Scientific Collaboration on the Internet, edited by Gary M. Olsen, Ann Zimmerman, and Nathan Bos. Cambridge, MA: The MIT Press, 2008. 432 pp. $45.00 Hardback. ISBN: 0262151200.

Reviewer: Kalpana Shankar, Indiana University

That science and scientific research has been increasingly physically dispersed, highly collaborative, and resource and data intensive, has been a given for several decades. These shifts in the nature of scientific practice have engendered the need for the development of tools for data management, as well as information and communication technologies to enhance collaboration. Furthermore, they have generated research interest in the practice of scientific collaboration, scientific networks, and the nature of knowledge production itself. The scope and scale of scientific collaboration tools have expanded to encompass home-grown technologies for small collaborations to very large scale cyberinfrastructure/e-science initiatives that support numerous, widely dispersed researchers that share tools, data, and services.

Many articles and several books (cf. Borgman 2007) have been published in recent years that document these shifts and reflect upon what these collaborations mean and how and why they work (or don't). These works seem to share interests that go beyond the design and deployment of systems for different disciplines (or multidisciplinary ventures). Instead, they argue, we are at a point where we need to ask how we can evaluate the success or failure of such systems and establish best practices. These steps, the authors argue, are necessary to incentivize the development and use of collaborative systems, overcome challenges, and understand the new kinds of management, both centralized and distributed, that are needed to manage these often world-wide projects.

This book, *Scientific Collaboration on the Internet*, tackles these themes by exploring them within an overarching question: Is there a science of collaboratories? It collects contributions from members of the *Science of Collaboratories* project, centered at the University of Michigan-Ann Arbor. The book is bracketed by two chapters written by the editors. The introduction discusses the history of collaboratories, identifies and disambiguates key terms, introduces the structure of the *Science of Collaboratories* project, and raise some questions that run throughout the rest of the book. The chapters of the book are organized in six parts: the contemporary collaboratory vision; perspectives on distributed, collaborative science; physical sciences; biological and health sciences; earth and environmental sciences; and the developing world. The first two sections provide overview concepts and theoretical perspectives; the others emphasize case studies of collaboratory development and use across the disciplines. The last section focuses on case studies of collaborative research in the Global South.

Although there are many terms of art that are used to describe collaborative systems, the authors refer throughout to them as "collaboratories." The editors of "Scientific Collaboration on the Internet" cite a 1989 report to define this term as "center[s] without walls in which researchers can work together regardless of physical location" (Wulf 1993). They acknowledge that this term has been considerably expanded and other terms have been more in vogue, such as e-Science (in Europe) and cyberinfrastructure (in the United States). These visions, they argue, are heavily influenced by the earlier vision of

collaboratories and that many of these issues raised in 1989 are still relevant today. The longevity of the term, its geographical neutrality, and its ability to encapsulate the many issues that still pertain make the term useful even now.

Although the focus of this volume is ostensibly on the science of collaboratories, the reader comes away with a sense that a successful collaboratory project is as much art as science. The editors and many of the chapters' authors acknowledge the usefulness of quantitative measures of success (numbers of petabytes stored/moved, articles published, researchers served, etc.), but emphasize that these are often not known until well after a project's completion. There are numerous other measures of success: seamless integration into daily work practice, new kinds of questions that can be asked and answered, and so on. For example, Olson et al draw upon work from the Science of Collaboratories project as well as other studies of collaboration to synthesize and describ factors that are important determinants of success in collaboratory development and use them to develop a theory of remote collaboration. They group these factors into several categories: effects on science itself, science careers, learning and education, inspiration to others, funding and public perception, and tool use. While much of the evaluation of collaboratories tends to focus on summative evaluation (i.e., evaluation of impact and outcomes), they argue for the need to conduct evaluations much earlier in collaboratory design and use. They suggest that the factors they enumerate are well-suited to conducting formative and in-process evaluations of collaboratories to pre-empt problems, and guide investments of time and other resources. Reading the specific case studies makes it very clear to the reader that the successes and failures of collaboratory deployment in different disciplines must be considered in the context of the discipline itself, its history of collaborative practice, funding trajectories, and the daily, mundane work practices of normal science.

For example, Hofer et al write about the community of researchers in high energy physics and the Large Hadron Collider collaborations. This field has long relied on very large project teams and specialized tools for collaboration and data analysis. Many of these researchers, they note, were not used to collaborating in the way that was made possible by e-science. The transition took decades and the technology was not deterministic; it co-evolved with the research community, making it possible for researchers to research new questions, even shaping the trajectory of the field itself. For example, the Large Hadron Collider is anticipated to provide over fifteen petabytes of data a year, with 1500 to 2000 physicists per experiment. The sheer scope and size of this project have shaped numerous other initiatives around the development of appropriate infrastructure for the high energy physics and other scientific communities, which in turn have used this infrastructure to create new collaborative projects around digital astronomy, molecular genetics, computer science, and other disciplines. This co-shaping of technology and scientific community is a theme that may not have been specifically articulated by the editors, but it is deeply important to understanding the science of collaboratories.

The development of these collaboratories and e-science has outstripped their evaluation, as Sonnenwald, Whitton, and Maglaughlin point out. Appropriate evaluative questions and methods pervade this book for several key reasons, some internal to the discipline using the collaboratory and others extending outward to other interested stakeholders.

Evaluation has political dimensions. These systems represent enormous layouts of public funds and other resources; as a result, justifying their expense to the holders of the purse strings is essential. Secondly, appropriate evaluation is necessary to examine which features and approaches are transferable across projects, thus minimizing waste. Third, evaluation is necessary to insure usability and usefulness of the collaboratory for the multiple roles it may be asked to play over time for its multiple stakeholder communities. Building a system that researchers are reluctant to use, or is only of short-term use, is not productive. Several of the chapters address these challenges.

These fundamental questions have to be asked at every stage of the collaborative project– design, initial use, and long-term re-use. Different methods and approaches to asking and answering questions are essential and the authors explore the range. For example, Sonnenwald and co-authors take on the task of evaluating scientific outcomes afforded by the nanoManipulator Collaboratory System through a mixed-methods approach. They use controlled experiments to assess the quality of scientific work produced via face-to-face collaboration vis-à-vis remote collaboration. Task performance measures, post-interviews with participants, and post-questionnaires resulted in a substantive evaluation from multiple perspectives. Interestingly, in this article, these measures did not support the hypothesis that working remotely was less effective and more difficult than face-to-face collaboration, supporting the position that there is potential for collaboratory adoption. Several articles include ethnographic approaches to evaluation in which the authors examine the integration of collaboratories into the daily work of the scientists using them. For example, Hackett et al studied the development and use of the National Center for Ecological Analysis and Synthesis (NCEAS) for seven years and collected 140 hours of ethnographic data, which provided extensive context for the documentary analysis, interviews, citation data, and surveys they also conducted. This longitudinal study afforded the Hackett and colleagues to observe the processes of multiple developments: the research organization, the data archives and other technologies that constitute the organization, and even the field of ecology itself.

Thematically, the authors included in this collection do not shy away from discussing the challenges, problems, and failures of collaboratories. They write of failures to engage the scientists using the systems, difficulties with organizational structures, funding concerns, pitfalls and perils of remote collaboration, and the many work-arounds that scientists adopt in working with systems not necessarily of their design or even choosing. For example, Spencer et al discuss many of the lessons learned in developing and deploying NEESgrid, an earthquake engineering cyberinfrastructure, as sources of "undocumented risk". They point to the need for awareness of professional cultures (which they call "cultural dissonance"), strong leadership, and tangible and periodic deliverables that prove to all stakeholders that needs are being met and that the cyberinfrastructure is meeting expressed needs. These findings dovetail with and instantiate Olson et al's call for collaboratory evaluation that is conducted at multiple stages of the project and not just at the conclusion.

Most of the authors do a masterful job of contextualizing and explaining these challenges without waving them away or worse, taking a technologically deterministic approach that is often so troubling in the e-science literature. However, the authors also excel for the

most part at placing these studies in a broader framework that is extensible, replicable, and implementable: in short, they make a great deal of progress in getting to the "science of collaboratories" that is the heart of this project. This balance is what makes this book of interest to many fields, since the successful implementation of large-scale information infrastructures, digital work tools, and the like are of interest to other communities as well. The cases generally point out that the challenges are not technical ones, but instead sociotechnical ones: of internal politics, work practices that are not harmonized with the new technologies, tacit knowledge and information practices, and on a macro-scale, funding and larger political priorities that are not necessarily aligned with the daily work needs of the scientists expected to use the collaboratories.

One of the other pieces that make this book unique is its inclusion of case studies of collaboratories in the developing world, including a partnership among AIDS resaerchers in the United States and South Africa and a chapter built upon interviews with researchers from China, Korea, Morocco, New Zealand, South Africa, and Taiwan. This section further opens up important questions and concerns of digital divides, international development, and the globalization of science. The chapters point to some successes for collaborations between developing and developed countries for improved access to tools, communication, and data. Nevertheless, numerous technical and cultural barriers and lack of institutional infrastructures are pointed to as obvious obstacles to successful collaborations. In a chapter on the effect collaboratories have on scientists from developing countries, Luo and Olson point to an ironic side-effect of technologically mediated collaboration: the full potential of collaboratories across borders are only realized when in-person conference travel and extended lab visits are possible.

The editors' conclusion provides an excellent capstone to this book. It serves as more than a summary of the themes and issues raised in the other chapters. Instead, it serves as a larger meditation. In addition to the challenges faced by those who wish to map the science of collaboratories and distill its important problematics, the authors use collaboratories to surface and foreground other topics that might well be called the political economy of knowledge. They use collaboratories to explore the intersections of diverse disciplines such as management and Science, Technology, and Society (STS); acknowledge perspectives contributed by emerging fields such as behavioral economics; reflect on the Information School (I-School) movement that is influential in computing and information education; and examine the policy, pedagogy, and pragmatic concerns that will be essential to address if collaboratories, e-science/cyberinfrastructure, etc. are to fulfill their expensive promises. Do these chapters answer the question, "Is there a science of collaboratories?" If one is expecting a clear-cut answer to that question, it is no. Instead, the authors and editors go further in getting to that answer than others: they develop a clear framework based on the disciplinary roots of this field, establish key research questions, and explore emergent trends.

References

Borgman, C.L. 2007. *Scholarship in the Digital Age.* Cambridge: MIT Press.

Wulf, W.A. 1993. The collaboratory opportunity. *Science* 261(5123): 854-855.