# A Bit More to It: Scholarly Communication Forums as Socio-Technical Interaction Networks

**Rob Kling, Geoffrey McKim, and Adam King**
*Center for Social Informatics, SLIS, Indiana University, Bloomington, IN 47405. E-mail: kling@indiana.edu*

**In this article, we examine the conceptual models that help us understand the development and sustainability of scholarly and professional communication forums on the Internet, such as conferences, pre-print servers, field-wide data sets, and collaboratories. We first present and document the information processing model that is implicitly advanced in most discussions about scholarly communications—the "Standard Model." Then we present an alternative model, one that considers information technologies as Socio-Technical Interaction Networks (STINs). STIN models provide a richer understanding of human behavior with online scholarly communications forums. They also help to further a more complete understanding of the conditions and activities that support the sustainability of these forums within a field than does the Standard Model. We illustrate the significance of STIN models with examples of scholarly communication forums drawn from the fields of high-energy physics, molecular biology, and information systems. The article also includes a method for modeling electronic forums as STINs.**

## Introduction

We are in a period where many analysts have high expectations of how information technologies (IT) can substantially improve scholarly communications, as well as communications with professionals and broader publics. Funding for the Internet was frequently justified in terms of speeding up and increasing access to various scholarly communication forums. Many of these expectations are based on conceptions of high-speed telecommunications, enabling information to move rapidly and relatively inexpensively "anywhere anytime"; thus enabling low cost and widely available electronic journals, pre-print servers, collaboratories, and so on.

Received May 18, 2000; revised November 29, 2001; accepted May 31, 2002



These expectations have both fostered and conditioned the development of a wide variety of new electronic scholarly communication forums. Many of these new electronic communication forums have been developed and even more are in the planning stages. These electronic communication forums take a wide variety of forms: electronic editions of paper journals, pure electronic journals, working article repositories, postpublication archives, pre-print servers, collaboratories, crosslinked Webs of resources, gene databases, etc. Later in this paper, we will discuss the architectures of four of these in some detail. The number of these new forums is growing rapidly, and the variety of their architectures is remarkable.

We refer to a broad family of communication forums as Scholarly Communication Forums (SCFs). The purposes of this term are: to abstract from any specific restrictions about the technical form of the forum (i.e., face-to-face meeting, paper journal, linked Web, central server-based repository, etc.) and the communicative role of the forum (i.e., data to be used in designing a study, teaching materials, unrefereed working papers, pre-prints, postpublication articles). Many SCFs are not purely digital; sensors may be physical devices with electronic controls, Web-cameras may be used to share real-time visuals of physical laboratory control rooms, etc. At the core of an electronic SCF (e-SCF), however, is some form of computer-mediated communication. In practice, many e-SCFs are not purely electronic either, as some of their participants may also work or meet face-to-face as well.

It should be not be surprising that some of the proposals for new e-SCFs, which are frequently replications of the well-known physics archive arXiv.org, have been controversial, and required radical restructuring.[1] Many e-SCFs have social properties that are not compatible with highly

[1] See Kling, Fortuna, and King (2001) for a detailed study of the case of the transformation of the original E-BioMed proposal into a radically different PubMed Central.

entrenched and durable scholarly communication practices, such as peer review.

In this empirically grounded theoretical exposition, we examine the kinds of design conceptions that help us better understand behavior within an e-SCF, and the sustainability of new e-SCFs within a scholarly field. First, we will present and document the model that is implicitly advanced in most discussions about electronic scholarly communications. We will explain why we find that this "Standard Model," because of its emphasis upon the information processing features of an e-SCF, provides an inadequate understanding of important human behavior with e-SCFs. Second, we will present an alternative kind of model, one that considers technologies as socio-technical networks. Socio-technical network models examine e-SCFs, identify key relationships between different technologies, social actors, resources (including money flows), and legal regulations. The term network is metaphorical, and refers to the structured relationships between these diverse elements in the use of a particular e-SCF. These networks cannot always be drawn as directed graphs. Subsequent sections of the article explain some heuristics for modeling an e-SCF as a socio-technical network, and illustrates the insights that come from STIN models with four empirical examples.

The article includes two methods sections: one methods section (Appendix B) describes our methods of data collection and analysis, whereas a second section (below) provides heuristics for modeling an e-SCF as a STIN. Comments from reviewers and participants in several formal talks about STIN models indicate that there is much more interest in learning how to model a situation as a STIN than there is in the specific empirical methods that helped us to develop the STIN models. Consequently, we have placed our empirical methods in Appendix B while we discuss STIN modeling methods in a subsequent section of this article.

We have found that these network models provide a richer understanding of human behavior with e-SCFs, as well as a more complete understanding of the conditions and activities that enhance the sustainability of an e-SCF within a scholarly field. We will contrast these socio-technical network models with a more limited, but much more commonly advanced, socio-technical systems models. We will conclude by illustrating the significance of a specific kind of socio-technical network model—Socio-Technical Interaction Networks (STINs)—with examples of e-SCFs drawn from the fields of high energy physics, molecular biology, and information systems. A Socio-Technical Interaction Network (STIN) is a network that includes people (including organizations), equipment, data, diverse resources (money, skill, status), documents and messages, legal arrangements and enforcement mechanisms, and resource flows. The elements of a STIN are heterogeneous. The network relationships between these elements include: social, economic, and political interactions.

Although our primary interest is in scholarly communication, and we have used STINs to examine social behavior in scientific collaboratories and in the development of electronic article repositories, the insights from STINs can also be extended to other electronic communication forums, including distance education electronic classrooms, professional development sites, and even community forums such as Weblogs or auction sites. STIN models help us to understand human behaviors in the use of technology-mediated social settings that would be anomalies in terms of the Standard Model of IT. Before describing our methods for modeling an e-SCF as a STIN, we will first provide a richer illustration of STINs by examining one type of e-SCF—a collaboratory.

### *Collaboratories as STINS*

We can begin to illustrate STINs by contrasting different ways to conceptualize collaboratories. The term "collaboratory" was coined by William Wulf, by combining the words "collaboration" and "laboratory." He defined a collaboratory as "a center without walls, in which users can perform their research without regard to geographical location—interacting with colleagues, accessing instrumentation, sharing data and computational resources, and accessing information in digital libraries" (National Research Council, 1993, p. 7). The standard model of e-SCFs treats collaboratories as collections of scientific instruments and the information technologies to enable their use and to support collaboration by people who are not co-located. Thus, a collaboratory is composed of a set of technologies; the *sociality* of collaboratories comes from the (collaborative) interaction of "the users" of the collaboratory with each other.

In contrast, Myers (1999) characterizes his Environmental Molecular Sciences Laboratory as one in which scientists who wish to use it have to understand the instrumentation, learn how to prepare samples for it, learn how to use it, and perhaps have the instruments reconfigured for their studies. Gaining this type of knowledge and skills requires help from scientists who have significant responsibility for selecting, configuring, and maintaining specific instruments. In Wulf's image, there is no one "in the collaboratory" before its users arrive and after they leave. In Myers' account, however, each major instrument has a scientist at its side before "users" come and after they leave. Further, to utilize instruments in a collaboratory, a scientist (or team) at a remote location has to develop social relationships, such as trust, with the scientists who know the instruments and who can be viewed as "inside" the collaboratory.

Another collaboratory director, N. Zaluzec (personal communication, December 3, 1999), stressed that the resident scientists who work "inside" a collaboratory are not just cognitive informants. Sometimes they act as evangelists who try to recruit other scientists to use their facilities. They work as collaborators, and often expect to be co-authors of major papers that are based on their instruments and expertise. Their helpfulness and skill as collaborators may be the lure that draws in researchers and maintains collaborations, even when their instruments can be found in other laboratories. In the view of practicing scientists such as Myers and Zaluzec, social behavior constitutes a collaboratory as much as its instrumentation.

In our terms, a collaboratory is constituted as a Socio-Technical Network: a network that brings together people and equipment in ways that are not meaningfully separable for understanding "how collaboratories work." Later on, we will refine this notion of a socio-technical network into STIN models to emphasize the importance of the character of interactions between people, between people and equipment, and even between sets of equipment. Some of these interactions may involve direct participants, and cannot be specified *a priori*. For example, the scientist who develops sustained trust is likely to work with a collaboratory very differently than one whom local experts are reluctant to work with. In extreme cases, a collaboratory evangelist who turns out to be a zealot can drive away other scientists who could have been potential collaborators (D. Atkins, personal communication, November 1998). The collaboratories are themselves "nodes" in a larger STIN of related (simultaneously competing and cooperating) laboratories, funding relationships, publishing opportunities, career advancement, and so on.

The value of an alternative to the Standard Model of e-SCFs is illustrated by the social interactions that energize collaboratory life that are briefly sketched in these accounts by Myers, Zaluzec, and others. These social and technical interactions seem to shape the work of collaboratories and their intellectual location in their own scholarly fields. They are anomalies relative to the standard model where information-processing features are central. But these kinds of socio-technical interactions are central to the STIN models that we will examine in this article.

### *The Roles of Standard Models and STIN Models of an e-SCF*

STIN models help us to more completely represent e-SCFs than do the standard models. These models are parts of explanations that enables us to understand the processes of technological change and their impacts upon our social, cultural, economic, political, and personal lives. Neither the standard models of an e-SCF nor STIN models are complete theories, which by themselves, lead to strong predictions. However, the standard models of an e-SCF in Information Science have been integral to visions of the possibilities of organizing e-forums to support distributed work, community life, education at a distance, and electronic publishing that is faster, cheaper, more accessible than paper-based alternatives, etc. (see, for example, Buchanan et al., 2000; Lesk, 1997; Sheldon, 1998; Varian, 2000).

However, these standard models of an e-SCF often mislead many analysts about such matters as estimating the costs and complexity of electronic journals, what may be problematic about collaboratories, the institutional complexity of e-print archives, etc. What is left out of the standard models are important features of very specific technologies and settings in which people try to use them, the organizational complexity in which IT-based services are provided and embedded. In addition, the standard models of an e-SCF are often elements of analyses that emphasize certain kinds of task rationality and cooperation and ignore the complexity and nuance of human motivations and relationships.

Many people who try to communicate via e-forums find that they are more cumbersome and time-consuming than they anticipate. Organizers of e-forums where participation is discretionary, have often found that participation rates are lower than they anticipate, that the costs and complexity of developing and maintaining innovative e-forums (or preexisting e-forums for new kinds of groups) is often much higher than expected. We need new theories and ontologies that can help us to explicitly understand phenomena like these.

STIN models help to map some of the key relationships between people and people, between people and technologies, between technologies and their infrastructures, and between technologies that constitute e-forums in use. The STIN models may be most concrete for representing some key elements and relationships for a specific e-forum (such as a particular e-journal). They can be used in a more heuristic way to represent these relationships of a prototypical e-forum (such as a typical e-journal).

## The Standard Model of Electronic Scholarly Communication Forums

There is a vast literature about the development, structuring, and use of e-SCFs in scholarly communication (Bailey, 2001). We have abstracted from an extensive reading of this literature, a series of underlying commonly held (i.e., frequent but not universal) assumptions about e-SCFs, and have synthesized these assumptions into a "Standard Model of e-SCFs." We then contrast this standard model with STIN models of e-SCFs. In a subsequent section, we will examine four major e-SCFs in light of the standard models and STIN models.

The Standard Model of e-SCFs is constituted from two key axioms:

### *Axiom 1*

Actor behavior is motivated by and/or constrained by the Information Processing (IP) features of the technology of an e-SCF.

TABLE 1A. Characteristics of the standard model for access to—and business models for—electronic publishing resources.

| Characteristic | Description | Example in the E-Publishing literature |
|---|---|---|
| INFRASTRUCTURE | | |
| Easy Access | Materials are available to readers 24 / 7, from any location in the world, via the Internet. However, access may be controlled if it is deemed desirable. Scholars living in far-flung and impoverished countries will benefit the most. | (Harnad, 1990) |
| RESOURCE FLOWS AND BUSINESS MODEL | | |
| Business Model | Business models are not clear, or clearly specified. Costs to the reader are emphasized (such as free or fee); however, the underlying business model is not. | |
| Low Costs of Production and Distribution | Lower production, publication and reproduction costs as archival space is inexpensive. Most costs will be covered by grants or institutions supporting the research; also, the increasing commercialization of the Internet will lower the costs of scholarly communication. The trend towards cheaper storage memory will facilitate disintermediation, with attendant increases in efficiency. It is cheaper to connect to the Internet than to acquire a library from scratch. Materials should eventually be free as S/L/P (subscriptions, licensing, and pay-per-view) is phased out. Most importantly, as the costs of computing and data storage approach zero, the distribution and transaction costs for electronic materials should also approach zero. | (Harnad, 1997; Harnad & Hemus, 1997; Odlyzko, 1995) |
| Fast Publication | Timely publication and display of ideas, content, and results. The quality of the final product is superior to some traditional paper-based journals. In addition, it is the only way to deal with the exponential increase in the creation and publication of scholarly articles. | (Ginsparg, 1994) |

*Axiom 2*

Actors can most usefully be considered as individual users who can choose to, or not to, use a specific e-SCF (as authors, readers, etc.).

Two important consequences of these axioms are reflected in the standard model. First, if actor behavior is motivated by the information-processing features of an e-forum, then discussions of the e-SCF will be dominated by discussion of these features and the efficiencies enabled by them (such as being able to publish more cheaply and quickly). Second, when the focus is only on the technology and the individual actor as a user of that technology, then the cultural work context of the actor, and the ecology of communications in which that actor is embedded in is relatively unimportant.

In Table 1A, B, and C, we use the literatures of electronic scholarly communications to explicate some of the characteristics of the standard model.

## Socio-Technical Network Models of Communication Forums

The standard models of information technologies characterize them as tools whose adoption by organizations is based on norms of rationality and technical efficiencies. Different ways of configuring technologies in practice are of relatively minor significance. In the case of e-SCFs, analyses that rely upon the standard models emphasize the rapidly decreasing price and performance of hardware, the declining size and weight of equipment, and the ubiquity of telecommunications to help people to readily move data within and across organizations. These analyses often tilt

TABLE 1B. Characteristics of the standard model for the production of electronic publishing resources.

| Characteristic | Description | Example in the E-Publishing literature |
|---|---|---|
| PRODUCTION | | |
| Materials | Documents are easy to update and markup. | (Ginsparg, 1994) |
| Storage | Materials are easy to archive and retrieve. Developments in technology will render these archives highly resistant to malfeasance and tampering, thus ensuring the safety and integrity of the contents. | (Harnad, 1990) |
| Formats | Materials can be accessed in wide variety of formats: PDF, gzip, TeX, LaTeX, ASCII, etc. | (Ginsparg, 1994) |
| Multimedia | Publications can include other types of media (audio files, charts, graphs, photos, executable files, etc.) | (Harnad & Hemus, 1997) |
| Extension | Extension of e-SCFs is generally unproblematic; extension comes in the form of greater numbers (of articles, of genes, etc.) and of more fields (high-energy physics to nuclear physics to chemical physics, etc.) | |

TABLE 1C. Characteristics of the standard model in the electronic publishing literature.

| Characteristic | Description | Example in the E-publishing literature |
|---|---|---|
| SOCIAL BEHAVIOR | | |
| Legitimacy and Control | Controls are not as restrictive as in traditional journals; peer review is not as crucial at some levels. Participation will be more open and democratic. | (Harnad, 1997) |
| Collaboration through Hyperlinking | Materials are available to a wider audience, including areas that are not necessarily related but that might have common research interests. It will promote a more collaborative form of research. | (Ginsparg, 1994) |
| Collaborative Communications | Information technologies provide tools to help organize and analyze data, and also to support communications amongst participants. Advances in the tools (such as improved real time communications through video conferencing and instrument controls) will make the collaboratories more productive | (Scott, 1999) |
| | Collaboratories differ in some key ways from many other kinds of e-SCFs, because they support symmetrical two-way communication between project participants. They may also include more heterogeneous kinds of information resources (such as whiteboards, measurement instruments, raw data sets, and analytical software). There is less emphasis upon "publication" than with many other e-SCFs. | |
| Social and Organizational Barriers | Good metadata will radically reduce many organizational (access control) and social barriers (linguistic). As long as data is properly coded (for example, as peer-reviewed, published in a journal, sponsored by a particular organization, etc.), then seamless integrated access to e-SCF's will be enabled. | Harnad, 1999) |

toward economic and technological determinisms. For example, some scientists believe that the high-energy physics working article (e-print) server at the Los Alamos National Laboratory (now called arXiv.org) is the model of publishing that will sooner or later be followed by all of the sciences: it is "just a matter of time" (e.g., Harnad, 1999).[2]

Empirical research studies about e-SCFs and other ICTs that have examined their configurations have found that "almost identical technologies" are often configured very differently in practice, and that these configurational differences can influence their use and uses (Fleck, 1994; Kling & Covi, 1995; Williams, 1997; Kling, 1999). Each social group may have to configure e-SCFs to use them most effectively. What are claimed as "best practices" may work well for some groups and organizations, but not others. Thus local R&D costs can remain relatively high and the overall costs of using new e-SCFs may not fall rapidly.

These theoretical differences are of major practical consequence. In the case of e-SCFs (broadly), analyses based on the standard models lead us to emphasize the rapidly increasing price and performance of hardware and to anticipate media convergence. Some go farther and "believe that the paper document is dead; we are just not aware of it yet" (Wulf, 1999). Further, the standard model suggests that one can expect that a few well-crafted pilot projects—done almost anywhere—can help to establish "best practices" that everyone else can follow. In this view, a first stage of social learning about new e-SCFs can be exploratory and costly; however, subsequent uses elsewhere can be imitative and relatively inexpensive.

### *Limited Socio-Technical Conceptions of Communication Forums: The Socio-Technical Systems Model*

Some analysts have been using the term "socio-technical" informally to understand e-SCFs and other IT applications. There are two common uses that differ considerably from ours. The first common use is that IT applications are "technologies" that have social consequences. Technologists, such as computer scientists, build the IT applications; social scientists then study their consequences for work, organizational forms, and other social behavior.[3] We will show how the STIN Models can be put to better use than this.

A second common use is reflected in some of the discussion of collaboratories (see Finholt, in press; National Research Council, 1993; Olson et al., 1998, 2001; Teasley & Wolinsky, 2001). In this view, e-SCFs generally, and collaboratories in particular, can be viewed as layered systems. The bottom layers are various technologies, such as computer networks and specific software applications. The "tool sets" of the collaboratory are "the technical layers."

[2] The efforts of NIH Director Hal Varmus in 1999, to develop an e-print server for the fields of biomedical research was extremely controversial; it had many supporters and many critics. This achitecture that would have allowed authors to post their own articles was abandoned for an architecture that placed control in the hands of journal editors and scientific societies. See Kling, Fortuna, and King (2001) for details. Some analysts argue that the arXiv.org approach is appropriate, but that the political economy of scholarly publishing must be seriously restructured (e.g., Guedon, 2001).

[3] A 1998 "socio-technical summit" about Internet2 at the University of Michigan was organized on this conception of socio-technical.

The "socio" arises when people use the e-SCF to communicate. The behavior of the participants should be understood as "socio-technical" because of the strengths and limitations of the tool sets at any given time. This conception can play a useful role for some purposes, but it also separates "socio" from "technical" by virtue of how the layers are conceptualized. Even so, this conception has undergirded some interesting and important research done under the rubric of Computer Supported Cooperative Work (CSCW) (Galegher, Kraut, & Egido, 1990). This socio-technical systems perspective also has a long heritage in the form of work and technology studies research coming out of the Tavistock Institute of Human Relations in the 1960s, which viewed work arrangements as being composed of separate social systems and technical systems, which had to be jointly optimized (Emery & Trist, 1960).

### *Socio-Technical Networks*

In our view, the concept of socio-technical behavior should be used to refer to more tightly integrated conceptions of the interaction between people and technologies. In particular, what are referred to as technologies are developed within a social world and supported by technicians and others with specialized skills.

Academics are familiar with oral forms of scholarly communication and its alteration by electronic communication. We will use this example to illustrate one view of Socio-Technical Networks.[4] Amplifiers in lecture halls, video conferencing, and videotape alter both the scope and nature of audiences that scholars can reach, and the relationships between those audiences and lecturers and speakers. These electronically enhanced forums do not simply provide more communication, but also alter the ways that people speak and interact. The speaker may have to work in a special conference room and be separated from local participants by complex equipment (thus altering local interactions). As the audience scales up in size, or moves out in space and time with real-time video or asynchronous videotape, the informal give-and-take between speakers and listeners becomes more difficult (in contrast with the smaller face-to-face seminar). On the other hand, people watching a videotape may privately replay sections to enhance their comprehension, whereas in a face-to-face meeting they may have to ask questions (that might also embarrass the speaker or questioner). Further, when videotape is used, there may be copyright constraints that restrict when and for whom the videotape may be played.

Voice-based face-to-face conferencing, video conferencing, and videotape are not simply equipment. They shape scholarly communications as Socio-Technical Networks in which social characteristics such as controls over access (via pricing and distribution channels), and social protocols for regulating discussions between speakers and audience also influence the character of scholarly communications. The use of the term "network" in this discussion is primarily metaphorical; the participants in socio-technical networks may or may not be "wired" via various technological or social networks. These socio-technical networks are heterogeneous because they bring together different kinds of social and technological elements—camera operators, their cameras, and speakers; editors and their technologies; copyright laws and perhaps even lawyers; funders and their budgets; producers and their time schedules—into a complex web.

The nature of videotape pricing and the distribution channels can lead to minor or huge expansions beyond the original conferees. Despite scholars' potentially broader access to conference talks via videotape distribution, a face-to-face conference is different from a videotape collection of its talks because of the diverse informal discussions and important social networking that conferences support. The face-to-face conference and the videotape collection are different scholarly communication forums with overlapping capabilities, but which also support very different forms of scholarly communication.

In a similar way, a scholarly journal can also be understood as the product of a socio-technical production and communications network. The publishing communication network includes both full-text materials (articles and books), and indexes and pointers to these materials (including book reviews, abstract sets, specialized bibliographies, and diverse catalogs). The network brings together authors, editors, reviewers, readers, publishing staffs, and others (see Fig. 1 for a simplified depiction).

In addition, the journal is embedded in other socio-technical networks: communication and reward networks of the fields and institutions in which its authors participate, and the libraries or archives that store copies, index them, abstract them, and so on (see Fig. 2 for a simplified depiction). The journal's viability will depend upon how it is positioned within this second network (i.e., whether it has the standing to attract high quality authors and readers). These (two) STINs that constitute the journal and the network in which it participates are directly linked. For example, the editorial board helps to constitute the journal. But the prestige of editorial board members is one influence on potential authors' perceptions of the journal's quality. This is especially important for new journals (Kling, 1999; Kling & Covi, 1995).

In Figs. 1 and 2, we have linked various participants and technologies in ways that constitute major parts of a STIN. But we have not labeled the links of these STINs to identify the nature of the social interactions, because they can be

[4] We use the term network rather than system because these configurations are open ended and not "designed." "A network, by contrast, is loosely organized; often imperfectly integrated; has nodes that may be part of one to many other networks as well; and can be reconfigured" (P. Edwards, personal communication, December 28, 1999).

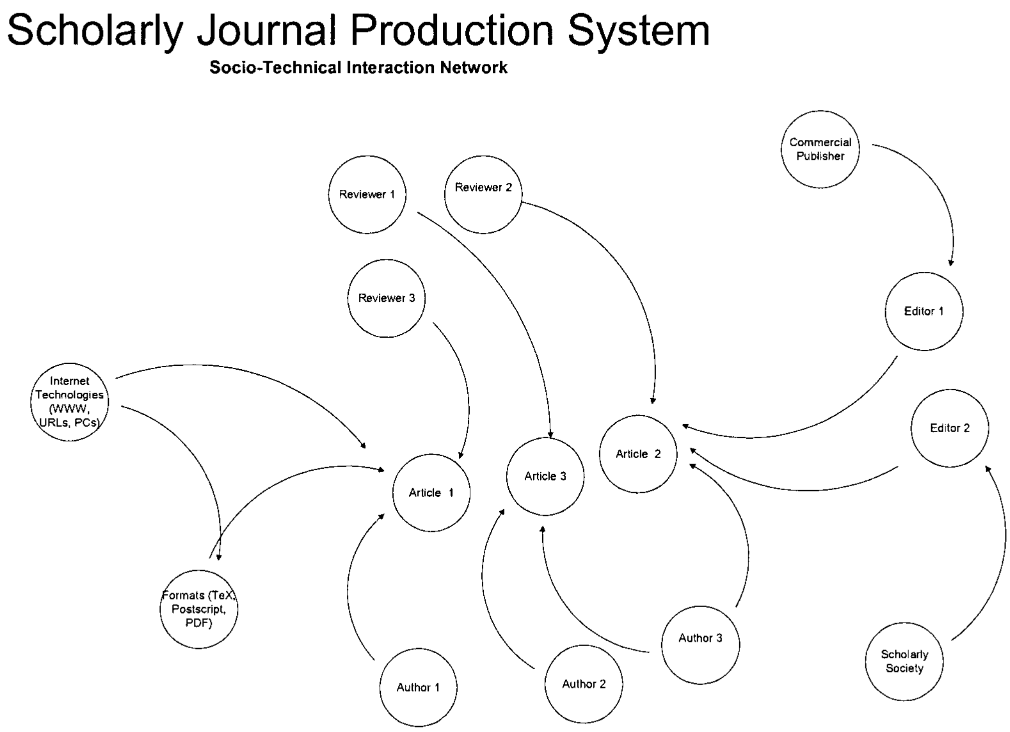


FIG. 1. A scholarly journal's production as a socio-technical (interaction) network.

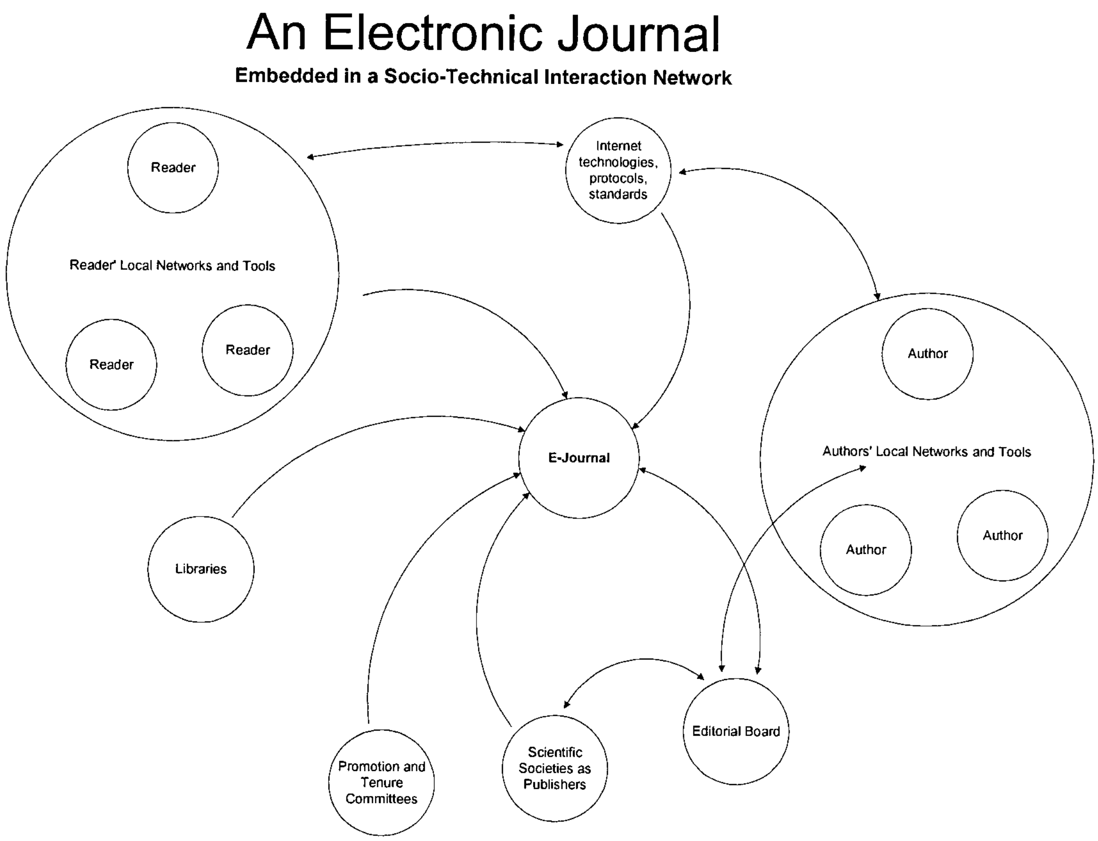


FIG. 2. An electronic journal embedded in a socio-technical interaction network.

complex.[5] For example, authors are often in a dependent position relative to the particular editors of a journal who are managing the reviews of their articles. However, some author–editor relationships can be more complicated if they have had prior collegial relationships or prior conflicts. In some cases, the extent to which an author has "hot results" can transform the relationship into one in which the editor tries to court the author, rather than simply administering an independent review of the author's manuscript. Further, the STIN in Fig. 2 includes participants who are not necessarily direct participants in creating or reading the e-journal, such as the members of Promotion and Tenure committees at an author's university. They are part of an e-journal's network because authors may decide whether or not to publish in an e-journal based on their expectations of the ways that such committees will (de)value their publications.[6]

In Figs. 1 and 2, we diagram several simple STINs. The term "network" must be understood as being metaphoric; not all relationships between elements in a STIN can be drawn in the form of a simple network diagram; some STINs may be too complex to represent graphically as a network. Graphical representations of STINs may reveal only some of the relevant interactions. For example, they do not represent social protocols, the character of legal contracts, actors' skills, etc. They help to develop or communicate an analysis rather than serve as a complete model of a setting or forum.

*Modeling an e-SCF as a Socio-Technical Interaction Network*

A significant problem faced by socio-technical analysts is that of how to figure out what belongs in the network and what does not—in other words, how to model an e-SCF as a network. The STIN model identifies two different classes of social interactions as being particularly useful and important in modeling STINs. These classes of social interactions are: resource dependency (or direct relationships) and account taking (or indirect relationships). Resource dependencies create networks of funders and grantees, employers and employees, and journal publishers and authors.

Constructing networks based on resource dependencies highlights several important themes, including the political economy of a forum, curation, and other forms of hidden work, and network extension through institutional linkages.

[5] This formulation of STINs has been significantly influenced by Latour's (1987) actor-network theory (or, perhaps more appropriately, ontology), as well as our own prior research about "Web models" of computerization (Kling, 1987, 1992; Kling & Scacchi, 1982). For more discussion about STINs and their relationship to Actor-Network Theory, please see Appendix A.

[6] The Promotion and Tenure committee in this example is a placeholder for any of the parties who may review an author's publications. An extended diagram might include department chairs and deans who set academic salaries, research grant review committees, etc.

However, the social relationships, like the nodes in the network, are heterogeneous; not all nodes in a STIN are linked by resource dependency. For example, money may flow in the STIN of an e-journal between the sponsoring organization and the Editor-in-Chief, but not directly between readers and authors.

Account taking refers to using other examples as reference points when making or defending decisions or proposing services. Account taking frequently takes the form of imitation (i.e., trying to make one system like another, probably successful, system), but can also take the form of differentiation (i.e., drawing out the differences to differentiate one system from another), or working around (i.e., trying to avoid a previous system failure). For example, when National Institutes of Health director Harold Varmus proposed the e-Biomed forum in 1999, he referenced arXiv.org as a model to imitate. This creates a strong linkage between e-Biomed and arXiv.org that is not based on resource dependency. We acknowledge, but do not explore in this paper, that this focus on account taking lends a strong neo-institutional dimension to the analysis (Powell & DiMaggio, 1991).

*The Social and the Technical in Socio-Technical Interaction Networks*

Before we examine some applications of STIN models of e-SCFs, we will discuss the nature of the relationship between the social and the technical in STINs.[7] The above characterization of STINs separated equipment (or technology) from social relationships and resources. This analytical separation between artifacts and social worlds is very common, even in social shaping analyses. In these approaches, as in the reinforcement politics theory, social relationships shape the kinds of artifacts selected, their configuration and their typical modes of use. But artifacts are conceptualized as "the products of engineering" and as 100% separable from social relationships.

The STIN model shares the views of many socio-technical theories: that technology-in-use and a social world are not separate entities—they co-constitute each other. That is, it is fundamental to STIN modeling that society and technology are seen as "highly" (but not completely) intertwined. Technology and society need not be seen as completely intertwined, because adherents of socio-technical approaches do not insist that this intertwining of technical and social elements is universal. Rather, the view of technology and society as highly intertwined is commonplace, and a good heuristic for inquiry, especially with complex technologies.

[7] For other accounts that examine socio-technical networks as complexes that intertwine social and technological elements as a complex admixture, see Bowker, Star, Gasser, and Turner (1997), MacKenzie and Wajcman (1999, pp. 22–24), Mansell and Silverstone (1995), and Wellman et al. (1996).

TABLE 2. Three models of e-SCFs.

| | Standard Model | Socio-Technical Systems Model | STIN Model |
|---|---|---|---|
| Analytical Focus | Use and Features: E-Forum features & users' interactions within the forum | Use: E-Forum & users' interactions within the forum | Ecological: E-Forum features, participation, participants' interactions in the E-Forums, and in teractions with other socio-technical networks and settings |
| Actors | Users (in relationship to each other) | Users (in relationship to each other) | Heterogeneous roles, scale and relationships: individual participants and groups and organizations that influence behavior in the E-Forum |
| Conceptions of Actors | Individuals | Individuals or on-line groups | Interactors (participating in multiple overlapping social & socio-technical networks & perhaps in different social settings |
| Treatment of IT | Cheap & easy & "standardized" | Features may matter (i.e. synchronous-asynchronous, anonymity) | Configurational—socially & by tech inscription |
| IT Infrastructure | Taken for granted | Taken for granted | Variable, sometimes can be problematic |
| Social Behavior | Can be easily reformed to take advantage of new conveniences, efficiencies and value. | Influenced strongly by features of the E-Forum & interactions within the E-Forum | Influenced strongly by interactions outside the E-Forum as well as within; E-Forum resources relative to other opportunities elsewhere |
| Resource Flows & Business Models | Taken for Granted | Taken for Granted | Examined (includes $ flows, regulatory regimes) |
| E-Forum Legitimacies | Taken for Granted | Taken for Granted | Mobilization of support treated as an accomplishment that requires work |

However, at a given level of analysis, analytical convenience may demand that, in accounts of technology, the technical and the social be treated as separate concepts for analytical convenience. For example, one might say, "Indiana University is using web-boards to support class discussions when the participants are not in class together." Indiana University and its classes would be treated as "social forms" and "Web-boards" as material "information technologies."

In the socio-technical network model, the Web-boards could be examined to see how they are constituted as socio-technical networks. For example, certain social relationships are inscribed into the Web-boards when they are used (such as access controls for who can read or write onto them) and are constituted in their supporting social protocols about legitimate content (to what extent are jokes or advertisements allowed in a specific class's Web board?).

Similarly, a social form such as Indiana University in Bloomington can be seen as co-constituted with diverse technologies.[8] Its routine operations rest on a complex set of building technologies, heating and cooling technologies, food acquisition and preparation technologies, and information and communication technologies. Without these (and other technologies) we would have 35,000 students, 15,000 faculty, and 2,000 staff milling around in the forested hills of Bloomington, foraging for food, and organizing themselves just by face-to-face conversation, word of mouth, and rumor! In contrast, the Indiana University of 1880 with about 300 students and a few dozen faculty was workable with much simpler technologies than those that are required for the much vaster contemporary university. Any means to record information about enrollments, courses, requirements, etc. would require some kind of information technology, however crude. In this sense, an organization such as Indiana University is constituted not just of people in social relationships, but also of diverse technologies. In fact, one can interpret many of the discussions of Internet-supported distance education as efforts to constitute new kinds of universities by changing their e-SCF infrastructures and pedagogies.

[8] This argument owes much to Strum and Latour (1999).

### *Comparison of STIN Models with the Standard and Socio-Technical Systems Models*

Table 2 compares STIN models with the Standard Model and the Socio-Technical Systems Models of e-SCFs across several key analytical dimensions. The most important differences lie in which contextual elements are explicitly examined as part of the model versus which contextual elements are taken for granted, or ignored by the model.

### *STIN Models in Use*

Explicit STIN models have been applied to understanding the IT support of scholarly research teams (Kling, 1992) and understanding the relative viability of early collaboratories within model organism molecular biology (Star & Ruhleder, 1996). Implicit STIN models have undergirded studies of IT applications failures (e.g., Markus & Keil, 1994).

Explicit STIN models have also been applied to understanding the character and development of electronic documents (Braa & Sandahl, 1998) and to the development of Internet standards (Hanseth & Monteiro, 1997; Monteiro, 1998; Monteiro, in press). These studies illustrate that STIN concepts are often understood informally in some of the professional IT communities. For example, Monteiro (1998) studied the processes by which the groups responsible for developing Internet infrastructure standards negotiated changes in the underlying IP address structure in the early 1990s. The character of several technical alternatives and debates between groups about them are outside the focus of this article. However, it is worth noting the language of the official "Request for Comments" (RFC) for new IP addressing schemes:

> The large and growing installed base of IP systems comprises people, as well as software and machines. The proposal should describe changes in understanding and procedures that are used by the people involved in internetworking. This should include new and/or changes in concepts, terminology, and organization. (Gross & Almquist, 1992, p. 19) Implicit in this passage is an understanding that the Internet is not simply a technology that is used by people. The "layer cake" conceptions of socio-technical systems would usually place the Internet as one of the lower layers in a model. In the view of the RFC authors, the Internet is infused with people, their concepts (of IP addressing), their procedures (for administering servers), and the various ways that they are organized.

Another illustration comes from Dr. James Myers' (1999) characterization of his Environmental Molecular Sciences Laboratory (EMSL):

> Before deciding which tools to use in their work, researchers first need to consider what occurs when they do science and how collaboration can help. Setting up a collaboratory is not simply a matter of running a remote experiment. Remote control software may let participants perform the experiment, but they also need access to the sample preparation procedures, instrument settings, and other information usually recorded in a local paper notebook today. Before the experiment can be considered, potential participants must discover the remote resource, understand its capabilities, contact the local researchers, develop trust, and perhaps receive training on a remote instrument. Even if the researchers decide to visit the EMSL to conduct the actual experiment, they can meet people, understand procedures, and learn about the instrument before they arrive. Remote researchers must also find effective techniques for analyzing the data and consulting with co-researchers in writing up publications. Because scholarly data are often complex and multidimensional, researchers will need to be able to confer with local researchers familiar with analysis of data from EMSL instruments.

In the Internet example, a STIN model might represent some elements, such as routers and servers as artifacts, and also represent the overall network of people, organizations, practices, and diverse devices as a STIN. In contrast, another STIN-based model might employ the heuristic that any artifact may be "opened up" to examine its social constitution, and any group may be examined to understand their co-constitution with various technologies. The STIN analyst does not open up all network nodes recursively; some must necessarily be left unexamined if any conclusions are ever to be drawn! But the analyst can ask how a router manages an activity like subnet addressing and how changes in the vendors' design teams' understanding and the vendors' marketing strategy will lead to different algorithms. Monteiro (1998) carefully examined the debates for the new IP address standard and found that many of the Internet engineers were especially sensitive to the "organizationally structured" character of many components as they sought strategies for a smooth transition to a new standard. In short, STIN analyses are not an esoteric social theory; they are part of the tacit understanding of many of the practicing computer scientists who have been developing Internet standards.

However, STIN analyses are more broadly applicable. Elsewhere, we have examined electronic journals and shown how their viability depends, not simply on the kinds of information processing features that the Standard Model foregrounds, but also on their location in the STINs of their respective scholarly communities (Kling, 1999; Kling & Covi, 1995). Now we will examine STIN models of more complex set of e-SCFs.

## Modeling an e-SCF as a STIN

STINs, like all networks, consist of nodes and links. Constructing a STIN is the complex process of determining what are the relevant nodes and links: what is important enough to matter, and what linkage is so tangential that its impact on the whole is negligible? In this section, we will discuss the methods and processes that we used in constructing STINs.

### *Assumptions*

Several fundamental assumptions underlie the application of the STIN methodology, and drive the methods we use to construct STINs. These assumptions are:

- The social and the technological are not meaningfully separable, at least not for the purpose of understanding how to design forums that are usable and sustainable.
- Theories of social behavior not only can, but should influence technical design choices (and thus the STIN methodology has a normative dimension as well).

- System participants[9] are embedded in multiple, overlapping, and nontechnologically mediated social relationships, and therefore may have multiple, often conflicting, commitments. Further, the system plays roles of varying importance in the social or professional lives of system interactors. The sustainability of a system will depend on other systems and communication forums that the interactors already participate in; interactors may be bound only weakly to the forum under discussion.
- Sustainability and routine operations are critical, and must play a key role in determining design.

*Modeling a STIN*

We have distilled the methodology for STIN modeling down to a series of discrete steps; however, these steps should be taken as illustrative, rather than strictly enumerative. That is, they are heuristics that we have found useful in analyzing design architectures for communication forums. Each of these steps will be discussed in more detail.

Heuristics for use in STIN modeling:

- Identify a relevant population of system interactors
- Identify core interactor groups
- Identify incentives
- Identify excluded actors and undesired interactios
- Identify existing communication forums
- Identify system architectural choice points
- Identify resource flows
- Map architectural choice points to socio-technical characteristics

*Identify system interactors.* In this step, we attempt to identify and characterize potential system interactors. In what ways and roles would the interactor participate in the system? What proportion of time would the interactors likely spend in the forum? Interactors are also not limited to "users" of the forum: funders and resource managers are also important interactors, and should have a role in determining the socio-technical architecture of a forum. This step is akin to a stakeholder analysis: identification of the likely actors, their roles, and their needs regarding the forum.

STIN models differs significantly from actor–network theories in that STIN methodology requires that the analyst make attempts to understand the characteristics and scope of the interactors ahead of time, rather than taking the interactors as they come and following them through use or development of the forum.

[9] We prefer the term "participants" or "actors" instead of "users" for several reasons. First, a user is characterized in relationship to a particular system (such as a user of Microsoft Word), whereas a participant or an actor may participate in multiple communications forums (such as an e-print server, a journal, and several conferences). Second, the term "user" generally implies a single type of relationship to the system, whereas a participant or actor may play multiple roles (such as reader, editor, critic, instructor, etc.)

*Identify core interactor groups.* In this step, we take the results of the previous step, and group interactors together with respect to their roles in the forum. The purpose of this step is to understand the differing roles potentially played in a forum by different subgroups. For example, if the forum under consideration is a scholarly journal, relevant interactor groups might be authors, reviewers, readers, grant funders, and departmental administrators. These groups may of course be overlapping, and their roles in a forum may even be conflicting.

*Identify incentive structures.* In this step, incentive structures are identified and made explicit to the degree possible: what kinds of incentives would energize appropriate communication in the forum. This is essentially equivalent to specifying the "business model" of a communication forum. The primary issue addressed by a business model—sustainability—is just as important in forums in which no money is exchanged or expected. After this phase, the creator of a STIN model should be able to specify why members of any of the relevant interactor groups identified in the previous phase would want to participate in the forum. Using an e-print server as an example, why would authors want to post their e-prints to the forum? Why would readers want to read?

In this phase of STIN-creation, a theory of social interaction is always implicit, and should be made explicit. This theory will help us to understand and specify why an interactor may want to have a particular type of interaction in the e-SCF.

*Identify excluded actors and undesired interactions.* Critical to determining sustainable forum architectures is understanding not only the types of interactions that interactors want to have in the system, but also the types of interactions they do not want to have. As a heuristic, unwanted interactions frequently involve various forms of monitoring or surveillance, as well as flaming. By conducting various types of social interactions in an e-SCF, participants may be allowing themselves to be monitored more closely than they might otherwise. Later in this article, we will discuss an example from the CONVEX experimental particle physics collaboration in which surveillance via a Web-camera is an important unwanted interaction. Another type of unwanted interaction is the avoidance of aggravation, or entanglement with bureaucracy. For example, developers of a system may prefer to develop at their own location, rather than at a customer location, because they want to avoid having to negotiate with the customer systems administration apparatus to obtain access codes, appropriate tools, etc.

*Identify existing communications systems.* In this step, we identify existing communications systems that interactors will already be using—that is, we attempt to characterize the communications ecology of the interactors. We also attempt to understand whether or not these communications will reinforce or compete with the e-SCF under consideration. For example, a new pure electronic journal might compete with a paper journal that covers the same topic. On the other hand, if a group works together closely, face-to-face at times (such as a particle physics collaboration), the use of a Web-based collaboratory may synergize well.

*Identify resource flows.* This step involves "following the money" or understanding how resources flow throughout the network. These resource flows can have both direct and indirect influence on interactions within the network. For example, Promotion and Tenure committees may have a direct resource control over individual scholars at their institution. This control may in turn influence the choice of publishing venue of the scholar, further impacting the viability of the new pure e-journal which the scholar may or may not publish in, depending on the views of the Promotion and Tenure committee that may review their scholarly careers.

*Identify architectural choice points.* Finally, after identifying the major socio-technical characteristics of the forum and its interactors, we identify the major systems architectural design choice points. A choice point refers to a technological feature or social arrangement in which the designer can select alternatives. However, some example choice points might be: Web-board (pull delivery) vs. mailing list (push delivery); moderated vs. unmoderated forum; restriction of ability to post or contribute (such as in a pure e-journal or even an organizational pre-print server); remote control of instrumentation vs. remote monitoring of instruments only. These choice points will clearly differ depending on the nature and architecture of the forum.

*Map socio-technical features to architectural choice points.* Finally, after having identified major socio-technical features of interactors and major architectural choice points of the e-SCFs, then we can begin to develop an understanding of which combinations and configurations of features are most compatible, viable, and sustainable. Viable configurations may involve tradeoffs between options.

## Examples of Four e-SCFS

In this section, we will examine some aspects of several important e-SCFs, in three different fields: high-energy physics, *Drosophila* (fruit fly) biology, and information systems. We will describe each of these forums in terms of the standard model and then again, using a STIN model. In each of these STIN descriptions, we will consider several themes that are foregrounded by the socio-technical network models. These will demonstrate how STIN models offer a more powerful understanding of the use, evolution, and sustainability of the forum. We selected these e-SCFs to illustrate a range of themes; however, there are many other e-SCFs we could have chosen. For example, the case of the proposed e-Biomed server (which ultimately became PubMed Central) illustrates the importance of scholarly societies as gatekeepers and policy influencers.[10]

STIN models foreground such socio-technical concepts as: content control, resource dependencies, the work required to make a system useful (articulation work), work and resources required to keep a system sustainable, translations used to mobilize resources, and the business model and governance structures. We thus emphasize many of these concepts in the following examples.

Our reporting of data in this article is somewhat unusual. Much of our data comes from an ethnographic interviewing (Appendix B). However, our reporting of data does not conform to one of the standard genres of ethnographic reporting. Our brief case examples do fit within a standard genre of realist ethnography (VanMaanen, 1988). But their role differs from standard realist ethnographic reports. "Realism is a mode of writing that seeks to represent the reality of the whole world or form of life. Realist ethnographies are written to allude to a whole by means of parts or foci of analytical attention which can constantly evoke a social and cultural totality" (Marcus & Fischer, 1986, p. 23). This article is not aimed at communicating the totality of life in a particular laboratory (for example, Knorr-Cetina, 1999; Traweek, 1992). Rather, we are advancing an ontology for understanding e-SCF's–STINs. We have selected aspects of scientists' and other professionals' interactions with or participation in their e-SCFs that best illustrate the character and value of STIN models.

### ArXiv.org and SPIRES-HEP

ArXiv.org (www.arXiv.org) is an electronic technical report server for physics. It was developed by Paul Ginsparg at the Los Alamos National Labs, and originally called XXX (xxx.lanl.gov). ArXiv.org has been one of the most successful and also one of the most notable electronic communication forums in scholarly communication, and is widely promoted as an exemplar for other disciplines to follow (Harnad, 1999: Kling, Fortuna, & King, 2001).

SPIRES-HEP is a Web-accessible bibliographic database of high-energy physics pre-prints, maintained by the Stanford Linear Accelerator Center (SLAC) Library (Kreitz, Addis, Galic, & Johnson, 1996) Although ArXiv.org is maintained at Los Alamos and SPIRES-HEP at SLAC, we are considering the two together, for reasons that will soon be made clear.

[10] We analyze this case in more detail in Kling, Fortuna, and King (2001).

*Standard models of arxiv.org and SPIRES-HEP.* The standard model could represent these two forums as crosslinked repositories with different attributes. ArXiv.org is a Web-accessible pre-print server. Authors post their own research articles in ArXiv.org, and readers may retrieve them, completely free of charge. Although the research articles submitted are pre-prints (they usually have not yet appeared in a journal at the time at which they are posted), authors can simply update their own articles on ArXiv.org with any changes. Although it plays a critical role in the physics communication system, and is widely used, the system was developed by physicist Paul Ginsparg in a short time, and for most of its life ran on a single computer sitting under Ginsparg's desk. Research articles are generally posted by authors either in LaTeX or PDF format. Although ArXiv.org was developed for physics, it is field-independent, and has sections for astrophysics, mathematics, nonlinear sciences, and computer science. Because all scholarly fields publish in journals, some advocates argue that its approach should be readily extended to other fields, except for the inertia of active scientists (who are afraid either of having their research scooped or plagiarized, or of having unrefereed articles made publicly available) and the vested interests of journal publishers (who fear losing revenue from journal subscriptions) (see, for example, Harnad, 1999). By April 2002, ArXiv.org contained about 195,000 articles and was growing rapidly in size.

SPIRES-HEP is an extensive bibliographic database of pre-prints in high-energy physics, containing more than 400,000 records. When a search is performed, a list of bibliographic records for high-energy physics preprints is retrieved. The search interface is elaborate, and allows both for simplified by-field search, and for a complex language-based search. When available, bibliographic entries link to the full texts of the pre-prints (by linking to the appropriate file in ArXiv.org). SPIRES-HEP also provides hyperlinks from retrieved records to records for both cited and citing articles.

*Socio-technical network models of arxiv.org/SPIRES-HEP.* The standard model would represent SPIRES-HEP and ArXiv.org as distinct repositories, maintained separately, but with many useful IT crosslinks between them. A STIN model would represent ArXiv.org and SPIRES-HEP as being linked through a matrix of social and economic relationships, as well as technologically. For example: SPIRES-HEP is linked through the acts of maintenance and curation not only to the SLAC library, but to the DESY[11] research facility in Hamburg. It is linked to ArXiv.org not only through HTML hyperlinks, but through the behind-the-scenes cooperation between Paul Ginsparg and the library staff at SLAC, who have developed a series of custom scripts to ensure that the links are maintained properly.

The staff of the SLAC library have collaborated with the Particle Data Group (an economic link, through funding) to link from appropriate summary data in the Lawrence Berkeley Labs Particle Data Group's Annual Review of Particle Physics (www-pdg.lbl.gov/) to the pre-print research that generated the summary data (bibliographic records which are stored in the SPIRES-HEP database). By mid-2002, there were over 8,000 links. Further, SPIRES-HEP collaborates with Durham University (UK) to link from pre-prints directly to particle reaction numerical data, which is maintained by the Durham Database Group in cooperation with the COMPAS group in Russia.[12] By mid-2002, there were almost 6,000 such links. This numerical data is gathered and presented by the Durham Database Group in a format that can be cut and pasted directly into the Physics Analysis Workstation, a software program developed at CERN, which can then be used to generate publishable plots of the data.[13]

ArXiv.org is linked to authors in the physics community through the act of submission to the database at the time of submission to a journal. It is linked to the high-energy physics community not only through direct searching and retrieval, but through SPIRES-HEP, which can serve as a "front-end" search engine to the particle physics-related working articles in ArXiv.org (see Figs. 3 and 4).

We will briefly identify two of the important types of relationships that are highlighted and foregrounded by adopting a STIN model.

*Curation.* SPIRES-HEP is highly curated by the SLAC library staff (O'Connell, 2000). First, the SLAC library staff screens all pre-prints and articles and rejects any that are not specifically relevant to high-energy physics or accelerator physics. Several of our HEP informants reported finding the results they received from searches in SPIRES-HEP to be more relevant than those they retrieved by searching ArXiv.org alone. We believe that the improved retrievals are a result of the selection performed by the SLAC library staff, our informants referred to the "better interface" for SPIRES-HEP rather than the curation, which is a form of invisible work to most of the scientists who use it as a search facility.[14]

---

[11] DESY, the "Deutsches Elektronen-Synchrotron," is one of the five largest experimental particle physics facilities in the world. In 2002, its experiments involved approximately 3,500 physicists from over 250 universities from 35 nations.

[12] See http://durpdg.dur.ac.uk/hepdata/reac.html for the HEPDATA: REACTION DATA Database.

[13] See http://wwwinfo.cern.ch/asd/paw/index.html for the Physics Analysis Workstation.

[14] One byproduct of this research project, and our conversations with SLAC librarians, is that they have explicitly identified the DESY physicists who add key words on a "key word help" page. See: http://www.slac.stanford.edu/spires/hep/keywords.shtml.

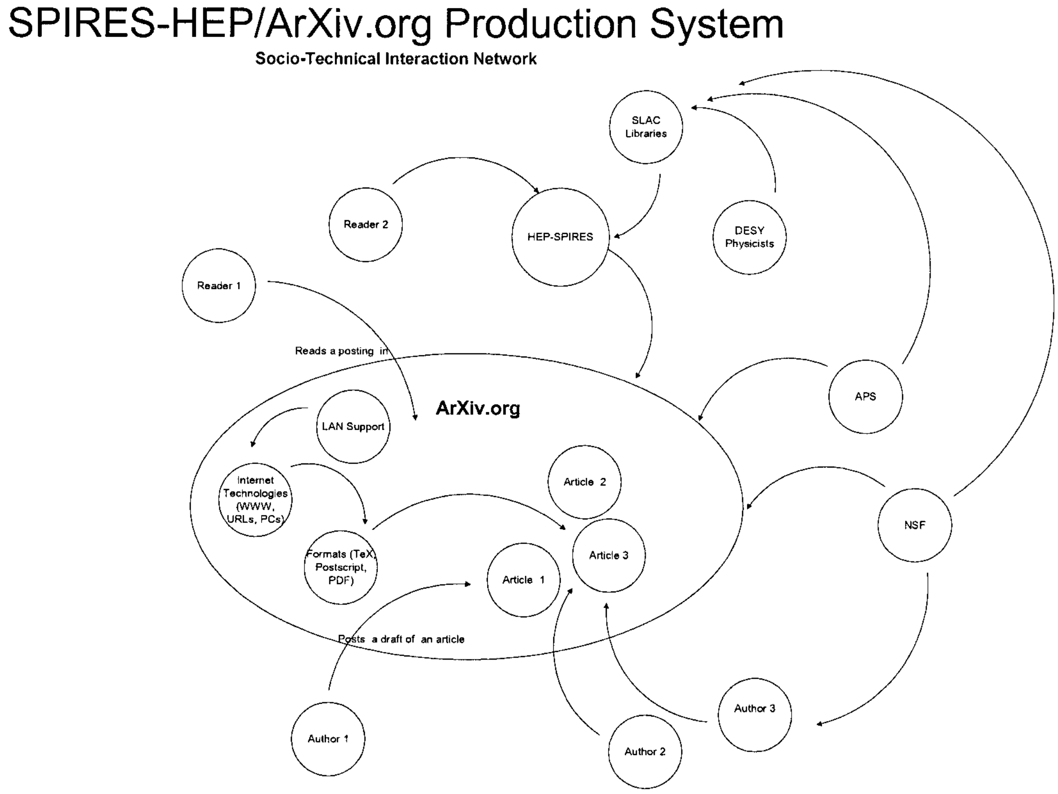


FIG. 3. SPIRES-HEP/ArXiv.org production system.

*Extension through institutional linkages.* Second, the SLAC library has entered into a partnership with DESY, the German particle physics laboratory. While the SLAC library makes selection decisions and enters new records into the database, DESY library staff enters descriptive key words for each article (from a controlled vocabulary). They

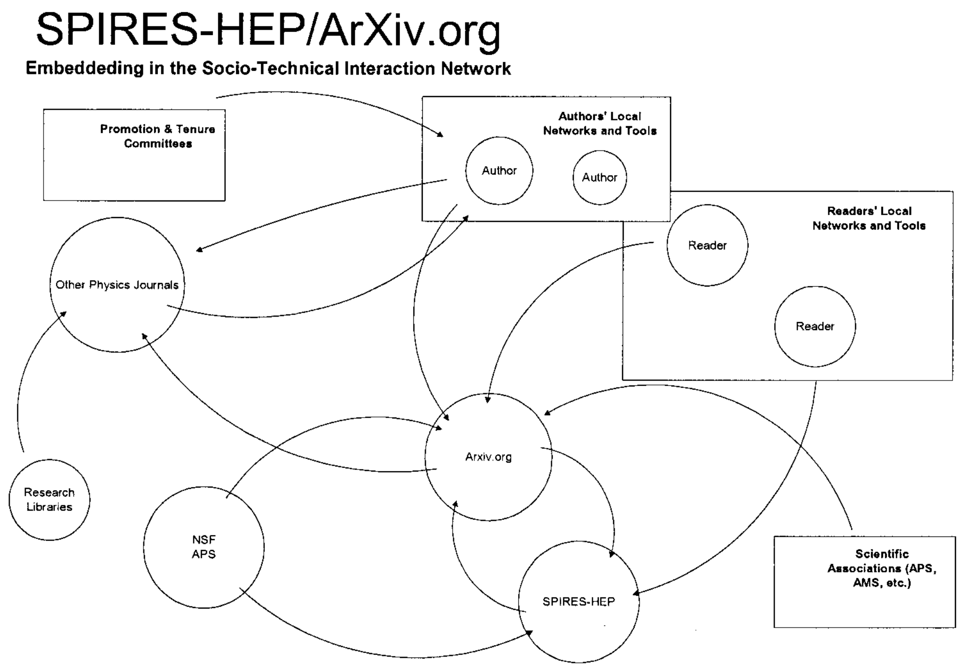


FIG. 4. SPIRES-HEP/ArXiv.org STIN.

also add records of some conference papers and journal papers. In the first half of 2002, they added about 3,500 references. Further, when an article eventually appears in a journal, DESY staff update the pre-print record with the journal citation (O'Connoll, 2000).

As purely a bibliographic database, SPIRES-HEP does not contain full-text documents. Rather, through arrangements with Paul Ginsparg, maintainer of arXiv.org, SLAC staff negotiated the ability for SPIRES-HEP to link directly to the full text when it is available by linking to the pre-print article in ArXiv.org. Through these linkages between SPIRES-HEP and arXiv.org, each forum is made more useful than it would be without the other. SPIRES-HEP gets full-text for many of its pre-prints, and arXiv.org gets a professionally developed interface, and a more heavily curated and maintained collection for high-energy physicists.

### *FlyBase*

*Standard model of flybase.* FlyBase is a gene database that stores genetic and molecular data about the *Drosophila* (fruit fly) genome. The database is funded by the National Science Foundation and is maintained by the FlyBase Consortium, a joint project with Indiana University, Harvard University, Cambridge University, the University of California, and the European Bioinformatics Institute. All materials are available free of charge to the end-user.

Materials available through FlyBase include: genomic clones, chromosomal aberrations, genetic and cytological map data, fly stocks descriptions, bibliographic references, data from the Berkeley and European gene projects, and addresses of *Drosophila* researchers. *Drosophila* researchers can do searches on a gene, retrieving not only its genetic content and location, but also data about the functions of the gene and bibliographic data on all published articles that reference the gene.

FlyBase's own documentation describes FlyBase, in standard model terms:

> FlyBase is a database of genetic and molecular data for *Drosophila*. It is available through FlyBase servers, as partial mirrors on other servers, as flat files, and as printed text (in special issues of *Drosophila Information Service*). The full database includes data files, documents, indices, forms, and images, and is maintained on a multi-protocol server that supports a variety of tools and formats. (FlyBase FAQ, http://flybase.bio.indiana.edu/.data/docs/refman/refman-A.1.html)

*Socio-technical interaction network models of flybase.* The standard model of FlyBase adequately characterizes some of its features. But it also fails to reveal several important things. It fails to specify the nature of the relationship between authors, readers, and FlyBase. More importantly, it fails to explain the incentives and forces that shape, and will shape the future evolution of the database.

First, FlyBase is embedded in the STIN of political relationships that make up its managing body—the FlyBase Consortium—and is linked through funding with the National Institutes of Health, its primary funding source. Each of the institutions that make up the FlyBase Consortium has a different role in maintaining the database. For example, Indiana University is responsible for curating data related to fly anatomy, and for maintaining the fly stock data. The fly stock data section of FlyBase is, in fact, used essentially as a stock inventory database by the maintainers of the fly stock lab at Indiana University.

There are actually three different major *Drosophila* databases: FlyBase, Berkeley Drosophila Genome Project database (BDGP), and the European *Drosophila* Genome Project database (EDGP). Each of them has extensive crosslinks to each other. However, the BDGP and EDGP reflect a different kind of relationship between authors and readers than does FlyBase. BDGP and EDGP actually contain genomic sequencing data from *Drosophila* that is produced from the respective sequencing projects. Authors from those projects deposit their data in the respective EDGP and BDGP servers; the data in these servers are not extensively curated. On the other hand, FlyBase data is derived entirely from the published *Drosophila* literature. Although this would seem to exclude direct submission to FlyBase, the maintainers of FlyBase actually encourage authors to submit both corrections to published data and new data that will not be published in a journal (for example, new data that is ancillary to the author's direct interest), which is then treated and curated as a "personal communication."

Currently, there is significant overlap between the three databases, as well as crosslinks between the three. On the other hand, although all data in each of the three databases are in electronic form, they are currently maintained in different formats, using different database management systems. The FlyBase Consortium is currently in the process of merging these three databases and integrating them into a single, more closely integrated system (http://flybase.bio.indiana.edu/docs/commercial.local/fb-nar99-doc.html).

*Curation.* All material in FlyBase is actively curated and is taken from the published *Drosophila* literature. This active curation is essential to the continued accuracy, and therefore usefulness of, the database.

*Extension through institutional linkages.* FlyBase is evolving, but not only in terms of numbers of items (genes, articles, etc.) added to the database. It is also evolving through the growth of institutional linkages. First, the databases managed by the members of the FlyBase Consortium, FlyBase, EDGP, and BDGP, are in the process of merging. Second, the FlyBase Consortium plans to integrate FlyBase more closely with other model organism databases. The purpose of this is to allow model organism researchers who are researching a particular gene to look for homologs (genes that play similar roles) in other model organism species.

*Skills.* Flybase's development and operation rests on a variety of technical skills in organizing a Web site, linking to other sites, curating genes, and even nagging researchers to share more meaningful portions of their experimental data. In addition, also essential to the maintenance of FlyBase, is political skill in coalition-building (by some of the senior investigators), to maintain the coalition between the various institutions.

*Political economy/funding.* FlyBase is not funded as an NIH line item—rather, the coalition must apply for new grants every 3–5 years. NIH research grants are generally not given for maintenance and operations, but rather for new developments. Therefore, the FlyBase coalition must continually propose the development of new features to receive funding for continued operation.

## ISWORLD

### *Standard Models of ISWORLD*

ISWORLD (www.isworld.org) is a Web-based collection of links to resources and information relevant to the information systems field, along with several LISTSERVs (a main list that is used for announcements, and a series of discussion lists that have developed in response to various topics, such as e-commerce). All materials are available free of charge to the reader. The links are divided into a several major sections: Research and Scholarship, Teaching and Learning, Professional Activities, and Countries and Regions.

ISWORLD contains such resources as a directory of information systems for scholars worldwide, a working paper server, lists of calls for papers for various conferences, and sections on research methods. However, unlike many of the other e-SCFs, such as arXiv.org, ISWORLD focuses on teaching and professional material, rather than research results.

### *Socio-Technical Interaction Network View of ISWORLD*

The feature that is most interesting about ISWORLD, according to the STIN model, is the business model. The business model is, in essence, that the e-SCF is free to readers, and is supported by the voluntary efforts of academics. Although ISWORLD has a single Editor-in-Chief (an unpaid position), maintenance of the pages (which includes solicitation of material for inclusion) falls to a series of section and division editors. The titles and positions where chosen to reward IS scholars for efforts to maintain various sections. One of the problems with gathering material for ISWORLD, however, is that, on the one hand, titles cannot be given out too freely or they become meaningless, and on the other hand, the editors are dependent on contributions from the "leaf-nodes" (i.e., individual contributors who are not section editors) that are not rewarded with titles.

One of the consequences of ISWORLD's business model is that the different sections of ISWORLD vary greatly in both their depth and their up-to-datedness. Another consequence of this business model is that scholars at less visible institutions (and in less visible regions of the world) can play a larger and more visible role in the information systems community by managing major sections of the ISWORLD site.

### *One HEP Collaboratory*

*Standard models of HEP collaboratories.* Experimental particle physics research is typically performed by large collaborations (from a few dozen to thousands of physicists). From before the inception of the Web, these physicists have developed Internet-based (and now, Web-based) collaboratories to facilitate various aspects of collaboration, such as communication of interim and final results, documentation of detector mechanics, and public outreach to remote instrumentation. Most, if not all, particle physics collaborations now have Web sites that are used by collaboration members for certain types of communication. Garrett and Ritchie (1995) provide an overview of the evolution of particle physics Web-based collaboratories, which, although they do not conceptualize it in these terms, implicitly use and exemplify STIN-based analysis.

The different collaboration Web sites are in various states of elaboration and are used for different purposes by different groups. We will use one collaboratory—the Web site for the CONVEX collaboration at HEPLAB[15] as a final example. The use of the STIN models in analyzing the Web site for the CONVEX collaboration in particle physics also highlights two other social relationships that have not yet been mentioned in this article: the link between communication and work practice, and the articulation of communicative boundaries.

CONVEX is an experimental particle physics collaboration whose massive equipment is based at HEPLAB. It consists of physicists at over a dozen institutions studying certain charm (quark) interactions. They have already gathered their data from the particle accelerator at HEPLAB. At the time of our field visit they were analyzing this data, but had not yet begun to publish the results.

### *Socio-Technical Interaction Network Models of the Convex HEP Collaboratories*

*Link between communication and work practice.* Most physics collaboration Web sites are used extensively for

[15] CONVEX and HEPLAB are pseudonyms.

documentary storage and retrieval. A few, including the CONVEX Web, are also used for remote instrumentation. In the CONVEX case, several physicists in the CONVEX collaboration developed a large suite of Web-based programs that allowed their collaborators to monitor remotely some of the instruments used in the data collection. During the 6 months of data taking, the CONVEX control room was staffed 7 days a week, 24 hours a day, by collaboration members. The work of the control room staff involves monitoring a number of activities, including the quality of several gasses in different parts of their detector, the quality of the positron beam entering the target area, and the number of particles that intercept the detector. Periodically, the gasses in a section of the detector degrade, and gas bottles have to be replaced.

The remote instruments allowed some collaboration members who were offsite to participate more actively in data taking. They could observe a number of data displays, which could be interpreted as indicating the quality of the positron beam, and could also observe indicators of the number of quark-related events that were being detected. The designers of the remote instrumentation wanted to enable the CONVEX collaborators not at HEPLAB at a given time to be able to observe the data taking and the experiment's progress.

However, on occasion, remote collaborators seemed to know more about the beamline quality than the physicists who were actually working control rooms shifts at the time. Physicists on the CONVEX collaboration reported that remote collaborators would occasionally phone in to let them know that some parameter was out of specification on the beamline. The physicists in the control room may have been doing other work, such as changing gas bottles, when the beamline degraded. We sensed a significant ambivalence about these interventions by remote collaborators. On the one hand, they helped to keep the experiment on track and allowed the CONVEX collaboration to gather much more data than anticipated during their scheduled beamtime. On the other hand, the physicists in the control room seemed to feel "caught short" and perhaps viewed as inattentive, when they were simply paying attention to another aspect of the experiment's complex operations.

Another way in which work practice shaped the development of the communication forum was in the choice of site to host the CONVEX collaboration Web. Most of the Web site was hosted at HEPLAB. However, several of the collaboration members who developed Web-based software for the collaboration actually developed and maintained certain applications on servers at their home institutions, rather than on the HEPLAB server. They reported that the development environment they were familiar with was already installed and functioning at their home institution, and it would have taken too much effort to reconfigure the HEPLAB server to work with their software development environment.

*Sharing vs. surveillance.* In the collaboratory literature, the primary mode of interpersonal interaction is sharing. The purpose of the collaboratories is to support the sharing of data, resources, and instruments among collaborators. However, a simple example from the CONVEX Web site illustrates that other modes of interaction besides sharing can be seen: in this case, sharing the control room during data taking also slips into surveillance.

The remote instrumentation mentioned in the previous section eventually culminated in the placement of a digital video camera in the beam control room, so that remote users could watch shift-workers during beam runs. However, after enough telephone calls from other collaborators letting shift workers know that something was not right with the beam, many shift workers began to see the camera as a form of surveillance. Eventually, many collaborators in the control room turned the camera to face the ceiling, so remote users would simply see a blank screen!

*Sharing vs. control: the negotiated boundaries between public and private.* Even when sharing is the primary mode of interaction, authors and maintainers of communication forums may be concerned with what is public and what should not be made public. These boundaries between public and private are also structured and shaped by STINs. For example, many collaborations are in competition with one to four other specific collaborations. This competition causes some collaborations (such as BaBar at SLAC which is competing with a group at the Japanese lab, KEK) to take fairly elaborate measures to ensure that data are shared only within the collaboration (from segmenting of access—to the use of Secure Sockets Layer (SSL) to ensure that the data channel itself is secure).

Further, concern for professional reputation both individually and for the collaboration causes most, if not all, collaborations to have developed systems whereby only research results that have been officially "blessed" by the collaboration may be shared by the world on the collaboration Web site. Articles in progress that have not yet been blessed by the collaboration are shared with other collaboration members via a section of the Web site that is password-protected from the outside world.

## Conclusions

We have crystallized some of the key assumptions of the Standard Model of e-SCFs, which emphasizes their information-processing features. We have articulated an alternative, STIN-based model that enables us to better understand some of their key aspects. One of the important consequences of adopting a STIN-based model is that it becomes clear that new technological developments will not wash away the issues of sustainability and integration into a social world. For example, a faster Internet, and better, more efficient Web protocols (or even more ubiquitous access to the Internet) would not have made the original proposals for the National Institutes of Health's PubMed Central more palatable to medical researchers. One key issue was not technology, but of gatekeeping—a crucial facet of the relationship between authors and readers that is foregrounded by STIN-based analyses. A second important consequence of STIN-based models is that STIN-based analyses inject social analysis into all phases of planning, development, configuration, use, and evolution of an e-SCF, rather than merely at the beginning (in determining user "requirements" of an e-SCF), and postdeployment (in determining it social "impacts").

The examples of arXiv.org/SPIRES-HEP, FlyBase, ISWORLD, and the HEP collaboratories help illustrate different types of social relationships foregrounded by a STIN-based analysis that are important to the use, sustainability, and evolution of e-SCFs. The case of arXiv.org is especially instructive because its value to high-energy physicists, derives in part, from its curation by librarians at SPIRES-HEP and the technical documentary analysis by a group of physicists at DESY. It is much more costly to develop and maintain than one can learn from the stories about a solo physicist "in the desert" developing arXiv.org (Hafner, 1998).

As we found with arXiv.org and HEP, FlyBase, and ISWORLD, business models are critical to understanding viability and evolution of field compilations. When the business model for maintaining the e-SCF rests, in part, on Federal R&D funding, then its likely developmental trajectory will tend toward increased institutional complexity and complexity on the back-end, but increased simplicity of access on the front-end.

In the case of ISWORLD, the "free to readers" business model depends upon the voluntary efforts of academics, and thus the up-to-datedness of portions of the compilation vary considerably. In both cases, the capabilities for developing and sustaining these e-SCFs are not only technical, but require the socio-political ability of forum developers to build durable alliances.

All of these behaviors would be hard to anticipate from the standard model of e-SCFs. We suggest that future discussions of e-SCFs, including collaboratories, should be informed by STIN models. The STIN heuristic of seeking the social elements of technical formations and the technical supports for social life opens up important lines of inquiry to enable better understanding of these complex practices. As we note in Table 2, STIN models are also ecological, insofar as they examine e-Forum features, participation, participants' interactions in the e-Forums, and interactions with other socio-technical networks and settings.

However, we do not view our exposition of STIN models in this article as definitive, Rather, we see them as an important advance in socio-technical analysis, and one that can serve as a basis for more refined and complete subsequent theorizing.

## Acknowledgments

Funding was provided in part by NSF Grant SBR-9872961 and with support from the SLIS at Indiana University. Jose Montalvo helped us to conceptualize the "Standard Model" of e-SCFs. This article benefited from helpful discussions with and comments from Ingemar Bohlin, Blaise Cronin, Geof Bowker, Elisabeth Davenport, Paul Edwards, Kristin Eschenfelder, Joanna Fortuna, Suzanne Iacono, Mark Napier, John Paolillo, Steve Sawyer, Lisa Spector, and Bill Turner. We benefited from discussions with various colleagues when we presented drafts of this article at an NSF-ESF Workshop on Digital Collaboration Technologies, the Organization of Scientific Work and the Economics of Knowledge Access held at IIASA in 1999; at the 2000 Meeting of the Association for Information Systems, at the 2000 Meeting of the Society for the Social Studies of Science, a lecture for ADEST-SFSIC at the Ecole des Mines in Paris, a talk at the Laboratory for Socio-Economie de la Communication Interactive of Limsi-CNRS (Université Paris Sud), and in two public lectures at Indiana University sponsored by SLIS and by the School of Education. In addition, we are most indebted to numerous scientists who have discussed their work practices with us, because they provided the primary data for this study. Heath O'Connell, as well as other members of the library staff at SLAC, generously provided data about their operations.

## References

Bailey, C.W., Jr. (2001). Scholarly electronic publishing bibliography (Version 38) [Online]. University of Houston, University Libraries. Available: http://info.lib.uh.edu/sepb/sepb.doc

Bowker, G., Star, S.L., Turner, W., & Gasser, L. (Eds.) (1997). Social science, technical systems and cooperative work: Beyond the great divide. Hillsdale, NJ: Erlbaum.

Braa, K., & Sandahl, T.I. (June 1998). Documents in infrastructure–Challenges for design. Proceedings of the 6th European Conference of Information Systems (ECIS) (pp. 1025–1041), Aix-en-Provence, France.

Buchanan, E., Brown, M., Casanova, J., Wolfram, D., and Xie, H. (2000). Web-based and traditional instruction: A systematic study of student and instructor perceptions from a graduate MLIS program. Teaching with Technology Today, Vol. 7, No. 1, September 15. Available: http://www.uwsa.edu/ttt/buchanan.htm

Dutton, W.H. (1995). The ecology of games and its enemies. Communication Theory, 5(4), 379–392.

Dutton, W.H., & Guthrie, K. (1991). An ecology of games: The political construction of Santa Monica's public electronic network. Informatization and the Public Sector, 1, 279–301.

Emery, F.E., & Trist, E.L. (1960). Socio-technical systems. In C. W. Churchman et al. (Eds.), Management sciences, models and techniques. London: Pergamon.

Finholt, T.A. (in press). Collaboratories. In B. Cronin (Ed.), Annual Review of Information Science and Technology, 36.

Fleck, J. (1994). Learning by trying: The implementation of configurational technology. Research Policy, 23, 637–652.

Galegher, J., Kraut, R., & Egido, C. (Eds.) (1990). Intellectual teamwork: Social and technological foundations of cooperative work. Hillsdale: Lawrence Erlbaum.

Garrett, P., & Ritchie, D.J. (1995). Collaborating over the Web: Libraries and laboratories or the librarian and the webmaster. AusWeb95 [Online].

Available: http://fncduh.fnal.gov/~ritchie/presentations/ausweb_95/AW02-03.html
Ginsparg, P. (1994). First steps towards electronic research communications. Computers in Physics, 8(4), 390–396.
Gross, P., & Almquist, P. (1992). IESG deliberations on routing and addressing. Request for comments (RFC) 1380 [Online]. Available: http://www.faqs.org/rfcs/rfc1380.html
Guedon, J. (2001). In Oldenburg's long shadow: Librarians, research scientists, publishers, and the control of scientific publishing. Washington, DC: Association of Research Libraries. Available: http://arl.cni.org/arl/proceedings/138/guedon.html
Hafner, K. (April 21, 1998). Physics on the Web is putting science journals on the line. New York Times.
Hanseth, O., & Monteiro, E. (1997). Inscribing behaviour in information infrastructure. Accounting, Management and Information Technologies, 7(4), 183–211. Available: http://www.ifi.uio.no/~oleha/Publications/siste.enkel.doc.html
Harnad S. (1990). Scholarly skywriting and the prepublication continuum of scholarly inquiry. Psychological Science, 1, 342–343.
Harnad, S. (1997). The paper house of cards (and why it is taking so long to collapse). Ariadne, 8, 6–7.
Harnad, S. (1999). Free at last: The future of peer-reviewed journals. D-Lib Magazine, 5(12). Available: http://www.dlib/december99/12harnad.html
Harnad, S., & Hemus, M. (1997). All or none: There are no stable hybrid or half-way solutions for launching the learned periodical literature in the postGutenberg galaxy. In I. Butterworth (Ed.), The impact of electronic publishing on the academic community. London: Portland Press.
Kling, R. (1987). Defining the boundaries of computing across complex organizations. In R. Boland & R. Hirschheim (Eds.), Critical issues in information systems (pp. 307–362). New York: John-Wiley.
Kling, R (1992). Behind the terminal: The critical role of computing infrastructure in effective information systems' development and use. In W. Cotterman & J. Senn (Eds.), Challenges and strategies for research in systems development (Chap. 10, pp. 153–201). New York: John Wiley. Available: http://www-slis.lib.indiana.edu/kling/pubs/webinfra.html
Kling, R. (1999). What is social informatics and why does it matter? [Online]. D-Lib Magazine, 5(1). Available: http://www.dlib.org:80/dlib/january99/kling/01kling.html
Kling, R., & Covi, L. (1995). Electronic journals and legitimate media in the systems of scholarly communication. The Information Society, 11(4), 261–271. Available: http://www.slis.indiana.edu/tis
Kling, R., & McKim, G.W. (2000). Not just a matter of time: Field differences and the shaping of electronic media in supporting scientific communication. Journal of the American Society for Information Science, 51(14), 1306–1320.
Kling, R., & Scacchi, W. (1982). The web of computing: Computing technology as social organization. Advances in computers (Vol. 21). New York: Academic Press.
Kling, R., Fortuna, J., & King, A. (2001). The real stakes of virtual publishing: The transformation of E-Biomed into PubMed Central [Online]. Working Paper No. 01-03, Indiana University, Center for Social Informatics. Available: http://www.slis.indiana.edu/csi/WP/wp01-03B.html
Knorr-Cetina, K. (1999). Epistemic cultures: How the sciences make knowledge. Cambridge MA: Harvard University Press.
Kreitz, P.A., Addis, L., Galic, H., & Johnson, T. (1996). The virtual library in action: Collaborative international control of high-energy physics pre-prints (SLAC—PUB—7110). Stanford Linear Accelerator Center, Stanford University.
Latour, B. (1987). Science in action: How to follow scientists and engineers through society. Cambridge, MA: Harvard University Press.
Law, J. (1996). Traduction/trahison—Notes on ANT [Online]. Department of Sociology, Lancaster University. Available: http://www.lancaster.ac.uk/sociology/stslaw2.htmL
Lesk, M. (1997). Practical digital libraries: Books, bytes, and bucks. San Francisco, California: Morgan Kaufmann.
MacKenzie, D.A., & Wajcman, J. (Eds.) (1999). The social shaping of technology (2nd ed.). Philadelphia: Open University Press.
Mansell, R., & Silverstone, R. (1995). Communication by design: The politics of information and communication technologies. New York: Oxford University Press.
Marcus, G.E., & Fischer, M.J. (1986). Anthropology as cultural critique: An experimental moment in the human sciences. Chicago: University of Chicago Press.
Markus, M.L., & Keil, M. (1994). If we build it, they will come: Designing information systems that people want to use. Sloan Management Review, 35(4), 11–25.
Monteiro, E. (1998). Scaling information infrastructure: The case of the next generation IP in Internet. The Information Society, 14(3), 229–245.
Monteiro, E. (in press). ANT and information infrastructure. In C. Ciborra (Ed.), The dynamics of global infrastructures. New York: Oxford University Press.
Myers, J.D. (1999). Tools for collaboration. In.North Carolina Board of Science and Technology and National Research Council. Collaboratories: Improving research capabilities in chemical and biomedical sciences: Proceedings of a multi-site electronic workshop (pp. 5–9). Washington, DC: National Academy Press.
National Research Council (1993). National collaboratories: Applying information technology for scientific research. Committee Toward a National Collaboratory: Establishing the User–Developer Partnership. Washington, DC: National Academy Press.
O'Connell, H.B. (2000). Physicists thriving with paperless publishing (SLAC—PUB—8357) [Online]. Stanford Linear Accelerator Center, Stanford University. Available: http://arxiv.org/abs/physics/0007040
Odlyzko, A.M. (1995). Tragic loss or good riddance? The impending demise of traditional scholarly journals. In R.P. Peek & G.B. Newby (Eds.), Electronic publishing confronts academia: The agenda for the year 2000. Cambridge, MA: MIT Press. Available: http://www.jucs.org/jucs_0_0/tragic_loss_or_good/html/paper.html
Olson, G.M., Atkins, D.E., Clauer, R., Finholt, T.A., Jahanian, F., Killeen, T.L., Prakash, A., & Weymouth, T. (1998). The upper atmospheric research collaboratory (UARC). Interactions, 5(3), 48–55.
Olson, G.M., Atkins, D., Clauer, R., Weymouth, T., Prakash, A., Finholt, T., Jahanian, F., & Rasmussen, C. (2001). The upper atmospheric research collaboratory (UARC). In G.M. Olson, T. Malone, & J. Smith (Eds.), Coordination theory and collaboration technology. Hillsdale, NJ: Lawrence Erlbaum Associates.
Powell, W., & DiMaggio, P. (1991). The new institutionalism in organizational analysis. Chicago: University of Chicago Press.
Scott, M.A. (1999). The DOE 2000 Project. In North Carolina Board of Science and Technology and National Research Council. Collaboratories: Improving research capabilities in chemical and biomedical sciences: Proceedings of a multi-site electronic workshop. Washington, DC: National Academy Press. Available: http://books.nap.edu/catalog/9465.html
Sheldon, B.E. (1998). Distance education in library and information science studies: Administrative issues and implications from a pilot study on power and status among recent graduates. 64th IFLA General Conference, August 16–21. Available: http://www.ifla.org/IV/ifla64/149-136e.htm
Star, S.L., & Ruhleder, K. (1996). Steps towards an ecology of infrastructure: Design and access for large-scale collaborative systems. Information Systems Research, 7, 111–138.
Strum, S., & Latour, B. (1999) Redefining the social link: From baboons to humans. In D. MacKenzie & J. Wajcman (Eds.), The social shaping of technology, (2nd ed.), Buckingham: Open University Press.
Teasley, S., & Wolinsky, S. (2001). Scientific collaborations at a distance. Science, 292(5525), 2254–2255.
Traweek, S. (1992). Beamtimes and lifetimes: The world of high energy physicists. Cambridge, MA: Harvard University Press.
Van Maanen, J. (1988). Tales of the field: On writing ethnography. Chicago: The University of Chicago Press.
Varian, H.R. (2000). Versioning information goods. In B. Kahin & H.R. Varian (Ed.), Internet Publishing and Beyond. Cambridge, MA: MIT Press.

Wellman, B., Salaff, J., Dimitrova, D., Garton, L., Gulia, M., & Haythornthwaite, C. (1996). Computer networks as social networks: Virtual community, computer supported cooperative work and telework. Annual Review of Sociology, 22, 213–238.

Williams, R. (1997). The social shaping of information and communications technologies. In H. Kubicek, W.H. Dutton, & R. Williams (Eds.), The social shaping of information superhighways: European and American roads to the information society (pp. 299–338). New York: St. Martin's Press.

Wulf, W.A. (1999). Improving research capabilities through collaboratories. In North Carolina Board of Science and Technology and National Research Council. Collaboratories: Improving research capabilities in chemical and biomedical sciences: Proceedings of a multi-site electronic workshop. Washington, DC: National Academy Press. Available: http://books.nap.edu/catalog/9465.html

## Appendix A: STINs and Actor-Network Theory

STIN models are derived extensively from insights incorporated into Actor-Network Theory (ANT), as elucidated in Latour (1987). Most importantly, in STIN models, as in actor-network theories and Web models, heterogeneous social and technical elements are brought together into complex networks that cannot be reduced either to technological determinism or to social determinism.

However, in developing STIN models, we diverge from ANT in several important ways, both epistemologically and methodologically. In the original formulation of ANT, for example, the primary driving social process is one in which some parties try to *enlist* others in a central project.[16] STIN models do not make a commitment to a single driving social process. The nature of the relationships and the dynamics of social action must be specified explicitly in addition to specifying the socio-technical network.

Incidentally, Latour (1987, pp. 203–205) provides an intriguing example about the ways that different social theories lead analysts to characterize different participants and their networks. In his example, he feels that a French colleague of his contributed to the drowning of an African informant who swam out to save the colleague who was intentionally swimming in dangerous waters. However, village elders who investigated the drowning ignored the French, drew very different social networks, and implicated an aunt of the deceased, who was not "directly" involved in the incident at all.

Another area in which STIN models and ANT diverge is in the area of *symmetry*. Latour's ANT also advances a "strong symmetry" principle in which nonhuman agents are as capable of the same range of actions as are human actors. STIN models support some symmetry between human and nonhuman agents. We are, however, much more conservative in attributing action to nonhuman agents. For example, technology journalist Katie Hafner (1998) wrote a story for the *New York Times* about the physics article repository, arXiv.org, that emphasized the heroism of Arxiv's founder, Paul Ginsparg. As we showed in this article, Arxiv is a byproduct of complex inter-institutional work. We do not see Arxiv as enlisting the *New York Times* to advance its cause (as Latour might argue). Rather, we see enthusiasts for scholarly self-publishing, such as Paul Ginsparg and Stevan Harnad, endeavoring to enlist a successful technology journalist to help advance their interests. In this case, there is enlistment; but the social dynamics are closer to those of an ecology of games (Dutton, 1995; Dutton & Guthrie, 1991).

We do acknowledge cases where nonhuman agents can be seen as actors. Interesting examples include the invocation of textual authority, such as high court rulings in court cases and bibles in religious discussions. Even here, human interpretations of such agents are commonplace. For example, the Koran has greater power as an agent in Tehran than in the Vatican.

A final important difference between the STIN models and the Actor-Network theories is methodological. The only way to model an actor network, as implied by Latour (1987), is to follow the key actors as they create the network. As our interest is in providing shapers, funders, and developers of new communications forums with a practical and useful analytical framework, we (and our intended adopters of STIN modeling) do not have the luxury of waiting until the network is fully constructed before attempting to model it. Rather, STIN modeling involves attempting to determine ahead of time who the relevant participants will be, and attempting to understand the socio-technical networks that these participants are embedded in.

[16] Law (1996) notes that researchers use several different theories under an ANT rubric.

## Appendix B: Research Methods

### *Data Collection Methods*

We used two different methods in performing the research for this article: documentary interpretation and semi-structured interviews with participants and organizers of e-SCFs. First, the research team read documentary materials exhaustively about the e-SCFs discussed in this article, as well as other field compilations that were not included. All of the e-SCFs discussed in this article are available free of charge to readers at Internet sites; therefore, we were able to read documentary material provided by the compilation maintainers, and also to examine most of the forums ourselves.

Further, there is an extensive literature on electronic publishing and scholarly communication (see Bailey, 2001). Many of the forums discussed in this article, such as the Los Alamos physics e-print archive, have been discussed and documented in this literature [see Table 1A, B, and C for specific references, and Kling and McKim (2000) for a more detailed discussion of these forums]. The Standard Model of e-SCF's has been abstracted from this literature.

During our research, we discovered an important methodological principle. Despite the wealth of documentary information available and the availability of these communication forums on the Web, we were unable to learn about many important facets of the e-SCFs—in particular, the business models and the institutional linkages needed to maintain these systems—from documentation. These elements turned out to be crucial pieces in our STIN model of e-SCFs, as we learned through interviews with e-SCF organizers.

The research team conducted in-depth, semi-structured interviews with some of the key shapers of these communication forums. We interviewed shapers of FlyBase, SPIRES-HEP, ISWORLD, and HEP collaboration Websites at HEPLAB (a pseudonym), as well as other e-SCFs, from March 1998 through October 1999. We conducted approximately 30 interviews with molecular biologists, high-energy physicists, information systems faculty, librarians, and technical support staff. We typically conducted these 1- to 2-hr interviews in our informants' offices. However, we also examined some of the laboratory facilities and interacted with each of these e-SCFs, except for the Webcam in the control room of the CONVEX collaboration.

In addition, we attended two information systems research conferences where we interviewed organizers of ISWORLD, and also attended sessions that were devoted to publishing practices and new pure electronic journals in the discipline. In addition, in 1998 and 1999, Kling participated in two workshops on the Internet and scientific communication that were sponsored by the American Association for the Advancement of Science, and in 1999 and 2001 participated in two workshops on collaboratories that were sponsored by the National Science Foundation. These workshops provided informal opportunities to discuss the ontologies of information technologies that informed discussions of the design, use, and value of e-SCFs with a variety of physicists, chemists, biologists, and computer scientists.

In our formal interviews, we discussed issues such as: support and funding for the e-SCF, governance structures, audiences (both targeted and actual), the role of the e-SCF within the communication system of the field, how the usefulness or success of the e-SCF is assessed, social activities to support the e-SCF, technical complexities of the e-SCF, and the opportunities and pressures that lead to new features for the e-SCF.

### *Data Analysis Methods*

The interviews were formally transcribed from our written notes and/or tape recordings. The STIN model was developed based on these data, but not in some algorithmic manner. Rather, we allowed the concepts to emerge and then tested them against our data.

We also refined the STIN model using other sources of data. For example, at one point while working on the model, we attended a workshop on electronic journals at an information systems conference. One of the concerns we heard from participant after participant at this workshop was that many of them (in particular, more junior scholars) were afraid to attempt to publish any of their works in newer e-SCFs because they were afraid that their Promotion and Tenure (P&T) committees would not count those publications or would disapprove of them in promotion decisions. This discussion led us to include P&T committees as a significant "node" in many STINs.

In general, the interviews provided critical sources of data about which people, groups, or organizations worked "behind the Web site" to develop, maintain, or enhance the e-SCF, and the technological and social relationships that characterized their work. In this paper, in the section "Modeling STINs," we discuss some of the heuristics we used and continue to use in modeling e-SCFs as STINs.