

Knowledge at scale unrecognizable at other scales



The Baroquerage

"Making your infrastructure baroque"

Is your infrastructure too direct?
Does the speed of finding knowledge render your users bored + unsatisfied?

OPEN A WORLD OF
WONDER
ALLOW US TO MAKE YOUR
INFRASTRUCTURE BARQUE!