



No. WP-01-02

A Bit More To IT: Scholarly Communication Forums as Socio-Technical Interaction Networks

[Rob Kling](#), [Geoffrey McKim](#), and [Adam King](#)

Center for Social Informatics
Indiana University
Bloomington, IN USA 47405

<http://www.slis.indiana.edu/CSI> and <http://www.slis.indiana.edu/SCIT>

Last Update: 5/27/2001

TABLE OF CONTENTS

Introduction
Methods <ul style="list-style-type: none">• Data Collection Methods• Data Analysis Methods
The Standard Model of Electronic Scholarly Communication Forums
Socio-Technical Network Models of Communication Forums <ul style="list-style-type: none">• Limited Socio-Technical Conceptions of Communication

<p>Forums: The Socio-technical Systems Model</p> <ul style="list-style-type: none"> • Socio-Technical Networks • Modeling the Socio-Technical Interaction Network • The Social and the Technical in Socio-Technical Interaction Networks • STIN Models in Use
<p><u>Modeling a STIN</u></p> <ul style="list-style-type: none"> • Assumptions • Modeling a STIN
<p><u>Case Studies: Four E-SCTs</u></p> <ul style="list-style-type: none"> • <u>Arxiv.org/SPIRES-HEP</u> <p>Standard Model of Arxiv.org and SPIRES-HEP Socio-technical Network Model of Arxiv.org/SPIRES-HEP</p> <ul style="list-style-type: none"> • <u>FlyBase</u> <p>Standard Model View of ISWORLD Socio-Technical Interaction Network Model of ISWORLD</p> <ul style="list-style-type: none"> • <u>ISWORLD</u> <p>Standard Model View of ISWORLD Socio-Technical Interaction Network Model of ISWORLD</p> <ul style="list-style-type: none"> • <u>HEP Collaboratories</u> <p>Standard Model of HEP Collaboratories Socio-Technical Interaction Network Model of ISWORLD</p>
<p><u>Conclusions</u></p>
<p><u>Acknowledgements</u></p>
<p><u>Appendix A: STINs and Actor-Network Theory</u></p>
<p>Notes</p>

INTRODUCTION

We are in a period where many analysts have high expectations of how IT can substantially improve scholarly communications, as well as communications with professionals and broader publics. From the beginnings of the Internet, funding for IT infrastructure is 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, preprint servers, collaboratories and so on.

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, post-publication archives, pre-print servers, cross-linked webs of resources, gene databases, etc. Later on we will discuss the architectures of four of these in some detail. But 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 purpose of this neologism is both to abstract from any specific assumptions about the technical form of the forum (i.e. face-to-face meeting, paper journal, linked web, central server-based repository, etc.) and about the communicative role of the forum (data to be used in designing a study, teaching materials, unrefereed working papers, pre-prints, post-publication articles, etc. Many of them are not purely electronic; for example, sensors may be physical devices with electronic controls and Web-cameras may be used to share real-time visuals of physical laboratory control rooms. 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 insofar as some of their participants may also work or meet face to face as well.

It should not be surprising that some of the proposals for new e-SCFs, such as the National Institutes of Health (NIH) proposal for a self-posting working paper archive of biomedical research, have been controversial (Bloom, 1999). Many e-SCFs have social properties that are not compatible with highly entrenched and durable scholarly communication practices, such as peer review, and the article as the primary unit of scientific research.

In this article, we examine the kinds of design conceptions that help us best understand both the behavior within an e-SCF and the sustainability of new e-SCFs within a 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 model, which we call the **standard model of e-SCFs**, provides an inadequate understanding of eSCFs. Second, we will present an alternative model, a model that considers technologies as **socio-technical networks**, and that we find provides a richer understanding of the e-SCFs, as well as a more complete understanding of the conditions and activities that enhance the sustainability of an e-SCF within a field. We will contrast this socio-technical network model with a more limited, but much more commonly advanced, **socio-technical systems** model. We will conclude by illustrating the significance of a specific kind of Socio-Technical Network model -- a Socio-Technical Interaction Network (STIN) -- with examples of e-SCFs drawn from the fields of high energy physics, molecular biology and information systems.

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 WWW sites, and even community forums such as Weblogs or auction sites.

We view STINs a framework that helps understand human behavior in the use of technology-mediated social settings.

Before moving on to a description of our methods, we will first provide a richer illustration of STIN models using one particular type of e-SCF – the collaboratory.

Collaboratories as Socio-technical Networks

We can begin to illustrate our Socio-Technical Network models by contrasting different approaches to 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.” 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) and Keating et. al (2000) characterize the Environmental Molecular Sciences Laboratory (EMSL) 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, in order 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 view as “inside” the collaboratory.

Another collaboratory director, Zaluzec (1999), stressed that the resident scientists who work “inside” a collaboratory are not just cognitive informants. Sometimes they act as evangelists, and 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. And 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, the 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 the STIN Model to emphasize the importance of the character of the 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 our research, we have also learned that “the collaboratory evangelist who turns out to be a zealot” can drive away other scientists, who could have been potential collaborators.

Further, the collaboratories are themselves “nodes” in a larger Socio-Technical Network of related (competing and cooperating) laboratories, funding relationships, publishing opportunities, and so on.

Our interests in framing an alternative to the Standard Model of e-SCFs are 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, but are central to the STIN models that we will examine later in this article.

METHODS

Data Collection Methods

We used two different methods in performing the research for this article: documentary interpretation and semi-structured participant interviews. First, the research team read substantial documentary materials 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 use some of the e-forums ourselves.

Further, there is an extensive literature on electronic publishing and scholarly communication (see Bailey, 2001). Many of the forums to be discussed in this article, such as the Los Alamos physics e-print archive, have been discussed and documented in this literature. The Standard Model that we present has been abstracted from this literature.

One of the things we discovered during our research has important methodological implications. 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 them – from documents alone. 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, HEP-SPIRES, ISWORLD, and HEP collaboration web-sites at HEPLAB (pseudonymous), (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. 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 our interviews we examined 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 STIN model was developed based on the data gathered in the interviews, which were transcribed. However, we did not use any mechanical technique for deriving concepts from the interview transcripts; 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 website" to develop, maintain, or enhance the e-SCF, and the technological and social relationships that characterized their work. Later in this paper, in the section “Modeling STINs”, we will discuss some of the heuristics we used and continue to use in modeling e-SCFs as STINs.

THE STANDARD MODEL OF ELECTRONIC SCHOLARLY COMMUNICATION FORUMS

There is a vast literature on 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. commonplace but not universal) assumptions about electronic e-SCFs, and have synthesized these assumptions into a “Standard Model of e-SCFs.” We then contrast this Standard Model with Socio-Technical Network models of electronic e-SCFs. In a subsequent section, we will examine four major electronic e-SCFs in light of the Standard Models and our STIN model.

The Standard Model focuses primarily upon the information-processing features of the technologies that constitute an e-SCF. We have listed the features of the Standard Model in order of decreasing importance, as reflected in the literatures of electronic scholarly communication that the model was abstracted from. While there are a few accounts of collaboratories (NRC, 1993; Ross-Flanigan, 1998; NRC, 1999; Olson, et. al, 1998), most of the literature focuses upon documentary compilations— where documents include data sets as well as memos, discussions, reports and articles. The following attributes, which emphasize the information processing features of the communication medium, comprise the Standard Model:

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.

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 communicationⁱ. 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 outⁱⁱ. Most importantly, as the costs of computing and data storage approach zero, the distribution and transaction costs for electronic materials should also approach zero.

FAST PUBLICATION: Timely publication and display of ideas, content, and results. The quality of the final product is superior to some traditional paper-based journalsⁱⁱⁱ. In addition, it is the only way to deal with the exponential increase in the creation and publication of scholarly articles^{iv}.

CONTENT 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^v.

MATERIALS: Documents are easy to update and markup^{vi}.

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^{vii}.

FORMATS: Materials can be accessed in wide variety of formats: PDF, gzip, TeX, LaTeX, ASCII, etc^{viii}.

MULTIMEDIA: Publications can include other types of media (audio files, charts, graphs, photos, executable files, etc.)^{ix}.

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.

CROSS LINKING: 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^x.

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.)

METADATA: Metadata is the solution to any organizational and social issues. As long as data is properly coded (for example, as peer-reviewed, published in a journal, sponsored by a particular organization, etc.), then social and organizational barriers disappear.

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.

SOCIO-TECHNICAL NETWORK MODELS OF COMMUNICATION FORUMS

Economic theories of technological innovation portray information technologies 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, the conventional analyses emphasize the rapidly decreasing price/performance of hardware, the declining size and weight of equipment, the ubiquity of telecommunications to help people to readily move data readily within and across organizations. The conventional theories tilt towards economic and technological determinisms. For example, some scientists believe that the experimental high energy physics working article (e-print) server at Los Alamos national labs (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).

Careful empirical research studies about e-SCFs and other ICTs have found that “almost identical technologies” are often configured very differently in practice. It is common for preexisting social arrangements to influence these configurations. A “social shaping of technology” perspective suggests caution in trusting deterministic claims. In addition, each social group may have to locally configure e-SCFs to use them most effectively. What are claimed as “best practices” may work well in some 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), the Standard Model leads analysts to emphasize the rapidly decreasing price/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, one can expect that a few well-crafted pilot projects – done almost anywhere -- can help to establish “best practices” that everyone else can follow. 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 which differ considerably from our own use. The first 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. We will show how the concept of STNs can be put to better use than this¹.

A second common use is reflected in some of the discussion of collaboratories (see NRC, 1999). 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 kinds of applications. The “tool sets” of the collaboratory are “the technical layers.” 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 and 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 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 Networks². 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/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 video tape, 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, while 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 to whom the videotape may be played.

Voice-based face-to-face conference, 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. It should be noted that 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 since they bring together different kinds of social and technological elements – cameramen, 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 usefully 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/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 Figure #1 for a simplified depiction).

In addition, the journal is embedded in other socio-technical networks, such as the 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 Figure #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 (see Kling and Covi, 1995; Kling, 2000).

In Figures #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, since they can be complex³. For example, authors are often in a dependent position relative to the particular editors of a journal who are managing the reviews of their own articles. However, some author-editor relationships can be more complicated if they 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 Figure #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 publications⁴.

At this point, we can put forward a definition of a STIN. 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 interaction.

The term “network” must be understood as being metaphoric. Although in Figures 1 and 2 we attempt to diagram several simple STINs, 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 as a directed graph of links and nodes. Graphical representations of STINs may reveal only some of the relevant interactions. For example, they do not represent social protocols, the character of contracts, and actors' skills. They help to develop or communicate an analysis rather than serve as a complete representation 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 decide what belongs in the network and what does not – in other words, how to model an e-SCF as a network. STIN models identify two different classes social interactions as being particularly useful and important in modeling STINs. These classes of social interactions are: resource dependency (or direct relationship) and account-taking (or indirect relationship). Resource dependencies link networks of funders and grantees, employers and employees, journal publishers and authors. Modeling 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. 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 in order 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. This creates a strong linkage between E-Biomed and Arxiv.org that is not based on resource dependency⁵.

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 STINs⁶. 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. In these approaches, 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 per cent separable from social relationships.

The STIN model holds that to understand technology-in-use, what is a social relationship or social resource and what is an artifact are not easily separable entities – they co-constitute each other. That is, it is fundamental to the STIN models that society and technology as seen as “highly” (but not completely) intertwined because its adherents do not insist that this intertwining of technical and social elements is universal. Rather, it is commonplace, and a good heuristic for inquiry, especially with complex technologies.

However, at a given level of analysis, analytical convenience may demand that, in accounts of technology, the technical and the social may be treated as separate concepts for analytical convenience. For example, one might say that “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.”

However, 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 technologies⁷. Its routine operations rest on a complex set of building technologies, heating/cooling technologies, food acquisition and preparation technologies, and information and communication technologies. Without these (and other technologies) we would have 35,000 students, 1,5000 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.

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 and 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 and Sandahl, n.d.) and to the development of Internet standards (Hanseth and Monteiro, 1997; Monteiro, 1998; Monteiro, 2000). 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” 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.

(RFC 1992, p. 19)

Implicit in this passage is an understanding that the Internet is not simply a technology that is used by people. The Socio-Technical Systems conceptions e-SCFs 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, and their concepts (of IP addressing), and 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 will 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, an STIN analysis might treat some elements, such as routers and servers as artifacts, but view the overall network of people, organizations, practices, and diverse devices as a STIN. In contrast, another STIN-based analysis 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 doesn’t 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 specially sensitive to the “organizationally structured” character of many components as they sought strategies for a smooth transition to a new standard. In short, stin analysis 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 & Covi, 1995; Kling, 2000). Now we will apply a STIN analysis to a more complex set of e-SCFs.

MODELING AN E-SCF AS A STIN

STINs, like all networks, consist of nodes and relationships between them. Modelling an e-SCF as a STIN is a complex process of determining what are the relevant nodes and relationships: what is important enough to matter, and what relationships are so tangential that its impact on the e-SCF is negligible? In this section, we will discuss the methods and processes that we used in modelig e-SCF's as STINs.

Assumptions

Several fundamental assumptions underlie the application of the STIN methodology, and drove the methods we used to model e-SCF's as 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).
- **E-SCF participants⁸ are embedded in multiple, overlapping, and non-technologically mediated social relationships**, and therefore may have multiple, often conflicting commitments. Further, the e-SCF plays roles of varying importance in the social or professional lives of its participants. The sustainability of an e-SCF will depend on other communication forums that the participants already participate in; participants may be bound only very 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 architectures for e-SCFs.

The steps are listed in a linear order; but in practice they are recursive. When we examine an existing e-SCF (by reading documentation and through interviews with participants), information that we garner to answer one of the questions listed below may also uncover new participants, technologies, relationships that modify our earlier analyses. If these steps are used to analyze the design of a prospective e-SCF, similar iterations are possible (even likely), as the range of participants, their relationships, technological opportunities and limitations, and other constraints becomes more clear. Each of these steps will be discussed in more detail below.

1. Identify a relevant population of participants
2. Identify core participant groups
3. Identify incentives
4. Identify undesired interactions and excluded participants
5. Identify existing communication forums
6. Identify e-SCF architectural choice points
7. Identify resource flows
8. Map architectural choice points to socio-technical characteristics

Identify e-SCF Participants

In this step, we attempt to identify and characterize potential e-SCF participants. In what ways and roles would the participant participate in the e-SCF? What proportion of time would the participants likely spend in the forum? Participants are also not limited to “users” of the forum: funders and resource managers are also important participants, 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 preferences regarding the forum⁹.

Identify Core Participant Groups

In this step, we take the results of the previous step, and essentially group together participants 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 participant 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, a STIN modeler should be able to specify why members of any of the relevant participant 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-modelling, a theory of social interaction is always implicit, and should be made explicit. This theory will help us to understand and specify **why** a participant may want to have a particular type of interaction in the e-SCF.

Identify Unwanted Interactions

Critical to determining sustainable forum architectures is understanding not only the types of interactions that participants **want** to have in the e-SCF, 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. By conducting various types of social interactions in a 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 avoidance of complication, conflict, or entanglement with bureaucracy. For example, developers of a software system may prefer to develop it at their own location, rather than at a customer location, because they want to avoid having to complication and conflict with the customer's systems administrators with respect to obtaining access codes, appropriate tools, etc.

Identify Existing Communications Networks

In this step, we identify existing communications networks that participants will already be using. 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 & 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 & Tenure committee.

Identify Architectural Choice Points

Finally, after identifying the major socio-technical characteristics of the forum and its participants, we identify the major e-SCF's 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.

Apply Socio-Technical Features to Architectural Choice Points

Finally, after having identified major socio-technical features of participants 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.

CASE STUDIES: FOUR E-SCTS

In this section, we will describe in detail several important e-SCFs, in three different fields: high-energy physics, *Drosophila* (fruit fly) biology, and information systems. First, we will describe each of these forums in terms of the Standard Model. Next, we will describe each e-SCF again, using a STIN model. In each of the STIN models, we will consider several themes that are foregrounded by the socio-technical network models. Then, we will show that the STIN models provides 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 others we could have chosen. For example, the case of the proposed E-Biomed server (which ultimately became Pub Med Central) illustrates the importance of scholarly societies as gatekeepers and advocates of science policies. We analyze this case in more detail in Kling, Fortuna and King (2001).

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 an e-SCF sustainable, translations used to mobilize resources, and business model and governance structures. We thus emphasize many of these concepts in the forthcoming case studies.

Arxiv.org/SPIRES-HEP

Arxiv.org (www.arxiv.org) is an electronic working article server in physics. 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 exemplified and widely promoted electronic communication forum in scholarly communication.

SPIRES-HEP is a Web-accessible bibliographic database of high-energy physics pre-prints, maintained by the Stanford Linear Accelerator Center (SLAC) Library. 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.

Standard Model of Arxiv.org and SPIRES-HEP

The Standard Model conceptualizes these two forums as cross-linked 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, arXiv.org 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, non-linear sciences, and computer science. Since all scholarly fields publish in journals, extension to other fields should be unproblematic, except for the inertia of active scientists (who are afraid either of having their research scooped or plagiarized, or of having unrefereed material made publically available) and the vested interests of journal publishers (who fear losing revenue from journal subscriptions). At the time this article was written, Arxiv.org contained approximately 120,000 articles in the repository, and was growing rapidly.

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, what is retrieved is a list of bibliographic records for high-energy physics preprints. 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 of SPIRES-HEP and Arxiv.org treats the two as distinct repositories, maintained separately, but with many useful crosslinks between them. Contrasting with this model is the STIN model. From this perspective, Arxiv.org and SPIRES-HEP are linked technologically. In addition, both are embedded in and energized by a matrix of social, technological, and economic relationships. For example: SPIRES-HEP is linked through the acts of maintenance and curation not only to the SLAC library, but to the DESY facility in Hamburg. It is linked to Arxiv.org not only through HTML hyperlinks, but through the behind-the-scenes cooperation of 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 is currently working on a joint grant proposal with the Particle Data Group (an economic link, through funding) in order 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). Further SPIRES-HEP enhancement plans include a plan to work 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. 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.

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.

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. 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 – this means that low-energy or medium-energy particle physics pre-prints are excluded from the SPIRES-HEP database. Several of our HEP informants reported finding the results they received from searches in SPIRES-HEP to be more relevant than those they got by searching Arxiv.org alone – this may be explained by the selection performed by the SLAC library staff.

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 enter descriptive keywords for each article (from a controlled vocabulary). Further, when an article eventually appears in a journal, DESY staff update the pre-print record with the journal citation.

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.

The documentation for FlyBase itself describes FlyBase, in Standard Model terms, as:

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 have extensive cross-links 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 data in the respective EDGP and BDGP servers; data in these servers is 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 cross-links 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 (by some of the senior investigators) in coalition-building, in order 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 three to five 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 in order to receive funding for continued operation.

ISWORLD

Standard Model of ISWORLD

ISWORLD is 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 scholars world-wide, 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 Models 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, but 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 were 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) which aren’t rewarded with titles. One of the consequences of this business model is that the different sections of ISWORLD vary greatly in both their depth and their up-to-datedness. Another consequences 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.

HEP Collaboratories

Standard Model of HEP Collaboratories

Experimental particle physics research is typically performed by large (from a few dozen to thousands of physicists) collaborations. From before the inception of the Web, these physicists have developed Internet-based (and now, Web-based) collaboratories in order 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 uses and exemplifies 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 HEPLAB¹⁰ 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 ISWORLD

Link Between Communication and Work Practice

Most physics collaboration Web sites are used extensively for 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 six months of data taking, the CONVEX control room was staffed seven 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 who were 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 even 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 is 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 practices whereby only research results that have been officially “blessed” by the collaboration may be shared by the world on the collaboration Web site. Work in progress that has not yet been blessed by the collaboration is shared with other collaboration members via the Web site, but is password-protected from the outside world.

CONCLUSIONS

We have crystallized some of the key assumptions of the standard (information processing) model of e-SCFs. 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 an STIN-based model is that it becomes clear that a rising tide of new technological developments will not wash away the issues of sustainability and integration into a social world. A faster Internet, and better, more efficient Web protocols would not have made the original proposals for PubMed Central more palatable to medical researchers, for example; the issue was not one of 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 for social scientists 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 a system), and post-deployment (in determining the social “impacts”) of the e-SCF.

The examples of Arxiv.org/SPIRES-HEP, FlyBase, ISWORLD, and the HEP laboratories help illustrate different types of social relationships foregrounded by an STIN-based analysis that are important to the use, sustainability, and evolution of e-SCFs. The case of Arxiv.org is especially instructive since 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 xxx.lanl.gov (Hafner, 1998).

As we found with Arxiv.org/HEP, FlyBase and ISWORLD, the business model is critical to understanding its 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 towards 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. Their 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.

ACKNOWLEDGEMENTS

Funding was provided in part by NSF Grant #SBR-9872961 and with support from the School of Library and Information Science 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 Blaise Cronin, Geof Bowker, Paul Edwards, Suzanne Iacono, Mark Napier, and Bill Turner. In addition, we are indebted to numerous scientists who have discussed their work practices with us.

REFERENCES

- Bailey, Charles W. Jr. .2001. Scholarly Electronic Publishing Bibliography. Version 35: February 16. University Libraries, University of Houston, Houston, TX at:
<http://info.lib.uh.edu/sepb/sepb.html>,<http://info.lib.uh.edu/sepb/sepb.doc> and
<http://info.lib.uh.edu/sepb/sepb.pdf>
- Bloom, Floyd. 1999. “Just a Minute Please (an editorial).” *Science*. (9 July). 285 (5425): 197.
- Braa, Kristin and Tone Irene Sandahl. n.d. “Documents In Infrastructure: Challenges For Design” The Internet project: <http://internet.adb.gu.se>
- Crawford, S. Y., Hurd, Julie, M. and Weller, A. C. (1996). *From Print to Electronic: The Transformation of Scholarly Communication*. Medford, N.J.: Information Today, Inc.
- Dutton, William H., 1995. "The Ecology of Games and Its Enemies," *Communication Theory*, 5(4)(November):379-392.
- Dutton, William H. and Kendall Guthrie. 1991. "An Ecology of Games: The Political Construction of Santa Monica's Public Electronic Network," *Informatization and the Public Sector*,: 279-301.
- Edwards, Paul N. 1998. "Y2K: Millennial Reflections on Computers as Infrastructure," *History and Technology*, Vol. 15, 7-29.

Edwards, Paul N. 1999. Personal communication. December 28.

Emery, F. E., and E. L. Trist. 1960. "Socio-technical Systems." In Management Sciences Models and Techniques , vol. 2. London.

Garrett, Paula and Ritchie, David J. 1995. "Collaborating Over the Web: Libraries and Laboratories or The Librarian and the Webmaster." AusWeb95.

http://fncduh.fnal.gov/~ritchie/presentations/ausweb_95/AW02-03.html.

Galegher, Jolene, Robert Kraut and Carmen Egido (eds.) 1990. Intellectual Teamwork: Social and Technological Foundations of Cooperative Work. Hillsdale: Lawrence Erlbaum

Hafner, Katie. 1998. Physics on the Web is Putting Science Journals on the Line. New York Times. April 21.

Hanseth, Ole and Eric Monteiro. 1997. Inscribing behaviour in information infrastructure. Accounting, Management and Information Technologies. 7(4):183 - 211, 1997

<http://www.ifi.uio.no/~oleha/Publications/siste.enkel.doc.html>

Hars, Alexander. 1999a. "A theory of scholarly online journals: The impact of IT on the publishing process." School of Business. University of Southern California. Available:

<http://cybrarium.usc.edu/cyb.dll/?object=1001950>

Keating, Kelly A., James D. Myers, Jeffrey G. Pelton, Raymond A. Bair, David E. Wemmer and Paul D. Ellis. 2000. Development and Use of a Virtual NMR Facility. Journal of Magnetic Resonance, vol. 143:172-183

Kling, Rob. 1987. "Defining the Boundaries of Computing Across Complex Organizations. pp. 307-362 in Critical Issues in Information Systems, R. Boland and R. Hirschheim (eds.). New York: John-Wiley.

Kling, Rob 1992. "Behind the Terminal: The Critical Role of Computing Infrastructure In Effective Information Systems' Development and Use." Chapter 10 (pp: 153-201) in Challenges and Strategies for Research in Systems Development. William Cotterman and James Senn (Eds.) New York, John Wiley. (Available from

<http://www.slis.lib.indiana.edu/kling/pubs/webinfra.html>).

Kling, Rob. 2000. "Learning about Information Technologies and Social Change: The Contribution of Social Informatics." The Information Society 16(3).

http://www.slis.indiana.edu/TIS/tables_of_contents/toc_16.html (July-Sept, 2000):217-232

Kling, R. & Covi, L. (1995). "Electronic Journals and Legitimate Media in the Systems of Scholarly Communication." The Information Society. 11(4):261-271. (Available from

<http://www.slis.indiana.edu/kling>)

Kling, Rob, Joanna Fortuna, and Adam King. 2001. "The Real Stakes of Virtual Publishing: The Transformation of E-Biomed into PubMed Central" SLIS, Indiana University Bloomington, IN.

Kling, Rob and Geoff McKim. 1999a. "The Shaping of Electronic Media in Supporting

Scholarly Communication: The Contribution of Social Informatics.” . in Electronic communication and research in Europe. Ian Butterworth and Erich Neuhold (Eds.) European Commission. . <http://academia.darmstadt.gmd.de/seeheim/thebook/index.html>

Kling, R., & McKim, G. W. 1999b. “Scholarly Communication and the Continuum of Electronic Publishing.” *Journal of the American Society for Information Science* 50(9): 890-906. (Available at: <http://www.slis.indiana.edu/SCIT>)

Kling, Rob and Walt Scacchi.1982. "The Web of Computing: Computing Technology as Social Organization", *Advances in Computers*. Vol. 21, Academic Press: New York.
Kreitz, P. A., Addis, L., Galic, H., & Johnson, T. (1996). “The Virtual Library in Action: Collaborative International Control of High-Energy Physics Pre-Prints.” Stanford Linear Accelerator Center, Stanford University, Stanford, CA.

Law, John. Nd. Traduction/Trahison - Notes on ANT. At: published by the Department of Sociology, Lancaster University at: <http://www.lancaster.ac.uk/sociology/stslaw2.html>

Latour, Bruno. 1988. *Science in Action. How to Follow Scientists and Engineers Through Society*. Harvard University Press.

MacKenzie, Donald A. and Judy Wajcman (Eds). 1999. *The Social Shaping of Technology* (2nd Edition). Open Univ Press. Philadelphia, Pa.

Mansell, Robin and Roger Silverstone. 1995. *Communication by Design: The Politics of Information and Communication Technologies*. New York. Oxford University Press.

Markus, M. Lynne and Mark Keil. 1994. “If we build it, they will come: Designing information systems that people want to use.” *Sloan Management Review* 35(4) (Summer):11?25.

Meyers, James D. 1999. “Tools For Collaboration” in NCBST and NRC.

Monteiro, Eric. 1998. *Scaling information infrastructure: The case of the next generation IP in Internet*. *The Information Society* . 14(3):229 – 245.

Monteiro, Eric. 2000. “ANT and Information Infrastructure.” In: Claudio Ciborra (ed.), *From Control To Drift: The Dynamics Of Corporate Information Infrastructures*, Oxford Univ. Press.

National Research Council. 1993.*National Collaboratories: Applying Information Technology for Scholarly Research* Committee Toward a National Collaboratory: Establishing the User-Developer Partnership, Washington, D.C. National Academy Press.

North Carolina Board of Science and Technology and National Research Council. 1999. *Collaboratories: Improving Research Capabilities in Chemical and Biomedical Sciences: Proceedings of a Multi-site Electronic Workshop*. Washington, D.C. National Academy Press. <http://books.nap.edu/catalog/9465.html>

Odlyzko, Andrew. M. (1996). *Tragic Loss or Good Riddance? The Impending Demise of Traditional Scholarly Journals*. In R. Peek & G. Newby (Eds.), *Scholarly Publishing: The*

Electronic Frontier (pp. 91-102). Cambridge, MA: MIT Press.

Olson, Gary M. Olson, Daniel E. Atkins, Robert Clauer, Thomas A. Finholt, Farnam Jahanian, Timothy L. Killeen, Atul Prakash and Terry Weymouth. 1998. "The Upper Atmospheric Research Collaboratory (UARC)," In *Interactions*, 5(3):48-55.

Powell, W. & DiMaggio, P. (1991). *The New Institutionalism in Organizational Analysis*. Chicago, IL: University of Chicago Press.

RFC 1992. IESG deliberations on routing and addressing. RFC 1380, IAB and IESG, November.

Ross-Flannigan, N. 1998. "The Virtues (and Vices) of Virtual Colleagues," in *Technology Review*, Cambridge, MA: MIT, March/April.

Scott, Mary Ann. 1999. "The Doe 2000 Project" in NCBST and NRC.

Star, Susan Leigh and Ruhleder, Karen. (1996). Steps towards an ecology of infrastructure: Design and access for large-scale collaborative systems. *Information Systems Research* 7: 111-138.

Wellman, Barry, Janet Salaff, Dimitrina Dimitrova, Laura Garton, Milena Gulia and Caroline Haythornthwaite. 1996. "Computer Networks as Social Networks: Virtual Community, Computer Supported Cooperative Work and Telework." *Annual Review of Sociology* 22: 213-38.

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, William A. 1999. "Improving Research Capabilities Through Collaboratories " in NCBST and NRC, 1999.

Zaluzec, Nestor. 1999. Comments at An ESF-IIASA-NSF Workshop December 3, at the International Institute for Applied System Analysis, Laxenburg, Austria

Zariski, Archie. 1997b. "'Knowledge Networks' or Discourse Communities?: Response to Hibbitts' Commentary on Electronic Journals," *First Monday*, 2(8) (August), http://www.firstmonday.dk/issues/issue2_8/zariski/index.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 (1988). 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 nor 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. 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 and in addition to specifying the socio-technical network.

Incidentally, Latour himself (1988: 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 non-human agents are as capable of the same range of actions as are human actors. STIN models support some symmetry between human and non-human agents. We are, however, much more conservative in attributing action to non-human 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 don’t 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 and Guthrie, 1991; Dutton, 1995).

We do acknowledge cases where non-human 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 is commonplace. For example, the Koran has greater power as an agent in Tehran than in the Vatican.

A final important difference between STIN models and Actor-Network models is methodological. The only way to model an actor-network, as implied by Latour (1988) is to follow the 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.

FIGURES

Figure #1 -- A Scholarly Journal's Production as a Socio-Technical (Interaction) Network

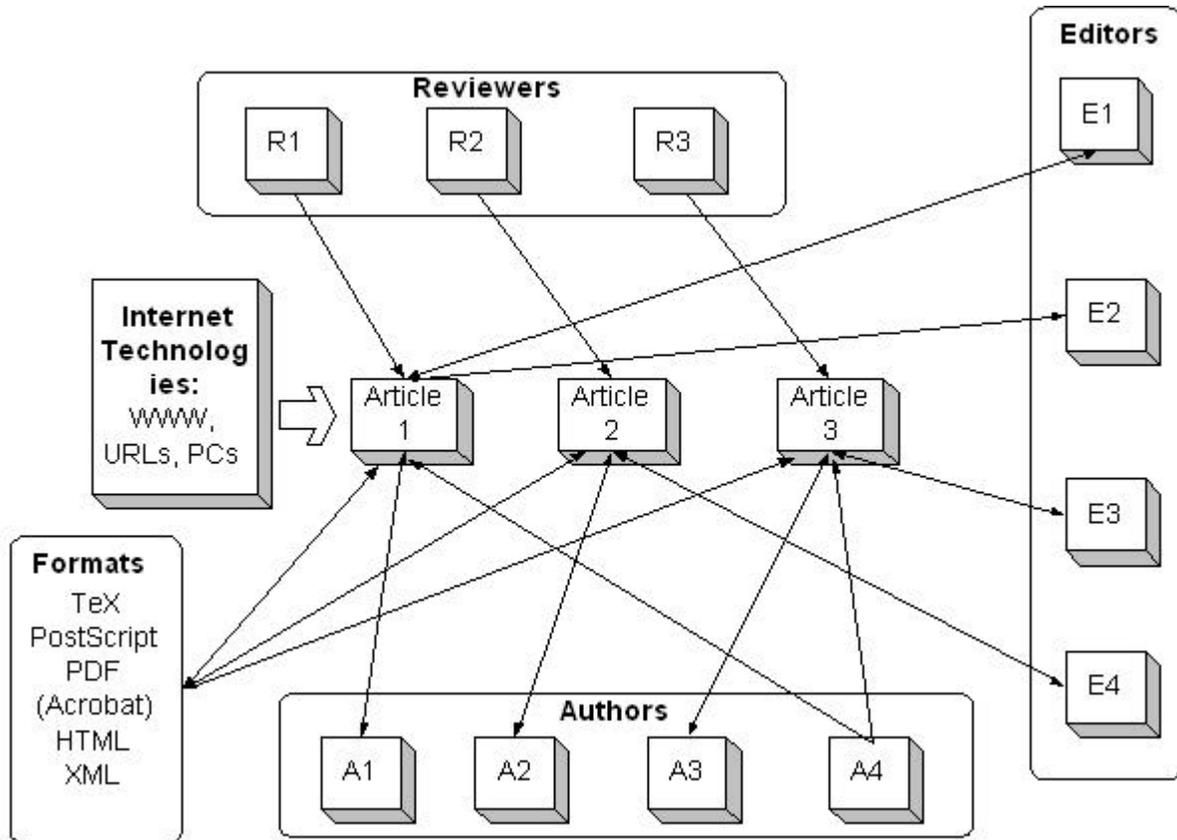


Figure #2 -- An Electronic Journal Embedded in a Socio-Technical Interaction Network

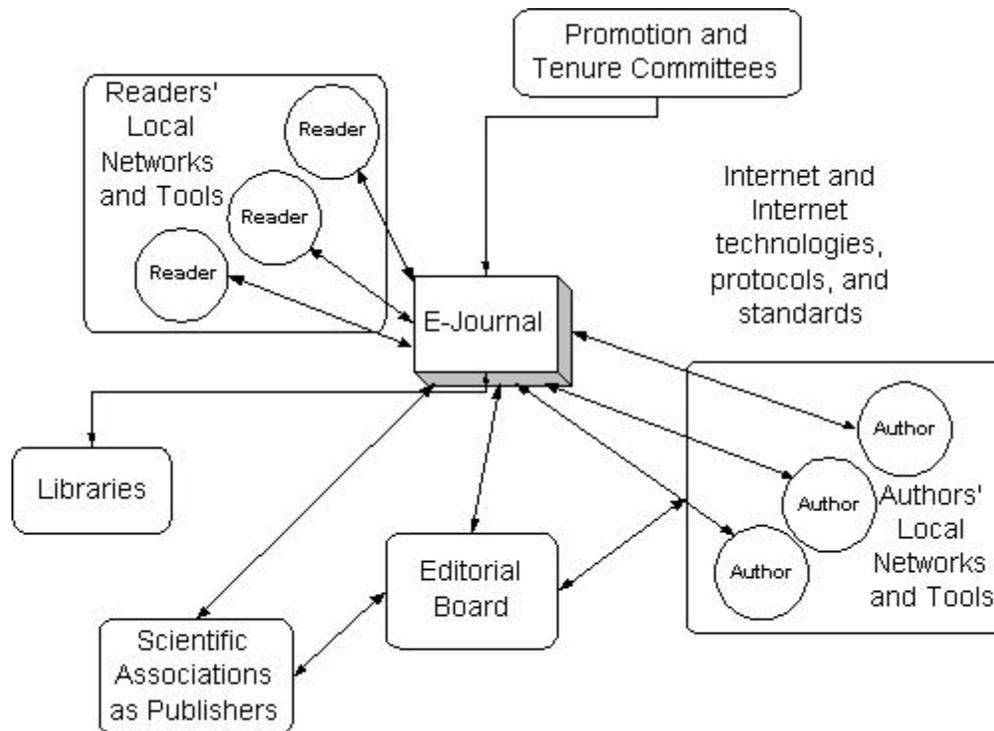
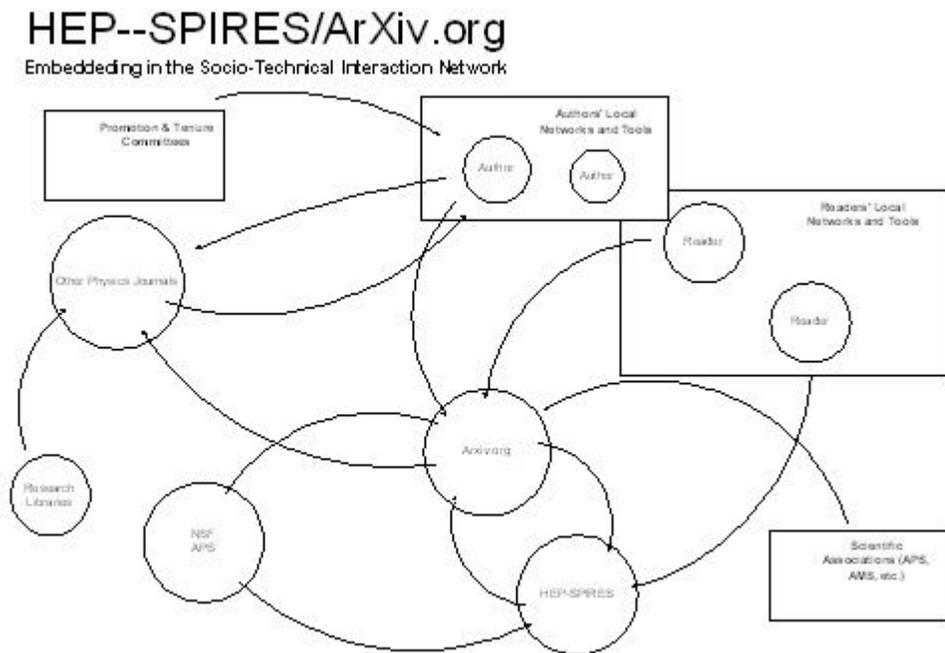


Figure 3: HEP-SPIRES/ArXiv.org Production System



1. A 1998 “socio-technical summit” about Internet2 was organized on this conception of socio-technical.

2. 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” (Edwards, 1999).

3. This formulation of STINs has been significantly influenced by Latour’s (1988) actor-network theory (or, perhaps more appropriately, ontology), as well as our own prior research about “web models” of computerization (Kling & Scacchi, 1982; Kling, 1987; 1992). For more discussion about STINs and their relationship to Actor-Network Theory, please see Appendix A below.

4. The Promotion and Tenure committee in this example is a place-holder 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.

5. This focus on account-taking lends a strong neoinstitutional dimension to the analysis (Powell & DiMaggio, 1991).

6. 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; Mansell and Silverstone, 1995; MacKenzie & Wajcman, 1999:22-24, Wellman, et. al., 1996.

7. This argument owes much to Strum and Latour, 1999.
8. 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), while 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, while a participant or actor may play multiple roles (such as reader, editor, critic, instructor, etc.)
9. This approach differs significantly from the actor-network approach in that STIN methodology requires that the analyst make attempts to understand ahead of time the characteristics and scope of the participants, rather than taking the participants as they come and following them through use or development of the forum.
10. CONVEX and HEPLAB are pseudonyms.

Notes

- i. Harnad, S.(1990) Scholarly skywriting and the prepublication continuum of scholarly inquiry, *Psychological Science* 1: pp. 342-343.
Harnad, S. and Hemus, M.(1997) All or none: No stable hybrid or half-way solutions for launching the learned periodical literature into the PostGutenberg Galaxy.
Ginsparg, P.(1995) First Steps towards Electronic Research Communications, *Computers in Physics* Vol.8, no. 4; July-August 1994, pp. 390-396.
Odlyzko, A., Tragic Loss Or Good Riddance? The Impending Demise Of Traditional Scholarly Journals.
- ii. Odlyzko, A., Tragic Loss Or Good Riddance? The Impending Demise Of Traditional Scholarly Journals.
- iii. Harnad, S.The Paper House Of Cards, ARIADNE.
Harnad, S. and Hemus, M.(1997) All or none: No stable hybrid or half-way solutions for launching the learned periodical literature into the PostGutenberg Galaxy.
Ginsparg, P.(1995) First Steps towards Electronic Research Communications, *Computers in Physics* Vol.8, no. 4; July-August 1994, pp. 390-396.
Odlyzko, A., Tragic Loss Or Good Riddance? The Impending Demise Of Traditional Scholarly Journals.
- iiii. Ginsparg, P.(1995) First Steps towards Electronic Research Communications, *Computers in Physics* Vol.8, no. 4; July-August 1994, pp. 390-396.
- iv. Odlyzko, A., Tragic Loss Or Good Riddance? The Impending Demise Of Traditional Scholarly Journals.
- v. Harnad, S.The Paper House Of Cards, ARIADNE.

Ginsparg, P.(1995) First Steps towards Electronic Research Communications, Computers in Physics Vol.8, no. 4; July-August 1994, pp.

390-396.

Odlyzko, A., Tragic Loss Or Good Riddance? The Impending Demise Of Traditional Scholarly Journals.

vi. Ginsparg, P.(1995) First Steps towards Electronic Research Communications, Computers in Physics Vol.8, no. 4; July-August 1994, pp.

390-396.

vii. Harnad, S.(1990) Scholarly skywriting and the prepublication continuum of scholarly inquiry, Psychological Science 1: pp. 342-343.

Harnad, S. and Hemus, M.(1997) All or none: No stable hybrid or half-way solutions for launching the learned periodical literature into the PostGutenberg Galaxy.

Ginsparg, P.(1995) First Steps towards Electronic Research Communications, Computers in Physics Vol.8, no. 4; July-August 1994, pp.

390-396.

viii. Ginsparg, P.(1995) First Steps towards Electronic Research Communications, Computers in Physics Vol.8, no. 4; July-August 1994, pp.

390-396.

ix. Harnad, S. and Hemus, M.(1997) All or none: No stable hybrid or half-way solutions for launching the learned periodical literature into the

PostGutenberg Galaxy.

x. Ginsparg, P.(1995) First Steps towards Electronic Research Communications, Computers in Physics Vol.8, no. 4; July-August 1994, pp.

390-396.

Odlyzko, A., Tragic Loss Or Good Riddance? The Impending Demise Of Traditional Scholarly Journals.

