Articulation of Information Systems[1]

Ted Ward

Information retrieval promises to provide conceptual and technological tools which will help partially to meet the data and documentation needs of an increasingly verbal culture. Scientific inquiry today knows no bounds; the interface between biological and behavioral sciences poses especially complex issues. Not only is the problem that of lateral communication across fields in which data are collected and reported in very different manners, but there is a problem of vertical communication involving several levels of verbal sophistication. In earlier days it was possible to design documentation centers and information systems with only a researcher's uses and perceptions in mind; not so in today's behavioral science arena where the multitude of practitioners constitute a major communication concern for the information system designer.

The Dawn (of a New Problem)

The advent of computer assistance in verbal information processing has been generally viewed as the dawn of a new day. It has been too easy to forget that information storage and retrieval is not new; the staggering problems which computers promise to solve and the glamour of computers themselves tend to obscure the fact that brave and lonesome men and women have been devoting themselves to library management, abstracting services, analysis centers, and document dissemination operations for a long time. The new wave of computer applications specialists has tended to dramatize what *can* be done rather than what *should* be done to solve the problems associated with burgeoning documentation.

A Viewpoint

It is only fair to state the biases of this paper. It proposes optimization rather than polar choices. The drawing of careful balances between or among the contrasting probabilities which affect a design or plan is the essence of sound planning.

Thus, to view the design of information systems as *library versus computer* is unscientific and unproductive. The dreamer sees what could come about if all circumstances are immediately configured like the dream; the realist sees the danger of abandoning the relative predictability of a present mode in favor of any unproven alternative. Optimizing requires a search for the intersection of the curves drawn by these two sets of values. The biases of the paper are activated by dissatisfaction with present modes and an anxiety about the potential calamity involved in arguing for the wrong dream.

---

[1] From "Concerns of Compatibility and Analysis of Research," presented to the Invitational Conference on Documentation and Informational Retrieval in Human Sensory Processes, (jointly sponsored by the American Foundation for the Blind and the University of Pittsburgh, at Pittsburgh, Pennsylvania, on June 14, 1968) by Dr. Ted Ward, Director, Learning Systems Institute, MSU, and Associate Director for Research, USOE/MSU Regional IMC for Handicapped Children and Youth.

Centers and Networks

Wild enthusiasm for defining a model information system along the lines of what *could* be (rather than what *should* be) produces a predictable tendency to think big. Thinking big leads to plans which place the highest values on centralizing, networking, and high-speed communicating of data that have been pre-conformed to compatible forms and formats.

Although it is a truism that information systems are effective to the extent that they meet real needs of real users, it is nonetheless easy to forget the user. What problems are created or solved for the user by centralization, networking, speed, and data format conformity? None of these is "good" *per se*. Thinking big can divert us from seeking timely solutions for the functional here-and-now problems. Such solutions could help to unite scientists; to delay in the hope of realizing greater dreams can aggravate the already serious backlog of unsolved scientific communication problems.

Distinguishing *centers* from *networks* could be a useful semantic device. *Center* would be the term used to indicate the individual information processing unit designed to meet a specific set of known needs for a known understood user group. Matters of geography, field of research, level of sophistication and habit (style) of information use are the principle variables which concern the designer of a center.

Network would then be used to indicate an articulated combination of component centers—the typical product of "thinking big." The major concerns of the design for the network are preservation of the essential characteristics of its component centers, optimization of network intercommunication, and appropriate transformation among levels of sophistication.

The Right to Meet Needs

In the face of super-system dreams and designs the rights of humanity must be set forth. These people are not really small in terms of their past or potential contributions to the world of information processing and dissemination—those who operate the many extant information processing operations: libraries, abstracting services, document centers and so forth. These men and women and their agencies not only have a right to survive, but they have the right, and the resources, to contribute substantially to the coordination and interlocking of information centers. They have a right to help structure the dreams of networks for the future; they are invaluable to plan the sorts of contributions present centers can best make. To overlook the experience and the aspirations of those who have made information centers the focus of their careers seems neither desirable nor tolerable. It could very well be that this oversight has been one cause for the demise of several of the abortive computerized information retrieval systems.

Those who are responsible for the maintenance of operating centers can serve a valuable purpose in keeping the dreams reasonably well tuned to reality. There are information needs already known and understood; there are user habits already defined and accommodated. The on-going information center has a right to continue meeting these needs. Planning of new systems or networks must account for continuance of services to established users.

Uniformity Without Conformity

Interlocked information centers require certain sorts of uniformity. There are two essentially different ways to approach the problem of uniformity:

1) To begin with a model of the ultimate network and work back toward the component and potential component centers, and to design toward a network. The first approach requires that constraints be imposed upon any pre-existing components of the network; these constraints will result in conforming the component centers to the design characteristic of the network.

The second approach seems preferable when pre-existing centers are to become part of a network, because it treats articulation as a secondary concern, allowing primary attention to remain focused on the several sets of users and their needs.

Monolithic conformity has not yet been shown to be an absolute requirement for articulated information centers. True, mandated conformity does offer a tempting alternative to the predictably tedious and difficult design problems common to the planning of uniform articulation among autonomous centers. The price of this tempting alternative is insidious; the required administrative bureaucracy leads to increased remoteness from the problems of users; which, in turn, leads to hopeless communication breakdowns.

Individual centers can be designed or redesigned to fit into an articulated network by accommodating these minimal guidelines:

1. Minimize dependency on peculiar hardware.
2. Minimize coding (maintain high level of meaning in the data entries and in the file search descriptors).
3. Develop standardized or compatible languages of indexing (and/or design transforming buffer functions for the computer to make file-to-file language transformations).
4. Develop standardized or compatible formats for the reduction of data through abstracting.
5. Design administrative procedures and communication structures to differentiate the labors of each center (to minimize duplicated effort in favor of a maximum breadth of coverage).
6. Develop computer programs for format transformations and thesaurus translation to provide center-to-center compatibility.

Challenging the Assumptions

The following list sets forth assumptions which seem to be implicit in some of the large-scale information networks being developed today:

1. Merely setting up an information network will bring potential information users into productive contact with important information which otherwise would not be available to them.

Comment: For scientists this is true only of peripheral information, for practitioners it is not true at all. Scientists have highly developed but largely informal systems for staying in contact with the central core of data which concerns their work. Practitioners need more than additional information systems—they need training in how (and why) to use outside information.

2. The existence (availability) of an information center or network will generate new habits of information use.

Comment: While this may contain borderline truth, it falls into the realm of wishful thinking. Either an information center makes contact with users through their established habits (minimally modified to accommodate procedural requirements) or else the center must engage in extensive training programs to develop the new habits required of users. It would be useful to test the hypothesis that

there is no predictable relationship between the existence or non-existence of a formal information system and the information-use behaviors of any professional group. (This hypothesis does not contend that there is no relationship with the information-seeking behaviors).

3. Effective nation-wide information networks can be designed *de novo*, while avoiding competition with existing information centers (abstracting services, microfilm, libraries and document dissemination centers).

Comment: The key word is "effective." It seems illogical to try to solve the problems of multiplication of information sources through the creation of one more source—unless the newly created source can eliminate the need for two or more others.

4. A network designed primarily for researchers cannot be useful for practitioners.

Comment: Experience to date supports this assumption, but computer classification and transformation of data may someday demonstrate that one network can meet the needs of several levels of information users. Meanwhile it would seem advisable not to sell several audiences on a system which will meet the needs of only one of those audiences.

5. Studying the user is a once-only job.

Comment: We can never know too much about the needs and characteristics of the clients of an information center. Effective use of the center is a function not only of design characteristics of a technological sort, but it is a function of the degree to which these design characteristics accommodate the human characteristics of users. Today users "jump lines" more easily and frequently; in fact, a major motivation for information centers and networks is the growth of interdisciplinary scholarship. Continuous study and systematic in-system feedback on user satisfaction! Frustration must be built into any sound plan.

6. There are vast quantities of information which never became available to potential users.

Comment: This assumption is clearly a statement of fact. The issue is whether or not this state of affairs should be changed. Researchers are not contending that this "fugitive" literature is needed in their work; they are, instead, inclined to trust professional reviewing and abstracting services, slow as they are, to serve as a sort of filter—sparing them from wasting time with inconsequential material. It is said, especially in education, that practitioners need access to this "fugitive" documentation. One wonders why, since it is apparent that teachers and school administrators are unsystematic users of professional literature—possibly a function of the disorganized and massive sprawl characteristic of professional and semi-professional literature in education. Thus when ERIC or any other network proposes to process "fugitive" documents and materials the issue is not so much whether the machines are capable of handling this mass, but whether the human processors will be able and willing (enough time? appropriate effort? discrimination skills?) to handle this quantity of disparate pieces.

7. What *can* be done *should* be done.

Comment: This is the great technological trap of our era. Technical capability teases us into building elaborate systems to accomplish missions of dubious worth. Research and development in information systems should be accompanied by sober philosophical examination of values and relative merits of different approaches to problems. Technological capabilities raise difficult value questions since human behavior is today responsive to and reconstructed by the time-effort skills demands of the technology.

Relating Computers to Information Problems

We still hear well-meaning people say, "Yes we have the computers, but we lack the software." I'm weary of this. It's neither hardware nor software that holds us back. What we lack are hard heads. We should stop arguing about characteristics of Utopian systems, and get about the business of applying information retrieval procedures to known needs in reference to know users.

Since we believe this at the Learning Systems Institute, our studies have proceeded through these steps:

1. Develop a view of the role of information use in the professional areas with which we are concerned.
2. Build a general system of computer programs capable of assisting in the various sorts of information use in the professional areas.
3. Generate applications of these programs for various information system configurations.
4. Field-test and refine procedures.
5. Evaluate the needs met as well as the needs not met by each configuration.
6. Redesign programs, applications, and configurations in terms of the feedback from the evaluation.

System design criteria which we have imposed on our own information system projects are presented here as a summary, not because they are right but because they reflect the best thinking our group of scientists and engineers has generated from careful study and analysis of contemporary needs.

What comes off the computer in response to a search must be verbally meaningful (not just a list of code words or serial numbers).

> On-line search operations must be available and must permit inquirers to use their own natural language.
>
> Abstracts of documents or data files must be available directly from the computer (to allow the inquirers to scan—or even study—as they participate with the computer in the narrowing down of their search questions).

Use the computer to solve problems, not to make them worse. Thus, before you begin to design, *make sure you understand the problems.*

Reference List

Committee on Scientific and Technological Information (COSATI) *Recommendations for National Document Handling Systems in Science and Technology*. Appendix A—A Background Study, U.S. Department of Commerce, 1965.

Menzel, H. *Review of Studies in the Flow of Information Among Scientists*. Bureau of Applied Social Research, Columbia University, 1960.

Menzel, H. Scientific Communication. Five Themes from Social Science Research. *American Psychologist*, 1966: Vol. 21, pp. 999-1004.

Rogers, E.M. *Diffusion of Innovations*. Free Press, Glencoe, 1962.