RESOURCE SHARING THROUGH NETWORKS: PROBLEMS AND POTENTIALS FOR THE SOCIAL SCIENCE COMMUNITY

Lorraine Borman

Northwestern University

ABSTRACT

This paper overviews the utilization of computer networks via interactive terminals for accessing and manipulating public data bases through a general purpose user-oriented information system. The major premise is that many of the problems inherent in such use could be alleviated with two mechanisms: first, an interaction monitor which would serve a dual purpose: 1) provide feedback to the designers of the information system to ensure continuing system responsiveness to the varying needs of the user community; and 2) provide data on how the system is used which could be transmitted to a tutorial serving both as a learning and reference tool. The second mechanism would be the introduction of a tutorial component into resource data bases accessible by the computer information system, which would provide browsing capabilities, context-sensitive and data base specific tutoring and diagnostic inquiry facilities. Without such capabilities, system design will not be truly responsive to user needs, utilization of information resources will not occur, and the potential of timesharing networks will not be reached.

INTRODUCTION

Utilization of networking facilities to gain access to large scale computer systems and to public data base services has been increasing rapidly in the Seventies. We might, in fact, call this period the age of the data base, with hundreds of data bases being generated by various agencies for public consumption. The extent and success of network access to these resources by social scientists may depend, however, on the unified efforts of computer and information scientists, educators, psychologists, linguists, data base producers, and network operators.

Our thesis in this paper is that while network access is indispensable for many purposes, it has not grown to its fullest potential for many reasons, some technological, but others behavioral and learning-oriented. We will not discuss the problems of hardware and communications linkages here except to say that they are rapidly being reduced. The information seeking behavior and characteristics of the potential user community, composed of both experienced and novice researchers, casual and frequent data base and computer system users, may, however, be a greater barrier to increased and satisfied usage.

The sharing of public data bases, accessed and manipulated by a general purpose data management system capable of providing transparent interfaces to specialized statistical, graphical and report generation packages, in a network environment, is the goal. A few problems immediately come to mind. The location of the terminal may be such that the user becomes separated from the computing center environment and from needed user aids. These may be reference materials, directories, and specialized technical documentation, or human contact with the computer operator, programmers, consultants, and subject specialists. Automated system support should be available in an interactive mode, or on request from the terminal. The network itself can be a useful communication medium to provide a summary of historical information of interest to the user.

User support in automated interactive systems in the network environment becomes increasingly important. Some capabilities for on- and off-line instruction and telephone support have been implemented, in some degree, on almost all operational networks. Several problem areas, though, emerge. Neumann (1973, p. 16) in a report prepared for the National Bureau of Standards notes that "proper design of support capabilities needs to be addressed in an integrated manner, providing support when and where needed on a highly individualistic basis, but at the same time considering the overall system economics of preservation of processing and storage capacity." Major interest, Neumann continues, should focus on interactive language design, tutorial design, integration of hard copy documentation and on-line printing, and further exploitation of user feedback. Two of these areas, tutorial design and user feedback, have been studied by the Information Systems and Services group at the Vogelback Computing

Center, Northwestern University. Tutorial design is closely related to the problem of computer-aided instruction. A comment often heard (Nickerson, 1969, p. 12) and true to a large extent, is that the need of the future is not so much for computer-oriented people as for people-oriented computers. Neumann (1973, p. 17) points out that many authors in the field of man-terminal interaction are still "computer-oriented" people and that there is a need for inter-disciplinary efforts including computer science, linguistics, psychology, and other human-oriented disciplines. The balance of this paper will look at work done at Northwestern University in the area of system performance and user feedback and will relate tutorials and feedback mechanisms to the problem of network access and resource sharing of public data bases.

The first barrier to successful utilization is the difficulty of acquiring the necessary knowledge required to use the system. Marcus et al. (1971) working with the INTREX system, have made some general observations regarding user behavior:

1. Users often fail to notice even the most explicit instructions.

2. There is not one single method applicable to all users.

3. If there are too many instructional options, they are all ignored; the user prefers to be given instructions only when needed.

4. Users do not like to spend time in preparation for system use; they would rather use the system.

5. Users are constrained by previous experience and by training. There is a barrier against learning anything new.

6. Some users fear the machine, either because they think they will appear foolish, or because they fear they may damage the machine (machine fear). Some users do not want to ask for advice (people fear).

7. Users are overawed by the complexity of the system, and assume they need not or cannot understand the system. Such attitudes impede learning.

The question, though, is how to translate these generalizations into a user support system, and this cannot be accomplished until the designers of systems know how the system is being used by the end user, which often is quite different from that envisioned by the designer. Any user-oriented system must be responsive to the continually varying needs of the user community and to the different and differing levels of familiarity with the system. Provision must be included in the network protocols and the information system to monitor the user/system interactions to generate feedback to the system designers. Study of such usage data will enable, when necessary, intelligent redesign of the user/system interface, and will provide the necessary data for the development of the comprehensive on-line tutorials necessary for network access (Borman and Dominick in progress).

THE USER COMMUNITY

We will define the anticipated user community of a nationally available educational network supporting access to social science oriented public data bases as composed of students, faculty, and staff of academic and research institutions. One example of such a network is EDUNET (EDUCOM, 1976). Within this community will be experienced/inexperienced users, casual/frequent users, programming/nonprogramming users, and others. The diversity of application areas can range from analysis of voting behavior in a recent local election to economic forecasting based on data gathered nationwide and spanning the previous 20 years.

THE NETWORK RESOURCES

We project a network which provides access to not only public data bases but which will also place additional content requirements on the data base before it is accepted as a network resource. These would include:

1. Comprehensive description of the data: its general content, its structure, and specific variable information.

2. References to both bibliographic citations and to previous users of the data base.

3. A tutorial, context-oriented, developed by the producers of the data base and reflecting anticipatory usage.

4. Examples of use: research selection, reporting, graphical representation of the data, etc.

5. The data.

In addition, the network resources would include: 1) an information system equipped with an interaction monitor and the necessary linkages to process the tutorial component of the data base, and 2) a simple and "understanding" set of access protocols to allow the user to enter the network, enter the information system data base, inquire about general information such as charging, scheduling, news events, etc.

THE CONCEPT OF USER INTERACTION MONITORING

All interactive information systems are originally designed according to some pre-established guidelines. These may include optimization for updating or searching or reporting, or ease of exportability from one hardware installation to another following the concepts of structured programming and modularity. Most will also proclaim their responsiveness to the user--that oft cited phrase--user-oriented.

The system designer begins with preconceptions based on the existing literature (almost negligible in the area of users interacting with numeric data bases and data management systems; larger, but often not more generally useful, regarding use of bibliographic data bases; and a few industrially oriented studies of the decision-making processes). The designer, is more often than not, not a user of information systems. Given the varying pressures upon design time vs. implementation time, design usually loses. Implementation proceeds, following the dictates of the easiest guidelines--those connected with hardware, software, and economics. The goal of "user-oriented" is nebulous, so features such as language, diagnostic messages, prompting--everything we think of as user/system interaction--are written into the system based on programmer dictates. In actuality, we design and implement interactive systems bsed on what we think they should do, rather than on what the user would expect them to do.

Given this situation, and we believe it to be widely true, the need for a feedback mechanism becomes apparent. An interaction monitor can provide various levels of information. Some have been implemented to report system-oriented data: length of session, central processor and peripheral processor time per search, amount of disk space used, job cost, etc. These data can then be used to evaluate the impact of the system on the total computing environment, on the use of capabilities within the system, etc. No monitor has yet been implemented to capture data on how the system is used by various types of users with the idea that a feedback mechanism would be used to both redesign for system efficiency and usability, and to provide the information necessary to produce a context-oriented, data base specific, learning and reference tutorial.

THE QUESTION OF CAI-TYPE TUTORIALS

The potential utility of CAI techniques within an information system environment is well recognized and has been described by Dominick and Borman (1976). The potential benefits can include:

1. Author controlled and user controlled sequences.

2. Dynamic instructional strategies which can adjust to the experience, performance, and information-seeking requirements of individual users.

3. Inclusion of various levels of preprogrammed spelling algorithms and synonym recognition which tend to give users the illusion of at least a minimum amount of system intelligence.

Given these potential benefits, why have information system tutorials failed to live up to expectations? The problem areas include:

1. Tutorials function only as stand-alone programs. Users, interacting with the tutorial sequence, cannot easily and

immediately transfer control to the information system to try what they just learned, returning back to the tutorial whenever they encounter problems.

2. Most tutorials provide only extremely verbose or extremely terse information with minimal user interaction.

3. Although many tutorials employ the concepts of teaching by example, it is not clear how effective such training is if the examples are not data base specific.

4. Most tutorials rely on user comments concerning user problems, errors, etc. to provide feedback to system designers. These are not sufficient to enable knowledgable evaluation and interface redesign.

RIQS REMOTE INFORMATION QUERY SYSTEM

An initial attempt to coordinate an interaction monitor, a comprehensive tutorial system and a feedback mechanism is exemplified in the RIQS system. The Remote Information Query System (Borman, et al., 1976) developed at Northwestern University, contains a comprehensive, automated on-line monitor for collecting data both on system efficiency and on user interactions. Some of the data collected include:

1. name and department of user
2. name and size of data base
3. time required for entry of query
4. time required to execute the query
5. frequency of use of system commands and capabilities
6. frequency and types of errors
7. full context of errors
8. full search text entered by user

Sixty five different elements of information are gathered for each query and total session. Three levels on monitoring are provided to avoid any invasion of data base privacy or user application privacy.

THE COMPLEAT TUTOR

In addition to the monitor, a tutorial in the form of a front-end processor has been developed. RIQSTUTOR is more than simply a tutorial; it also provides public file browsing capabilities, and automatic retrieval and display of sample file descriptions and sample records, the enforcement of data base password specifications and other front-end processing functions.

The RIQSTUTOR has tried to alleviate some of the problems previously listed. It has four modes: browsing, tutorial, searching, and selective inquiry. Browsing mode provides capabilities whereby anyone, regardless of experience, can browse through available public data bases via simple list or menu selection. Users can select general information such as the size of the data base, a description of the contents and suggested uses, and application areas and disciplines to which the data base may be especially relevant. Also available are the names and descriptions of the data elements and a sample record from the data base.

Tutorial mode represents a full tutorial for initially training users in the syntax and semantics of the data base system's query language, and provides examples of search strategies and tips for searching. Searching mode provides the internal linkage from the tutor to the processor. Users can search and analyze data in either their own data bases or any public data bases. For each of the public data bases, sample searches are provided which illustrate common usage. Selective inquiry mode is invoked whenever a user enters a HELP command. The information system diagnostics are coordinated with tutorial and error recovery sequences to allow automatic branching to the appropriate error recovery information and to allow presentation of context-sensitive tutorial information.

SUMMARY AND CONCLUSIONS

The concept of user interaction monitoring, multi-mode tutoring, and a feedback mechanism to provide individual user, and individual data base information, to the tutorial has been presented. In order to interface a diverse user community utilizing a computer network with public data bases via sophisticated processing systems, it appears that such a three-way interface system must be adopted. The monitor, serving both as a recording device and a feedback mechanism, can ensure system responsiveness to the varying needs of the user community; the tutorial can provide file browsing capabilities, context-sensitive training, selective inquiry facilities linked to diagnostics, and internal linkages to the data base management system itself. Only by providing such personalized, and necessary, services will utilization of public data bases via time-sharing networks reach their full potential.

REFERENCES

Borman, Lorraine; Chalice, Robert; Dillaman, Donald; Dominick, Wayne; and Kobbe, Ruth. RIQS Remote Information Query System. (Evanston, Ill.: Vogelback Computing Center, Northwestern University, 1976)

-----, and Dominick, Wayne. An analysis of data captured by an interaction monitor. National Science Foundation Grant # DIS 76-19481, in progress. Northwestern University, Evanston, Ill.

Dominick, Wayne and Borman, Lorraine. User/System Interfacing in an Interactive Retrieval, Statistical and Graphical Analysis Environment. Proceedings of the 5th ASIS Mid-Year Meeting. Vanderbilt University, Nashville, TN, May, 1976, pp. 34-49.

EDUCOM. Description of EDUNET. Bulletin of the Inter-university Communications Council (EDUCOM). 11 (3, Fall, 1976).

Holmes, D. C. Computers in Oil 1967-1987. Computer Yearbook and Directory (2nd Ed.) (ed. F. H. Grille) (Detroit: American Data Processing)

Marcus, R. S.; Benenfeld, A. R.; and Kugel, P. The User Interface for the INTREX Retrieval System. Interactive Bibliographic Search (ed. D. E. Walker) (Montvale, NJ: AFIPS Press, 1971), 159-201.

Nickerson, R. S. Man-Computer Interaction: A Challenge for Human Factor Research. Ergonomics 12, 4 (July, 1969).

Neumann, A. J. Network User Information Support National Bureau of Standard Technical Note 802 (Washington, D. C.: U. S. Government Printing Office, 1973).