The Global Electronic Community: Modeling the Sociology of Self-Organization

Joan Roca
Memorial Library
Mankato State University
Mankato, MN 56002-8400, USA

Bill Ammentorp
University of Minnesota
Minneapolis, MN USA

Tom Morgan
Augsburg College
Minneapolis, MN USA

Abstract

The revolution in telecommunications of the past decade has brought about fundamental changes in the patterns of professional communication. Internet and other electronic networks make it possible for geographically-dispersed individuals to engage in meaningful dialog concerning common problems. These exchanges result in a Global Electronic Community which is unique in human experience. This is a community that is truly **self-organizing** in that it can take on new forms momentarily. Consequently, it is a dynamic community with continually-changing membership and structure; a society populated with both human and electronic agents. It is, as a result, more complex in the roles its members can assume. The sociology of this community is both novel and open to inquiry in depth - in ways never before imagined by students of human society.

This paper explores the sociology of the Internet as a **self-organizing** system. The Internet is represented as a "electronic landscape" where users and their agents seek collections of relevant information. The "landscape" is defined by the accessibility, connectedness and relevance of data files. As users navigate across this "landscape", they add both connectedness and relevance. This creates collections of useful information and welds users into Electronic Communities held together by shared models and paradigms.

The authors present a System Dynamics model of the prototypic Global Electronic Community. Using data drawn from Internet file servers, the authors identify the model and show how its structure and behavior conform to the principles of **self-organization.** Simulation results are used to outline the sociology of such Communities and the implications for professional contributions to sustainable development.

Internet: An Information Landscape

As the Internet grows in size and complexity, it is apparent that users experience an ever-greater difficulty in finding relevant information. This is a particularly vexing problem for the inexperienced user. Consider the following observation resulting from an extensive review of Internet resources.

> *As the number of resources grow, it is increasingly difficult for a user to identify and locate potentially relevant resources to satisfy a need for information in a specific area, particularly if the user is not a specialist in that area.* (Lynch and Preston 1990, 282).

The challenge posed by the explosive growth of networked information utilities is one of *organization.* As Gregorian (1993) states,

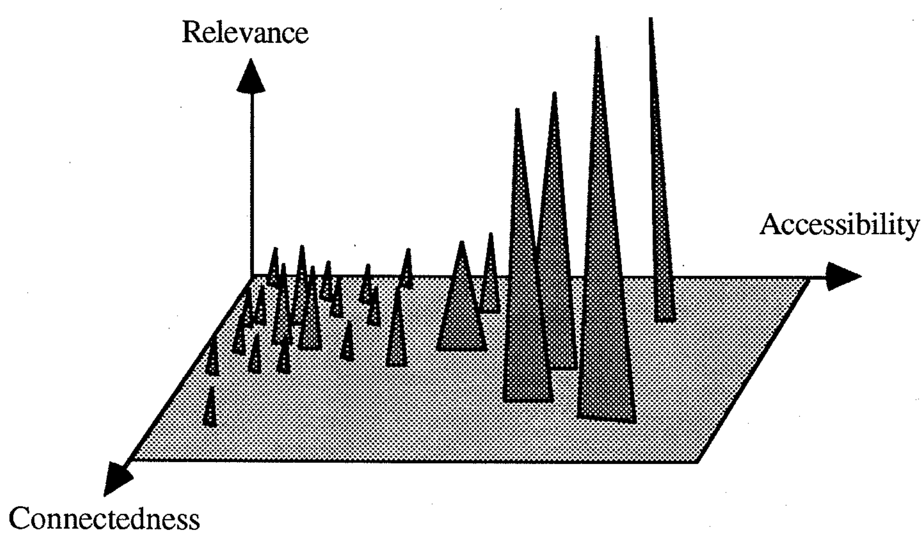
> *We must rise above the obsession with quantity of information and speed of transmission, and recognize that the key issue for us is to organize this information once it has been amassed - to assimilate it, find meaning in it, and assure its survival for use by generations to come* (5).

To date, networked information systems have been largely driven by telecommunications and computing technology. Ever larger data bases have been made accessible on a global scale and endless links have been created to enable users to "surf" the sea of information. However, there is no theoretical foundation which describes how users are navigating this sea - nor are there general principles whereby relevant information might be made readily accessible (Brett 1993).

The Infoscape Model:

The authors have attempted to address this shortcoming by proposing (Ammentorp, et. al. 1994) and testing (Roca, 1995) a model of the *information landscape (Infoscape)* which locates data bases in a three dimensional space (Figure 1).

Figure 1. The Infoscape Model.

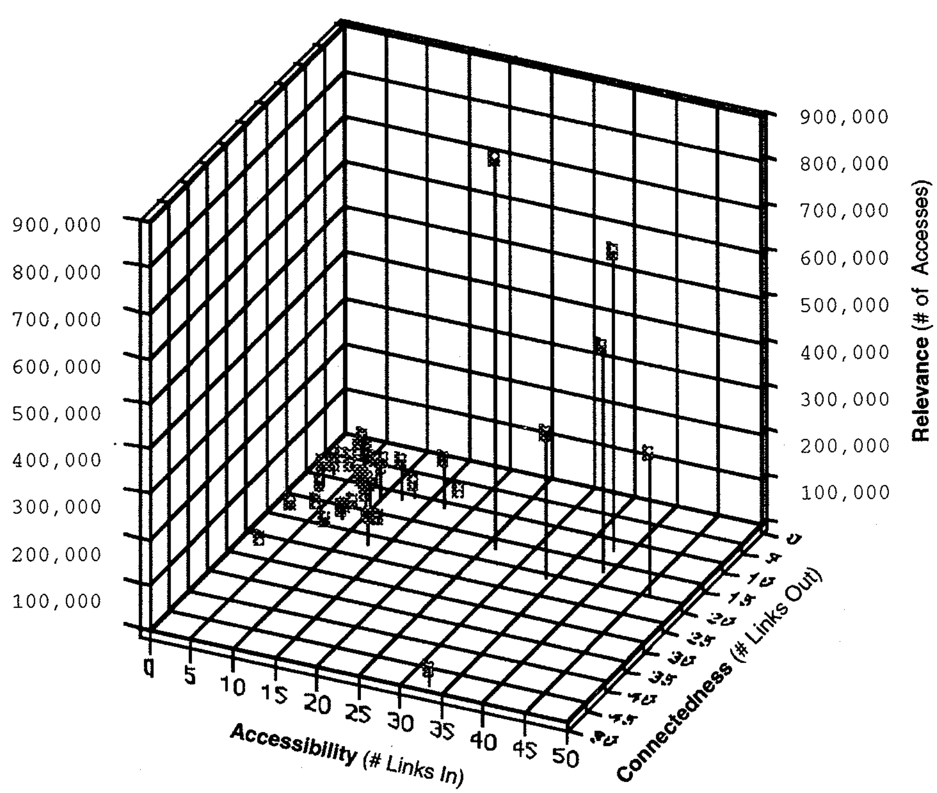


Each dimension of the *Infoscape* describes a critical feature of electronic information. **Accessibility** refers to the extent to which the "common speech" of users can lead them to media of interest. **Connectedness** refers to links among media which enable users to collect information from a variety of sources. **Relevance** is a measure of the value of information to a community of users. As these measures are applied to a particular set of electronic information, the result is a grouping of media in "islands" on the landscape.

Two of these dimensions, **accessibility** and **connectedness**, are drawn from bibliometric studies of scholarly communications. **Accessibility** is measured by the link between materials in a particular database and the keywords of the thesaurus used as an index. This is an instance of the "common speech" of a particular field or profession in that those who assign keywords to materials have a similar understanding of the thesaurus, as do those who use it to access the database. **Connectedness** is determined by the extent to which data bases "point" to one another - thereby assisting the user in navigating across the landscape.

Roca (1995) has attempted to assess the construct validity of this model by positioning library information servers in a three-dimensional space. **Accessibility** of a given server was measured by the number of servers that pointed to it. **Connectedness** was defined as the number of outgoing pointers emanating from a particular server. **Relevance** was determined by the total number of user accesses in a given time period. When the 92 servers in this study are plotted on these dimensions, we find a landscape that is remarkably like that proposed in the *Infoscape* theory (Figure 2).

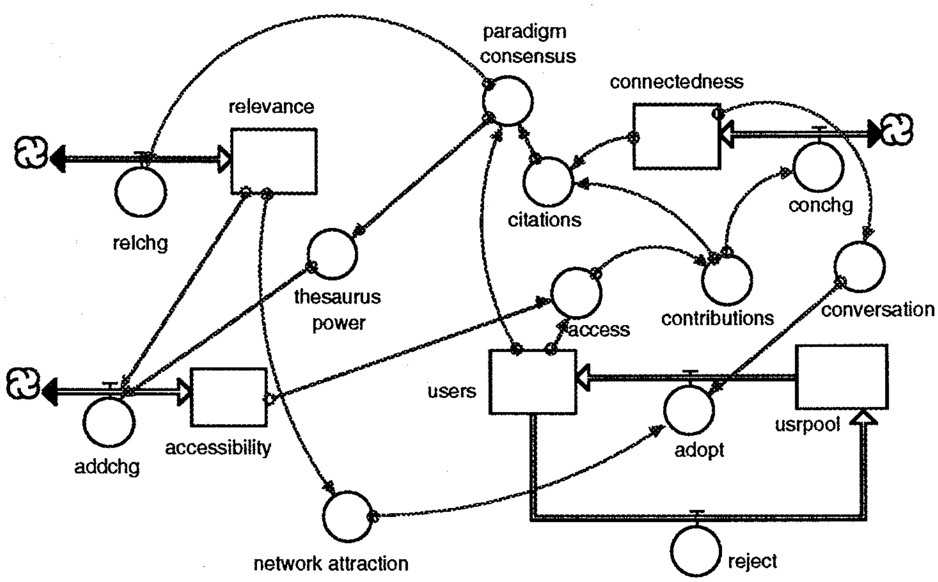
Figure 2. Library Gophers on Infoscape.



The main feature of this figure is the extent to which highly **accessible, connected** servers also have high **relevance**. It is also clear that the converse holds such that servers with low **relevance** are only marginally integrated into the information network.

These findings suggest a model where the interactions among the three dimensions are detailed (Figure 3).

Figure 3. Dynamic Infoscape Model



In this model, the key Infoscape variables **relevance, connectedness**, and **accessibility**, are shown as 'boxes'. We have added two additional 'boxes' to capture the effect of **users** and the **usrpool** on model dynamics. Values in the 'boxes' may increase or decrease through the variables *relchg, conchg, addchg, adopt* and *reject* which act on the flow 'pipes' connected to the 'boxes'. The seven 'circles', network attraction, citation, access, paradigm consensus, thesaurus power, conversation, and contributions are "links" which tie the major factors into a coherent model. For example, paradigm consensus is a function of contributions, citations and **users**. In turn, paradigm consensus causes changes in **relevance**. At the same time, through access and contributions **users** contribute to the **connectedness** of the data base and, ultimately, to paradigm consensus. This is the fundamental dynamic of the model in that it uses the prevailing patterns of professional thought to organize the Infoscape.

The system has the potential to be self-organizing due to the feedback loops shown above. As **users** increase paradigm consensus through citations and contributions, their numbers increase due to feedback of effect from **relevance** through the network attraction variable. There is also a key feedback loop which strengthens **accessibility** as the result of changes in **relevance**. This is "fed back" to further increase paradigm consensus as **accessibility** improves access.

The system is open to outside influence through the introduction of new **users**. They enter the system from the **usrpool** at the lower right-hand side of the above schematic. Their access and contributions add an increment of effect into the feedback and, over time, the system can increase in scope of use and **users** can be drawn into the ruling paradigm. The model also shows that **users** can *reject* the network and return to the **usrpool.**

Model Foundation:

The Infoscape Model grows out of two major research traditions: informetrics and innovation studies. Each of the variables in the Model is defined using conventions developed in these lines of work and links among variables are derived from selected research studies as follows.

Accessibility: In distributed information systems, access is determined by the organization of data bases and the key words or terms used in indices. It is especially critical in electronic information networks due to the "location independence" of resources (Tomer 1992). [Measure: Number of nodes in the information network.]

Connectedness: Information resources do not stand alone in cyberspace. They are connected through information managing utilities and pointers. These pointers provide a road map for navigating the network and for finding distributed information (Kappe, et.al. 1993). [Measure: Number of links among network nodes.]

Relevance: As information becomes accessible and connected, it merges into clusters where new ideas build on existing theory and research. Such clustering or "mapping" increases the value of information and its relevance to users (McCain 1990). [Measure: A product of paradigm consensus and **connectedness.**]

Users: Professional and scholarly groups are the reference points for the study of the sociology of electronic communities. They are the "invisible colleges" where roles are taken and symbolic social behavior is carried out (Crane 1972). [Measure: Actual count of network members.]

Usrpool: The professional group under study is defined by an area of practice and/or scholarship. (American Physical Society 1991) [Measure: Membership in a professional society or other independent assessment of potential use.]

Access: Users are drawn into electronic communities by the perceived value of information. As the currency of information increases in the society, the probability of access in a given user community increases (Drucker 1993). [Measure: **Accessibility / users.**]

Contributions: Electronic communities can be passively explored, or they can be actively engaged. This variable is a measure of the probability that users will offer to include their opinions or products in network data bases (Dordick and Wang 1993). [Measure: **Users** * probability of contribution.]

Citations: Contributions do not stand alone; they are linked to other resources in the data base through citations. Thus, citation is a measure of the extent to which the data base is organized (Egghe and Rousseau 1990). [Measure: Contributions * **Connectedness** * probability of citation.]

Conversation: By participating in informal exchanges, users clarify contributions and relate them to their personal needs and positions (Mulkay 1991). [Measure: **Users** * probability of Listserve participation]

Paradigm Consensus: The orientation of a user to the electronic community is shaped by the ruling paradigm of his/her field. As more users are drawn to the paradigm, it becomes a more powerful organizer of information (Sterman 1985). [Measure: Number of co-citations = citations * probability of co-citations.]

Thesaurus Power: The sway of the paradigm helps to organize information via the language of the profession. This, in turn, defines the keywords and terms of the indexing thesaurus (Sutton and Davis 1992). [Measure: Paradigm consensus * keyword generation rate.]

Network Attraction: The degree to which relevant information brings new users into the network. This is a function of current levels of relevance plus an effect due to the number of users already in the network (Rogers 1974). [Measure: **Relevance** * attractiveness multiplier.]

The Sociology of Infoscape:

The behavior of users of information networks takes shape in the content and direction of communication. Although the initial steps toward network involvement may be based on an individual's membership in a non-electronic group, his/her involvement in electronic communications is based on the benefits of such exchanges (Blau 1964). Interaction in the Infoscape is governed by three principles which are, in turn, representative of feedback loops in the Infoscape Model.

1) User participation in relevant exchanges leads to substantive contributions linked to other users and resources. These are, at first, one-way with the user as receiver. Later, they become two-way exchanges among users. And, finally, they are multi-way as the user contributes to the data base. This is a modern (and simplified) version of the processes described by Merton (1973).

2) Paradigm consensus develops through shared information and results in a "common speech" in general use among users.

3) User populations will grow in size according to logistic innovation patterns in information-relevant environments.

Propositions 2 and 3 make it possible for the model to be self-organizing as users build relevance through paradigm consensus and, ultimately, draw new users into the network (Prigogine 1993).

References:

American Physical Society. 1991. "Report of the APS Task Force on Electronic Information Systems," *Bulletin of the American Physical Society*, 36 (4): 1119-1151.

Ammentorp, W., Morgan, T. and Michels, D. 1994. *The Electronic Community* Paper presented at the annual meeting of the University Council for Educational Administration. Philadelphia, PA.

Blau, P. 1964. *Exchange and Power in Social Life*. New York: J. Wiley.

Brett, G. 1992. "Networked Information Retrieval Tools in the Academic Environment: Towards a Cybernetic Library," *Internet Research*, 3(3) 26-36.

Crane, D. 1972. *Invisible Colleges: Diffusion of Knowledge in Intellectual Space*. Chicago: University of Chicago Press.

Dordick, H. and Wang, G. 1993. *The Information Society*. Newbury Park, CA: Sage.

Drucker, P. 1993. "The Rise of the Knowledge Society," *The Wilson Quarterly*, 17 (2): 52-71.

Egghe, L. and Rousseau, R. 1990. *Introduction to Informetrics.* New York: Elsevier.

Kappe, F., Pani, G. and Schnabel, F. 1993. "The Architecture of a Massively Distributed Hypermedia System," *Internet Research*, 3(1): 10-24.

Kuhn, T. 1973. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lynch, C. and Preston, C. 1990. "Internet Access to Information Resources," in M. Williams (Ed.), *Annual Review of Information Science and Technology*, 25, 263-312.

McCain, K. 1990. "Mapping Authors in Intellectual Space," *Journal of the American Society for Information Science,* 41(6): 433-443.

Merton, R. 1973. *The Sociology of Science*. Chicago, IL: University of Chicago Press.

Mulkay, M. 1991. *Sociology of Science*. Bloomington, IND: University of Indiana Press.

Prigogine, I. 1993. Time, Dynamics and Chaos. In J. Holte (Ed.) *Chaos: The New Science*. St. Peter, MN: Gustavus Adolphus College: Lanham, MD: University Press of America.

Roca, J. 1995. *Road Signs in Infoscape: A Study of the Links Among Academic Library Gophers*. Unpublished doctoral dissertation, University of Minnesota, Minneapolis, MN.

Schultz, T. 1992. "A World Physics Information System," *Serials Review,* 18 (Spring/Summer): 45-48.

Shannon, C. and Weaver, W. 1949. *The Mathematical Theory of Communication.* Champaign-Urbana, IL: University of Illinois Press.

Simsek, H. and Ammentorp, W. 1993. *The Paradigm Shifts of the 1990's.* Paper presented at American Vocational Association annual meeting. Nashville, TN.

Sterman, J. 1985. "The Growth of Knowledge: Testing a Theory of Scientific Revolutions," *Technological Forecasting and Social Change*, 28:93-122.

Sutton, B. and Davis, C. (Eds.) *Networks, Open Access and Virtual Libraries*. Graduate School of Library and Information Science, University of Illinois at Champaign-Urbana.

Tomer, C. 1992. "Emerging Electronic Library Services and the Idea of Location Independence," *Journal of Computing in Higher Education*, 4(1): 88-121.