The mess has always been with us: Organizing for distributed scientific collaboration[1]

Steve Sawyer (ssawyer@syr.edu)[2], Sarika Sharma[2], Matt Willis[3], Carsten Østerlund[2]

[2] iSchool, Syracuse University, Syracuse NY USA
[3] Oxford Internet Institute, University of Oxford, England

Abstract:

We advance an empirically-grounded set of four thematics that reflect what we are learning about how distributed scientists collaborate with each other, and the ways in which various information and communication technologies (ICT) are being bound up in this activity. We do so because the scientific enterprise is increasingly a distributed, collaborative, and digitally-reliant activity. The four thematics are: (1) the messiness of doing science, (2) the importance of managing talent, (3) attention to collaborative parsimony, and (3) reliance on steadily evolving collections of digital tools.

Acknowledgments: This material is based upon work supported by the National Science Foundation under Grant No. OCI-1221945. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. We thank the many participants who have so generously given of their time and expertise.

Keywords: distributed collaboration, science, cyberinfrastructure, ICT, organizing

INTRODUCTION

The contribution of this paper is to is to advance an empirically-grounded set of concepts that reflect what it is that distributed scientists are doing with each other, how they organize their work, and the ways in which various information and communication technologies (ICTs) are being bound up in this organizing. The empirical material supporting our theorizing is drawn from an ongoing project whose focus is to the ways in which distributed collaborative science teams draw on ICTs to support their science practices, with a particular emphasis on the ways in which documents help to structure and shape such activities. Moreover, this study focuses on distributed science projects of a scale of four to eight collaborators: common, but not typically of a size to warrant purpose-built cyberinfrastructure (CI) (Ribes and Lee, 2010, Pollock and Williams, 2012). And, given the paucity of attention to social science, we have focused primarily on social science practices.

DISTRIBUTED SCIENTIFIC COLLABORATION

We begin by noting that the contemporary scientific enterprise is increasingly a distributed, collaborative, and digitally-reliant activity. This shift elevates the organizational and institutional aspects of these arrangements (Atkins, 2003; Edwards, Jackson, Bowker, Knobel, 2007; Jirotka, Lee, Olson, 2013; Ribes and Lee, 2010). As Haythornthwaite (2006, p. 144-145), states so well:

[1] **Prepared from and please cite as:** Sawyer, S., Sharma, S., Willis, M. and Østerlund, C. (2016) The mess has always been with us: Organizing for distributed scientific collaboration, Paper prepared for the International Federation of Information Processing, Working Group 8.2: Information Systems in Organizations and Society, Working Conference, 7-9 December, 2016, Dublin, Ireland.

> *"As Knorr-Cetina (1999) has remarked about scientific teams, current research initiatives often bring together quite disparate disciplines, locations and technologies, often leaving a single researcher of such teams woefully at a disadvantage to understand a team's work. Studying contemporary teams requires a more comprehensive examination than is commonly employed, encompassing interdisciplinary processes, group interaction, institutional practices, career interests and uses of information and communication technology. Moreover, contemporary views that consider technology as providing the solution to the "problem" of collaboration—e.g., through faster connection, seamless integration of geographically distributed people and projects and new information and communication technology infrastructures—fail to acknowledge the negotiation of practices and the coevolution of practices and technology that are involved. Collaborations involve dealing with existing embedded practices, as well as emergent ones that take time and effort to evolve."*

Particular to the empirical approach we pursue, we know that the ways in which a work is structured both shapes and is shaped by the ways in which it is organized (e.g., Hackman and Wageman, 1995). And, specifically for our work here, scientific collaboration is "the interaction taking place with in a social context among two or more scientists that facilitates the sharing of meaning and completion of tasks with respect to a mutually shared, subordinate goal (Sonnenwald, 2008)."

Traditionally, the practice of collaborative science was dependent on the physical proximity of the researchers (Finholt 2002). Researchers were bound to work with those at their institution and the research that was conducted also correlated to the resources that were specific to the location (Finholt and Olson, 1997). The institution became the place to organize scientific work. That began to change, slowly at first, with the emergence of the Internet and information and communications technologies, and more rapidly of late (e.g., Olson, Zimmerman, Bos, 2008; Ribes and Lee, 2010). The very nature of scientific collaboration and collaborating has changed through the interconnection of people and resources online.

For scientists, collaborating virtually is challenging for several reasons. For instance, scientific knowledge is difficult to aggregate, scientists work independently and they have to work harder to maintain open communication, to adopt and adapt to common toolsets, and to keep groups focused on a common goal (Bos, Zimmerman, Olson, Yew, Yerkie, Dahl, Olson, 2007). Jackson, Ribes, Buyuktar, Bowker (2011) describe the four types of rhythms that collaborative teams work around: organizational, infrastructural, biographical, and phenomenal. Phenomenal rhythms reflect the characteristics of the focus of study: the demands of studying how trauma units in a hospital react to emergencies leads to a different set of working arrangements than does the study of traffic flow in urban settings. Biographical rhythms reflect the training, pedigree, attitudes, skills and knowledge of the participants. The infrastructural and organizational arrangements reflect decisions regarding the tools being used and the way work is to be done. And, yet, despite the barriers to collaboration, there is a growing sentiment and vision that networked technologies will give rise to digital platforms (Atkins, 2003 and Bowker, Baker, Millerand, Ribes, 2009).

As noted, the increasing reliance on CI provides for new forms of organizing for science (e.g., Bowker, Baker, Millerand, Ribes, 2009; Lee, Dourish, and Marks, 2006) and support "pervasive and enabling technologies" (Bowker, Baker, Millerand, and Ribes, 2009). Studies that look at the CI-enabled science projects focus on how they organize for long time scales (e.g., Bowker, 2015). These studies also focus on aspects of CI design and development (Stein, 2008; Ribes and Lee, 2010; Bietz, Baumer, Lee, 2010), maintenance (Lee, Dourish, and Marks 2006; Bietz, Ferro, Lee, 2012), governance (Barjak, Eccles, Meyer, Robinson, Schroeder, 2013), sustainability (Ribes and Finholt, 2009) temporality (Steinhardt and Jackson, 2015; Karasti, Millerand and Baker, 2010), and repair (Jackson, 2015).

Most studies of CI go on to highlight the organizational aspects of science. For example, Ribes and Finholt, (2009) specifically outline the organizational, institutional, and technological tensions that must be thought through in order for the science infrastructure to sustain for the long-term. Without understanding the practices of how resources are used shared and leveraged, a key equation is missing on how CIs are used by scholars in their scientific knowledge creation (Lee, Bietz, Thayer, 2010; Wang and Armstrong, 2009; Lane, Hues, Mulcahy, 2008).

RESEARCH APPROACH

The study from which the data are drawn is designed as a multi-phase, multi-method project. The first phase was exploratory and combined secondary data collection, a survey, and interviews with 26 scholars involved in distributed scientific collaborations. The interview and survey were designed to be paired, with the survey done online before (or sometimes during) the interview. The interview was designed as a semi-structured protocol to be done either face-to-face or via some digital mediation (and about half were done each way). Interviews lasted about an hour and were de-identified and transcribed for analysis.

During interviews participants also talked in detail about other group members work practices in their collaboration. The highest number of collaborators on a single project was over 30 and the lowest was two. This meant that interviews with our primary informants gave is insights into the work practices of over 170 collaborators and project staff. Interview participants complete a survey that further detailed their project collaborations and the work of their collaborations. The survey also features a "check sheet" or inventory of the software, hardware, physical tools, and objects that enable the project collaboration. This provided us a means to better understand the specific arrangement of ICTs that each person assembles.

In the second phase of the study, we pursued reflective understanding. Following a round of analyses, we conducted two sets of focus groups with four different groups of academics, with each focus group consisting of between four and seven people. For these focus groups, we developed multiple scenarios – synthetic situations that were derived from the phase one interviews, surveys and secondary data collection. The purpose of these scenarios was to provide an empirically-grounded situation to generate conversation among the focus group participants. The intent was to use these synthetic scenarios to generate discussion regarding shared experiences with distributed scientific collaborations, common problems and breakdowns in collaboration and using ICTS, and to uncover further idiosyncrasies of independent but collaborative work practices. Focus groups lasted about 90 minutes, were very interactive, and we were able to secure permission to record these. We transcribed the recordings for analysis, and also had notes from one of the research team members (who sat in as an observer on each of the focus groups).

The third phase of the project is a more detailed study of document-sharing and uses, with a more sustained analysis of document-related practices and uses of ICT. We are not focusing on this phase here (as it is not yet completed). However, participants in the third phase also participated in the first phase of the project.

What we present here builds from interim analyses of the interview transcripts and focus groups (per Miles and Huberman, 1994). The interim analyses and ensuing research team discussion are structured by the concepts embedded in the survey, semi-structured interview guide and the scenarios that frame the focus groups. The next two sections provide a summary of the insights from these analyses.

FOUR THEMATICS OF DISTRIBUTED SCIENTIFIC COLLABORATION

We highlight four thematics that arise from the interim analysis about the doing of distributed scientific collaboration. These are not orthogonal constructs or fully developed conceptual commitments. Rather, these represent at attempt to synthesize what we are learning from the study of distributed collaborative scientific practice. Our focus is to small research teams and our thematics are tentative insights into the patterns of these arrangements. As such they deserve more careful scrutiny and to be subjected to the empirical light of larger-scale collaborations and a broader range of scientific practices. Our practice-based lens inspires us, however, to speculate on the likelihood that these are more common themes than we our data provides us basis.

The mess has always been with us

The first thematic reflects the simple insight that the "ground truth" or reality of doing distributed collaborative science is not reflected in the carefully sterilized cleanliness of a methods chapter or research section. In the words of one senior focus group participant: "The mess has always been with us." What we learn from our participants is that so much of what they do is focused on making sense of what they are doing: of relearning and redoing. Our participants are active across several projects and these are woven into the other parts of their work lives, and often their non-work lives. The collaborative working arrangements are constantly renegotiated to reflect learning (and forgetting), changes in the research team makeup, evolving interests, new data and opportunities, and the realities of research approaches oriented towards discovery and learning (rather than production and repetition). As we note in quote above, this is not new: scholars of science have long noted this messiness – and the desires to rationalize it for others (e.g., Latour, 1987). What we find in our studies of distributed collaboration is that this continues, and is possible a bit more complicated because of the virtual working arrangements.

Managing talent v. managing roles

The second thematic we identify is the difference between the approaches to work structure and project management extolled in most formal organizations relative to what we find in distributed scientific efforts. Instead of carefully structured control and reporting systems, role delineation, and job descriptions – the staple of modern organizing, the collaborative research projects we study are managed informally. They are of a size and level of complexity that informal arrangements seem to work reasonably well. Participants work together both to identify goals, deadlines, and tasks and to resolve problems or explore opportunities. Most participants are working on multiple projects and are pursuing multiple goals such as different papers, particular data, and new opportunities. Because of this, project structures are focused on matching talents, not assigning roles. Such an approach demands negotiation, ongoing coordination and communication, and a level of shared interest and trust that is hard to replicate in the contemporary work organization. In fact, certain power-related social structures, such as academic seniority, topical expertise, disciplinary differences, and institutional status are accommodated both explicitly and implicitly. And, while we have not focused on gender issues, it is clear these are also part of the mix in the loose project management approaches pursued.

Collaborative parsimony

The third thematic begins by noting that participants want their distributed scientific collaborations to succeed. Most are involved in multiple collaborations, with different groups of people, and expect that their collaborators are doing the same. There is substantial good will provided in support of pursuing multiple projects and an expectation that collaborating is how science is to be done. This acceptance, expectation, and orientation are part of what new scholars seem to be learning in their graduate programs.

We observe that each of the collaborative team develop their own set of operational practices and norms relative to the ways in what that project moves forward. This includes things as fundamental as document naming and storage conventions, the norms of communicating (uses of social media, how fast one replies to emails, whom to cc), decision-making, task delegation and deadline-setting, writing and doing analysis). Such norms build from personal approaches to a common arrangement: something that works. Often this is informal and implicit: it emerges. As such, the arrangements often reflect modest adaptations of each participant's personal approaches. That is, part of the discussion on collaboration is how to accommodate the most viable shared approaches while honoring the primacy of individual and personalized work styles. This is what we are calling collaborative parsimony: the desire to maximize the success of working together while minimizing the amount of changes to any one participant's more well-established personal work processes.

These shared approaches are often, if not typically, evolving in response to changes in personnel or in the face of problems or issues. These negotiated evolutions are done with care, humor, understanding and demarcations between the shared / collaborative spaces and private spaces -- a recognition that both exist for good reasons. As noted, since most scientists are involved in more than one of these distributed collaborative efforts, they have multiple informal and modestly different arrangements. So, it is not unusual for a collaboration's arrangements to be relearned and reconfigured, updated and evolved. This is particularly true for projects that tend to have bursts of activity followed by periods of quiet.

Constantly-evolving collections of ICT

The fourth thematic is the differences across the large and evolving collections of ICTs used by participants and teams. One of the most impressive insights from the data on ICT uses is that participants in this study use an amalgam of software and platforms for the express purpose of collaborating with other team members and to accomplish the goals and objectives of their professional work. We identified 75 different software packaged and platforms in use. This includes multiple email clients, Internet browsers, calendars, note-taking software, word-processing software, document storage and analysis software. Each of these categories has between four and ten specific products. Many of these software products are used in tandem with each other and it's not uncommon to use two or more software products in the same category. In the Internet browser category, we observed used of Internet Explorer, Safari, Chrome, Firefox, and Opera. Many participants used more than one browser depending on the type of work, location, web resources needed, and other factors. This was the case with every software category.

In addition to using bundles of commodified software products to support work practices, we discovered different commitments or allegiances to certain software platforms and ecosystems. Most of our data speak to participants fully invested in either a Google, Microsoft, or Apple ecosystem. It was not uncommon to see the use of a platform like Google for word processing, email, calendar, contacts, and other functionality. This ecosystem allegiance played out in-group meetings through persuasion of other group members to adopt a particular workflow, in the case above to write collaboratively in Google docs. This phenomenon occurred intermittently throughout the group process. We noted periods of negotiation or checking in on the tools and practices that group members used. For example, planning to write an article by articulating what set of tools (word or Google docs) to use when producing the article. Additionally, who would use what software and when for what work: such as Dropbox for an archive and email for active documents. These decisions would persist throughout the group's collaborative work, or the group might revisit how they use these tools if tension occurred in the collaborative work practices. For example, not using the same filing convention or placing files in the wrong location.

NEXT STEPS

We are keen to expand from the current study's focus on social science approaches to study humanities scholars and those pursuing physical, natural, and artificial sciences. As noted, we expect to see differences in the ways in which these scientific enterprises are being pursued. But, we are also expecting many of the themes we have identified to also be visible. For example, Woolgar (1988) focuses on how bench scientists dealt with the management of methodological horrors as practical reasoning: How they made messes make sense.

The second area for continued attention is to emphasize the ways in which principles of organizing and managing are being adapted and drawn in to science. As noted, the ways in which people and roles, goals and controls, and the very nature of organizing for work, seems to be tuned differently in science. Whether this continues to be the case, and what it means to be working across multiple boundaries and multiple institutions, deserves greater attention. Because of the conceptual premise -- of managing a common resource with competing interests -- we are drawn to precepts of networked (or collaborative) governance that has been advanced by Provan and Kenis (2007; Provan, Fish and Sydow, 2007). This approach takes as a given that all governance is shared (there is no singular decision-maker) and that all participants are expected to be pursuing multiple goals. Given the ways in which the distributed scientific collaborations are formed, networked governance (as a social concept, not a technical arrangement) has conceptual heft.

The third area for continued attention is a focus on the ways in which the collections of digital technologies and ICT become a proto-infrastructure. While infrastructure connotes stability and service and sharing, what we find is a more flexible, steadily adapting and evolving assemblage of elements that Sawyer, Wigand and Crowston, 2014, call a digital assemblage and Pollock and Williams (2012) call e-infrastructures

With these as opportunities, with this paper we have advanced an empirically-grounded set of four thematics that reflect what it is that distributed scientists are doing with each other, how they organize their work, and the ways in which various information and communication technologies are being bound up in this activity. In doing this we have both characterized distributed scientific collaborations as a form of virtual organizing and identified from our data the messiness of doing science, the importance of managing talent, the attention to collaborative parsimony, and the impressive complex, steadily evolving collections of digital tools used to support these scientific enterprises.

REFERENCES

Atkins, D. (2003). Revolutionizing science and engineering through cyberinfrastructure: Report of the National Science Foundation blue-ribbon advisory panel on cyberinfrastructure. National Science Foundation, Arlington, VA.

Bietz, M. J., Baumer, E. P., & Lee, C. P. (2010). Synergizing in cyberinfrastructure development. Computer Supported Cooperative Work (CSCW), 19(3-4), 245-281.

Bos, N., Zimmerman, A., Olson, J., Yew, J., Yerkie, J., Dahl, E., & Olson, G. (2007). From shared databases to communities of practice: A taxonomy of collaboratories. *Journal of Computer- Mediated Communication*, *12*(2), 652-672.

Bowker, G. C., Baker, K., Millerand, F., & Ribes, D. (2009). Toward information infrastructure studies: Ways of knowing in a networked environment. In *International handbook of Internet research* (pp. 97-117). Springer Netherlands.

Chudoba, K., & Watson-Mannheim, M. (2007). Exploring the Virtual Work Environment. In A. Salazar & S. Sawyer (Eds.), *Handbook of Information Technology in Organizations and Electronic Markets* (pp. 57-76). Singapore: World Scientific Publishing

Chudoba, K., Wynn, E., Lu, M., & Watson-Mannheim, M. (2005). How Virtual are We? Measuring Virtuality and Understanding its Impact in a Global Organization. *Information Systems Research, 15*(3), 279-306.

Ciborra, C. (2004) Digital Technologies and the Duality of Risk. CARR, LSE, London.

Cummings, J., Finholt, T., Foster, I., Kesselman, C., & Lawrence, K. (2008). *Beyond Being There: A Blueprint for Advancing the Design, Development, and Evaluation of Virtual Organizations*. from http://www.ci.uchicago.edu/events/VirtOrg2008/VO_report.pdf.

DeSanctis, G., & Monge, P. (1999). Communication Processes for Virtual Organizations. *Organization Science, 10*(6), 693-703.

Edwards, P. N., Jackson, S. J., Bowker, G. C., & Knobel, C. (2007). Understanding infrastructure: dynamics, tensions, and design, NSF Report of a Workshop: History and theory of infrastructure: lessons for new scientific cyberinfrastructures.

Grabowski, M., & Roberts, K. (1999). Risk Mitigation in Virtual Organizations. *Organization Science, 10*(6), 704-717.

Geiger, S. and D. Ribes (2011). Trace Ethnography: Following coordination through documentary practices. Proceedings to the 44nd Annual Hawaii International Conference on System Science (HICSS-44), Hawaii, HI, IEEE Computer Society Press.

Griffith, T., Sawyer, J., & Neale, M. (2003). Virtualness and Knowledge in Teams: Managing the Love Triangle of Organizations, Individuals and Teams. *MIS Quarterly, 27*(3), 265-287.

Hackman, J. R., & Wageman, R. (1995). Total quality management: Empirical, conceptual, and practical issues. *Administrative Science Quarterly, 40*(2), 309-342.

Haythornthwaite, C. (2001). Exploring Multiplexity: Social Network Structures In A Computer-Supported Distance Learning Class. *The Information Society, 17*(3), 211-226.

Haythornthwaite, C. (2006). Articulating Divides in Distributed Knowledge Practice. *Information, Communication and Society, 9*(6), 761-780.

Haythornthwaite, C. (2007). Social networks and online community. In A. Joinson, K. McKenna, U. Reips & T. Postmes (Eds.), *Oxford Handbook of Internet Psychology* (pp. 121-136). Oxford: Oxford University Press.

Haythornthwaite, C. (2009). *Crowds and communities: Light and heavyweight models of peer production.* Paper presented at the 42nd Hawaii International Conference on System Sciences, Los Alamitos, and CA:.

Haythornthwaite, C. (2009). Social networks and information transfer. In M. J. Bates & M. N. Maack (Eds.), *Encyclopedia of Library and Information Sciences*. NY: Taylor & Francis.

Haythornthwaite, C., Lunsford, K. J., Bowker, G. C., & Bruce, B. (2006). Challenges for research and practice in distributed, interdisciplinary, collaboration. In C. Hine (Ed.), *New Infrastructures for Science Knowledge Production* (pp. 143-166). Hershey, PA: Idea Group.

Haythornthwaite, C., & Wellman, B. (1998). Work, Friendship, and Media Use for Information Exchange in a Networked Organization. *Journal of the American Society for Information Science, 49*(12), 1101-1114.

Hine, C. (2006). *New Infrastructures for Science Knowledge Production: Understanding E-Science*. Hershey, PA: Idea Group.

Karasti, H., Baker, K. S., & Millerand, F. (2010). Infrastructure time: Long-term matters in collaborative development. *Computer Supported Cooperative Work (CSCW)*, *19*(3-4), 377-415.

Kaziunas, E., Sawyer, S. and Østerlund, C. (2011) Social Scientists, Documents, and Cyberinfrastructure: The Cobbler's Children or the Missing Masses? Proceedings of the iConference 2011, February 8-11, Seattle, WA.

Knorr Cetina, K. (1999). *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge: Harvard University Press.

Jackson, S. J., Ribes, D., Buyuktur, A., & Bowker, G. C. (2011, March). Collaborative rhythm: temporal dissonance and alignment in collaborative scientific work. In Proceedings of the ACM 2011 conference on Computer supported cooperative work (pp. 245-254). ACM.

Jirotka, M., Lee, C. P., & Olson, G. M. (2013). Supporting scientific collaboration: Methods, tools and concepts. *Computer Supported Cooperative Work (CSCW)*, *22*(4-6), 667-715.

Lonsdale, H., Jensen, C., Wynn, E., & Dedual, N. J. (2010). *Cutting and Pasting Up: "Documents" and Provenance in a Complex Work Environment.* Paper presented at the 43nd Annual Hawaii International Conference on System Science (HICSS-43), Hawaii, HI.

Lee, C. P., Bietz, M. J., & Thayer, A. (2010, May). Research-driven stakeholders in cyberinfrastructure use and development. In *Collaborative Technologies and Systems (CTS), 2010 International Symposium on* (pp. 163-172). IEEE.

Lee, C. P., Dourish, P., & Mark, G. (2006, November). The human infrastructure of cyberinfrastructure. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work* (pp. 483-492). ACM.

Miles, M, & Huberman, A. (1994). *Qualitative data analysis: An expanded source book.* Thousand Oaks, CA: Sage Publications.

O'Leary, M. B., & Cummings, J. N. (2007). The Spatial, Temporal, and Configurational Characteristics of Geographic Dispersion in Teams. *MIS Quarterly, 32*(2 (June)), xx-xx.

Olson, G., Zimmerman, A., & Bos, N. (2008). *Scientific Collaboration on the Internet*. Cambridge, MA: MIT Press.

Olson, G. M., Atkins, D. E., Clauer, R., Finholt, T. A., Jahanian, F., Killeen, T. L., ... & Weymouth, T. (1998). The upper atmospheric research collaboratory (UARC). *Interactions*, *5*(3), 48-55.

Østerlund, C. (2008a). Documents in Place: Demarcating Places for Collaboration in Healthcare Settings. Computer Supported Cooperative Work (CSCW), 17(2-3), 195-225.

Østerlund, C. (2008b). The Materiality of Communicative Practices: The boundaries and objects of an emergency room genre. Scandinavian Journal of Information Systems, 20(1), 7-40.

Østerlund, C., & Boland, D. (2009, January 5-8, 2009). Document Cycles: Knowledge Flows in Heterogeneous Healthcare Information System Environments. Paper presented at the 42nd Annual Hawaii International Conference on System Science (HICSS-423), Hawaii, HI.

Østerlund, C., Sawyer, S. (2010). Using Documents to Study Virtual Organizing: Then Materiality, Provenance, and Context Matter. Presented at Academy of Management (AOM), Montreal, Canada. August 11-17.

Østerlund, C., Sawyer, S. and Kazianus, E. (2010) Documenting Work: A Methodological Window into Coordination in Action, Presented at 26th Conference of the European Group for Organizational Studies, 3-6 July, Lisbon, Portugal.

Østerlund, C. & Crowston, K. (2011). "What characterize documents that bridge boundaries compared to documents that do not? An exploratory study of documentation in FLOSS teams" Proceedings of HICSS-44, Hawaii, January 4-8, 2011.

Pollock, N. and Williams, R. 2010. E-infrastructures: How do we know and understand them? Strategic ethnography and the biography of artefacts. *Computer Supported Cooperative Work*, (*19*:6), 521–556.

Price, D. (1963). *Little science, big science*. New York: Columbia University Press.

Provan, K., Fish, A. and Sydow, J. (2007). Inter-organizational networks at the network level: A review of the empirical literature on whole networks. *Journal of Management, (33*:3), 479–516.

Provan, K. and Kenis, L. (2007). Modes of network governance: Structure, management, and effectiveness. *Journal of Public Administration Research and Theory, (18*:2), 229–252.

Ribes, D., & Lee, C. P. (2010). Sociotechnical studies of cyberinfrastructure and e-research: Current themes and future trajectories. *Computer Supported Cooperative Work (CSCW)*, *19*(3-4), 231-244.

Sawyer, S., Østerlund, C. and Kaziunas, E, (2012) "Social Scientists and Cyberinfrastructure: Insights from a Document Perspective," Mark, G., Riedl, J. and Grudin, J. (eds.) *Proceedings of the 2012 Conference on Computer-Supported Collaborative Work*, February, 11-15, Seattle, WA, ACM press.

Sawyer, J., & Tapia, A. (2007). From Findings to Theories: Institutionalizing Social Informatics. *The Information Society, 23*(4), 263 - 275.

Sawyer, S. (2008). Data Wealth, Data Poverty, Science, and Cyberinfrastructure. *Prometheus, 26*(4), 357-371.

Sawyer, S., & Southwick, R. (2002). Temporal Issues in ICT-Enabled Organizational Change: Scenes from an ERP Implementation. *The Information Society, 18*(4), 263-280.

Sawyer, S, Wigand, R. and Crowston, K. (2014) "Digital Assemblages: Evidence and Theorizing from a Study of Residential Real Estate," *New Technology, Work, and Employment*, *29*(1), 40-54

Sismondo, S. (2008). Science and technology studies and an engaged program. *The handbook of science and technology studies*, *3*,13-32.

Sonnenwald, D. H. (2007). Scientific collaboration. *Annual review of information science and technology*, *41*(1), 643-681.

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.

Steinhardt, S. B., & Jackson, S. J. (2014, February). Reconciling rhythms: plans and temporal alignment in collaborative scientific work. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing* (pp. 134-145). ACM.

Stein, L. D. (2008). Towards a cyberinfrastructure for the biological sciences: progress, visions and challenges. Nature Reviews Genetics, 9(9), 678-688.

Wang, S., & Armstrong, M. P. (2009). A theoretical approach to the use of cyberinfrastructure in geographical analysis. *International Journal of Geographical Information Science*, *23*(2), 169-193.

Watson-Manheim, M., Chudoba, K., & Crowston, K. (2002). Discontinuities and Continuities: A New Way to Understand Virtual Work. *Information, Technology, and People, 15*(3), 191-209.

Weick, K. (1995). What theory is not, theorizing is. *Administrative Science Quarterly*, 385-390.