

# Social Scientists and Cyberinfrastructure: Insights from a Document Perspective

Steve Sawyer  
School of Information Studies  
Syracuse University  
Syracuse NY 13244-4100 USA  
+1-315-443-6147  
ssawyer@syr.edu

Elizabeth Kaziunas  
School of Information Studies  
Syracuse University  
Syracuse NY 13244-4100 USA  
+1-315-443-2911  
eskaziun@syr.edu

Carsten Østerlund  
School of Information Studies  
Syracuse University  
Syracuse NY 13244-4100 USA  
+1-315-443-8773  
costerlu@syr.edu

## ABSTRACT

Contemporary cyberinfrastructure (CI) efforts are most often directed to meeting the needs of physical and natural scientists. We wonder: might the CI needs of social scientists be different? To provide insight into this question we report on a pilot study of social scientist's distributed and collaborative work practices. To frame this study we drew from the literatures of science studies and e-science and followed a mixed-methodological research design. Findings suggest focusing on documenting practices provides novel insights into the distributed work practices and the CI uses of social scientists.

## Author Keywords

Cyberinfrastructure, social scientists, work practices, documents.

## ACM Classification Keywords

H.4 [Information Systems Applications] K.4.3. [Computers and Society]: Organizational Impacts – *computer-supported collaborative work* J.4 [Computer Applications]: Social and Behavioral Sciences – *sociology*.

## INTRODUCTION

We report on a pilot study of social scientist's distributed collaborative practice as a means of gaining insight into the needs of these scholars relative to cyberinfrastructure (CI). We observe that much of the current uses of CI are oriented towards supporting natural and physical scientists. And, the emphasis of these CI have been to supporting increased access to large data sets, providing greater computing resources and broadening the uses of a range of scarce or unique instruments and sensors [2]. We further note that CI provides a means for more scientists to work together across time and space while pursuing larger scientific problems. We wonder: do the same needs and motivations

hold true for social scientists? [e.g. 15][14]

We continue with a too-brief introduction to CI and its roles in the social sciences. We then review literature from several intellectual communities regarding the work practices of social scientists. In the third section we outline and highlight insights from our pilot study, situating our findings within the current CI discourse. In the final section we offer tentative next steps.

## Cyberinfrastructure and Social Science

Cyberinfrastructure encompasses a range of digital resources such as computational capacity; access to data sources; tools to support access, use and analysis of this data; and the needed networking infrastructure to support access and use. Seen this way, CI is: "the sharing of computational resources, distributed access to massive datasets, and the development of tools for collaboration and communication" [5]. Often CI is advanced as a neutral technological platform imbued with the potential to extend research practices beyond any one laboratory and to transform scientific practice [1] [7] [15].

Contemporary evaluations of social scientists' uses of CI often echo the excitement of these ICT-enabled platforms as a new frontier for social science research [29] [14]. However, these studies also highlight a number of socio-technical issues that impede the widespread adoption of CI in the social sciences. Others find that many social scientists often resist sharing data and collaborating – key assumptions in the design of CI – as there are disincentives in their intellectual communities for doing so [3][19][28].

While CI proponents argue such obstacles can be overcome with better design, more usability testing and training, others are less sanguine. Some argue CI are far from being generic research platforms as they emphasize computational and quantitative approaches to doing science [5] [10] [24]. One implication of this emphasis is that simply extending CI principles for social scientists might fail to lead to the kinds of advances these infrastructures are seen as enabling. In response, we and other social science-oriented CI advocates call for an empirically-driven view of CI usage.

## Social Scientists

For this paper social science are those intellectual communities whose central phenomenon of interest is the

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CSCW 2012, February 11–15, 2012, Seattle, Washington.

Copyright 2011 ACM XXX-X-XXXXX-XXX-X/XX/XX...\$5.00.

human condition. This definition encompasses sociology, economics, anthropology, political science, linguistics, geography and others (and community inclusion is socially negotiated) [26]. One characteristic which helps define the social sciences is that knowledge is built from facts or evidence in support of theories (that are partial explanations of the world) and not typically represented as laws.

While a rich body of literature details the working lives of scientists in the laboratory (e.g., [16] [17] [33]), and more recently the virtual collaboration of scientists in e-science collaboratories [9] [21], much less has been written on the work practices of social scientists; distributed or otherwise. In fact, with the exception of Ellis' work on the information seeking behavior of social scientists, "...studies investigating academic social scientists have been in steady decline since the mid-1970s" [20]. And, there exists little research focused on the design and evaluation of collaboratories for pursuing social science [15] [25] [31].

This lack of empirical insight into social scientists' work practices, combined with the practical problems being routinely encountered in CI-oriented initiatives, leads to: 1) odd extrapolations being made from studies of current CI to the needs of social science [25] [31]; and 2) nascent social science CI efforts falling short of becoming the kind of transformative research envisioned. To help achieve transformative uses of CI by social scientists, we are wise to better understand their distributed collaborative practices.

#### **STUDYING DISTRIBUTED SOCIAL SCIENCE PRACTICE**

To help advance our understanding, we conducted a pilot or "micro" study of distributed collaboration among four social scientists. These scholars collaborated across three different time zones and never met face-to-face during the 13 months we studied them across 2009-2010. To do this study we employed a document-centric methodology that involved mapping their uses of physical and digital documents, personal and shared repositories, and shared online interactions. We used software to track uses of desktop and online repositories; gathered participant-generated images of physical documents and desktops; collected documents; conducted interview and behavioral inquiries; and did participant observation.

Following this intensive micro-study, we did follow-up interviews with a convenience sample of another seven social scientists involved in distributed collaborative projects in order to explore in more detail our initial findings. Here we present our observations from this data relative to the five elements of social scientist's work practices that are drawn from the literature.

#### **SOCIAL SCIENTISTS' WORK PRACTICES**

To advance our understanding of social scientists' distributed collaborative work practices, we build on the current literature. In particular we build from Bealieu and Wouters [5] [36] who theorize social science research practices as characterized by five inter-connected elements:

1. Epistemic diversity
2. Lack of consensus
3. Low technology (ICT) research environment
4. Historically grounded solitary research practices
5. Specificity of writing and reading as features of knowledge creation

Relative to the specificity of writing and reading we focus on documents. Our premise is that documents are at the center of most scientists' work. This centrality can be seen in a range of practices: from reading and writing research to doing email with colleagues. And, document-centered interactions are increasingly (but not exclusively) reliant on digital forms. We define documents broadly as material, geographic, temporal and socially constructed communication. This can include traditional, paper-based, material forms such as memos, reports. This can also include publications, online documents, web pages, and text-based communications such as email.

Documents provide a tangible entree for observing both the dynamic nature of coordination work in virtual collaborations and the uses of digital infrastructure. This approach takes into consideration: 1) the nature and uses of social ties in virtual organization 2) the web-like arrangements of documents and material and digital artifacts (i.e. computers, sheets of paper, whiteboards, software) along with the rules and norms that govern them, and 3) the everyday work practices that produce and reproduce socio-technical structures [13][23].

#### **Negotiating Epistemic Diversity**

Articulated by Karin Knorr-Cetina, "epistemic culture" is an analytic framework highlighting the unique ways researchers "maintain and adapt a particular culture that distinguishes them from other researchers" [36]. The four scholars were trained in and pursue their craft in different intellectual areas: information systems, organizational studies, information science, and social networks. Examining documents from the study highlights that this negotiation continued throughout the collaboration. Short-but-frequent email exchanges, messy document drafts, and the sharing of numerous whiteboard sketches suggest a constant state of negotiation and boundary-spanning.

#### **Documents of Consensus and Uncertainty**

Social science is often characterized by a lack of consensus among participants and a high-level of task uncertainty [35]. Collins [8] argues the basic pattern of intellectual life among the social sciences is disagreement. One consequence is that in "...many hard areas research takes place, and can be reported, within a known framework of assumptions; in soft areas, the context has often to be separately elaborated" [35]. Our findings support Collin's [8] and Whitley's [35] claims that social scientists often disagree as to what counts as empirical knowledge and how best to frame their research efforts, given their constant attention to negotiating meanings, goals and issues.

## **Technology (ICT) Use**

For technology we focus on ICT use, noting use is common and there are shared patterns. But, ICT uses are distinctly individualized. The most common ICT are commodified (e.g. laptops and computers, word processing, statistical and network visualization software). This does not differentiate social scientists from others. While CI often is seen as helping physical and natural scientists share sophisticated technological equipment such as earthquake sensor networks, radio telescopes, and computational resources, this is less common in the social sciences and nothing like this was seen in our single group of scholars.

It may be the lessened need for shared instrumentation via CI means is because the analysis of social science data is more individualized (i.e. done by a researcher at their office or home instead of the laboratory) [18] [8] [29]. We observe two aspects of this in our pilot study: (a) the pervasive and enduring uses of email and (2) the importance of pre-existing and personalized infrastructure in structuring work practices.

### *The Enduring Power of Email*

Despite the many ICT-based tools available, and as expected, the scientists in our pilot study used email as the primary way of sharing information and communicating. This seems to be a “lowest common denominator” as email was a routine part of everyone’s day-to-day work practices. Documents arrived into each other’s email inbox: a push-in reminder of workflow, tasks to-be-done and deadlines. Scholars typically ‘cc’d’ collaborators on any email involving the project to keep everyone in the communication loop. Emails amongst collaborators also tended take the form of numerous short threads. Many of these were “negotiating” emails such as: Where do I fit in? What are my responsibilities? How can I help? These emails served to inform and provide feedback about a specific issue rather than engage a wider project discussion.

### *Shifts in Work Practice/Work Structures*

The scholars in our pilot study explored various ICT to better support collaboration on project tasks. For example, at the study’s outset the researchers used a university-provided document management system (Sharepoint). Use made clear this was a static repository of the group’s materials rather than an active place for collaboration. This led them to explore (and adopt) “Dropbox” (see <http://www.dropbox.com>) as a more useful means to share and store documents because its design allowed for more indications of use and interactivity through desktop notifications of changes to the shared document repository (a form of social translucence to support distributed collaboration [e.g., 16]).

## **Independent Research Practices**

We know social scientists are trained to be critical and independent scholars [35]. So, co-authorship is more common in the natural and physical sciences than in the

social sciences [6] [19]. The research practices of study participants reflect this individuality. The four worked individually and stored project documents (physically and digitally) in highly personalized ways, following their own provenance, storage and retrieval systems. Shared digital resources were secondary. This made it easier to make ICT changes (like the Sharepoint-to-Dropbox swap) as these were not critical to any one scholar’s personalized work practices. This suggests collaborative practices are layered on, or are secondary too, individual efforts.

Evidence suggested to us that the social scientists’ digital infrastructure evolves organically over time and is largely premised on providing ways to link individual work patterns. In doing this social scientists are entrepreneurial, drawing on commodified technologies and organizational ICT infrastructures at hand and layering these onto their existing work practices through exploration and debate.

## **Documents in the Reading and Writing Practices**

In the natural and physical sciences, collaborative work is more common, research teams are typically larger, and publications tend to emphasize succinctness and precision of language. Ambiguity is reduced via formalized and standardized reporting procedures such as lab notes and short articles [35]. In contrast, the literature and writing practices in the social sciences are typically “... diverse, unsettled and open to interpretation” [4]. The scholars in our pilot study reflect this. They typically chose to write alone and followed personalized work practices regarding the uses of digital and material documents. Their writing was often directed towards publishing venues which differed from their collaborators. This led each to emphasize different points and to develop these in rhetorically different styles.

What became clear from the pilot study is that explicitly focusing on document and documenting practices made it easier to discern social scientists’ distributed and collaborative work practices. Moreover, we found that by focusing on documenting practices, we were able to learn about ICT uses, communication patterns, and these scholar’s approaches to pursuing shared scientific efforts. In returning to the literature, we find that contemporary CI studies that mention documents typically do so only in passing [19][25][28]. This may be due to the perceptions of the roles documents play in the physical and natural sciences, or it may be a reflection of the importance of documents for social scientist’s work practices.

## **NEXT STEPS**

Initial findings, albeit from a small pilot study, suggest that a document-centric approach is useful for studying distributed collaborative work. These findings also suggest CI needs for social scientists may be different than that which is currently provided or envisioned [27][19][16]. A larger-scale empirical effort is both needed and next.

## **ACKNOWLEDGMENTS**

We thank the participants in our pilot study for their willingness to allow us to study their work practices. We acknowledge financial support of Syracuse University's Faculty Assistantship Program.

## REFERENCES

- [1] American Council of Learned Societies. 2006. *Our cultural commonwealth*. New York: ACLS. DOI=<http://www.acls.org/cyberinfrastructure/acis.ci.report.pdf>.
- [2] Atkins, D. 2003. Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure. The National Science Foundation, Arlington VA. DOI=<http://www.nsf.gov/od/oci/reports/toc.jsp>.
- [3] Barjak, F., Lane, J., Kertcher, Z., Poschen, M., and Robinson, S. 2009. Case studies of e-infrastructure adoption. *Soc. Sci. Comput. Rev.* 27, 583-600.
- [4] Bazerman, C. 1981. What written knowledge does: three examples of academic discourse. *Phi. Soc. Sci.* 11, 335-359.
- [5] Beaulieu, A. and Wouters, P. 2009. E-research as intervention. In *E-research: Transformations in Scholarly Practices* N.W. Jankowski, Ed. Routledge, New York, NY
- [6] Beecher, T. and Trowler, P. 2001. *Academic Tribes and Territories: Intellectual Inquiry and the Culture of Disciplines* (2<sup>nd</sup> Ed.). Open U. Press, Milton Keynes, UK.
- [7] Berman, F. and Brady, H. 2005. *Final Report: NSF SBE-CISE Workshop on Cyberinfrastructure and the Social Sciences*. DOI=<http://www.sdsc.edu/sbe/reports/SBE-CISE-FINAL.pdf>
- [8] Collins, R. 1994. Why the social sciences won't become high-consensus, rapid discovery science. *Soc. Forum.* 9, 2, 155-177.
- [9] Cummings, J. and Kiesler, S. 2003. Collaborative research across disciplinary and organizational boundaries. In *Sci. Collaboration on the Internet*, G. Olson, A. Zimmerman, and N.Bos, Eds. The MIT Press, Cambridge, MA.
- [10] Dutton, W.H. 2007. Reconfiguring access to information and expertise in the social sciences: the social shaping and implications of cyberinfrastructure. In *Proc. 3rd Intrn'l. Conf. on e-Social Science*. (Ann Arbor, Michigan, October 7-9).
- [11] Dutton, W.H. and Meyer, E.T. 2008. E-social science as an experience technology: distance from, and attitudes toward, e-research. In *Proc. 4<sup>th</sup> Intrn'l Conf. on e-Social Science* (Manchester, UK, June 18-19).
- [12] Erickson, T. and Kellogg, W. 2000. Social translucence: An approach to designing systems that support social processes. *Transactions on Computer-Human Interaction*, 7(1), 59-83.
- [13] Fleck, J. 1994. Learning by trying: the implementation of configurational technology. *Res. Policy.* 23, 1, 637-652.
- [14] Fry, J. 2006. Coordination and control of research practices across scientific fields: implications for a differentiated e-science. In *New Infrastructures for Knowledge Production: Understanding e-Science*, C. Hine, Ed. Info. Sci. Pub., Hershey, PA.
- [15] Fry, J. and Schroeder, R. 2009. Towards a sociology of e-research. In *E-Research: Transformations in Scholarly Practices*, N.W. Jankowski, Ed. Routledge, New York, NY.
- [16] Hine, C. 2006. Computerization movements and scientific disciplines: the reflexive potential of new technologies. In *New Infrastructures for Knowledge Production: Understanding e-Science*, C. Hine, Ed. Info. Sci. Pub., Hershey, PA.
- [17] Knorr-Cetina, K. 1999. *Epistemic Cultures: How the Sciences Make Knowledge*. Harvard Press, Cambridge, MA.
- [18] Latour, B. and Woolgar, S. 1986. *Laboratory Life: The Social Construction of Scientific Facts*. Sage, Beverly Hills, CA.
- [19] Lin, Y.-W., Procter, R., Halfpenny, P., Voss, A., and Baird, K. 2007. An action-oriented ethnography of interdisciplinary social scientific work. *Proc. Of the 3<sup>rd</sup> e-Social Science Conf.* (Ann Arbor, MI, October 7-9).
- [20] Meho, L. and Tibbo, H. 2003. Modeling the information-seeking behavior of social scientists: Ellis' study revisited. *J. Am. Soc. Inf. Sci. Tec.* 54, 6, 570-587.
- [21] Merz, M. 2006. Embedding digital infrastructure in epistemic culture. In *New Infrastructures for Knowledge Production: Understanding e-Science*, C. Hine, Ed. Info. Sci. Pub., Hershey, PA.
- [22] Olson, J., Hofer, E., Bos, N., Zimmerman, A., Olson, G., Cooney, D. and Faniel, I. 2003. A theory of remote scientific collaboration. In *Scientific Collaboration on the Internet*, G. Olson, A. Zimmerman, and N.Bos, Eds. The MIT Press, MA.
- [23] Østerlund, C. 2008. Documents in place: demarcating places for collaboration in healthcare settings. *CSCW.* 17, 195-225.
- [24] Pieri, E. 2009. Sociology of expectation and the e-social science agenda. *Info., Comm. & Society.* 12, 7, 1103-1118.
- [25] Ponti, M. 2010. *Actors in Collaboration: Socio-technical Influence on Practice-Research Collaboration*. Doctoral Thesis. University of Gothenburg. DOI=<http://bada.hb.se/bitstream/2320/5458/2/Avhandling.pdf>.
- [26] Rosenberg, A. 2008. *Philosophy of Social Science, 3<sup>rd</sup> Edition*, Westview Press, Boulder, CO.
- [27] Sawyer, S. and Tapia, A. 2007. From findings to theories: institutionalizing social informatics. *Info. Soc.* 23 4, 263-275.
- [28] Schroeder, R., and Fry, J. 2007. Social science approaches to e-Science: Framing an agenda. *JCMC.* 12, 2, 229-248.
- [29] Schroeder, R. and Spencer, D. 2009. Social scientists and the domestication of e-research tools. In *Proc. 5<sup>th</sup> International Conf. on e-Social Science* (Cologne, Germany, June 24-26).
- [30] Smith, D. 1990. *Texts, Facts, and Femininity: Exploring the Relations of Ruling*. Routledge, London, UK.
- [31] Sonnenwald, D.H., Lassi, M., Olson, N., Ponti, M., and Axelsson, A-S. 2009. Exploring new ways of working using virtual research environments in library and information science. *Libr. Hi. Tech.* 27, 2, 191-204.
- [32] Star, S.L. and Ruhleder, K. 1996. Steps toward an ecology infrastructure: design and access for large information spaces. *Inform. Syst. Res.* 7, 1, 111-134.
- [33] Traweek, S. 1988. *Beamtimes and Lifetimes: The World of High Energy Physicists*. Harvard University Press, MA.
- [34] Wessels, B. and Craglia, M. 2007. Situated innovation of e-social science: integrating infrastructure, collaboration, and knowledge in developing e-social science. *JCMC* 12, 358-7.
- [35] Whitley, R. 2000. *The Intellectual and Social Organization of the Sciences* (2<sup>nd</sup> Ed.). Oxford U. Press, Oxford, UK.
- [36] Wouters, P. and Beaulieu, A. 2006. Imagining e-science beyond computation. In C. Hine (Ed.), *New Infrastructures for Knowledge Production: Understanding E-Science*. Information Science Publishing, Hershey, PA.