



STANFORD RESEARCH INSTITUTE
Menlo Park, California 94025 · U.S.A.

Annual Report

Covering the Period 15 April 1973 through 31 December 1973

December 1973

SRI-ARC 21453

EXPERIMENTAL DEVELOPMENT OF A SMALL COMPUTER-AUGMENTED INFORMATION SYSTEM

*By: JEANNE B. NORTH, Research Analyst
Augmentation Research Center*

Prepared for:

INFORMATION SYSTEMS BRANCH
OFFICE OF NAVAL RESEARCH
DEPARTMENT OF THE NAVY
ARLINGTON, VIRGINIA 22217
Attention: MR. M. DENICOFF

CONTRACT N00014-70-C-0302

SRI Project 8622

Approved by:

D. C. ENGELBART, *Director*
Augmentation Research Center

BONNAR COX, *Executive Director*
Information Science and Engineering Division

Copy No.

Abstract

A system of information-handling for use in an Office or by a Community of knowledge-workers, using computer-based tools, is under design at the Augmentation Research Center. Support from ONR during the period April 15 to December 31, 1973, was directed toward study of the tools as currently used at ARC and on the ARPA Network, to determine their feasibility for early transfer to another Office, and to produce procedural tools to facilitate use of the system in an Office.

The components of the information-handling system that were examined were: the Online Journal System, the Ident System, User Programs, Personal Files, and Retrieval Provisions.

Procedures were designed for capturing, storing and retrieving research information in an Office, and a manual was written giving procedures for computer-based operations and contrasting them with procedures for non-computer-based operations for the same activities.

Other procedural tools were designed and are included in the report: a design for an online correspondence log system, a scenario for producing subject bibliographies, and a scenario for using the Network Journal.

CONTENTS

Abstract

INTRODUCTION Section 1

SUMMARY OF RINS ACTIVITY Section 2

CAPTURING, STORING, AND RETRIEVING RESEARCH INFORMATION Section 3

CRITIQUES OF TOOLS FOR AUGMENTATION Section 4

Journal System	4a
Ident System	4b
User Programs	4c
Personal Files	4d
Retrieval Provisions	4e

APPENDICES

- A. Procedures for Capturing, Storing and Retrieving Research Information
- B. Design for a Correspondence Log System
- C. Scenario for Producing Subject Bibliographies
- D. Scenario for Using the Network Journal

INTRODUCTION

1

A small computer-augmented system is being developed in the Augmentation Research Center at SRI, based on online, timesharing components, to provide research workers with a means for handling information. The set of techniques developed for this use has been termed the Research Information System (RINS).

1a

The computer-based tools on which the Research Information System is built are being developed at the Augmentation Research Center (ARC) of Stanford Research Institute in a long-range program supported by several sponsors. With this multiple support, ARC has developed hardware and software by which information in text form can be input, edited, analyzed, recorded, disseminated, and retrieved by individuals and groups working in geographically-distributed situations.

1b

Support from ONR is directed toward study of means of incorporating the tools into the handling of research information by knowledge workers, in the knowledge workshop concept (14724,), pursuing their work as a Group and as members of information-handling Communities. Through study of the needs of knowledge workers and of the extent to which the experimental tools meet these needs, the further development of these tools is given direction.

1c

This report documents some of the study carried on to examine the use of these tools in the Workshop environment at ARC, where these tools are used in the Group and in a Network Community, and to examine the suitability of these tools for transfer to other Offices which function as Workshops. The examination tends to search out the weaknesses and the features that may need improvement, to counterbalance the general enthusiasm for the unquestioned potential of these tools with a hard look at problems which the user faces at the present stage of development.

1d

SUMMARY OF RINS ACTIVITY

2

During the current period, April 15 to December 31, 1973, various analyses and module designs were completed. The subjects chosen were selected for their apparent relevance at this point in the ongoing development of a Research Information System.

2a

The main objectives of the project work remained the same as those stated at the outset of the project:

2b

(1) Develop an operational information system specifically to support the needs of a computer-systems research and development community.

2b1

(2) Provide as much computer augmentation as possible to all phases of the operation of such a system -- with a practical orientation toward learning how users will get maximum cost effectiveness in these research-information functions within the systems development environment of several years hence.

2b2

(3) Stimulate the building of technical information data bases of intrinsic value to the systems development community, by leading other people into cooperative use of these ONR-supported tools and services, with the ultimate objective of having the other people pay part of the costs and share the data bases they develop with a larger community.

2b3

During the contract period, development of the ARC Utility raised the possibility of an imminent introduction of the tools to an outside group (see footnote), an "Office" in the Knowledge Workshop concept (14724,). To accelerate preparation for such an introduction, emphasis of the ONR work was placed on study of the RINS tools as currently used and of means for transfer of present technology to such an Office. Plans for transfer resulted in a proposal for 1974 work submitted to ONR in November (19938,). Emphasis of work during the period changed somewhat from design for an indefinite future to study of the suitability of present tools for an early transfer, particularly of aspects which could be improved to make transfer most satisfactory. The choice of aspects to be studied and the conclusions drawn are the author's and are not to be considered as representing ARC or ARPANET consensus.

2c

[Since this report was prepared, Offices in the following have become users of the Utility: Naval Ship Research and Development Center, Rome Air Development Center, ARPA, Ballistics Research Labs, Bell Canada, Hudson Institute, MIT Seismic Program, and Educational Testing Services.]

2d

Two approaches were taken to the design of an information system to be implemented in the near future for an Office of knowledge workers. 2e

I. A dual system was designed for handling research information; it included a design for handling information in a non-computer-based mode, and a design based on computer use where feasible. This report contains a juxtaposition of the steps in the two modes, with some comments on the results and on the merits of each. 2e1

II. A series of critiques were made of ARC computer-based tools as currently used by ARC and by the ARPA Network Community. 2e2

The design builds on the April 1973 examination of the various information-handling needs of a group of knowledge workers (16737,) that appeared in the previous report on this project (16508,). The dual system was designed in the form of a set of policies and procedures covering many of the information-handling activities presumed to be essential to an Office. The activities covered were: 2f

Information Contained in External Publications

Information Generated by the Individual in His Work

Information Generated by Dialogue in the Office

Information Issued in Formal Office Reports

Information Communicated by Correspondence

Information Communicated by Phone

Information Gathered from Visits and Visitors

Information Collected by the Individual for Personal Use

The other phase of the project work was a series of critiques of tools useful to RINS as they have been experimentally implemented in the ARC environment. The ongoing evolutionary nature of the tools makes any assessment of their successes and drawbacks a reflection of a temporary state. However, assuming the possibility of an imminent export of these tools, it seemed there would be value in examining their recent use for any implications this use would have for further development. 2g

The RINS design considerations study referred to above (16737,) looked at the handling of research information as a conceptual system whose components are collection, capture, processing for

dissemination, means of dissemination, processing for retrieval, and means of retrieval. The efforts reported here view information handling from a different perspective. Various tools used for one or a combination of the conceptual components are placed in primary focus, and experience with them is reported. Then for each tool a set of conclusions is derived as to the ways in which and degree to which it meets the needs it was designed to meet, and suggestions for improvement are made where applicable.

2h

The computer-based tools which are examined are:

2i

The Journal System

The Ident System

User Programs

Personal Files

Retrieval Provisions

Each of these tools is the subject of a separate section in this report.

2j

One area of information handling had not been addressed before, the handling of correspondence. A design for handling correspondence in the Office is included as Appendix B.

2k

A scenario was prepared for constructing bibliographies, as an alternative to making annotated catalogs of citations for documents, and this scenario (17884,) is included as Appendix C.

2l

A scenario was prepared to aid the user in use of the Network Journal, and this scenario (22383,) is included as Appendix D.

2m

It is important to note that the tools examined represent the then-current state of parts of a whole Knowledge workshop system that is under constant evolution. The comparisons that are made, advantages and disadvantages that are brought out, and the conclusions reached, are important contributions to an emergent discipline, the design and analysis of augmented Knowledge Workshops; but many of the specific points will soon be obsolete as evaluative measures. New assessments will need to be made as the tools and system evolve.

2n

It should also be noted that no methodology had been developed for studying these kinds of systems. The analyses made here are numerically accurate, and the data are representative of the periods studied, but the selection of aspects to be studied was based on intuition rather than scientifically established principles. There is a clear need for research in the means by which use of computer-augmented communication tools should be studied, and for continuing analyses based on this research. A prime consideration should be the dynamic nature of the phenomena being studied: the evolution of the tools, and the concurrent changes in user behavior.

20

BACKGROUND DOCUMENTS

2p

Engelbart, Douglas C., Richard W. Watson and James C. Norton. The Augmented Knowledge workshop. Presented at National Computer Conference, New York, June 1973. Stanford Research Institute. Augmentation Research Center. ARC Journal 14724. 1 March 1973. 19p.

Engelbart, Douglas C. Coordinated Information Services for a Discipline- or Mission-Oriented Community. Presented at Second Annual Computer Communications Conference in San Jose, California, January 24, 1973. Stanford Research Institute. Augmentation Research Center. ARC Journal 12445. 12 December 1973. 7p.

A knowledge Workshop for the Navy: An Experiment in Technology Transfer. Proposal for Research SRI ISU 73-175. Part 1 - Technical Proposal. Stanford Research Institute. Augmentation Research Center. ARC Journal 19938. 28 November 1973. 19p.

North, Jeanne B. Experimental Development of a Small Computer-Augmented Information System. Annual Report, Contract N00014-70-C-0302, April 1973. Stanford Research Institute. Augmentation Research Center. ARC Journal 16508. April 1973.

North, Jeanne B. A Status Report on RINS Design Considerations. Stanford Research Institute. Augmentation Research Center. ARC Journal 16737. April 1973. (Included as Appendix A in 16508.)

USER GUIDES

29

ARC Primer. Stanford Research Institute. Augmentation Research Center. 3 January 1974. 10p.

SRI-ARC / Office-1 TENEX Users' Guide. Stanford Research Institute. Augmentation Research Center. 22 March 1974. 39p.

TNLS Users' Guide. Stanford Research Institute. Augmentation Research Center. ARC Journal 19200. 27 November 1973.

DEX Primer. Stanford Research Institute. Augmentation Research Center. No date. (Online document, continuously revised.)

DNLS Users' Guide. Stanford Research Institute. Augmentation Research Center. ARC Journal 20400. 4 December 1973.

Journal Users' Guide. Stanford Research Institute. Augmentation Research Center. No date. (Online document, continuously revised.)

North, Jeanne B. Scenario for Using the Network Journal. Stanford Research Institute. Augmentation Research Center. ARC Journal 22383. 27 March 1974. (Included as Appendix D in this report.)

Network Journal Submission and Delivery. Stanford Research Institute. Augmentation Research Center. ARC Journal 17777. 13 July 1973. 8p.

User Programs. Stanford Research Institute. Augmentation Research Center. No date. (Online document, continuously revised.)

L10 Programming Guide. Section 7. Content Analysis and Sequence Generator Programs. p.69-73. Section 8. Invocation of User Filters and Programs. p.75-87. Stanford Research Institute. Augmentation Research Center. ARC Journal 9246. 4 April 1972.

North, Jeanne B. Scenario for Producing Subject Bibliographies. Stanford Research Institute. Augmentation Research Center. ARC Journal 17884. 17 July 1973. 6p. (Included as Appendix C in this report.)

North, J. B. Codes Used in the Master Catalog. Stanford Research Institute. Augmentation Research Center. ARC Journal 10937. July 1972. 3p.

Output Processor Users' Guide. Stanford Research Institute. Augmentation Research Center. ARC Journal 12209. 23 August 1973. 97p.

CAPTURING, STORING, AND RETRIEVING RESEARCH INFORMATION

INTRODUCTION

3a

The product of the knowledge worker is more knowledge, and the aim of the workshop is to maximize the quality and quantity of the product. In manufacturing operations for most products it is possible to achieve efficiency of production by measuring input and output of materials and assuring that waste is kept to a minimum. In the production of knowledge, the principal raw material is information. Measurement of the raw material is inexact and the means of moving and processing it are loose, and it can be assumed that much is wasted.

3a1

To increase the amount and availability of information to be used in the production process, methods must be devised to capture it and to contain it where it can be drawn upon. Sources must be recognized and tapped, and the material acquired must be stored for timely and easy retrieval.

3a2

This section discusses the means for capturing, containing and retrieving the raw material of knowledge from several sources. The sources are:

3a3

- Information Contained in External Publications
- Information Generated by the Individual in His Work
- Information Generated by Dialogue in the Office
- Information Issued in Formal Office Reports
- Information Communicated by Correspondence
- Information Communicated by Phone
- Information Gathered from Visits and Visitors
- Information Collected by the Individual for Personal Use

For each topic of this report, a set of procedures for handling this information without the computer was designed and then a set of procedures by which augmenting tools could be used to handle it was outlined, the outline being structured to allow a comparison of the two.

3a4

Comparison was made on the basis of processes to be carried out, and on the differences in results to be expected under the best procedures now evident. For the computerbased operations, the processes were annotated with references and links to instruction manuals and other background material, so that the process description could serve as a procedure manual. This procedure manual appears as Appendix A of this report.

3a5

One important fact should be emphasized here. New tools not only make it possible to perform old tasks more efficiently, but as the

CAPTURING, STOPING, AND RETRIEVING RESEARCH INFORMATION

user gains experience, new uses and new methods are perceived, and the tools affect the nature of the tasks themselves. 3a6

The availability of online text-handling and text-communication tools has opened new possibilities in the ways the worker can use information. It is possible to foresee only a portion of the evolution which will take place in the handling of research information as new tools become familiar to the users. In the procedures described here it will be evident that some processes have already changed in nature so that comparison of old and new methods is difficult. Some old advantages or products may be lost temporarily in the changes which bring new advantages or products. This fact is not an argument for continuing old methods, but such losses should be noted so that they can be compensated for as the new systems evolve. On the other hand, the advantages of new tools are not as evident as the losses, because only as the worker gains experience with the whole of a new system, can he perceive its further potential and improve the tools to realize their full power. 3a7

SOURCES 3b

Information Contained in External Publications 3b1

The difficulties encountered in the receipt, announcement, circulation, indexing, and recovery of books, periodicals and reports which are of interest to an Office are too well-known to need reiteration. Even a small Office which lacks a formal library finds that control of such materials is necessary and is hard to accomplish.

In Appendix A, detailed procedures for setting up workable controls are designed, for both manual and machine systems. Comparison of the steps in each shows that the use of computer-based tools for handling this information changes the actual work performed, but often substitutes one time-consuming task for another. The degree of importance of such information to an Office will be a guide as to the amount of effort and the type of task which will fill its needs.

Information Generated by the Individual in His Work 3b2

Workers vary in their capacity for recording and transmitting the results of work in progress. Some are facilitated in their thinking by the process of formulating their thoughts for a record. Others feel diverted from productive activity when they must document their results. Some are publication-minded and publish and distribute even informal documents of their

CAPTURING, STORING, AND RETRIEVING RESEARCH INFORMATION

work. Others, while not secretive, will not think to do so in the absence of a formal policy or vehicle to accomplish this.

Tools which maximize the return from effort, by reducing the labor of transcribing and of editing, by creating a clean record to file, by making dissemination a normal procedure, or by providing a filing system which aids in retrieval, will tend to improve the quality and quantity of the record of individual contributions to knowledge.

Information Generated by Dialogue in the Office

3b3

The total effort of groups or teams is facilitated by awareness of their colleagues' advances, and by building on each others' ideas. In an Office much such information is transmitted orally.

Sometimes a fertile subject is pursued for a time and then pushed aside by more immediate goals. When the subject comes up again, a large amount of rethinking occurs which could be bypassed by a record of previous discussions. Capture, dissemination and retrieval of the cooperative thinking in an Office are desirable, both to stimulate current efforts and to eliminate redundant work. Application of computer-based tools to augment such communication appears to allow dialogue to be carried on in a medium which simultaneously transmits, records, and provides for retrieval.

Information Issued in Formal Office Reports

3b4

Requirements for formal reports vary in different Offices, and where reporting is done infrequently, the substance of each report may be in everyone's mind for a year or two at least. However, in many Offices it is the practice to issue reports each month, or on the completion of a short investigation or project, and the contents of these is less well remembered, as occurs after a period of time for even a small report output.

The preparation of formal reports to sponsors or customers or colleagues requires publication effort as well as the consumption of a considerable amount of creative effort by the workers. These efforts often seem to be diversions from the principal thrust of Office projects. Means of shortening the time and effort necessary by tools that aid the worker in the creation of the text, and means of augmenting the publication and distribution of the intellectual product are to be sought.

CAPTURING, STORING, AND RETRIEVING RESEARCH INFORMATION

Information Communicated by Correspondence

3b5

Correspondence provides a fruitful source for information which is pertinent in subject and timely in content. The degree to which the work of an Office is integrated with the work of other Offices or Communities and the nature of this integration affect the relative importance of correspondence as an information source and determine the appropriate efforts to be made for its control by indexing, storage and retrieval. Conventionally, correspondence files are kept in chronological order either according to the sender's name, when a continuing dialogue is carried on, or grouped by subject of the correspondence when letters in relation to a particular topic are received from numerous sources. Appendix B deals at length with points to be considered in controlling correspondence. The alternative procedures for offline and online correspondence systems are set forth in the Manual in Appendix A.

Information Communicated by Phone

3b6

Phone communication plays a large part in interoffice information transfer but provision is seldom made for its recording. Conversation tends to be informal, redundant, and even inconsequential. Where records are kept, these may be in the form of a chronological log, sometimes with notes on the topics dealt with and the recorder's impressions of any points made or agreements reached. When appropriate, correspondence about agreements may be generated.

The growing trend toward audio recording of phone calls may be acting to decrease the intimacy in what has been a characteristically direct means of interpersonal and interoffice communication. Phone communication may be increasingly inhibited by the knowledge that a record may be being created. Even more inhibiting would seem to be the possibility that the record can be manipulated to distort the intention of a communicator. The particular need in capturing the information from phone contacts is the capability for verification of the record. It seems desirable that the interactive nature of the communication itself be accompanied by the capability for creation of a real-time interactive record which can be private and also permanent. The concurrent use of phone and shared online terminal facilities by which a record can be created, agreed upon, recorded and protected offers an attractive potential.

CAPTURING, STORING, AND RETRIEVING RESEARCH INFORMATION

Information Gathered from Visits and Visitors

3b7

The purpose of personal visits between Offices is usually to gather or disseminate information about the activities of either Office for the benefit of the other. To meet this purpose it is important that the information transferred be made available to other people in the respective Offices. Sometimes a meeting is held for formal verbal transfer. It is common practice to prepare a trip report or visitor memo to record information transferred. These may be long or brief, and are often regarded as a formality, rather than as an information transfer tool. Trip reports particularly may suffer from delay in preparation. In general, the process of transfer of such information and the deposit of it are likely to be haphazard.

A practice of creating immediate, online records, which can be discussed with the visitor or Office visited, and disseminated discriminately but immediately, with the provision for retrieval by online search, offers possibilities which can be explored.

Information Collected by the Individual for Personal Use

3b8

Such information will be received from the same sources as those discussed in the preceding categories. The distinction is in the use intended. These files are those which an individual creates without considerations about communication to others. They may contain such items as appointment notes, random thoughts, links to documents he intends to read, and actions he intends to take. When these files are created in a computer-augmented system, the whole may resemble a combination of the user's desk file drawer, bulletin board and calendar pad. An important difference is that at present the system can store only text, and only text which has been key input, so that such files are likely to be less comprehensive than offline files and to consist largely of leads to other files, many of which may be accessible directly through online links.

The potential of computer-augmented personal files has only begun to be explored, and by only a few individuals. However, it is a prime use for which the system was envisioned. Individual thinking patterns and proclivities, and the freedom from protocol requirements, will make this one of the most creative uses of the system.

CAPTURING, STORING, AND RETRIEVING RESEARCH INFORMATION

CONCLUSIONS

3c

General

3c1

These conclusions are drawn from general information-handling experience, that of the author and that of others as reported in the extensive literature available, and from the results of using the procedures as developed in the Manual. (Appendix A). The Manual is placed in an Appendix for convenience, but a careful reading of the alternative procedures is important to an understanding of the author's conclusions here.

Use of computer-based tools for capturing, storing and retrieving research information has evident advantages for a Worker who is accustomed to receiving and sending information online and wishes to tie as many of his tasks into his online mode as possible.

There is no assurance that any saving in total effort involved in information capture, storage and retrieval in one Office will result from use of computer-based tools to handle information gathered by the Knowledge Worker. There are tradeoffs in human and machine effort, and when the system as designed does not work optimally, as when machine speed or capacity is less than planned, or machine features require new and unanticipated effort, total effort may be increased or result impaired. Use of the machine means dependence on the machine also.

It can be presumed that if other Offices can use the products (data bases) produced by an Office, that savings in effort will result. The extent of re-use of machine-disseminated and machine-stored dialogue needs to be studied, before cost-benefits can be established.

Because the possession of new tools changes the nature of the tasks to which they are applicable, it is important to look for benefits derived from performance of tasks in new ways, and from new tasks which were not attempted heretofore because tools were not adequate.

A communication system has benefits for inter-Office communication which are not evidenced in a local situation. Analogously, a telephone system installed for a single Office might not improve communication in that Office, but may facilitate communication with other Offices.

CAPTURING, STORING, AND RETRIEVING RESEARCH INFORMATION

Increased communication traffic may accompany use of computer-based tools, but before defining the increase in traffic as an equivalent increase in communication and before ascribing the increase to the use of particular tools, the full communication framework in the Office would need to be examined.

Increased communication traffic may have a large noise component, and may not represent a true gain in communicated information. The processes of computer-based information transfer need to be studied under controlled conditions and compared with the results achieved by use of face-to-face dialogue for ephemeral messages, and of typewriters and photocopiers and hardcopy files for distribution and retrieval of substantive material.

Many of the limiting factors in usefulness of gathered information are machine-independent, and result from the application of semantic and bibliographic efforts not yet demonstrably improved by machines.

Information Gathered from External Publications

3c2

No advantage is seen for a small Office to put online its records of the acquisition or subsequent storage or circulation of external documents. The small number of records, the storage and access provisions necessary, and the lack of any secondary use of these records pose no need for online records, and the offline procedures are clearly easier.

Online citations to external documents are advantageous for remote retrieval if continual updating is practiced; to keep offline records as current as it is possible to maintain online records would be exceedingly costly in effort and materials. The deciding factor in value of online catalogs is the extent of use of these citations by remote Offices. The online catalog of the NIC (ARPA Network Information Center) never developed to a state where its format, content or currency made it very useful. A time, cost, and use study could be made to establish the value of an online catalog in an Office. For the NIC, an ad hoc, intuitive decision was that it was not an important use of resources.

The national efforts now under way to provide improved public access to large federal bibliographic data bases may eventually make it possible for any Office to copy prepared citations from these bases and would then eliminate some of both offline and online work in the Office as described here. A new area of

CAPTURING, STORING, AND RETRIEVING RESEARCH INFORMATION

work in retrieving and manipulating the citations would develop.

Information Generated by the Individual in His Work

3c3

For the individual who has a terminal readily available, the recording of his thinking on a subject or problem may be facilitated by input of his thoughts in a file he can access, modify, build on, and print out clearly with or without the aid of a secretary.

It has been observed that an individual's thinking process is often stimulated when he is shown the clean text being produced and experiences the ease with which it can be modified.

The worker can have a file archived when his present use of it is ended or suspended, and by keeping an online description of such files, can have one retrieved from archive if or when he wants it back.

Upon completion of an analysis or study, the Worker can submit his file to the Journal, which serves not only to announce his efforts to his co-workers, but to document his contribution. In many Offices, no offline means of capturing this work is provided, although it could be.

Information Generated by Dialogue in the Office

3c4

Having created an online draft, the individual can send a message to a co-worker telling the location of the file and asking for comments. By making a copy of the file, the co-worker can insert comments, additions, or corrections, and let the author know where the copy is, to read or recopy. This can occur offline, of course, but the practice of this activity may be encouraged by the online Journal.

The practice of creating and distributing opinions, and requests for opinions, on a topic of interest is encouraged by the online Journal. Any matter of Office interest has been found to be discussed in Journal dialogue, to an extent which is unlikely by any other written means.

From the Journal it is possible to reconstruct the thinking that went on while an issue was being resolved or a design decision made.

CAPTURING, STOPPING, AND RETRIEVING RESEARCH INFORMATION

Information Issued in Formal Office Reports

3c5

Online input provides speed and flexibility in preparation of reports which are not possible with simple pencil and typewriter methods. It has an advantage even over mag tape typing in quicker feedback. Whether the deferred input method designed at ARC, called DEX, has advantages over all mag tape systems is not known.

The convenience with which major changes can be made in text and format is an aid in producing consistent, attractive report products. Use of COM (Computer Output Microfilm) for final output to print-ready film allows a great variety in typeface and format. One balancing factor at the present time is the experience needed to produce an acceptable product in COM; an inexperienced user can spend a great deal of time and money experimenting before the result is satisfactory. It can be assumed that any Office using COM will need at least one consultant knowledgeable in typographic design and proficient in use of the output devices.

Information Communicated by Correspondence

3c6

In an online environment, in which the recipient is reading his mail online, it is clearly an advantage for him to be able to make notes in the same medium. The use of citation numbers and links enables him to connect his notes with the originals, and to send an original, his notes, or an answer immediately to others online.

A carefully worked out system is required to duplicate the record-keeping capability of offline originals, on which notes can be made, and copies filed in as many contexts as desired.

Information Communicated by Phone

3c7

Using a headphone, the listener can make online notes during the conversation as easily as he could write them, and they are in a reusable form. If the other party is also online, by linking terminals the listener can show the notes to the other party for mutual agreement.

As a habit, the taking of notes as phone conversations occur would tend to result in more records, and more accurate recording.

An online log, standardizing online records, presents the same challenge in design as a correspondence log.

CAPTURING, STORING, AND RETRIEVING RESEARCH INFORMATION

Information Gathered from Visits and Visitors

3c8

Making notes online has the great advantage of getting instant verification of notes, if the visitor can see the terminal. It can facilitate the personal dialogue; if the visitor finds that a point is not noted, he can give more information to fill out the topic.

Taping a visit will make a more complete, and possibly less intrusive, record and can be used in conjunction with online note-taking when the topics are complex and detailed. Tapes can be transcribed and input as part of the document, or as separate linked documents. This is a fairly expensive process.

The online record of the visit need not be made during the visitor's stay, and can be made online by the person visited, or may be dictated and transcribed online or in DEX later.

The online log, which is independent of the method of input, has the advantages of being accessible by some or all of the others in the Office, and of being retrievable by the same means as are other online indexes.

Information Collected by the Individual for Personal Use

3c9

The principal drawbacks to the keeping of personal files of references that serve to organize records and to aid in locating them for individual use are the same for offline and online systems. Either means will only be as good as the design for the record-keeping, and most individuals don't have the expertise or the patience to set up and maintain workable record-keeping systems.

The added limitations currently imposed on storage size and program speed mean the user is sometimes reluctant to depend on online files for keys to his personal information. When these limitations are not present, the online collection of personal references will clearly serve the purposes of the user who works habitually in the online mode.

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

4a

INTRODUCTION

4a1

The Journal system is the basic tool that supports the dialogue seen to be necessary to a discipline- or mission-oriented Office or Community. Since its inauguration by ARC in June 1971 it has been enormously successful in stimulating and recording dialogue between members of the Office group that constitutes ARC and between members of the Community group represented by the ARPA Network.

4a1a

Ref: Journal User Guide (userguides,journal-guide,)

Capabilities provided in the ARC Journal Design are:

4a1b

- Online dialogue between individuals, including those geographically separate.
- Recording of the dialogue by serial numbers to allow unique reference.
- Online and offline retrievability of text of past dialogue.
- Online and offline subject and author retrievability.
- Creation of networks of documents connected by online links.

For the RINS project, analyses were made of the experimental use of the Journal by the ARC Office and by the ARPANET Community. These analyses do not do justice to the future use of the Journal, in that they do not indicate the eventual potential of the Journal system in an Office or Community of experienced users who take full advantage of the features. These studies were made to reflect the actual usage during a period of experimentation, and as such indicate some results to be expected in technology transfer to new Offices. They record interim problems encountered and point to possible solutions which can be considered as the system is evolved. An outline of the sections of the critique is as follows:

4a1c

4a2. How the Journal Serves an Office and a Community

4a3. Analysis of the Present ARC Journal as Used by ARC

This data represents use by an Office of experienced workers, who use many of the present Journal capabilities regularly.

4a4. Analysis of the Present Journal System as Used by the ARPANET Community

This data represents use by a Community of users, many of whom have received some training in the techniques but who vary greatly in their experience and expertise.

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

4a5. A Critical Look at the Journal Delivery Announcement

The format of the announcement which the Journal delivers to addressees is considered.

4a6. A Critical Look at Hardcopy Output Provisions

Online reading and reference to the Journal is the preferred mode for the Office, but because of present constraints on size of online storage, and because of the certainty that for a long time in the future it will be impossible to communicate online with the entire outside world, hardcopy printout is not only useful but necessary. The forms in which this is currently available were examined.

4a7. Analysis of the Output of Subject Indexing Programs

Because the success of retrieval by subject depends on the input as much if not more than on the indexing process, this aspect was analyzed here.

4a8. Modifications Deemed to be Desirable

From the analyses and critical examinations, a list of suggested modifications is given.

4a9. Recommendation for Use in a Workshop Office

This summary discusses the implementation of the Journal as a tool in a new environment.

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

HOW THE JOURNAL SERVES AN OFFICE AND A COMMUNITY

4a2

With the growth of online network communication, various mail systems have been designed by which messages can be sent between workers at the same or separate Offices. The most widely used mail system is SNDMSG, which quickly transmits messages to individuals recognized as users at their particular installation. While such mail systems are useful for transitory dialogue they lack universality and the capability for reference and retrieval by specific item. The Journal system at ARC was designed to satisfy the range of needs perceived for support of continuing dialogue. Within ARC, as a prototype Office, it is used in the online system NLS. In a Community, such as the ARPANET, it is used in the ARC online system, NLS, and in other systems through the Network. The Network Journal allows input by SNDMSG or other local mail system, and augments these with the special features for dissemination, reference, and retrieval which the Journal provides. The Network Journal was not in general use until after this study was made, and its use is not reported here. 4a2a

If the Journal provided no more than a running record of work and dialogue concerning the work, it would be an extremely useful tool. The capability of capturing and archiving a substantial portion of the thoughts and actions of a group of workers as they design a system is of great value, in their interaction and for later reference 4a2b

In addition to creating such a record, the Journal system allows the dissemination of questions, replies, ideas, and reports to selected recipients with economy of sender effort. A worker can send his messages with ease, assured that they will reach the addressees, whether an addressee is in the same room or across the country, whether the addressee is online or has only U.S. mail service, whether or not his address is known to the sender, whether the sender is personally acquainted with an addressee or is even unaware that the person is a member of a Community with similar interests. Mail is addressed to groups of people by using an identification for the group. This ident is entered in the Identfile (see Section 4b) which is used by the Journal to accomplish all distribution. For example, the Journal will distribute to individual members of the Network Graphics Group when the address "NGG" is given. 4a2c

After initial dissemination, the item can be referred to by a unique number, and even particular paragraphs can be referred to by a system of marginal subnumbers, such as those in this report, allowing easy reference in subsequent dialogue. 4a2d

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

The online recipient of Journal documents and messages is given the capability of storing the item or the announcement which links to the actual item. He can file items in any arrangement he likes, and can put copies of an announcement in several places if he wishes. 4a2e

The Journal system is designed to allow instant online retrieval of the text of the message when the number is known. The linking provision allows the Worker to follow paths of dialogue, as succeeding input includes file names that are linked online to earlier dialogue items. 4a2f

With content analysis programs, the item can be retrieved when the author or subject is correctly searched for. Online content searching through large sections of the recorded dialogue is not now practicable, but is expected to be a useful capability in the future. 4a2g

The Journal provides hardcopy that can be filed in any way desired. 4a2h

Index programs operate on the files to produce online or hardcopy indexes for general online or offline retrieval. 4a2i

The Journal system, then, in principle functions as a good information-handling system. Over the two years it has been in use by ARC, detailed design features have been given critical use and improvements have been made. The system has proved itself in user acceptance as a means of acquiring, capturing, disseminating, and retrieving research information in an Office, and is appropriate for use in other Offices. 4a2j

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

STUDY OF THE ARC JOURNAL AS USED BY ARC

4a3

The ARC Journal contains all items sent through it by ARC and ARPANET users. An analysis of the Journal for the months September through November 1973 showed approximately 1300 items sent. Of these about 50% were authored by non-ARC users. This analysis considered only the items authored by ARC members.

4a3a

Activity

4a3b

Specific attention was paid in the analysis to the nature of the dialogue carried on by ARC users between themselves, to look at the nature of the dialogue among workers in an Office. However, an initial count was made of the submissions of all kinds by the 31 ARC members during the 3-month period examined, which showed the following:

No. of Items Sent	No. of ARC People Sending this No.
None	2
2- 7	8
11-16	7
20-28	4
31-37	4
42-49	4
58	1
93	1

Addressees

4a3c

The Journal system allows addressing to individuals and to groups of individuals. Much of the dialogue on the Network is addressed to groups, but inside ARC most is addressed to specific individuals. Analysis of the addressees, reduced to percentages shows:

Addressees per Message	Percent Addressed to this No.
All of ARC	25% of Messages
10+ Addressees	8%
4 - 9	12%
3	12%
2	8%
1	35%

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

Length of Item

4a3d

As might be expected if the Journal is an informal mode of communication, the majority of items are less than a page long; about 45% were only a paragraph or two. Another 30% were one-page entries. Almost 25% were longer than two pages, these averaging about five pages each.

Subject Matter

4a3e

Types of subject which can be identified, and the percentage of the dialogue which they represent are:

Design suggestions, and responses in ARC	30%
Formal reports	15%
Notices of visits and other events	15%
ARC and SRI internal business	10%
Dialogue from ARC to Network users	20%
Correspondence from ARC to outsiders	5%

Dialogue

4a3f

The Journal can be seen upon casual reading to be a true dialogue. There are items journalized for their record value which are not internally linked to other items. However, most messages refer to previously discussed subjects, and many items which are direct responses to earlier items use the capability of the online link to make access to references easier for the reader. An analysis of the links actually inserted in the messages shows the use made by ARC of the linking capability. About 30% of the items sent by ARC members to each other contained online links to other documents. This percent may be expected to increase as users become more familiar with the feature and as more inter-related documents are entered into the Journal.

The link capability (Userguides,tnls-address,entry:w) allows the user to insert a reference in the parenthetical form shown in this sentence, and this insertion, or link, can be addressed and will retrieve the file or particular section of the file denoted.

Links occasionally fail in their immediate-retrieval function. One reason is that storage capacity is too limited at this time to keep all old files online, and the link may lead only to a message that the file is archived, in which case the user must wait a

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

short time for the file to be retrieved by an operator. Another reason is that the link format is used by some to refer to documents which were never online, for which only a reference has been cataloged, in which case the link can lead only to a coded citation, and currently it is not able to do even this. A further reason is occasional carelessness of the author in typing a link, so that the number is wrong and leads to the wrong document or nowhere. In this case augmentation is again helpful, in that the author can try his links to assure they are correct, if he will think to do so.

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

ANALYSIS OF THE PRESENT JOURNAL SYSTEM AS USED BY THE ARPANET COMMUNITY 4a4

Introduction

4a4a

The Journal system has been in use by a small subset of the ARPANET Community for over two years. To provide a basis for any conclusions drawn, and to serve as an indicator of aspects to be considered, various appropriate samples of the Number and Author indexes to the NIC Journal were analyzed to obtain data on the use of the Journal and to critique the use made of the Journal. A representative page of the Author Index is shown on the next page.

Samples Studied

4a4b

Number Index

About 300 NIC items were retrieved by the catalog program from those items entered in two month period 1 July to 1 September 1973, items between 17593 and 18796. These 300 items were analyzed and compared with the entire collection of Journal hardcopy in binders for this period, with the following results:

Content of Item -- 44 were trials or handshakes which could be deleted for lack of substance; for most of these the hardcopy had not been filed. (See Conclusions 1). 33 were transmittal letters which could be better handled in a correspondence log than in the Journal.

Content of Citation -- 63 other than the above 44 were untitled, and therefore not usefully informative here, nor subject-retrievable in the Titleword Index.

Content of Index -- 60 items in the hardcopy Journal binders, items sent by NIC to Network people not RADC, were not in the NIC Journal index. These omissions were the result of input manipulation by the Journal, not the result of indexing programs.

Author Index

The Author Index gives the title, date and number of all items sent by each author. Because of limitations in the amount of information which can be fitted on one line, it does not give addressees, which would often be useful.

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

The Author Index was analyzed to determine the extent of use by the Community represented by the ARPANET. It was analyzed for data on use by individuals over the entire period of the Journal's existence, with attention given to the frequency and recency of use by individual authors.

Users

Number of Users -- 170 non-NIC people used the Journal at least once, in the period June 1971 to August 1973, often during a TNLS course.

Number of Continuing Users -- The pattern observed is shown in the next table.

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

ANALYSIS OF CONTINUING USE OF THE JOURNAL BY ARPANET USERS

The following table indicates patterns in continuing use by members of the ARPANET Community. It appears that a number of users tried the Journal only experimentally, sending only 1 to 5 messages and their last input was some time ago. Others, named in the table, have sent continuing dialogue over the Network and apparently find it very useful. (See observations in following section).

Breakdown by number of messages, and last date of input:

Total	Users With	Date of Last Input by Specific User				
Msgs.	This Total	12/71	12/72	4/73	6/73	8/73

1	44	1	12	13	10	8
2	35	2	9	5	13	7
3	25	2	1	9	7	7
4	15		1	1	5	8
5	7		1	2	2	2
6	5	(Sandum, Lindamood, Bruffey, Bouknight, Cornell)				
7	6	(Mitchell, Kelley, Kahn, Fikes, Cutler, Cosell)				
8	3	(Strollo, Licklider, Kampe)				
9	2	(Day, Stern)				
10	2	(Walden, Rockweg)				
11	1	(Bressler)				
12	3	(Barden, Karp, Silberski)				
14	1	(Metcalfe)				
15	1	(Lieberman)				
19	1	(Krilanovich)				
21	3	(Cotton, Dolan -12/72, O'Sullivan)				
22	1	(Masinter)				
23	1	(Stoughton)				
28	2	(Kline, Padlipsky)				
29	1	(Forman)				
34	1	(Bhushan)				
41	1	(Levin)				
43	1	(Owen)				
51	1	(Pickens)				
58	1	(Cerf)				
72	1	(Iseli)				
96	1	(Deutsch)				
131	1	(Neigus)				
162	1	(McKenzie)				
176	1	(Postel)				
256	1	(Dave Crocker)				

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

Observations on the Data

4a4c

1. About 15% of the items in the NIC Journal Index are not substantive, and another 11% are routine transmittals. These add to the bulk of the Index without contributing much to its usefulness.
2. About 20% of the items are entered, carelessly or purposely, without titles, making their retrieval by subject impossible.
3. About 20% of the existing Network Journal dialogue is not being captured by the NIC Journal Index. This occurs because all messages originating at Sites other than RADC are automatically coded for NIC in the subcollection field, but messages from NIC to Sites are only coded as NIC at the instigation of the sender, and the sender does not know how to or does not remember to take the action to do this.
4. Only 44 non-ARC people had used the Journal for more five messages. Yet some non-ARC people use it heavily and continuously.

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

A CRITICAL LOOK AT JOURNAL DELIVERY ANNOUNCEMENT

4a5

The following is a sample of Journal delivery as seen by the worker
in his initial file:

4a5a

PR 11-DEC-73 10:15 20845
Bibliography on the Economics of Information
Location: (MJOURNAL, 20845, 1:w)

Comments: This bibliography was given to me by Mike Rothkopf.
He told me that I can publish it in our Journal System.

PR 11-DEC-73 09:22 20844
An Interesting Weekend Seminar
Location: (MJOURNAL, 20844, 1:w)

MLK 10-DEC-73 16:52 20835
Hark! ARC!
Location: (MJOURNAL, 20835, 1:w)

Comments: Please read this important announcement!
JCN 9-DEC-73 15:21 20819
Note to RADC Users: Transfer of AKW Online Services to OFFICE-1
Location: (MJOURNAL, 20819, 1:w)

Several improvements have been suggested for the format of these
items. They include:

4a5b

Give the full name of the author(s).
Incorporate the third line into the first line.
Remove the empty line between the item and any comments, so that
the relation of the comment is seen to be to the item above.
Place the title line first, so that a one-line view is more
informative.

A reworking of the preceding sample according to these suggestions is
shown on the next page.

4a5c

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

Bibliography on the Economics of Information

Paul Rech 11-DEC-73 10:15 (J20845,)

Comments: This bibliography was given to me by Mike
Rothkopf. He told me that I can publish it in our Journal.

An Interesting Weekend Seminar

Paul Rech 11-DEC-73 09:22 (J20844,)

Hark! ARC!

Marcia L. Keeney 10-DEC-73 16:52 (J20855,)

Comments: Please read this important announcement!

Note to RADC Users: Transfer of AKW Online Services to OFFICE-1

James C. Norton 9-DEC-73 15:21 (J20819,)

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

A CRITICAL LOOK AT HARDCOPY OUTPUT PROVISIONS

4a6

Hardcopy is available to the user in four forms:

4a6a

- Output Quickprint
- Output Device Teletype
- Output Device Printer
- Hardcopy Journal Delivery

Evaluations of the adequacies of these forms are:

4a6b

Output Quickprint gives the user a copy of the item with all information as to sender, date, and other recipients printed on the terminal or on the printer. This is the preferred mode for informal dialogue and is eminently satisfactory.

Output Device Teletype prints the item on his terminal using all formatting directives supplied by the Journal and those supplied by the user which are not overridden by the Journal. This mode is required for output of formal, formatted items such as reports. Its continuing drawback is the frequent conflict between format directives supplied by the author and those imposed by the Journal, so that highly formatted tables and text designed to nonstandard line width or page length are deformed by the Journal formatting program. Intimate familiarity by the author with provisions for circumventing these problems is required to prevent unpleasant surprises in the output.

Output Device Printer provides high-speed printer output using the formatting directives as above.

When the user's ident information marks him as a recipient for hardcopy, he receives under present conditions a copy printed either on the ARC printer or through use of an IBM 360 and an off-site printer, and this copy includes an outside page containing the address for mailing. This form is used for hardcopy mail to non-ARC recipients. It currently has the drawback that it entails a delay of from two days to a week, caused by the processes it must undergo, of tape conversion, IBM 360 processing, printing, bursting and collating. These delays are not intrinsic to the Journal system, but result from a decision not to overload the Office printer with this bulk of output. A drawback which is intrinsic is the format program which spreads the information over three or four sheets: the first sheet containing the text with a heading consisting of the ident of the sender and the date and Journal number, a second sheet giving the full name of the author and full names of all addressees, and a

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

third sheet formatted to give the full name and address of the specific addressee so that it can be folded and mailed. A fourth sheet is generated if any comments on the message are included by the sender. To keep the complete record all sheets must be retained. For filing, this presents a problem of bulk of paper and of staples used to attach them.

Despite the criticism detailed above, it should be appreciated that no other system has the variety of distribution modes of the Journal.

4a6c

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

ANALYSIS OF THE OUTPUT OF SUBJECT INDEXING PROGRAMS

4a7

Retrieval by Subject from the Journal

4a7a

Three means are provided for retrieval of items by subject:

Keywords can be supplied by the person Journalizing.

There is no prompt for this action, and few people use this facility. At present no programs are run to produce general indexes on sender-supplied keywords.

Subcollections can be specified by the person Journalizing.

The Journal by default categorizes input into certain subcollections according to the sender's address (e.g. SRI-ARC, NIC, RADC) and according to any groups addressed (e.g., SRI-ARC, SUR, NCG, PRG, NLS). Alternatively, the sender can specify existing or invented subcollections, which allows a sender to categorize his input according to subjects named as subcollections. Index programs are run for the large subcollections, to provide ARC, NIC, and RADC indexes, and the same capability exists to retrieve any subcollection, but is seldom used at this time.

Titlewords are used as subject terms for Journal indexes run at the times that Author and Number Indexes are produced.

Since the Titleword Index is the principal means for subject retrieval at present, an analysis was made of the output in the ARC Titleword Index dated 6 November 1973, which covered the input of August through October 1973.

Analysis of ARC Journal Titleword Index for August-October 1973.

4a7b

An incremental issue of the Titleword Index of the ARC Journal was analyzed to determine its makeup, to draw any conclusions as to its benefits, and to look for ways in which it could be improved. In the reporting of results, these terms are used:

Items, i.e., messages and files submitted

Terms, i.e., words recognized by index programs

Citations, i.e., all one-line references to items as retrieved by terms

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

Citations appearing in the Index

Number of Journal items submitted: 645
Number of keyword citations generated: 2620
Citations per item, average: 4

Terms appearing in the Index

Number of terms: 1099
Number of citations generated: 2620
Citations per term, average: 2.4

Terms correlated to citations

Personal last names and idents, as terms: 91
Citations under personal names and idents: 143
Citations per name or ident, median: 1

Proper names of organizations and groups, as terms: 41
" " " " " " citations: 131
Proper names of systems, equipment, etc., as terms: 27
" " " " " " citations: 119
Citations per proper name, average: 3.7

Subject terms other than names, excluding plurals: 360
Subject terms occurring as plurals: 70
Citations under subject terms: 1233
Citations per term, average: 3.4

Misspelled words appearing as terms: 16
Nonsignificant terms, remainder: 510
Citations under nonsignificant terms: 994

Proportion of potentially useful to nonsignificant citations

Citations by personal and proper names: 393
Citations by significant subject: 1233
Citations by nonsignificant or erroneous terms: 994
Percentage of content which is nonsignificant: 37%

Proportion of potentially useful to nonsignificant terms

Terms for personal and proper names: 159
Terms for subjects: 360
Terms, nonsignificant and erroneous: 580
Percentage of terms which are nonsignificant: 53%

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

Benefits and Drawbacks to the Use of Keyword-from-Title Indexes. 4a7c

Benefits:

The indexing process is mechanized so that no judgment, i.e., human labor, is required.

Once programs are written and a file exists, new input can be processed at little added cost.

Drawbacks:

Retrieval effectiveness depends on how well the title reflects the information content of the item. If a key term is not in the title, the item is not retrievable from that aspect.

No cross-references are produced to remind the searcher of other forms of the word, or of synonymous or related terms under which he would find citations he would want. For example: "Online" and "Line" appear, but not "On-Line" nor "On Line" because "On" is suppressed.

Without editing, the resultant index will contain sizable proportions of citations under terms that no one would look for.

In addition to the generic benefits and drawbacks above, there are specific considerations in the use of titleword indexing programs as currently used in indexing the Journal data base. These are:

A title is sometimes omitted by the sender, and the item is not retrievable. Usually, but not always, the lack of a title occurs on items which carry little or no information likely to be desired for retrieval, where there is no substantial loss.

Many items in the Journal are informally written and unedited, and titles for these tend to be informal, even flippant, and to contain unpredictable, nonsignificant, and misspelled words which introduce garbage into the index. The sample page of the NIC catalog will show the extent of this occurrence.

The length of the titleword program output for even a few months' content is currently too great to keep in one file, so the program must be run on files which are arranged by date, then the output, which is arranged alphabetically, must be merged with passes over three files covering A-E, and so forth. At present, incremental indexes cannot be merged with existing

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

basic indexes, because there is no way to suppress the repetition of marginal words at proper places in the printout. The result is that the mass of old raw data must be reprocessed with the new increment for a new index. This reprocessing becomes proportionately greater with each new index.

A shortcoming of all the Journal indexes is the lack of a format program to produce an online index. At present, the indexing program format is oversize. It is designed for production of a printed index suitable for reduction, so as to save on quantity in reproduction and yet include a maximum amount of information in one line. This output is used also for online viewing, and the wraparound is confusing. For online viewing, a citation formatted for two lines is needed.

Conclusions Regarding Titleword Index

4a7d

Only half the terms currently generated by Titleword Index programs on the ARC Journal are likely to be useful for retrieval.

About two-thirds of the citations generated are under terms which may be useful for retrieval. No check was made of the contents of the Journal items to find whether they were as useful as the Index made them appear. It is observable that there are random nonsignificant occurrences of words which could not be suppressed because they would be significant in other contexts. Words in the latter class are: "Help", "Query", "Message", "Text".

A conclusion difficult to document, but easy to observe, is the use of Author and Number Indexes to retrieve by scanning, where the Titleword Index fails, and where a good subject index would better serve the purpose.

There is no evidence that users cannot eventually find items they remember. A study would be needed to establish what is found and what is missed. Studies made on other data bases have shown keyword indexes to be far inferior to human-edited subject indexes, and a study of the adequacy of this retrieval tool could point to the need for more sophisticated programs or for human editing to fill out the record for complete retrieval capability.

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

MODIFICATIONS DEEMED TO BE DESIRABLE

4a8

The following modifications are summarized briefly, based on the foregoing discussion and conclusions.

4a8a

Modifications to Input Provisions

4a8b

1. The Journal needs better automatic prompts and feedback for input, specifically a reminder for all needed items. Input procedures should include a reminder, if not a requirement, that a title be supplied. Lack of a title makes retrieval so difficult as to be impractical and usually not attempted. Other items desirable for input, but which the system does not demand, are special subcollections, keywords, and comments.
2. NIC as a subcollection entry should be made automatic for all items logged from NIC to Sites over the ARC utility, unless the design for the utility makes subcollection NIC a redundancy. In case automatic entry is the option, it should be echoed and allow overriding. The user is often not aware of the provisions for and the present results of subcollection entry, and subcollection indexes therefore do not contain all items they could be expected to.
3. A correspondence log system should be implemented to work in conjunction with the Journal, to provide desirable links to offline dialogue of similar content and to remove routine transmittals from a record of substantive dialogue.
4. Numbers should always be assigned and fed back to the user at the time of input, so that the sender can use them for referral immediately. The number can be obtained on request, but should not have to be specially requested.
5. Provision should be made for two dates to be recorded: the date of writing and the date of Journalizing, because these are often not the same. Older files which are Journalized at a much later date should so indicate.

Modifications to Delivery

4a8c

6. A temporary note, designated as such, should be placed in the author's initial file at the time of submission.
7. The format of the delivery notice should be improved, perhaps as suggested above.

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

Modifications to Hardcopy Provisions

4a8d

8. Journal formatting programs should be more easily overridden. An item submitted to the Journal should appear in the exact format the author designed, with the Journal information supplied in the upper and right-hand margins outside the formatted areas. At present it is not possible to set arbitrary statement numbers to conform to an outline, with the result that an Introduction becomes statement 1, and Section 1 becomes statement 2 and leads to confusion.

9. Hardcopy for mailing should be redesigned to reduce the unnecessary bulk and to present all information about sender, recipients, comments, and the first page of the message.

10. Two formats of hardcopy output should be available and presented to the user as options.

A formal document such as a report would be preceded by a one-page transmittal sheet:

To: Name(s) of recipient(s)

DATE NUMBER

From: Sender

Title:

Comments:

A informal file or message such as a memo would be supplied with a heading of the above, preceding the file or message.

The journalizer could be given the format options:

REPORT format (separate transmittal letter)

MEMO format (To, From, heading)

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

Modifications to Subject Indexing Programs and Procedures

4a8e

11. Index programs need a slightly longer stoplist, perhaps 40 or more terms, to suppress the most common useless terms. Beyond these, the anticipation of useless terms is so difficult for this text as to be impractical.

12. Because the input will always be uncontrolled, and because the Journal files themselves should be unmodifiable, the output from index programs will always contain useless and redundant and incomplete entries. Therefore, the output from index programs should be human-edited to remove and insert material which will increase the usefulness of the indexes.

13. A format for online viewing of the finished index needs to be designed.

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

RECOMMENDATION FOR USE IN A WORKSHOP OFFICE

4a9

The Journal system is essential to any augmented Office. As conceptually designed it is an ideal vehicle to support the dialogue in research information. As currently implemented it has some characteristics which detract from its complete satisfactoriness and these can be remedied.

4a9a

Additional user aids (scenarios, improved prompts, more frequent and more readable indexes) are all needed to make use of the Journal easier and more useful to the worker, and with these aids the Journal can gain wider acceptance.

4a9b

Several features are known to be needed which are not now implemented, principally because of resource limitations:

4a9c

1. Capability of annotating or commenting upon existing Journal items in situ. It is not now possible to add comments in the logical place, i.e., following the Journal item commented on.
2. References in an existing document to later documents which may modify or update it. While Journalized items themselves should not be modifiable, there should exist the capability to append such information. This feature, referred to as back-linking, is needed to alert the reader to the continuing dialogue.
3. Facilities for privacy. The Journal can be browsed by anyone, making the contents of all communications public to others than the intended recipients. Elemental privacy facilities are to be implemented very soon.
4. Secondary distribution, the sending of previously Journalized items to addressees, does not allow comments, which are often desired to tell why such distribution was made.
5. A Boolean search capability is needed to make retrieval fully satisfactory.
6. Better interface with SNDMSG is needed. SNDMSG messages could be automatically translated into NLS for receipt with the rest of Journal mail. Also, the user should be facilitated in entering SNDMSG items he has received into the Journal.
7. The sender should be able to specify whether his submission is to be delivered in its entirety or as a citation only.

JOURNAL SYSTEM

Critique of the Journal and Its Indexes

8. The new or occasional recipient of citations could have available a standard set of instruction for retrieval of the full text.

9. The Journal should allow immediate delivery when desired, as SNDMSG does, so that the present delay of a half hour or more can be bypassed.

The Journal is a tool whose full potential will be realized only as Workers become so comfortable with it that they use it continually. With increased use will come demands for further Journal features not now implemented and surely some not now even envisioned. 4a9d

IDENT SYSTEM

Critique of the Ident System and Its Uses

4b

INTRODUCTION

4b1

The Identification system is basic to the use of any online tools in the Workshop. It allows the online system to recognize the user so that he can build and access files, and is fundamental to the Journal system which supports online dialogue. 4b1a

Ref: Journal Users' Guide (userguides,journal-guide,2f) and
(userguides,journal-guide,4)

This critique of the Identification System deals with:

4b1b

User Views

The Ident System as seen by the user who seeks ident information on himself or others

The Ident System as seen by the user who desires to enter ident information on himself

The Ident System as seen by the user who desires to enter ident information on others so he can address them

Maintenance of the Ident File by an Information Center

Input of new idents

Modification of existing idents

Directory Programs Using the Ident File

Present programs

Program changes needed

New programs needed

After detailing specific difficulties, a list of modifications deemed to be desirable is given: 4b1c

User access procedure changes

User input procedure changes

Directory program changes

IDENT SYSTEM

Critique of the Ident System and Its Uses

FEATURES OF THE IDENTIFICATION SYSTEM

4b2

User Views

4b2a

The Ident System as seen by the user who seeks ident information on himself or others:

The facts are readily available to him if he knows the ident. Typing a "?" brings him information on querying for names and namestrings. If the user doesn't know the Ident, he will usually fall into a trap from which he cannot extricate himself. The response of the system is often slow, and comes in spurts which make the user uncertain that the response is completed.

The Ident System as seen by the user who desires to enter ident information on himself:

A prospective user can get his Ident entered only by someone else; guests cannot enter idents.

For a user, entering his address properly is difficult or impossible if he does not know the Site acronym already. [See Mod 1]. Entering other information such as phone and comments is not open to him until after the Ident system recognizes him. He can then enter as himself and and modify his record, but this is not straightforward.

The Ident System as seen by the user who desires to enter ident information on others so he can address them:

If he fails to spell the name of the other properly in the inquiry stage above, and concludes the other is not in the file, and proceeds to enter a name, he will have erroneously made a second, incorrect entry for an individual. [See Mod 3]

If he does not know the acronym for the individual's site, he will find it difficult or impossible to discover it, because the listing he is shown is not in alpha order, and he will then create a new, erroneous address. [See Mod 1]

Maintenance of the Ident File by an Information Center

4b2b

Input of new idents

New ident information is entered at the NIC by a NIC agent from handwritten messages from phone calls, and from Journal

IDENT SYSTEM

Critique of the Ident System and Its Uses

messages and U.S. Mail. There is no verification step for the user. Through clerical error at some step, an entry may be wrong for some time before it is caught. [See Mod 5]

Modification of existing idents

Individuals are able to modify their own idents without notice to NIC, leaving the offline mailing lists unchanged. worse, this allows the entry of dirty data which will interfere with the programs used for Directories. Provision for a branch into which all changed idents would be placed has not been implemented since it was agreed upon several months ago. [See Mod 4]

Directory Programs Using the Ident File

4b2c

Present programs

Programs now available format the output for a large page to be reduced. It is necessary for the printed directory that some such format be produced. It is important for the online version that a normal page be produced, and for Group lists this is accomplished by manual input. [See Mod 6, Mod 7]

Program changes needed

A distinction between Hosts and other organizations needs to be made, and different types of indexes made for the two. [See Mod 7]

A program revision is needed to provide a display or tty page to be produced through output processor. [See Mod 6, Mod 7]

New programs needed

For Hosts, a new program is needed to format information which is not pertinent to non-Host organizations.

There is currently no way to generate a list of Groups to which an individual belongs, to aid in changing the offline mailing lists of Groups when an individual member changes his address. [See Mod 2]

IDENT SYSTEM

Critique of the Ident System and Its Uses

MODIFICATIONS DEEMED TO BE DESIRABLE

4b3

User access procedure changes

4b3a

Mod 1. Alpha order list of organizations. Names of organizations should be entered in alphabetic order, rather than in site acronym order, and should be offered to the enterer in alphabetic blocks.

Mod 2. Group membership indication. Status of each individual should show the groups of which he is a member.

User input procedure changes

4b3b

Mod 3. Editorial assistance. When a user attempts to enter a name, a bell on a reference terminal should signal, and a human editor should come on, to assist in entry. The editor should be able to do content analysis online, to determine the status of individual and organization names.

Mod 4. No modification of the file by the user. Notice of desired modification should be sent to the NIC; modification should not be allowed by the individual, even of his own ident information. (At present, the individual is not allowed to modify any but his own ident information, but the message regarding this occurs after he has attempted a change, such as address or phone, and should occur at the time he types the ident he wishes to modify.)

Mod 5. User verification. Whether entered by online cooperation between user and agent, or entered by the agent from a note or message, the entry should be verified by the user. An online message requesting an ok, or a message mailed to the user should conclude the entry procedure.

Directory program changes

4b3c

Mod 6. Format for Individuals Section.

The information could be formatted so that the single column would be seen online, and the full page could be constructed with COM or by pasteup of the printout.

Mod 7. Format for Organizations Section.

The information could be formatted so that the single column would be seen online, and the full page could be constructed with COM or by pasteup of the printout.

IDENT SYSTEM

Critique of the Ident System and Its Uses

DISCUSSION

4b4

In mechanization, often an early design is produced with the expectation that one process or file can be created to serve multiple uses. In many cases this proves a disappointment, in that the different needs cannot be met by exactly the same design. Finding this to be the case, modifications or branching processes are added to accommodate the different needs, and the final design may not be as satisfactory for any of the purposes as a completely separate design for each.

4b4a

The Ident file has gone through some changes caused by the desire to produce a file and process for multiple purposes. It is still in evolution, lack of resources having hindered designers from implementing some of the features they see to be desirable, and it remains to be seen how further modifications will remedy the present difficulties.

4b4b

The Identfile was conceived for four purposes: to allow system identification of users of NLS, to allow rapid sending and receiving of online Journal mail, to print U.S. mail addresses for offline delivery of Journal mail, and as added value, to serve as a base for programs to produce directories of individuals and groups. At first, the conflict between the necessary provisions for each of these uses was not recognized.

4b4c

To meet the need for rapid sending of online mail it was necessary to allow free input of addresses by the sender, and the first release of the Ident system encouraged the user to submit ident information for himself and those whom he wished to address for whom he did not find existing idents. This resulted in hurried and incomplete address information without consistency for addresses at the same Host, and for duplicate entries for an individual where an inaccurate search failed to turn up an existing ident. For the online mail system the presence of this garbage was not critically detrimental.

4b4d

There is a basic incompatibility between free input and the consistency necessary for use of a data base for accurate and formatted retrieval. Gradually, some controls over input have been added, and the necessity of further restriction has become recognized, leading to reprogramming and more formal input control.

4b4e

The deficiencies in conceptual design, and the degrees to which these have been dealt with, are expanded here.

4b4f

IDENT SYSTEM

Critique of the Ident System and Its Uses

Faulty and incomplete entries

4b4g

It was soon apparent that the ordinary user could not cope with the intricacies of entry procedures, and beyond that, cannot be depended on to enter even his own address or phone number correctly. The experience of the Identfile overseer at ARC was that not one entry "made by someone on the NET has not needed to be corrected in some way" (12770,) to make it accurate or satisfactory for the hardcopy Directory.

The first control added was provision that ident's added by Network users (which were available for use for online mailing immediately and achieved their purpose to the degree that the input was acceptable) were sent also to a branch called NEWIDS, where an ARC member could examine and correct them as needed before they turned up as problems for the U.S. mail or for directory programs.

It appears probable that the most satisfactory means of adding to or changing an Identfile will be to have input controlled by a central operator who is available by phone at all times a change might be needed.

Handwork necessary to make cross-references

4b4h

Until recently, the entering of an ident for an individual did not achieve the relation of his ident to his organization and this relation was done separately by ARC, with some resultant discrepancies. At the present time, membership in groups is not an element in the entry for an individual. Of more consequence, U.S. mail to groups is accomplished through a separate, punched card system for producing mailing labels, and the records for these can be corrected only by examining listing of membership for all groups to determine where an individual is a member and resubmitting the information for a new card and new labels, with resultant work, delay, and possibility of error.

Size of the file

4b4i

The use of one file for entry of all individuals of which anyone wants to make a record, including spurious or practice entries, has created a file of over 1100 entries by this time. Manipulation of this file becomes increasingly more difficult for programs.

Use of the file for NLS identification is slowed by the presence of the many entries and quantities of material unnecessary for this purpose.

IDENT SYSTEM

Critique of the Ident System and Its Uses

Use of the file for online delivery is hampered by the presence of the material unnecessary for this purpose.

In concept, the existence of one file for all person identification purposes is fine, but in practice, duplicate subfiles would be more usable.

Lack of protection

4b4j

Group coordinators can change group membership, without notice to an editor for verification or for mailing list purposes.

Spurious and practice entries can be made by anyone, and require tedious hand deletion.

UnVerified material and unmanageable entries mess up the results of directory programs each time they are run.

Random arrangement

4b4k

Straightforward inquiry status for a valid ident proceeds smoothly, but entering the Identfile for retrieval by incomplete information is so time-consuming as to be impractical. In the operator's userguide (14288,) are instructions for creating a sorted file for editing purposes, but this is not practical under most operating conditions.

Obsolete idents

4b4l

Carrying deadwood slows down daily processing.

Purging the file of obsolete idents must be done by a computer operator upon special request of the Identfile overseer. Purging of obsolete idents of individuals who have written Journal mail causes author indexing programs to fail.

Here again, multiple purposes of the file cause conflict in the use for any one purpose.

IDENT SYSTEM

Critique of the Ident System and Its Uses

RECOMMENDATIONS FOR USE IN A WORKSHOP OFFICE

4b5

An online Identfile is a necessity for use of an online Journal system for intra-office communication, and for online communication between the office and other online offices using a common online Journal system. For such uses, it is important only that the system be reliable and accurate.

4b5a

The following conditions limit its present usefulness:

4b5b

The file will rapidly grow to contain hundreds of entries, and require more time to use than an offline file serving the same purposes.

The effort of building such a file is greater than is warranted by the predictable use of individual entries it contains.

The practice of using initials as ids is awkward because these sets of initials are duplicated and triplicated. The present accommodation for this is to add numerals 2, 3, and so forth. Such ids are confusing to users, and seem demeaning to the owners. Some other systems use initials and last names.

Expertise and close familiarity is required to input and access the file, making it a task for a particular person rather than a general tool.

At present, no capability exists to quickly transfer a retrieval address to a document of correspondence being created. A user program is required to place the address in the address area of a letter being written. The process of retrieval and copying is too unwieldy and time-consuming to be a practical way of creating correspondence.

Hardcopy correspondence created from online input is not yet formatted acceptably. The output is rough in appearance, spreads the information over several sheets, and lacks an appropriate letter heading.

In concept, the Identfile is a tool with important, needed, capabilities. A workshop has need of:

Finding the full name, address, and phone number of an individual previously contacted.

Finding the address of an organization to whom correspondence is to be sent.

IDENT SYSTEM

Critique of the Ident System and Its Uses

In addition, an augmented workshop has use for:

Automatically supplying the full address in creating correspondence with outside people and organizations.

In further addition, an augmented Workshop has use for:

Automatically sending correspondence online to other organizations.

These needs the present Identfile is conceived to meet.

Difficulties in its current use result particularly from:

Lack of recognition of the importance of controlling sources, quantity, and format of input.

Lack of good format for output to hardcopy.

Constrictions on its purging and changing procedures by need of its use for Journal file retrieval.

As soon as resources are available, it seems important to develop the Ident system to a level more consistent with its potential.

USER PROGRAMS

Critique of the User Program Capability

4c

INTRODUCTION

4c1

As a user gains experience with NLS, he often finds that he has repeated need for a particular series of actions which he could perform with less effort if he combined them into a small program. The need that prompted the writing is often a need common to others, and users are encouraged not only to write these programs, but to place them in a file called (user-progs,-contents,) to make them available to others. ARC also has written a guide to use of these programs (user-progs,-userguide,). It appeared productive to investigate the extent of use of such programs and of their documentation, and a questionnaire was designed and given to 15 ARC members at random to provide information on program use. The questionnaire is given here, with cumulated results from the questioning. The directory of programs is given also.

4c1a

USER PROGRAMS QUESTIONNAIRE

4c2

1. What User Programs Have You Used, and How Often?

4c2a

Respondent	Programs used	Times per month
(analyst)	append inrun	unsatisfactory not online
(analyst)	inmes seqgraph	30 4
(secy)	letter	8
(sys progr)	append	very seldom
(sys progr)	changed filter (other personal)	1
(sys progr)	trace (many others)	30
(sys progr)	letter trace	1 a lot, in bursts
(analyst)	inmes	30
(analyst)	oeldir inmes	seldom
(t writer)	address	5

USER PROGRAMS

Critique of the User Program Capability

	addtext	4
	append	4
	deldir	2
	inseqh	6
	jform2	20
	notabs	3
	sortnocase	1
(operator)	--	
(programmer)	append	1
	letter	1
	trace	2
(user progr)	(many special)	often
(user progr)	addtext	6
	append	4
	deldir	6
	format	12
	inmes	2
	showdir	20
	toc	4
(t writer)	addname	uses another
	addtext	3
	deldir	2
	format	8
	letter	1
	showdir	3

3. How Did You Learn to Use These?

4c2b

(Numbers are quantities of respondents, out of 15 total.)

6 learned from demonstration only

5 from a combination of demonstration and documentation

4 wrote the programs themselves

0 learned from documentation alone

4. Have You Read the Documentation?

4c2c

4 reported there was no documentation on ones they used

USER PROGRAMS

Critique of the User Program Capability

4 did not indicate knowledge of location of documentation

3 had placed instructions in their initial files

3 indicated approach through <user-guides> or <documentation>

2 named the correct directory <user-progs>

5. Have You Shown Anyone Else How to Use a User Program?

4c2d

10 have shown others how to use a User Program

6. Have You Written Any User Programs?

4c2e

8 have written no User Programs themselves

Most programmers had written one or more for specific purposes.

About half of these programs were written for general use and were included in the <user-progs> directory. About half were undocumented and not submitted to the directory.

7. Did You Write Them Alone or With Help?

4c2f

2 reported having written a User Program without assistance

8. Comments and Opinions

4c2g

"Can be a real help when you have a repeated process to do."
(analyst)

"Found the documentation impossible to implement; when was shown how to use a commonly-referred-to program, found the demonstrator had difficulty and the documentation was agreed to be inaccurate."
(analyst)

"Not sure whether it is a time saver but feel LETTER is a hopeful feature for future use when idents are known to the system."
(secretary)

"Currently good for professional programmers only since L10 knowledge is needed; also poor because there is no documentation on procedures available; minimal diagnostics and error checking."
(systems programmer)

"Do a lot of work with the use of User Programs, testing part of the system and in research on additions or modifications, simulation of changes." (systems programmer)

USER PROGRAMS

Critique of the User Program Capability

"Don't really know what's available" (systems programmer)

"Think they're great" (analyst)

"Can rarely run them in the daytime; many don't work satisfactorily. There is no clear mechanism to use to get the program written and maintained." (analyst)

"Have found them very helpful and greatly improved with the REL file (to keep a compiled file around)." (programmer)

"Fantastic potential; augmented accessing would be nice; L10 should be getting more usable, as userguide, sysqd, etc. develops; some programs aren't used for a long time, then are absolutely necessary; the use of the system of user programs depends on having the right program at the right time, archiving them has discouraged people in the past; I have got repeated requests to make all user programs permanently available." (user programmer)

"It's more useful to write one for a particular purpose than to use the general ones, however the format ones are really great." (technical writer)

DIRECTORY OF USER PROGRAMS

4c3

PROGRAM	FUNCTION
addname	Adds name to nameless stmts from first word in stmt
address	Asks for ident, inserts the address at the bug
addtext	Adds text to front/back of stmts in plx/brnch/grp/st
append	Sequentially appends stmts in group, text between
appendlist	Like APPEND, but leaves substructure
changed	Marks statements changed since a given date
delcol	Deletes bugged column, assuming next col lined up
deldir	Deletes Output Processor directives
delname	Deletes statement names
delsp	Deletes leading spaces from statements

USER PROGRAMS

Critique of the User Program Capability

format	Add print directives to a file
index	Creates a word index for st/br/plex/group
inmes	Inputs all of message.txt file into NLS file
inseqh	Does a sophisticated Input Sequential file
inrun	Inserts TENEX RUNOFF file into NLS file
jform1	Reformats journal references
jform2	Reformats journal references
letter	Puts file in letter form, adds dear-- & sincerely--
lowercase	recovers from an erroneous XSET UPPER CASE
makeref	Scans for journal links and makes ref branch
notabs	Replace tab keys by spaces in plex
showdir	Shows only stmnts with Output Processor directives
sortnmskp	Sort key extractor: as usual but disregards st names
sortnocase	Sort key extractor: alphabetic regardless of case
sortnum	Sort key extractor: sorts by first number in stmt
sortrev	Sort key extractor: exactly the reverse of usual
sriform	Puts in O/P dirs and spaces to SRI standard format
sublist	Does substitutions on list of files given links
tblpts	Adds periods to end of st out to given col
toc	Generates Table of Contents with stmt num refs
trace	NLS call return tracing system for micro-analysis
truncate	Truncates st/br/plx/grp to one line-assume 3/lev ind
wordcount	Counts visibles in st/br/group/plex

4c3aq

USER PROGRAMS

Critique of the User Program Capability

CONCLUSIONS

4c4

The capability for the user to write small programs for repeated use is a potentially important one. The concept of sharing these programs with others to assist them or to prevent their having to write their own is clearly in line with the principle of resource-sharing and with the whole concept of dialogue support in a Group or Community. 4c4a

The results of this survey of half the members, selected randomly, of a fairly experienced Group showed: 4c4b

Even a program meeting a common situation, such as "inmes" to process SNDMSG files for online reading and retention, and "append" for turning sequential files into NLS files, was not used by more than one third of the Group. The existence of a program for a common situation doesn't result in its widespread use. 4c4b1

Most users did not access the documentation in learning or in reminding themselves of the programs they used; demonstration is the present learning mode. It may be assumed that the documentation alone is inadequate for training inexperienced users. 4c4b2

Half of the Group had written programs themselves, but almost all had had assistance in doing so. It appears that writing of user programs is not a common habit, even for experienced users. 4c4b3

Enthusiasm for present user programs was reported by 6 of the users, 4 responders were critical of present shortcomings, almost all appeared to appreciate the potential of such programs. 4c4b4

A factor not clearly brought out in the questionnaire is the difficulty of invoking user programs. The user must give separate commands to get the program, set buffer size, and run the program, and sometimes additional commands are needed. The possible steps needed to employ user programs are numerous and hard for the novice to execute successfully. 4c4c

USER PROGRAMS

Critique of the User Program Capability

RECOMMENDATIONS FOR USE IN A WORKSHOP OFFICE

4c5

A single command should load and run a user program.

4c5a

Existing user programs and the procedures for using them should be documented in explanations and scenarios written for the non-programmer. Then they should be publicized to bring their existence and usefulness to the attention of the Group.

4c5b

Those people with programming ability should be encouraged to write programs for general needs, and to have them described for use by all who might want them.

4c5c

Every means should be explored to make the writing of user programs, as well as their use, easy for any Office member.

4c5d

PERSONAL FILES

Critique of Their Use in an Office

4d

INTRODUCTION

4d1

Any person working with information feels a constant need to organize the information he expects to use so that any piece of it can be recalled when it would be useful.

4d1a

To achieve this recall, most individuals create some kind of ordered collection in which to store items for future use. Usually, the person assembles a collection of books, papers, notes, and correspondence, and as it grows larger, designs a filing system, simple or complex, for arranging these items where memory will remind him to look. As the filing system becomes unwieldy, he makes an index to it. Usually, he finds the system less than satisfactory, and redesigns it to better fit what he now sees he needs. If it is complex, he finds input tedious and output inadequate. He would really like an automatic system by which he would be relieved of the labor of marking, recording, and filing, and by which he could usually locate items he vaguely remembers. Using a computer, which is more orderly and less fallible than a human brain, is an attractive possibility.

4d1b

Various systems have been designed to provide a framework within which a user can create his personal collection of references from which he expects to retrieve what he needs. No system is currently widely used, and few are known to be praised by other than their designers. A bibliography of items on personal file making is appended.

4d1c

The three components of a file system are the body of information to be stored and retrieved, the tags or keys by which this information is identified, and the means by which these tags are manipulated. A system will be satisfactory to the extent that these components are well-designed and the links between them are strong. The information stored should be the complete information, not citations for which the original information must be located in less controlled storage. The keys and tags, such as subject terms and names, should be reflective of the stored information. The means of manipulating the keys and tags should be flexible, available, and understood by the user. Introducing numbers into the system to make the keys and tags more manipulable or to simplify filing of the stored information must be seen as placing additional links in the retrieval chain, and their usefulness must be examined to be sure it outweighs the potential handicap they present. Numbers used as surrogates seldom carry any intrinsic information, and need to be used invisibly, as in computer manipulation, or used in conjunction with information-bearing items.

4d1d

As dialogue grows, the the individual feels a need to find ways to sort and store references to this dialogue to keep track of the

PERSONAL FILES

Critique of Their Use in an Office

information he has received and to flag some items for action. He is prompted to create personal ways of storing these references. 4d1e

In (17394,), Outline of a Personal Information Management System, Paul Rech at ARC has defined the problem and suggested the features of a system using the ARC online system, NLS. Features include: provision for general categorization of items, provision for personal categorization of items, flagging by subject and date for purposes of retrieval and deletion, and automatic deletion of older items not accessed. 4d1f

NLS is well-suited to the creation of such reference files by its provision for links between files and by its provision for hierarchical grouping of categories of references. There exists also the capability for identifying and deleting materials which are unused for a specified period. 4d1g

To investigate the experience of ARC users in creating and using such files a survey was made of 15 randomly-selected members, using a questionnaire. The tallied answers are given, with annotations. 4d1h

PERSONAL FILES

Critique of Their Use in an Office

PERSONAL FILES QUESTIONNAIRE

4d2

The same 15 ARC users surveyed for their use of User Programs in the previous critique were questioned about their use of personal files. (Numbers shown in these tallies are quantities of correspondents.) 4d2a

1. What files do you maintain for your personal record of: 4d2b

a. Journal items received and sent?

All those questioned receive Journal citations online in their initial files. The questions relate to their disposition of these citations.

How are these files organized?

7 leave in form and order in which Journal system transmitted them.

6 rearrange items by topics; 5 by subject, 1 by action needed.

2 delete all Journal items after reading.

How often are these updated?

4 who reorganize their Journal item files update them daily, the others less often.

How often are these used for reference?

6 seldom or never use the Journal item citations saved.

2 who reorganize their items report they refer to them less than weekly, but find the files extremely helpful.

7 refer to Journal citation files occasionally, probably less than weekly.

b. Notes of schedules to be met, actions to be taken, actions taken?

9 keep a branch in their initial file for tracking actions.

How are these files organized?

1 uses an elaborate scheme for his total activities, and

PERSONAL FILES

Critique of Their Use in an Office

updates it daily. One person started a detailed scheme for recording status of work-in-progress, but does not keep it current.

2 keep in random arrangement.

Methods of organization range from "to do" and "done" branches, to highly categorized and task-related branches.

How often are these updated?

2 update theirs daily; others updated irregularly.

How often are these used for reference?

3 reported more than daily reference to these files.

c. Reports of visitors?

4 make a record of visitors.

How are these files organized?

2 keep a file with a branch for each visitor.

1 keeps in branch of initial file.

1 records in Journal only

How often are these updated?

Secretary reported a file to be filled in by others, and hence updated by them.

How often are these used for reference?

All 4 reported these were rarely used for reference, one stated was for record only.

d. Notes on personal reading, bibliographic references?

4 people keep such files, 1 keeps two files.

How are these files organized?

1 keeps a branch in initial file

PERSONAL FILES

Critique of Their Use in an Office

3 keep separate files, by subject.

How often are these updated?

Approximately weekly.

How often are these used for reference?

Approximately weekly.

e. Other personal record files?

9 keep other personal files.

How are these files organized?

3 keep idea files

2 keep files of sendmessages.

1 keeps file of user programs, 1 keeps record of relevant files created, 1 keeps phone directory.

How often are these updated?

Input to these files was fairly steady, deletion was not studied.

How often are these used for reference?

Varied according to type of content; most seemed to be active.

2. Have you started such files and abandoned them?

4d2c

10 reported having started personal files and abandoned them, of these, 2 did not now maintain personal files of citations or ideas.

Reasons for abandoning them?

Reasons given:

"too much of a problem with scarce system resources"

"too much work to keep up; system limited availability"

"unavailability of system"

"cumbersome"

PERSONAL FILES

Critique of Their Use in an Office

"partly because of lack of time to maintain and partly because of inconvenience of trying to keep track of on and off line files"

"too much trouble to keep up"

"too much trouble"

"terminal time is too valuable to spend on maintaining online files of personal information"

"most if not all other schemes have been traded off for new elaborate scheme (abandoned old schemes for new one)"

"lack of usefulness"

"change in interest, archive process, etc."

3. What do you do with Journal references in your initial file, if you do not reorganize them or move them?

4d2d

All respondents deleted these references occasionally, some deleted nonsignificant items as they were read.

PERSONAL FILES

Critique of Their Use in an Office

ANALYSIS OF THE RESULTS OF THE QUESTIONNAIRE

4d3

Extent of Use Made of Capability

4d3a

Less than half of the sample group, 6 of 15, utilize the citations received from the Journal for other purposes than announcement. Some who move or reformat their Journal citations report they seldom refer to them later.

Over half, 9 of 15, maintain a branch in their initial file for keeping track of personal work and schedules.

The secretary and 3 others of the 15 surveyed recorded some information about visitors.

Only 4 people reported keeping online references to articles, books, or documents read.

Reasons Given for Lack of Use

4d3b

The most common reason given, by 5 of 12 respondents, for having dropped previous efforts in personal file maintenance was that the work was too much trouble. It is the common experience that creating personal files of information is not worth the effort, in the benefit derived from the time-cost. In no case was it indicated that any other feature of online use had made personal files unnecessary; conversely, the files of user programs and phone directories were created where they would seem to be unnecessary in view of the directories of user programs and the Identfile.

The second most common reason given, by 4 out of 12, was the limitations imposed by the system, particularly the uncertainty that files would be online when desired. This is presumed to be a temporary hindrance.

Several explanations for disinclination to keep online references to materials were explored in follow-up discussions with the responders, although no tallies were made:

They find no added value in keeping references or notes online, when they do not lead to online files.

Literature is not of much importance in their work.

Lack of space, coupled with a low priority placed on such files, discourages preparation of such files.

PERSONAL FILES

Critique of Their Use in an Office

RECOMMENDATIONS FOR USE OF PERSONAL FILE SYSTEMS IN AN OFFICE

4d4

The capability to prepare and maintain personal files of information, both that received online and information input by the worker for his own memory aid, is clearly an important one. That few workers seem to be using the capability to its full power seems due in part to present system limitations. Lack of use also seems to be the result of lack of personal inclination and to difficulties in constructing adequate retrieval systems. It is also possible that the need for files of personal information may not exist for most people in this environment, doing the kind of work they do.

4d4a

To create a climate in which the full usefulness of the personal file capability would be realized, it would be possible to:

4d4b

Suggest formats for such files

Prepare thesauri to suggest subject classes for the interests of the Office

Make lists of bibliographic citations available

Give clerical support to maintain the files according to instructions from the worker.

PERSONAL FILES

Critique of Their Use in an Office

BIBLIOGRAPHY ON PERSONAL FILE BUILDING

- Bannur, B. B., and G. M. Purandare. Personal documentation: a study. *Annals of Library Science and Documentation*, vol. 14, no. 4, December 1967. p.182-186.
- Bennett, John L. The user interface in interactive systems. In: *Annual Review of Information Science and Technology*, vol. 7, 1972. p.159-196.
Personal user-programs, p.173+
- Burton, Hilary D. Personal information systems; implications for libraries. *Special Libraries*, vol. 64, no. 1, January 1973. p.7-11.
- Burton, Hilary D. and Theodor B. Yerke. FAMULUS: a computer-based system for augmenting personal documentation efforts. In: *ASIS Proceedings*, vol. 6, 1969. p.53-56.
- Engelbart, Douglas C. et al. Advanced Intellect-Augmentation Techniques. SRI Project 7079. Stanford Research Institute, Menlo Park, California. July 1970. 198p. (Final Rpt. NAS 1-7897).
- Glantz, Richard S. SHOEBOX - a personal file handling system for textual data. In: *AFIPS Fall Joint Computer Conference*, 1970. *Proceedings*, vol. 37. p.535-545 (ED 041 612). Same title: MITRE Corp., Bedford, Mass. Rpt. MTP-116. 1970. 32p.
- Hersey, David R., Willis R. Foster, Ernest W. Stalder, William T. Carlson. Comparison of online retrieval using free text words and scientist indexing. *ASIS Proceedings*, vol. 7, 1970. p.265-267.
- Holm, Bart E. *How to Manage Your Information*. Reinhold Book Corporation, 1968. 292p.
- Jahoda, Gerald. *Information Storage and Retrieval Systems for Individual Researchers*. Wiley-Interscience. 1970. 135p. Cat. No. 9670.
- Jahoda, G., Ronald D. Hutchins, and Robert R. Galford. Characteristics and use of personal indexes maintained by scientists and engineers in one university. *American Documentation*, vol. 17, April 1966. p.71-75.
See annotation, *Annual Review*, vol. 2. p.9-10.

PERSONAL FILES

Critique of Their Use in an Office

- Jahoda, G., Ronald D. Hutchins, and D. M. Miller. Analysis of case histories of personal index use. In: ADI Proceedings, 1966. Adrienne Press, Woodland Hills, California. p.245-254.
- Lancaster, F. Wilfred. Searching natural language data bases. In: Vocabulary Control for Information Retrieval. Information Resources Press. 1972. p.135-152.
- Linn, William E. Jr., and Walter Reitman. Referential communication in AUTONOTE, a personal information retrieval system. In: ACM Annual Conference Proceedings, 1971. p.67-81.
- McAllister, Caryl. A Study and Model of Machine-Like Indexing Behavior of Human Indexers. IBM Advanced Systems Development Division, Los Gatos, California. Lab. rpt. 16,205, Nov. 1971. 139p.
- McEwen, J. R. PAL: A Design for a Personal Automated Library. University of Pennsylvania, 1969. 264p. Cat. No. 4920.
- Mittman, Benjamin, and Gilbert Krulee. Development of a remote information management system--RIMS. In: ASIS Proceedings, vol. 6, 1969. p.199-206.
- Rech, Paul. Toward a Design of an Information Management System. ARC Journal 17394. 21 June 1973. 3p.
- Sabine, J. E. and W. J. Dane. Do-it-yourself indexing. New Jersey Librarian, vol. 3, Spring 1970. p.36-39.
- Siegmán, R. M. Information Systems in Universities. Georgia Institute of Technology, School of Information Science, Atlanta. 1969. 81p.
- Stenfel, L. E. Sequential adaptation of retrieval systems based on user inputs. Information Storage and Retrieval, vol. 7, August 1971. p.69-78.
- Vallee, J., J. Hyneck, G. Ray and P. Wolf. The organization of research data banks: experience with DIRAC-based information systems. In: ASIS Annual Meeting Proceedings, vol. 8, 1971. p.387-394.
- Williams, J. H. BROWSER: An Automatic Indexing On-Line Text Retrieval System. IBM Federal Systems Division, Gaithersburg, Md., 1969. 30p. NTIS AD 693 143.

RETRIEVAL OF INFORMATION BY THE KNOWLEDGE WORKER IN HIS OFFICE

4e

INTRODUCTION

4e1

The Augmented Knowledge Workshop facilitates the capture and storage of information from all sources, as seen in Section 3.

Capture methods are designed to interface with the particular sources: external publications, the human mind, Office dialogue, formal reports, correspondence, phone, and visits.

Storage methods are designed to place items in manageable units and modules, based generally on the type of item, and having distinctive data elements:

Online storage

Journal, for individual records, for dialogue, internal reports, phone and visit records, all in full text

Document catalog for citations to external documents

Correspondence log for citations to incoming correspondence

Hardcopy storage

Journal hardcopy, in binders

External reports and articles, in vertical files

Books, in author arrangement on shelves

Correspondence, in vertical files

Each piece of input is tagged with numbers supplied, individually or in batches, from a sequence of unique numbers.

Thus the input follows branching paths, easily selected for the source or format, into categories, each of which is consistent within itself.

These procedures clearly suit the convenience of an information center in accomplishing capture and storage. And they were designed to facilitate the production of various indexes for retrieval and to assist in retrieval of identified hardcopy.

RETRIEVAL OF INFORMATION BY THE KNOWLEDGE WORKER IN HIS OFFICE

However, they will be satisfactory for a Workshop to the degree that they can be integrated into a system that meets the total retrieval needs of the worker as he functions in his Office. In this section, the retrieval needs of the worker are examined and the means by which they can be met are discussed.

RETRIEVAL OF INFORMATION BY THE KNOWLEDGE WORKER IN HIS OFFICE

IDENTIFYING THE RETRIEVAL NEEDS OF THE WORKER

4e2

As the worker logs in at his terminal, he receives citations for his recent mail. If he gets Network delivery of his mail, he receives citations in EXEC for mail that has been sent him. He also receives SNDMSG items that he would like to transfer out of his message file into separate files.

NEED 1. Means to retrieve full text of files sent him as mail in EXEC.

In his initial file he receives citations for mail sent him through the Online Journal.

NEED 2. Means to retrieve full text of mail received in NLS.

He would like to save the citation for later reference, but he wants to keep the mail branch of his initial file cleaned up.

NEED 3. Means to classify and store useful references to Journal mail for later reference.

In the body of a Journal file he is reading, there are references to other numbered items, with links that give him the expectation they are online.

NEED 4. Means to retrieve items by number, without other identification of the item.

He remembers that he has seen other dialogue on the subject of the item he is currently reading, but does not know the file numbers of the dialogue items.

NEED 5. Means to retrieve items by subject from the Journal dialogue.

He wishes to reply, and to keep the online letter in sight for reference, and perhaps to look at other files for reference.

NEED 6. Means to hold one file in view for reference while creating another.

While looking at a document, he needs to know what other documents reference it, update it, supersede it, or make it obsolete.

RETRIEVAL OF INFORMATION BY THE KNOWLEDGE WORKER IN HIS OFFICE

NEED 7. Means of placing links in an existing document when new information relevant to the document is created.

He recalls a letter sent to him which he would like to see again.

NEED 8. Means to retrieve offline correspondence.

He is ready to answer a letter he received in the mail, and wants to refer to the notes he made about it when he received it.

NEED 9. Means to tie the record of incoming mail to notes made about it, to retrieve both at once.

He remembers having seen a report issued by National Science Foundation which contained data he would like to quote, and he would like to get hold of the document.

NEED 10. Means to locate an actual document which exists only in hardcopy, knowing only the issuer and the general subject.

He wishes to cite the NSF document in a report he is writing.

NEED 11. Means to retrieve a proper bibliographic citation for external documents.

RETRIEVAL OF INFORMATION BY THE KNOWLEDGE WORKER IN HIS OFFICE

HOW THE RETRIEVAL NEEDS OF THE WORKER ARE MET

4e3

NEED 1. Means to retrieve full text of files sent him as mail in EXEC.

Provision: Using FTP according to the scenario in Appendix D, the recipient can create a copy of a Journal file received in his Network mail which can then be printed out on the line printer.

Provision: By the user program INMES, the recipient can convert a SNDMSG file into an NLS file which can be read online or printed out.

NEED 2. Means to retrieve full text of mail received in NLS.

Provision: The recipient can jump to the file referenced by the link given in the citation, and read or print the file.

Provision: He can give the command to print Journal mail, and get formatted printout of all files in the Journal mail branch of his initial file.

NEED 3. Means to classify and store useful references to Journal mail for later reference.

Provision: He can selectively move the citations to a different branch of his initial file or to another file and can use the sort program or place them where he likes. He can add his comments about them. Later he can retrieve from the stored file by content analysis.

NEED 4. Means to retrieve items by number, without other identification of the item.

Provision: Online, the Worker gives the command to jump to the link which, if the item is currently online, brings him the complete text of the file, for reading or printing.

Note: If, due to lack of storage, the file has been deleted, and resides on tape, he can interrogate archive and ask to have the computer operator read it in for him, after a short delay. If he then finds it is not of interest, this is wasted effort.

RETRIEVAL OF INFORMATION BY THE KNOWLEDGE WORKER IN HIS OFFICE

Note: Finding a number in link format, i.e., "(12345,)", does not guarantee that the full text of the item referred to was ever online. It may be a reference to a hardcopy document which was cataloged, in which case the number appears as part of a statement name "A12345" in a catalog file of coded input and essentially is not retrievable. It may be a number selected for a document that has not yet been journalized, in which case only the number system knows to whom the number was assigned. With the institution of a correspondence log, it may be a number assigned to hardcopy correspondence.

Provision: He can assume the number is for a Journal item and can refer to the hardcopy of Journal items in binders.

Provision: He can load the online number indexes to the Journal, to get author, title and date. If he finds the citation, this may tell him whether he wants to pursue the item.

Provision: He can search by statement name in the document indexes, if they are online, to discover whether the number was assigned to an external document.

Note: Statement names of external document citations are the number prefixed by "A".

Added Provision Needed: When presented with a number in link format, the worker should be able to query the system with one command, to find out what the number refers to.

NEED 5. Means to retrieve items by subject from the Journal dialogue.

Provision: The user can access an online copy of the Titleword Index run on the recent content of the Journal. He can search by asking for likely terms as statement names, or he can do a content search on the titles for words or for word stems. From the citations found he can select numbers of items he would like to retrieve, and he can load the files or print them out.

RETRIEVAL OF INFORMATION BY THE KNOWLEDGE WORKER IN HIS OFFICE

Note: He can only retrieve items which contain the subject term in the title.

Provision: He can run a program to retrieve by keywords inserted by the sender.

Note: At present, few keywords are input.

Provision: He can run a program to retrieve by subcollection, when the subcollection capability is utilized to form a set of items by subject.

Note: No practice exists at present to facilitate use of this capability except by the individual user; he may make subject groups of his own input, but input by others will be put into such a subcollection only by mutual agreement.

Added Provision Needed: A thesaurus of selected keywords should be prepared, for the user to refer to in assigning keywords that will be most useful for retrieval. This can be constructed, with thought and care, based on the keywords used by national data bases and by local catalogers in coding the external documents, and taking into consideration the titlewords found in Journal dialogue.

Added Provision Needed: Catalog files for external documents need to be accessible as a class, just as Journal links lead to Journal items across directories.

Added Provision Needed: All files which are indicated to be online should be available online, without needing to be retrieved from archive.

NEED 6. Means to hold one file in view for reference while creating another.

Provision: Using the "split screen" capability, he can divide his screen into two or more parts, and may load a file for reference while he creates another file, or sends a message or file to the Journal.

NEED 7. Means of placing links in an existing document when new information relevant to the document is created.

RETRIEVAL OF INFORMATION BY THE KNOWLEDGE WORKER IN HIS OFFICE

Added Provision Needed: There is currently no provision for adding such important information to a document which has been Journalized.

NEED 8. Means to retrieve offline correspondence.

Provision: Online correspondence logs, for the individual and for the Office, will contain citations, including abstracts, to correspondence received only in hardcopy. The user can retrieve from these logs by content search, or by running a content analysis program to filter the citations for items likely to meet his desire. From numbers in the citations, the user can go to his own or the Office hardcopy correspondence files and retrieve the items.

Note: Content analysis may be slow on these files, as on other large files.

Provision: Indexes to Office correspondence can be run, and kept online.

NEED 9. Means to tie the record of incoming mail to notes made about it, to retrieve both at once.

Provision: In both the Office correspondence log and the individual correspondence log, comments may be added as substatements.

Note: It will be necessary to include the number of the item in the comment, because content retrieval programs at present will retrieve only a single statement, and the substatement will be individually retrieved.

Added Provision Needed: For retrieval from catalog and correspondence files, and possibly other files of two-level formats, programs are needed that will search by content from the second-level and retrieve both levels, or retrieve items such as statement name from the upper level.

NEED 10. Means to locate an actual document which exists only in hardcopy, knowing only the issuer and the general subject.

RETRIEVAL OF INFORMATION BY THE KNOWLEDGE WORKER IN HIS OFFICE

Provision: Online Titleword Indexes may help, if the subject term is in the title.

Provision: If the topic was picked up in the abstract, or as a keyword, a content search of the online catalog files containing the full citation with abstracts and keywords will find the citation.

Note: Online catalog files in coded input form are difficult to read, and their whereabouts is not generally known; they are not public files. Online files of full citations in formatted form are not now prepared, and are too long to keep online when storage is limited. This provision is therefore only theoretically useful at this time.

Added Provision Needed: Indexes by issuing agency are needed, based on input of agency name according to a standard form, so that USC, for example, does not occur under four or five forms of the name under "U", "University", "Southern", and "Department".

NEED 11. Means to retrieve a proper bibliographic citation for external documents.

Provision: For documents that the Office has under bibliographic control, there can exist files of citations in standard format.

Note: Citations as formatted for the online indexes are not adequate for the purpose.

Provision: Online access to large national bibliographic data bases can provide a source of precise citations to use in referencing external documents and periodicals.

Note: Capability to access these data bases is still a separate operation, but efforts are being made to tie these in to an Office utility.

Added Provision Needed: Programs are needed to reformat coded bibliographic input into citations suitable for lists of references, in addition to the

RETRIEVAL OF INFORMATION BY THE KNOWLEDGE WORKER IN HIS OFFICE

programs now existing which format for indexes
and for full catalog-card format.

Expansion of index concepts included in the Needs above.

It may appear initially attractive to merge citations to all forms of material into a combined index for the classes of item handled by the user: online files, offline hardcopy of documents, and offline correspondence. Making separate indexes for these categories is of benefit to the user in that:

The information to be retrieved for each of these differs enough so that the content of the citation should differ. For example, the addressee of a letter is important, the recipient of a document is not significant. Programs to extract the appropriate elements will run better if the base on which they operate is consistently formatted. Also, the output will be more readily understood if the citations are consistent.

The expectation of the searcher is different for different categories. He does not expect to find an overall discussion of a subject in a letter as he might in a document.

Documents and books may be treated as public knowledge, Journal dialogue and correspondence may not be assumed to be public.

Because files have practical limits, and must be divided in some way, division by category is more significant than an arbitrary division based only on length of file.

From an online index to the Journal, the online text can be retrieved; an online index to correspondence or documents is a dead end as far as online retrieval is concerned. Mixing references that lead to online retrieval with those that do not is not as helpful to the searcher as separating them.

Rather than integrate citations for various forms into one index, the form of the material should be clear from the number used as address, and separate indexes should be produced. Combined indexes can always be produced from merging the separate indexes if and when this is desired.

RETRIEVAL OF INFORMATION BY THE KNOWLEDGE WORKER IN HIS OFFICE

DISCUSSION OF ADDED PROVISIONS NEEDED

4e4

Added Provision Needed: When presented with a number in link format, the Worker should be able to query the system with one command, to find out what the number refers to.

Numbers work well as addresses, but they have no intrinsic information value. When the Worker commands a jump to link, he needs to be presented with the file for which this is the address, or he needs to be given information to allow him otherwise to find the item. This is particularly necessary if the numbered item is not in online form at all. The following are suggested provisions to meet the need.

All numbers could be given a letter prefix to indicate their form and location. As a number beginning with "A" refers to a hardcopy document which has been cataloged, Journal items should be prefixed with "J", correspondence with "C", and so on as any other categories are set up. "M" would be used for "Microfiche" when that form is machine-retrievable.

If a file has been archived, as for lack of space, the user should receive a citation to it as a response. This system should respond "Journal file" if that is what it is and give him the full information needed to retrieve it.

If a number refers to an external document, he should receive the response "Catalog file" and a citation to it from the Catalog files which indicates whether it exists in hardcopy only.

If a number refers to incoming correspondence, the message should say "Correspondence" and give an appropriate citation.

Added Provision Needed: A thesaurus of selected keywords should be prepared, for the user to refer to in assigning keywords that will be most useful for retrieval. This can be constructed, with thought and care, based on the keywords used by national data bases and by local catalogers in coding the external documents, and taking into consideration the titlewords found in Journal dialogue.

Random user-selected keywords have never proved satisfactory for good retrieval. Control of vocabulary, so that input and retrieval can be well-matched, can be achieved only by

RETRIEVAL OF INFORMATION BY THE KNOWLEDGE WORKER IN HIS OFFICE

preparing a standard thesaurus and making it easily available. The basis for the construction of such a thesaurus on subjects related to networking exists in the citations in Catalog files. A thesaurus for other Office and Community interests could be built as the files of information accumulate.

Added Provision Needed: Catalog files for external documents need to be accessible as a class, just as Journal links lead to Journal items across directories.

Although Journal items reside in separate directories, when a link is given, the directory need not be specified. To achieve good retrieval of citations, a similar capability is needed; a link referring to a citation should lead to the citation wherever it resides.

Added Provision Needed: All files that are indicated to be online should be available online, without needing to be retrieved from archive.

This will be possible only with increased storage capability.

Added Provision Needed: For retrieval from catalog and correspondence files, and possibly other files of two-level formats, programs are needed that will search by content from the second-level and retrieve both levels, or retrieve items such as statement name from the upper level.

Citations for documents would be more useful if the abstract could be inserted as a substatement. And provision could be made for insertion of other specified information in even lower statements. Then programs are needed to extract the upper-level statement, the basic citation, at the time the content search locates a match in the second-level statement.

Added Provision Needed: Indexes by issuing agency are needed, based on input of agency name according to a standard form, so that USC, for example, does not occur under four or five forms of the name under "U", "University", "Southern", and "Department".

A list of standard agency names needs to be prepared, so that bibliographic input can follow the standard name and so that index programs run on the Catalog, Journal, and Correspondence data bases can put related items together.

RETRIEVAL OF INFORMATION BY THE KNOWLEDGE WORKER IN HIS OFFICE

Eventually, to reduce redundancy in input, an Identfile-like table might be built to allow conversion of agency idents to output of full agency name.

Added Provision Needed: Programs are needed to reformat coded bibliographic input into citations suitable for lists of references, in addition to the programs now existing that format for indexes and for full catalog-card format.

RETRIEVAL OF INFORMATION BY THE KNOWLEDGE WORKER IN HIS OFFICE

RECOMMENDATIONS FOR USE IN A WORKSHOP OFFICE

4e5

The online retrieval aids available to the Knowledge Workshop make retrieval of online dialogue a capability of great value and even greater potential. At present, there are missing components which when supplied will make the Office Worker's terminal his point of retrieval for all information which has come to his attention. Added provisions are needed as described in this section.

EXPERIMENTAL DEVELOPMENT OF A
SMALL COMPUTER-AUGMENTED INFORMATION SYSTEM

SRI-ARC 31 DEC 73 21453
Annual Report to ONR

Appendix A

Procedures for Capturing, Storing, and Retrieving
Research Information

Non-computer-based Procedures

INFORMATION GATHERED FROM EXTERNAL PUBLICATIONS

ORDERING AND RECEIVING PUBLICATIONS

301

----- periodical Subscriptions

301a

Premise:

301a1

Subscriptions will be placed for the Office, from a list selected by the staff. Weekly and daily periodicals and newspapers need not be checked in as received; periodicals of less frequency will be checked in.

Procedures:

301a2

1. Compile a list of titles currently received by the Office or by individual staff members.
2. From a subject listing of periodical titles, select additional titles which would appear to be of interest in the work of the Office.
3. Circulate the list to all members, requesting them to indicate titles they a) desire to have available in the Office, b) desire to have routed to them.

Note: If the Office is part of an organization with a strong library, this service may not need to be an independent operation; the library may already have arrangements for browsing or routing.

4. Make a card file of subscriptions placed, with the names of those who requested and of those to whom issues are to be routed. These cards should have boxes for entering the dates of issues as received.
5. When an issue arrives, mark the subscription card with the issue number and the date received.
6. Stamp the issue with a stamp indicating it is the property of the Office.
7. Attach a routing slip to any issue of a periodical for which routing has been requested. No circulation card is necessary.
8. Forward to cataloger.

Computer-based Procedures

INFORMATION GATHERED FROM EXTERNAL PUBLICATIONS

ORDERING AND RECEIVING PUBLICATIONS

periodical subscriptions

3b1

Premise:

3c1a

3c1a1

(Same)

In addition, the computer-based system assumes use of the ONLINE JOURNAL SYSTEM and the IDENT SYSTEM.

See: Journal User Guide (userguides, journal-guide,) and (userguides, journal-guide,4) and Sections 4a and 4b.

Procedures:

3b1a2

1. (Same)
2. (Same)
3. ANNOUNCEMENT of the list, and how to indicate selections from it, should be SENT ONLINE AS A JOURNAL MESSAGE.
4. The LIST, with publisher and price information as substatements, may be KEPT ONLINE, AND IDENTs OF THOSE APPROVING OR DESIRING ROUTING MAY BE ADDED TO EACH. THE ONLINE LIST COULD BE EXPANDED IN PURPOSE TO INCLUDE NOTATION OF ISSUES AS RECEIVED. (This procedure in itself has no advantage over a card file, in which all of this information can be easily noted and more easily read.
5. ENTER AS A SUBSTATEMENT THE DATE OF ISSUE RECEIVED, AND THE DATE RECEIVED.
6. (Same)
7. (Same)
8. (Same)

Non-Computer-based Procedures

INFORMATION GATHERED FROM EXTERNAL PUBLICATIONS

Report Orders

301b

Premise:

301b1

Various indexes will be scanned for selection of individual reports to be ordered. Orders will be placed for such reports and for reports from other sources, as desired by the Office members.

Procedures:

301b2

1. Have each issue of Government Reports Announcements scanned by a designated Office staff member to select reports of interest. Other indexing and announcement bulletins are also to be scanned.
2. Place orders for selected reports, in hardcopy and/or microfiche, with NTIS and other document agencies as appropriate, through procedures satisfactory to the Office's parent organization.

Note: Files of reports in microfiche may be built at relatively small expense in money and space, in anticipation of use. Reports of immediate interest may be ordered in both forms, and the hardcopy discarded after immediate use.

3. Make a copy of each order and the announcement from which it is selected, and file alphabetically by title in an Order folder. A duplicate may be filed by agency.
4. When a report arrives, mark received and the date received on the copy of the order, and move order to a Received folder.
5. Stamp the hardcopy report as property of the Office.
6. Make a circulation card, giving bibliographic information.
7. Forward to cataloger.

Report Distribution Requests

301c

Premise:

301c1

The Office will request to be placed on the distribution lists of other Offices and Laboratories doing work of interest to this Office.

Procedures:

301c2

1. Send form letter to those government and academic groups whose technical and progress reports are of continuing interest.
2. File a copy of the form letter to each addressee in an Order folder, by addressee named. File copies of report publication announcements from these groups in the same folder.
3. As requests are granted or refused, file these replies with the appropriate copies.

Computer-based Procedures

INFORMATION GATHERED FROM EXTERNAL PUBLICATIONS

Report Orders

3b1b

Premise:

3b1b1

(Same)

Procedures:

3b1b2

1. (Same)

2. (Same)

3. Periodically, probably weekly, BIBLIOGRAPHIC INFORMATION FOR EACH REPORT, IN A STANDARD FORMAT, MAY BE INPUT TO AN ONLINE FILE OF REPORTS ON ORDER.

ABSTRACTS MAY BE ENTERED AS SUBSTATEMENTS, WITH AN EMPTY STATEMENT IF NO ABSTRACT IS AVAILABLE AT THIS TIME. INSERT THE NAME OF THE REQUESTOR AS A THIRD-LEVEL SUBSTATEMENT. THESE REFERENCES MAY BE ACCESSED ONLINE BY SCANNING OR BY CONTENT ANALYSIS. THIS FILE MAY BE PRINTED OUT IN DUPLICATE, and each reference cut and pasted on a card to be filed in a hardcopy file by title and by agency. In this way, two useful points of entry may be maintained. THE ONLINE FILE MAY BE SORTED BY ENTRY. (Online entry has the advantage of creating a file which may be used in CREATING ONLINE BIBLIOGRAPHIES AND CATALOGS. (See Appendix B/)).

4. When a report is received, the ONLINE ENTRY MAY BE MODIFIED TO GIVE DATE OF RECEIPT AS A FOURTH-LEVEL SUBSTATEMENT,

5. (Same)

6. A CIRCULATION CARD MAY BE GENERATED BY PRINTING THE REFERENCE. (See 3 above/. If an ONLINE CIRCULATION SYSTEM is part of the system, no cards would be needed. (See Announcement and Dissemination/).

7. (Same)

Report Distribution Requests

3b1c

Premise:

3b1c1

(Same)

Procedures:

3b1c2

1. THE FORM LETTER MAY BE GENERATED BY THE LETTER USER PROGRAM.

(Ref: (user-progs,-contents,19))

2. A COPY OF EACH LETTER MAY BE KEPT ONLINE, AND A FOLLOW-UP PROCEDURE INSTITUTED, BASED ON DATE OF ENTRY.

3. REPLIES MAY BE INDICATED IN THE ONLINE FILE. (This procedure would be burdensome, and is not recommended).

Note: Executing these procedures online has no particular advantage over using a form letter and filing a copy. Such a file has no purpose for communication, retrieval, or reformatting, and does not tie into any other processes.

Non-computer-based Procedures

INFORMATION GATHERED FROM EXTERNAL PUBLICATIONS

BOOKS

3b1a

Premise:

3b1d1

Publishers' catalogs and publicity notices, and book reviews from periodicals will be brought to the attention of Office members. Their selections will be ordered by procedures satisfactory to the parent organization.

Procedures:

3b1d2

1. Post or circulate announcements and reviews to interested members.
2. Place orders either directly or through a library, according to organizational procedures.
3. File copies of the orders and of the announcements in the order folder with report orders, alphabetically by title. Indicate on the copy the name of the requestor.
4. When a book is received, mark the order record as received and give date of receipt.
5. Stamp the book as property of the Office.
6. Make a circulation card, giving bibliographic information.
7. Forward to cataloger.

Computer-based Procedures

INFORMATION GATHERED FROM EXTERNAL PUBLICATIONS

- Books: 3b1d
- Premise: 3b1d1
- (Same)
- Procedures: 3b1d2
1. (Same)
 2. (Same)
 3. Periodically, probably weekly, BIBLIOGRAPHIC INFORMATION FOR EACH BOOK, IN A STANDARD FORMAT, MAY BE INPUT TO AN ONLINE FILE OF BOOKS ON ORDER. ABSTRACTS SHOULD BE ENTERED AS SUBSTATEMENTS. AN EMPTY STATEMENT SHOULD BE ENTERED IF NO ABSTRACT IS AVAILABLE AT THIS TIME. These references may be ACCESSED ONLINE BY SCANNING OR BY CONTENT ANALYSIS. This file may be PRINTED OUT IN DUPLICATE, and each reference cut and pasted on a card to be filed in a hardcopy file by title and by author. THE ONLINE FILE MAY BE SORTED BY ENTRY. IT MAY BE MERGED WITH THE ONLINE REPORT FILE if neither is large. (Online entry has the advantage of creating a file that may be used in CREATING ONLINE BIBLIOGRAPHIES AND CATALOGS.
 4. When a book is received, the ONLINE ENTRY SHOULD BE MODIFIED TO GIVE DATE OF RECEIPT AS A FOURTH-LEVEL STATEMENT.
 5. (Same)
 6. A CIRCULATION CARD MAY BE GENERATED BY PRINTING THE REFERENCE. (See 3 above/.
 7. (Same)

Non-computer-based Procedures

INFORMATION GATHERED FROM EXTERNAL PUBLICATIONS

CATALOGING

Periodicals

3b2

Premise:

3b2a

3b2a1

Articles pertinent to the work of the Office should be captured for announcement and retrieval.

The level of effort should be determined by estimation of benefit to the Office. Factors to be considered are: extent of use of such literature by the Office, availability of a large technical library and information service, and availability of an online search service, such as DIALOG, ORBIT or BASIS-70, by which large external bibliographic data bases can be accessed.

Procedures:

3b2a2

1. One or more Office members will scan the new issues for pertinent articles.
2. Make a photocopy of the article and a photocopy of the contents page of the periodical and place as a cover sheet, with a colored mark indicating the article attached. Only if the periodical collection is complete and all issues are saved for several years is it advisable to depend on the original issue for retrieval.
3. Write the author's name or an accession number on the upper left corner, depending on whether the filing system chosen is to be by author or by number.
4. Select elements of the citation: author, title, periodical, date, and subject words, to be recorded in step 5.
5. Type a card in standard citation format, with subject words listed at the bottom.
6. On a photocopy machine, make as many copies of the card as necessary to supply one for each catalog entry point.
7. On each set of cards, underline in red on each card one of the author, title, or subject entries.
8. (Not necessary).
9. Each week, sort the accumulation of cards according to the underlined entry on each. The accumulated set is then to be interfiled in the card catalog.
10. (Not necessary).
11. (No comparable product).
12. (No comparable product).

Computer-based Procedures

INFORMATION GATHERED FROM EXTERNAL PUBLICATIONS

CATALOGING

3b2

periodicals

3b2a

Premise:

3b2a1

(Same)

In addition, an ONLINE FILE OF CITATIONS WILL BE PRODUCED AND KEPT
UP-TO-DATE.

Procedures:

3b2a2

1. (Same).
2. (Same).
3. A number from a preassigned block will be written on the upper right corner.
4. On a coding sheet designed for the purpose enter the appropriate data elements for the citation. It may seem a saving to enter the citation online without this step, but experience will prove that including this step is time-saving.
5. From the coding sheet, TYPE ONLINE THE CITATION in the selected input format, i.e., IN CODED OR IN CITATION FORMAT.
6. (Not necessary).
7. (Not necessary).
8. Each week, or when 50 citations have been entered, RUN PROOFING PROGRAMS, and check that entries are correctly entered and spelled.
9. At selected intervals, RUN THE CATALOG PROGRAMS TO FORMAT THE ENTRIES INTO A BOOK-TYPE CATALOG AND TO PRODUCE AUTHOR AND KEYWORD INDEXES, AND possibly a SUBJECT INDEX to number entries. Make these catalogs available online.
10. Edit the output for errors and inconsistencies which appear.
11. PRINT OUT and issue A HARDCOPY INDEX.
12. Make the CATALOG AVAILABLE ONLINE.

Non-computer-based Procedures

INFORMATION GATHERED FROM EXTERNAL PUBLICATIONS

Reports

3b2b

Premise:

3b2b1

Selected reports will be cataloged, to be accessed along with periodical articles and books.

Procedures:

3b2b2

1. Reports which arrive gratuitously and reports ordered for their presumed interest will be examined for their probable relevance to the work of the Office.
2. Filing entry--usually a serial number, a coded notation, or a label giving originating agency and partial title, according to the arrangement selected for the report file--will be affixed to the upper left corner.
3. Elements of the citation: originating agency, author, title, date, abstract and subject words, will be selected, to be recorded in step 4.
- 4 - 6. (Same as 5 - 9 for periodicals).

Books

3b2c

Premise:

3b2c1

Books acquired on Office funds, for particular use of one worker or general use by the Office, will be cataloged and processed for the Office.

Procedures:

3b2c2

1. Elements of the citation: author, title, publisher, date, and abstract, will be selected, to be recorded in step 2.
- 2 - 6. (Same as 5-9 for periodicals).

Slides, Photographs, any Other Special materials

3b2d

Premise:

3b2d1

Any and all forms of information should be marked, cataloged and made retrievable.

Procedures:

3b2d2

Any and all forms of hardcopy information can be controlled by adapting procedures used for materials above.

Computer-based Procedures

INFORMATION GATHERED FROM EXTERNAL PUBLICATIONS

Reports	3b2b
Premise:	3b2b1
(Same).	
In addition: an ONLINE FILE OF CITATIONS WILL BE PRODUCED AND KEPT UP-TO-DATE.	
Procedures:	3b2b2
1. (Same).	
2 - 10. (Same as 3 - 11 for periodicals).	
Books	3b2c
Premise:	3b2c1
(Same).	
In addition, an ONLINE FILE OF CITATIONS WILL BE PRODUCED AND KEPT UP-TO-DATE.	
Procedures:	3b2c2
1. (Same).	
2 - 10. (Same as 3- 11 for periodicals).	
slides, Photographs, Any Other Special Materials	3b2d
Premise:	3b2d1
(Same)	
Procedures:	3b2d2
1. (Same).	
2 - 12. special codes will be used, as determined by the nature of the collection.	

Non-computer-based Procedures

INFORMATION GATHERED FROM EXTERNAL PUBLICATIONS

ANNOUNCEMENT AND DISSEMINATION

3b3

Premise:

3b3a

Publications received in the Office will be given Office-wide publicity, and will be made available quickly and widely.

3b3a1

Procedures:

3b3b

1. Photocopy and post or display in a binder the contents pages of all new professional periodicals.

3b3b1

2. Photocopy and post or display in a binder the title and abstract pages of all new reports judged to be pertinent.

3b3b2

3. Route all new journal issues as requested.

3b3b3

4. Circulate all books and hardcopy reports to requestors.

3b3b4

Computer-based Procedures

INFORMATION GATHERED FROM EXTERNAL PUBLICATIONS

ANNOUNCEMENT AND DISSEMINATION

	3b3
premise:	3b3a
(same)	3b3a1
procedures:	3b3b
1. (same)	
AN ONLINE JOURNAL ANNOUNCEMENT SHOULD BE SENT TO THE OFFICE STAFF, GIVING TITLES AND ISSUE DATES OF NEW ISSUES AS RECEIVED.	3b3b1
2. EACH WEEK THE REFERENCES ADDED FOR NEW REPORTS AND BOOKS SHOULD BE MADE AVAILABLE TO ALL ONLINE READERS.	3b3b2
3. (same)	3b3b3
4. (same)	3b3b4

Non-computer-based Procedures

INFORMATION GATHERED FROM EXTERNAL PUBLICATIONS

RETRIEVAL

Premise:

All items cataloged should be retrievable.

Procedures:

1. Retrieval of citations may be accomplished by looking in the card catalog under author, title or subject work. 3b4
3b4a1
 2. Retrieval of a periodical article may be accomplished by noting the author's name, or the serial number if that is the filing mode (See Section 3b2a2), and looking in the collection for this item. The item may be borrowed from the collection by filling out an "Out" card and substituting it in the collection. 3b4b2
 3. Retrieval of a document may be accomplished by noting the serial number or coded notation (See Section 3b2b) and looking in the collection for the item. The item may be borrowed by the same procedure as 2. 3b4b3
 4. Retrieval of a book may be accomplished by noting the author, and looking in the collection for the item. It may be borrowed by filling out a card as in 2. 3b4b4
 5. Retrieval of a photograph, slide, or other item may be accomplished by accessing special collection and removing by placing a record of the removal. 3b4b5
- Note: The success of this system depends on keeping these collections under surveillance. The honor system will not work. 3b4b6

Computer-based Procedures

INFORMATION GATHERED FROM EXTERNAL PUBLICATIONS

RETRIEVAL

-----	304
Premise:	304a
(Same).	304a1
Procedures:	304b
1. Retrieval of citations may be accomplished by use of OFFLINE PRINTED INDEXES AND ONLINE INDEXES GENERATED BY INDEXING PROGRAMS. Such indexes will include: AUTHOR, TITLEWORD, AND SUBJECT TERM INDEXES. (See also Section 4e).	304b1
2 - 5. (Same).	304b2
Note: (Same). In addition, success of retrieval of citations depends on frequent running of indexing programs, to keep the indexes up-to-date.	304b3

Non-computer-based Procedures

INTELLIGENCE GENERATED BY THE INDIVIDUAL IN HIS WORK

3c

Premise:

3c1

All ideas, plans, and research results produced by the Office worker should be captured, some to be retained for his personal use, and some to be released for dialogue with other workers. He should be provided privacy for such intelligence until he wishes to release it.

Procedures:

3c2

1. The individual should make handwritten, typed or dictated notes of all ideas to be explored, plans to be developed, and research results produced?
2. Notes are to be transcribed as desired.
3. These notes and preliminary reports can be filed in binders or folders by a scheme devised by the worker.
4. The individual can discuss the topics with coworkers, and give them copies of such items as would be of interest to them.
5. Some means of subject annotation in addition to the filing arrangement in 3 can be devised. This may be a subject cross-index which the individual maintains.

See: Personal Files Section 4d

6. He may access his collection of notes and papers by referring to his subject index and pulling out the referenced notes.
7. Such collections of notes can be pooled for use of the group.

Note: When this mode of operation is not explicitly encouraged, it tends to be neglected, and fruitful ideas, plans, and results are not captured. When well worked out and maintained, a manual system can achieve the same result for the individual as a computer-based system in this application. Group use of such collections is rare.

Computer-based Procedures

INTELLIGENCE GENERATED BY THE INDIVIDUAL IN HIS WORK

3c

Premise:

3c1

(Same)

In addition: ONLINE FILE CREATION for recording the worker's output is provided and its use encouraged.

Procedures:

3c2

1. On a TERMINAL available to him at all times, the Office worker can INPUT his ideas, plans, and results as they occur to him. He will be CREATING FILES IN HIS PERSONAL DIRECTORY.

Ref: TNLS User Guide (userguides, tnls-beginners, 0:x))

or: (ifjournal,19200,) or ARC 19200 in hardcopy

or: DNLS User Guide (userguides, dnls-intro, 1:wn)

2. His ROUGH INPUT CAN BE EDITED AND FORMATTED ONLINE, by a clerk if he desires. INPUT ITSELF CAN BE PERFORMED by a clerk FROM HIS WRITTEN NOTES, USING THE OFFLINE SYSTEM "DEX".

Ref: DEX User Guide (userguides, dex-primer, 1:w)

3. ADDITIONS AND CORRECTIONS MAY BE MADE TO THE ONLINE WORKING FILES, AND THEY MAY BE SEPARATED, COMBINED, AND LINKED to conform to any filing arrangement.

See: Personal Files -- Section 3i

Note: This mode of operation encourages the recording and updating of the worker's intellectual effort, at any level of personal participation or clerical support he desires. This mode requires training and continued use of the online system.

4. The worker should JOURNALIZE SUCH A FILE; whether he distributes it to others or not, he will have made a RECORD which is PLACED IN HIS INITIAL FILE, to which he can REFER ONLINE, without the necessity of keeping it in his online working space.

Ref: Journal User Guide (userguides,journal-guide,))

5. In Journalizing, he may use the KEYWORDS provision, by which he can ATTACH SUBJECT TERMS TO THE RECORD.

Ref: Journal User Guide (userguides, journal-guide, 0:sdv)

6. He can USE CONTENT ANALYSIS OF HIS JOURNAL CITATIONS TO SELECTIVELY RETRIEVE FILES HE HAS JOURNALIZED. CONTENT ANALYSIS RETRIEVES FROM ALL TEXT IN THE CITATION: TITLWORDS AND KEYWORDS would be used for the present purpose.

Ref: Content Analyzer Primer (userguides, llo-guide, 3a:w)

7. Others in the office can USE CONTENT ANALYSIS TO RETRIEVE FROM THE JOURNAL RECORD.

In addition, CATALOG PROGRAMS can be used to PRODUCE INDEXES OF ALL JOURNALIZED ITEMS OR SUBSETS OF THESE.

Ref: (user-progs, -userguide, 1:x)

Note: The Online Journal encourages formalizing notes of individual work, and provides a mechanism FOR THE WORKER AND OTHERS TO AUGMENT THE USE OF INTELLIGENCE WHICH OTHERWISE MAY ESCAPE CAPTURE.

Non-computer-based Procedures

INTELLIGENCE GENERATED BY DIALOGUE IN THE OFFICE

3a

Premise:

3d1

To gain the greatest symbiotic advantage of the Office workers' efforts, the sharing of ideas and results should be encouraged.

Procedures:

3d2

1. Within the constraints of secrecy or confidentiality, workers discuss their work with their coworkers.
2. Meetings of teams, groups and departments convene, for which it is useful to prepare agenda.
3. Minutes of the meetings can be recorded and distributed.
4. Periodic progress reports can be prepared and disseminated internally.
5. Minutes and reports can be indexed by the same procedures as are used for external documents.

See: External reports Section 3b2b.

6. Notes and comments may be added to the group file copies to record reactions or actions taken.

Computer-based Procedures

INTELLIGENCE GENERATED BY DIALOGUE IN THE OFFICE

3d

Premise:

3d1

(Same)

In addition: Workers should carry on continuing dialogue between individuals and among teams utilizing the ONLINE JOURNAL SYSTEM in place of verbal dialogue and handwritten or typed notes.

Use of the ONLINE JOURNAL SYSTEM to distribute agenda, to immediately record thoughts presented and conclusions reached during the meeting, and to distribute these records should be provided and use encouraged.

Procedures:

3d2

1. Individual WORKERS, as they develop ideas, may RECORD THEIR IDEAS IN FILES AND JOURNALIZE THE FILES, DISTRIBUTING THEM ONLINE to colleagues in the office.

COLLEAGUES may then RECORD THEIR COMMENTS AND JOURNALIZE THEM, WITH LINKS to tie the items together.

Ref: TNLS User Guide (userguides,tnls-address,ld)

2. AGENDA for a meeting may be prepared, and INPUT ONLINE AND DISTRIBUTED BY THE ONLINE JOURNAL SYSTEM TO INDIVIDUALS OR GROUPS who are to attend the meeting and to others who should know of the meeting.

Ref: Journal User Guide (userguides,journal-guide,)

3. At the meeting, a DISPLAY TERMINAL, and if possible a projector to project the display, may be USED TO ALLOW PARTICIPANTS TO FOLLOW THE AGENDA. As items are discussed THOUGHTS AND RESULTS can be RECORDED IN THE OUTLINE by a designated meeting secretary, allowing participants to follow these as they are recorded, so that consensus is assured. The FINAL VERSION at the conclusion of the meeting can be PRINTED OUT for the participants, AND JOURNALIZED for others and for the record.
4. INFORMAL WORKING PAPERS AND REPORTS prepared by one or more workers can be INPUT AND JOURNALIZED to keep office colleagues informed of developments in the work.

Non-computer-based Procedures

INTELLIGENCE ISSUED IN FORMAL OFFICE REPORTS

3e

Premise:

3e1

formal reports must be issued by the Office at regular intervals, to satisfy the terms of contracts and as a product of the research and development carried on by the Office.

Procedures:

3e2

1. Drafts of a report or sections of a report are prepared by the workers, usually in handwritten form.
2. Experienced typists transcribe the drafts into typewritten form.
3. The writers mark the typed drafts with emendations.
4. The typists retype the drafts.
5. An editor marks the drafts for changes.
6. The typist retypes the drafts, as many times as substantial changes are made.
7. The final copy is retyped in a reports preparation group on a special typewriter.
8. The final copy is reproduced in quantity usually by a separate reports preparation group, and given a cover if desired.
9. The reproduced copies are mailed.
10. A stock of reproduced copies may be filed for continued access and distribution.
11. Each report may be cataloged and subject indexed by the procedures used for external reports.

See: External reports Section 3b2b.

Computer-based Procedures

INTELLIGENCE ISSUED IN FORMAL OFFICE REPORTS

3e

Premise:

3e1

(Same).

Procedures:

3e2

1. Drafts of a report or sections of a report can be INPUT BY WORKERS IN ONLINE FILES of their own, OR IN ONE ONLINE FILE WORKED ON JOINTLY AT VARIOUS TIMES.
2. If desired, handwritten portions may be given to a TYPIST who MAY TRANSCRIBE IN AN OFFLINE MODE CALLED DEX WHICH IS LATER FED INTO AN ONLINE FILE.
See: DEX User Guide (userguides,dex-primer,)
3. WORKERS WILL EMEND THE FILE ONLINE.
4. (Eliminated by 3).
5. EDITOR CAN EDIT THE FILES ONLINE, ADDING PRINT DIRECTIVES FOR OUTPUT PROCESSOR OR DIRECTIVES FOR COM.
See: (userguides,op-intro,)
6. (Eliminated by 5).
7. FINAL COPY FOR REPRODUCTION CAN BE PRODUCED BY THE ONLINE PRINTER OR BY COM.
8. (Same).
9. (Same).
Also: The final ONLINE COPY CAN BE JOURNALIZED AND DISTRIBUTED ONLINE TO THOSE WHO CAN RECEIVE IT ONLINE.
See: (userguides, journal-guide,)
10. (Same).
Also: THE JOURNALIZED COPY CAN BE ACCESSED ONLINE FOR READING AND REPRODUCTION AT ANY TIME.
11. A copy of the report may be cataloged by the procedures for external reports, and the entry ENTERED INTO AN ONLINE BIBLIOGRAPHY.
See Section 3b2.

Non-computer-based Procedures

INTELLIGENCE COMMUNICATED BY CORRESPONDENCE

3f

Premise:

3f1

All research intelligence contained in correspondence generated by the Office and gleaned from correspondence received by the Office should be captured and processed for retrieval.

Procedures:

3f2

1. Each item of incoming correspondence can be given a number by a central secretary.
2. A copy of each incoming letter can be made and given to the addressee.
3. A copy of each incoming letter can be coded for a subset of data elements applicable to correspondence (number, date, author, author's organization, topics, action to be taken).
4. Originals can be filed by number.
5. Any enclosures can be routed to a cataloger to be processed like other external reports.
6. The number, date, and author of each incoming letter can be posted on cards in a card file arranged by names of author or by authors' organizations.
7. The number, date, and author of each letter can be posted on cards in a card file of topics.
8. Each item of outgoing correspondence can be given a number.
9. A copy of each outgoing letter can be filed by number.
10. An expendable copy of each outgoing letter can be coded for a subset of data elements applicable to correspondence (number, date, author, recipient's organization, topics).
11. Number, date, and author can be posted on cards in the file (See 6) of cards by recipient's organization.
12. Number, date, author, and recipient's organization can be posted in card file (See 7) by topic.
13. When desirable, searches can be made in either card file to locate the numbers of letters to be retrieved, and the cited letters read or copied from the files of originals and copies.

Computer-based Procedures

INTELLIGENCE COMMUNICATED BY CORRESPONDENCE

3f

Premise:

3f1

(Same).

Procedures:

3f2

1. (Same). The number should be one from the ONLINE NUMBER SYSTEM, to be used when the record is journalized.

2. (Same).

3. (Same).

See: Correspondence Log System, Appendix A

4. (Same).

5. (Same).

6. Number, date, author, author's organization, addressee, and so forth, can be ENTERED IN AN ONLINE LOG OF CITATIONS.

7. At the same time, KEYWORDS CAN BE ENTERED for each item.

8. Each outgoing letter can be TRANSMITTED TO ADDRESSEES ONLINE, using SNDMSG or other online mail system, AND BE JOURNALIZED in the process.

Ref: Scenario for Network Journal (22383,) Appendix B

9. JOURNALIZING WILL CREATE AN ONLINE COPY.

10. In JOURNALIZING, a set of DATA ELEMENT FIELDS are indicated, INCLUDING KEYWORD INPUT.

11. These ELEMENTS can be ENTERED IN AN ONLINE FILE OF CITATIONS.

Non-computer-based Procedures

INTELLIGENCE COMMUNICATED BY PHONE

38

Premise:

381

Information transferred by phone between Offices is often critical to the course of work in an Office, and a standard means for recording should be provided.

Procedures:

382

1. Important points and decisions reached during the course of a phone call should be recorded in a memo.
2. A copy of the memo can be filed in a chronological file, with an accession number.
3. A copy of the memo can be filed in any project folders it applies to.
4. A copy of the memo can be coded for a subset of data elements applicable to it (number, date, time, office member called or calling, outside person called or calling, outside person's organization, topics, agreements, actions to be taken).
5. Data elements from the memo can be posted on cards in a card file by name of outside person or his organization.
6. Data elements from the memo can be posted on cards in a card file arranged by topics.
7. Information may be retrieved by accessing the name or topic files and noting the wanted numbers, and pulling the memo copies from the number file of memos.

Computer-based Procedures

INTELLIGENCE COMMUNICATED BY PHONE

38

Premise:

381

(Same).

Procedures:

382

1. Important points and decisions reached during the course of a phone call may be RECORDED ONLINE BY THE OFFICE WORKER.
2. This ONLINE RECORD may be kept IN AN ONLINE PHONE LOG.
3. (Not necessary, see 4).
4. The online record may be JOURNALIZED, WHICH AUTOMATICALLY RECORDS DATA ELEMENTS OF DATE, TIME, AUTHOR, NUMBER. DATA ELEMENTS for topics, contracts, persons contacted, and so forth, CAN BE ENTERED AS KEYWORDS during journalization.
5. (Not necessary, see 4).
6. (Not necessary, see 4).
7. Information may be retrieved by ACCESSING THE ONLINE RECORD BY CONTENT ANALYSIS AND LINKING TO THE FULL MEMO.

Non-computer-based Procedures

INTELLIGENCE GATHERED FROM VISITS AND VISITORS

3h

Premise:

3h1

Information gathered by Office workers on visits to other organizations, and information learned from visitors to the Office is often useful to the Office. Standard procedures should be devised to capture such information and make it selectively available.

Procedures:

3h2

1. On returning from a visit, the worker can write a trip report.
2. After entertaining a visitor, any interesting information on the work of the visitor, or any suggestions made by the visitor regarding the work of the office, can be recorded.
3. Such a report can be typed, copied, and distributed to those interested.
4. A copy can be numbered and filed to allow reference through an indexing system?
5. The report can be indexed by the visitor's name, by organization, and by topic.
6. Numbers can be posted in the same file used to index correspondence.
7. When desirable, searches can be made by locating the numbers of reports to be retrieved, and the reports can be read or copied.

Computer-based Procedures

INTELLIGENCE GATHERED FROM VISITS AND VISITORS

3n

Premise:

3n1

(Same).

Procedures:

3n2

1. On returning from a visit (or even during a visit trip, if the Office visited has online access) the worker may ENTER ONLINE HIS NOTES ON THE VISIT.

Ref: TNLS User Guide (userguides, primer,)

Ref: TELNET User Guide

Ref: Journal User Guide (userguides, journal-guide,)

2. After entertaining a visitor (or during the visit, if practical) the worker should ENTER ONLINE ANY INTERESTING INFORMATION RECEIVED from the visitor.

Ref: See 1 above

3. The FILE should be JOURNALIZED and may be DISTRIBUTED to colleagues.

4. SUBCOLLECTION AND KEYWORD ENTRIES may be USED to provide for retrieval.

Non-computer-based Procedures

INTELLIGENCE COLLECTED BY THE INDIVIDUAL FOR PERSONAL USE

31

Premise:

311

The individual in an Office will collect miscellaneous bits and pieces of information in various forms, and will find it expedient to group these items into files that are unique and personal. Shelf and drawer space will be provided him to store the documents, articles, books, notebooks and folders in which this information exists.

Procedures:

312

1. The individual will accumulate a stack of books, clippings, letters, notes and so forth, until he finds its use unwieldy and retrieval unsuccessful.
2. The individual will sort the stack and reassemble it into folders or boxes or notebooks in some classes which appear to fit the accumulation and his present interests. If he repeatedly cannot locate items he recalls having, and finds he is uncertain in which of several places he could have filed it, he may try making an index to the collection. If some of the material is published items he may turn them over to an assistant to be indexed, at which time the material is subject to procedures for external documents.
3. The most common means of indexing is to attach a serial number to each item, so that when he calls for that number, he will get that item.
4. He will then select terms he expects he would want to retrieve by, and make a list or a card file, entering the numbers under the appropriate terms.
5. He may find he is using synonyms in his term-creation, and start making cross-references, to keep numbers of like subject together.
6. He may file the numbered items by number, for retrieval.
7. He may weed the collection periodically, by examining items under early numbers.

Note: Difficulties he may run into include:

- He will get behind in this effort, and still accumulate a mass of unlocatable material.
- He will want to retrieve by other terms than he used, and will not be able to recall the terms he used.
- He will pull a selection of hardcopies from his shelf for one purpose and then not know where an item is when he looks for it by number for another purpose, and knowing only the number, will not be able to guess what he might have pulled it for or where he might have put it.
- He will want to make a record of some items which he cannot retain in hardcopy, such as borrowed books or reports, and he will find a number is an inadequate clue for retrieval.

Computer-based Procedures

INTELLIGENCE COLLECTED BY THE INDIVIDUAL FOR PERSONAL USE 31

Premise: 311

(Same).

In addition, the Worker is assumed to have ONLINE MATERIALS which he will want to keep available for retrieval.

Procedures: 312

1. (Same). In addition, the individual may COLLECT in his initial file some ONLINE MESSAGES AND REFERENCES TO ONLINE ITEMS which he expects to want to refer to.
2. (Same). In addition, he may ARRANGE HIS ONLINE MATERIAL in some manner he feels will facilitate his retrieving it at will, by MOVING REFERENCES INTO DISTINCT BRANCHES.
3. Using his initial file, he may ENTER CITATIONS as a record for hardcopy also. He may COPY CITATIONS PREPARED BY THE CATALOGER for catalog or indexes.
4. He can keep a RETRIEVAL PATTERN to be used to make A CONTENT ANALYSIS SEARCH of his online records by inserting his selected terms in the pattern. He can INSERT SUBJECT TERMS AS STATEMENT NAMES, to anticipate retrieval by statement name.
5. He can PREPARE his own ONLINE THESAURUS, WITH LINKS between terms, AND LINKS TO CITATIONS OR BRANCHES IN HIS FILE.
6. (Same).
7. (Same). In addition, he may WEED HIS ONLINE REFERENCES, both online and offline items, BY A PERIODIC SEARCH BY DATE OF ITEM OR DATE OF ENTRY, AND DELETE ONLINE CITATIONS he no longer expects to need. Also, he may ARCHIVE A SET OF OLDER REFERENCES.

Note: The same difficulties may impede the online user as the offline user; he may have added difficulties, such as:

As his file grows larger, content searches may become too slow to be practical.

He may find he must print out his whole record file in order to scan it closely when he doesn't retrieve what he expects to find. Such printouts will have to be recent; an older printout will not contain current items.

Advantages he will gain include:

For online documents, he can generate a new hardcopy when he can't locate a copy he has filed.

By content analysis, he can make searches in titles or abstracts for words he did not anticipate when he filed the items.

He can copy citations into new documents he is preparing, without retyping.

EXPERIMENTAL DEVELOPMENT OF A
SMALL COMPUTER-AUGMENTED INFORMATION SYSTEM

SRI-ARC 31 DEC 73 21453
Annual Report to ONR

Appendix B

Design for a Correspondence Log System

INTRODUCTION

1

This document sets forth the specific provisions and alternative provisions of a system designed to make available the research information contained in correspondence, defined as communication by letter. The system is a module in a total design for managing research information, as would be desirable for a knowledge workshop. 1a

In the knowledge worker's relations with his Office co-workers and with his colleagues in his discipline- or mission-oriented Communities, it is desirable for him to be able not only to retrieve but to reuse and recommunicate information gathered from correspondence. This design deals not only with the means by which the information in correspondence may be gathered, stored, and retrieved, but also explores means by which it can be reused and redisseminated. 1b

ITEMS AND INFORMATION TO BE HANDLED IN A CORRESPONDENCE CONTROL SYSTEM

2

In a department of 20 or more people carrying out project work, published and unpublished information of many types and forms is received daily. Forms in which the information arrives are typically: 2a

U.S. Mail

Official correspondence from contracting or funding agencies

Personal letters, sometimes with enclosures, some of which are publications

Letters to the department, requesting information, reports, or action

Internal Mail

Memos from parent organization to individual or classes of individuals

Official project-related reports and memos from the parent organization

Personal memos from individuals

Routed memos and reports

One class of mail typically received in a department is not dealt with here. Purchase orders and invoices are judged not to contain research information and are assumed to be subject to a separate control system. 2b

Other classes of mail not of principal concern here are reports and periodicals received on distribution or by subscription, and catalogs, brochures, and meeting announcements. These items clearly bear research information; however, except when they are received as enclosures to personal letters, thereby suggesting further dialogue by correspondence, they are outside the scope of this system and are covered by a document control system.

2c

RELATION OF CORRESPONDENCE TO OTHER ITEMS BEARING RESEARCH INTELLIGENCE

3

Research information gleaned from correspondence will be used in conjunction with information received from many other sources, so that systems of capturing, storing, and retrieving it should be designed with these uses in mind, and designed in relation to systems set up to handle information from the other sources. A correspondence control system ideally should allow the receiver and communicator of research information to move freely from records for one medium (correspondence) to records of others (telephone dialogue, his personal notes, periodicals, reports, etc.) with a common set of procedures.

3a

CHARACTERISTICS OF CORRESPONDENCE

4

Correspondence presents interesting problems in information receipt, processing, storage and retrieval. Written correspondence as a medium for research information has several characteristics which make it both easier and more difficult to handle than other written media. The characteristics are:

4a

No standard reference, such as the standard references existing for books and journals

Small volume of distribution, so that reference is not meaningful to the general public

Ephemeral nature of some items, requiring purging

Time value of content, requiring quick handling

Dialogue nature of content, requiring (1) retrospective reference and (2) suspense until process continues

Unpredictability of receipt

Worthlessness of some items, requiring selectivity

Relationship of content to enclosures which may have independent value

Informational value of items outside text body, such as sender's title, address, phone number

DESIRABLE FEATURES OF A CORRESPONDENCE CONTROL SYSTEM

5

From these characteristics, certain desirable features of a correspondence control system may be derived. Features desirable for a system are:

5a

Selectivity rules

Quick recording procedures to avoid delay in transmission

Concise, unambiguous reference to each item

Control of original item, for quick retrieval

Provision for separating enclosures while retaining evidence of relationship

Links between item and any predecessors and successors in dialogue

Reminder system to insure required or desired responses are made

Provision for recording thoughts stimulated or actions taken as a result of the item

Retrieval by personal and corporate author, by specific subject, and by general class or set

Purging procedures

Archiving procedures

PRINCIPLES OF INFORMATION RETRIEVAL SYSTEMS AS APPLIED TO CORRESPONDENCE 6

The basic principles identified by Holm ["How to Manage Your Information", by Bart E. Holm, Reinhold, 1968] as essential for an effective information retrieval system apply to handling a correspondence data base. A look at these principles as they relate specifically to correspondence is presented here.

6a

1. Balance input and output effort.

6b

The tendency in designing a recording system for correspondence is to make a brief log of date and sender. Too little information captured at any input step causes retrieval to be difficult or impossible. In an effort at control it is possible to design a complex indexing system which takes more time and effort than some items are likely to be worth and tends to cause delay in transmission and to break down under its own weight. Each class of item handled should be considered at each step to make sure the expected retrieval value and envisioned retrieval steps are in balance with the effort at that step. In a complex system with

many steps it is also important to design input procedures so that each element of input is handled at the most efficient step. 6b1

2. Evaluate single entry and multiple entry files. 6c

The basic assumption in designing a retrieval system for research information in correspondence is that a single type of entry, such as sender, organization, or single subject is not adequate for retrieval for the many types of information contained in the file. On the other hand, some types of information, such as telephone numbers, will need to be retrieved only in connection with personal names, and an entry point by telephone number itself is valueless. 6c1

3. Describe items fully. 6d

Often a reference which seems adequate when the item is freshly in mind is found to be inadequate when the context is out-of-mind. To save retrieval of a bulk of items in order to find an appropriate one or few, enough information should be captured to make specific items identifiable. Lacking specifics, the searcher may be faced with retrieving and reading whole text of several letters to find the one that deals with a particular topic. Isolated incidents of this type will occur, but well-planned description will make such effort the exception. 6d1

4. Control the vocabulary. 6e

This principle is the single biggest pitfall of retrieval systems. The effort of constructing a controlled vocabulary with cross-references, and possibly in thesaurus form, is difficult and time-consuming, and the shortcut of using keywords in a free vocabulary is often tried. When any item from a class will be as satisfactory as any other, then uncontrolled vocabulary often will serve to bring the searcher an adequate return. However, when a searcher wants a particular letter and searches for it by the term "packet-switching" when it was indexed under "store-and-forward" with no cross-reference to remind him to ask for this term, he will fail to find it. Comprehensiveness and consistency in indexing are essential. 6e1

5. Know the subject. 6f

It is important that the preparation of the vocabulary for indexing, the indexing itself, and the retrieval be done by people knowledgeable of the subjects covered. Ideally, the user will have a hand in all three of these processes. 6f1

6. Select appropriate storage form.

6g

Various decisions on storage form must be made:

6g1

Where will the original items be physically stored, and how will they be both accessed and protected?

Will duplicate files of copies of originals be worth the effort of maintaining them, and where will they be stored?

How will the online references be stored, to make them accessible for added input while secure from unauthorized changes, accessible to proper users while secure from unwanted readers.

How will the online files be linked to other online information files?

How will online references be linked to filed hardcopy originals?

Will hardcopy of indexing tools be useful, and how will such hardcopy be kept up-to-date?

What storage forms for the text that will allow retransmission are possible and feasible?

ACTIVITIES IN HANDLING CORRESPONDENCE

7

Sorting out Types of Incoming Mail

7a

For convenience in describing the handling of mail, it will be assumed that mail addressed to the department and all business mail addressed to individuals as well, is delivered to the department secretary. From this miscellaneous flow of receipts the secretary separates the items to be entered into the system. 7a1

Letters addressed to individuals

7a2

In the prototype system, all business mail addressed to individuals in the department is assumed to have the potential of containing research information. In a research group, this mail constitutes a dialogue between members of a discipline-oriented Community (the invisible college concept) and serves to stimulate research in the same way as do technical articles read by the worker. The value of such letters may be high, and their capture be important.

To preserve the privacy of correspondence, the design allows each individual a personal record, with pooling of entries to be accomplished at the option of the addressee. Mail to

individuals may be opened by the secretary and logged at that time or may be delivered to the individual who will bring back to the secretary items to be logged for his file only or coded to be available in a department file.

Letters to individuals often contain enclosures which are reports or reprints published elsewhere, or which have at least a separate identity, and are to be related to the letter but should be individually logged.

Letters to the Group, requesting reports or information

7a3

Because of its project work, a department receives mail from other individuals and organizations requesting information about the work. Sometimes these requests will be addressed to individuals whose relation to the work is known. Whether opened by the secretary or the individual, these letters may be handled as department letters.

These letters may request specific documents, ask general questions answerable with documents, or specific questions requiring specific answers. The research information in these requests is usually very slight, and any importance may rest in the fact that they represent a contact in the Community.

Citations to these letters, and replies or citations of replies, will be entered into the system.

Official correspondence from contracting or funding agencies

7a4

Incoming mail of this class includes requests for proposals, notice of acceptance of proposals, requirements of contracts, and occasionally carries enclosures which are independently useful. Mail of this class requires a procedure for follow-up.

Internal mail

7a5

Within an organization, transmission of informal notes, of formal organizational items and of copies of internal reports, and retransmission of external documents occurs. Some of these items are so ephemeral that there will be no need to record them. Several types are of great value. Representative types which will be put into the system are:

Policy documents from the parent organization

Technical reports from related departments

Periodic reports of the parent organization and of other departments in the organization

Reports and periodicals

7a6

All copies of journals received on subscription and all reports without transmittal letters are put aside. These will all go to the person who functions as order clerk, and are routed through the department librarian for examination and selection of articles and reports to be coded for the document catalog and for items to be brought to the attention of individuals in the department. All items of these classes are considered out of the scope of the correspondence files.

Catalogs, meeting announcements, and advertisements

7a7

These are out of the scope of the system, and are sent to the librarian for posting, routing, or filing. Any person wishing to have an online record made of any of these may make an entry in his personal file or may request the librarian to make a citation in the document catalog files.

Purchase orders, invoices, and other equipment and service mail 7a8

A separate purchase order online record would be set up in an augmented Office. No plan for this is included here.

Recording the Correspondence in an Online Log

7b

Numbering

7b1

The recording and retrieval system for correspondence employs the numbering conventions used in recording online dialogue and hardcopy documents. A common series of numbers is available in the Journal system and is drawn upon for unique serial numbers for all types of information captured for the Office. [Ref: Journal User Guide (userguides,journal-guide,5)].

For the purposes of the correspondence system, a group of numbers are extracted as "preassigned numbers". The Journal system records these at this time as being assigned to the person taking them, and no record is made of the purpose to which they will be put. A number acquires meaning only when the coding information is entered with it and the record is "Journalized". The acts of taking a number and Journalizing can be accomplished at the same time, but in practice the two operations are usually separated, because often it is not convenient to complete the record at the moment that a number is desired to provide a reference point for an item.

The number to be assigned to the record of the item is written on the item, so that the item and any copies made of it will show under what number it is recorded. Reference to the number

of the item in subsequent dialogue will then make recall of the item's record possible.

Coding

7b2

On this manual step depends the success of the whole procedure. It is important to capture the essential information from the item so that subsequent references to the item by number will retrieve meaningful information, and so that indexing of the item is possible.

Data elements (see Section 10 of this report) selected for this file are a subset of those used for cataloging items for the document collection and for online dialogue in the Journal system.

Entering the Citation

7b3

The citation elements may be entered in any format selected as a standard, depending on the method of retrieval planned. There are two modes of retrieval to be provided for.

1. A concise citation of the correspondence item is needed for immediate use by the recipient in making notes and in referring to it for further dialogue.
2. A full citation is needed, to be used in running indexes for retrospective retrieval.

It is convenient at this time to use the input format shown here, because there are programs to operate on this format to prepare indexes for Journal items, and these same programs can be used to prepare similar indexes for these citations

(Cnumber) *a1 Writer #1 Writer's job title #2 Writer's organization #3 Organization address #4 City #5 Zip *b5 Addressee #1 Addressee's job title and so forth for all data elements for which there is information after coding.

However, this format is unsuited to need (1) and therefore the prime need is for files of citations input in a paragraph for easy reading. Such format is equally open to content analysis retrieval, and has the advantage that no formatting programs are needed before they are copied as references. Programs could and should be written to convert a citation input in the form below to the various formatted indexes. An example of such input is:

To: Addressee (His organization)
From: writer (His organization)
Date: date written Number: Cnumber
Re: Title
Action called for:
Action taken:

The log will be constructed in segments, so that a set of updated citations can be issued daily by the secretary, and added as an increment to the files which may have been modified by individuals since the last update.

Accessing and Using the Log

7c

Accessing the Log

7c1

For privacy reasons, the correspondence log probably will not be entirely open to all readers. In this case, subsets of the log should be made by the secretary for various individuals and classes of readers.

The worker can access his version of the log to read the citations there and to use them for further dialogue, as indicated in the following activities.

Making Notes

7c2

The correspondence log as seen by the Office worker will be open for him to write in his comments and action he takes in response to letters. These will remain in the file and be unaffected by updates.

Copying Citations

7c3

If the worker keeps personal files, he will be likely to copy citations from the log into his files. He may of course construct links instead.

Using for Follow-up

7c4

An important use for the log is to insure that correspondence is answered as called-for. One means to do this is to have the secretary check the Action line at intervals to follow up on actions indicated.

Charting the Flow in the Correspondence Control System

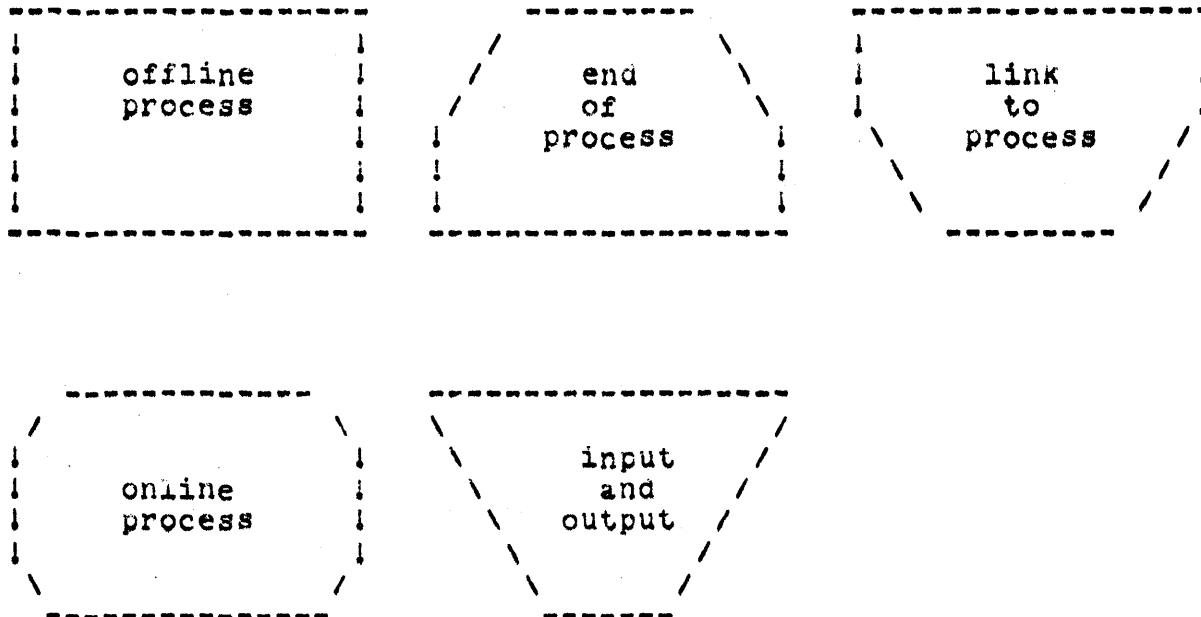
7d

Based on the foregoing analysis, a correspondence control system has been charted. Explanatory text follows the flow charts, to briefly describe the operations charted.

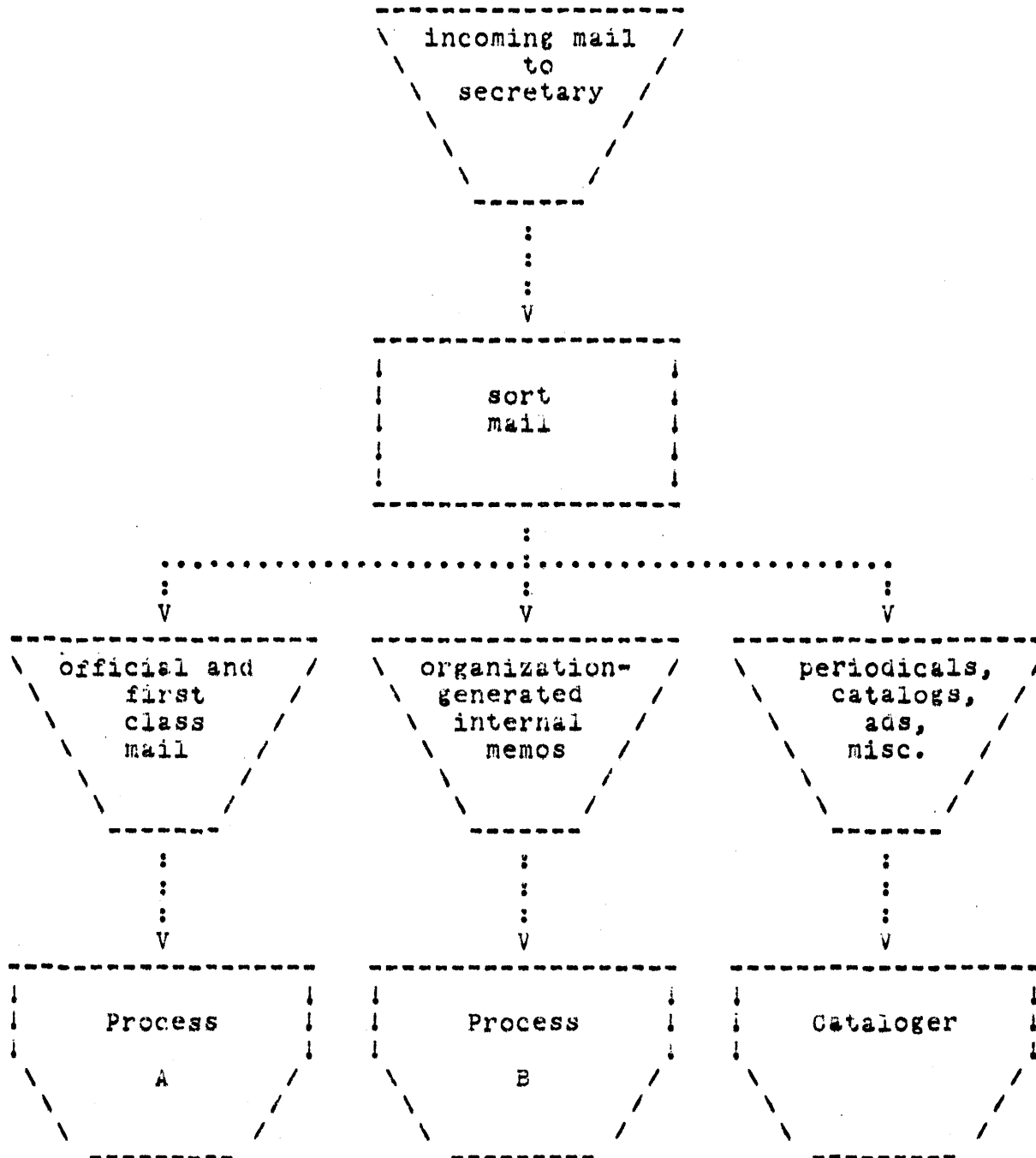
7d1

FLOW CHARTS FOR CORRESPONDENCE

SYMBOLS USED IN CHARTS



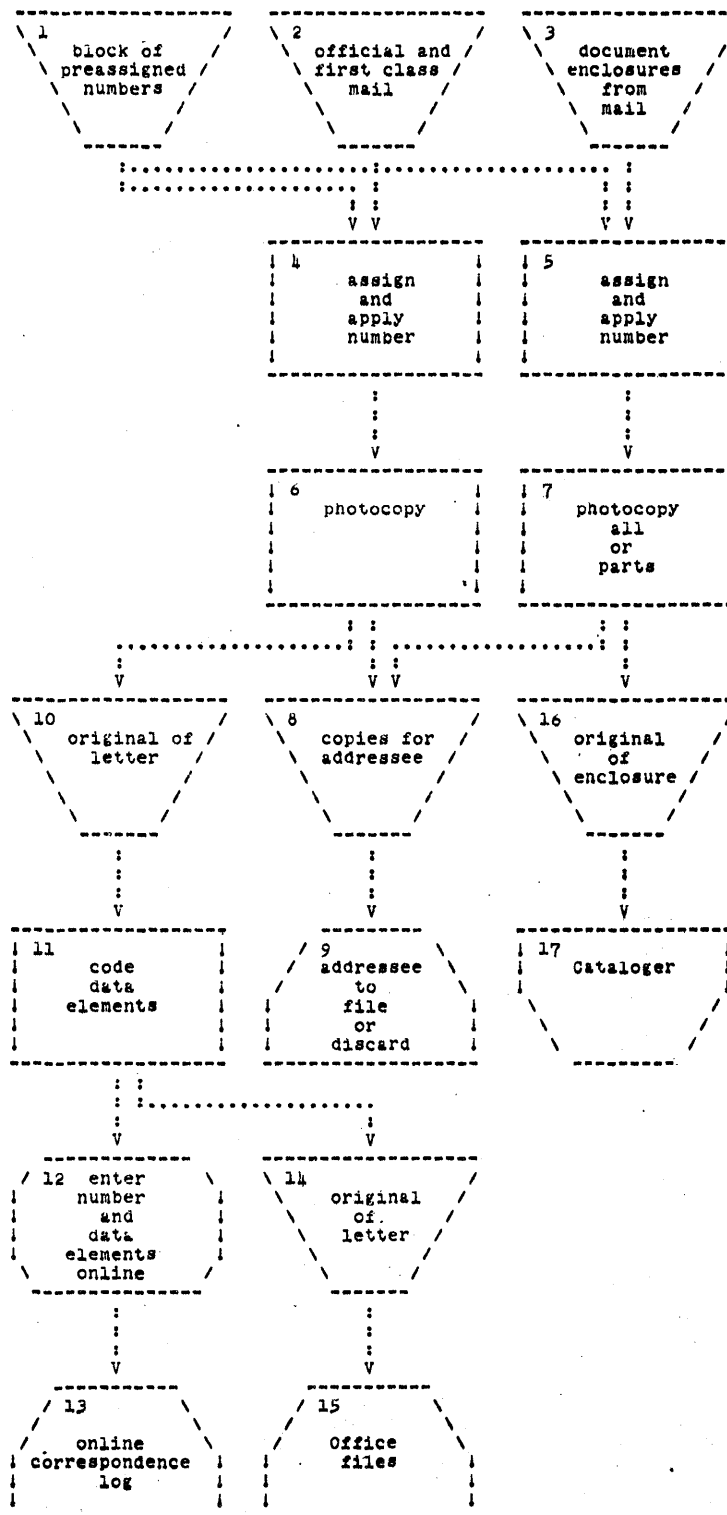
FLOW CHARTS FOR CORRESPONDENCE



FLOW CHARTS FOR CORRESPONDENCE

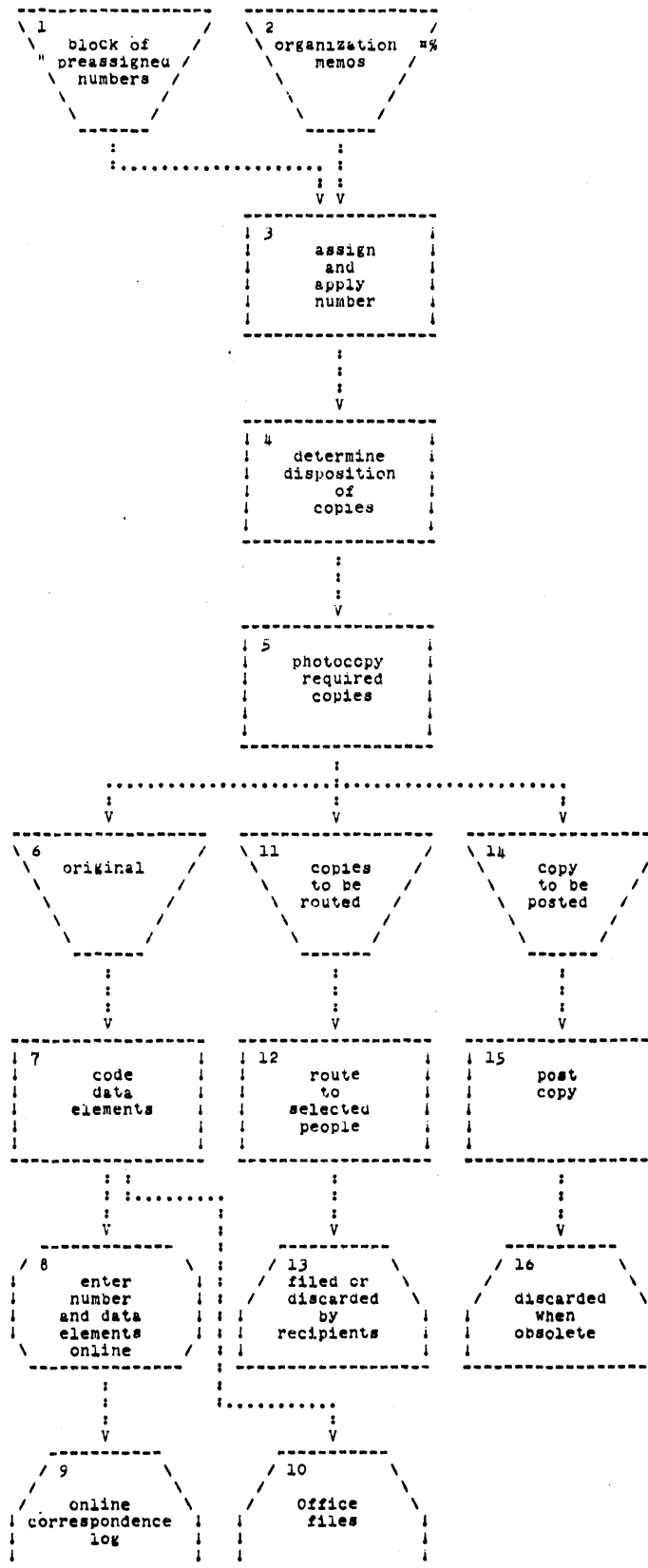
PROCESS A

Numbers in boxes refer to procedure notes which follow the charts.



FLOW CHARTS FOR CORRESPONDENCE

PROCESS B



INCOMING CORRESPONDENCE

8

PROCESS A

8a

The following procedures presuppose an interest of the Office in keeping track of all hardcopy correspondence that is received as official business and that is received by individuals in their Office capacity. It is adaptable for use in recording all correspondence received by an individual, in which case the originals would go into a separate file for the individual, and the citations would be placed in an online file to be accessed by him alone.

1. Access the online number system and request a block of preassigned numbers, perhaps as many as will be needed for a week. Make an offline record, and check off as used.
2. Examine the letter to assure it is a separate codable entity. It may be a formal contract document, a letter to an individual in his Office capacity, a letter from another professional, or even a personal interoffice communication from a colleague in the same organization.
3. Examine any enclosure to decide whether it should be coded. Enclosures which are significant documents in themselves should be treated as are other external documents, but linked in the data elements of the citation to note their connection to correspondence.
4. Assign and apply a number to the letter.
5. Assign a number, preferably an adjoining number to the number assigned the accompanying letter, and apply the number to the enclosure document.
6. Photocopy the letter in entirety.
7. If practical, photocopy the entire enclosure, so that the original can be retained for the Office collection. If not practical, photocopy enough to indicate the contents.
8. Give the copies of letters and documents to the addressee.
9. Addressee retains, forwards, or discards his copies, with the knowledge the originals have been retained by the Secretary.
- 10-11. From the original letter, code the data elements applicable to the letter. (See: Data Elements for Correspondence, Section 10 of this report).
12. Enter the selected information into an ongoing file.

13. This file may be accessible by the recipients of the correspondence, for links to the citations, and for adding comments of their own.
- 14-15. File the original of the letter, by number, in the Office files, or, file the correspondence of an individual in a separate file accessible to him.
- 16-17. Code the document with data elements appropriate to external documents.
18. Enter the citation in the external document catalog with citations for other external documents not related to correspondence, adding the indication of its connection to correspondence.
19. The online document catalog will contain citations to other externally-generated publications. The assumption is that the value of the document is independent of the accident of its transmittal. In fact, a copy of the document may have been received and cataloged previously, in which case the new document may be indicated as a second copy and the earlier number used.
- 20-22. The document itself may be filed in the hardcopy document collection, or charged out to the user with an OUT card, or in the case of a second copy, given to the original addressee. It will be a matter for local practice as to how much control is maintained over documents that are actually the property of the original addressee.

PROCESS B

8b

This process is designed to control written material that is issued within the parent organization for its own people, and that is expected to be needed for reference for internal dialogue.

1. Use the same block of preassigned numbers obtained for Process A.
2. Examine memos, management reports, policy documents, and all other types of written communications received in the Office from the parent organization or other Offices in it. Some will be seen to be ephemeral and not worth recording. Others will be intended for the Office management only. Others will be of interest to several of the Office workers. Local practice will determine application of procedures.
3. Assign number to each such communication of more than ephemeral interest.
4. Determine whether the communication is to be routed and if it is to be posted.
5. Make photocopy or copies as indicated by 4.
- 6-7. Code the original with applicable data elements.
- 8-9. Enter the citation into Office online correspondence record files, including information of routing if any.
10. File original in Office hardcopy files.
- 11-12. Route copies as desirable.
13. Copies may be disposed of as desired by recipients.
- 14-16. Post copy if this action is indicated, and discard when out-of-date.

OTHER MATERIALS RECEIVED

9

Some externally-produced documents will be received as Office mail, rather than in response to orders or on automatic distribution. These may be addressed to the Office director or simply to the Office. They differ from these latter two cases in that discretion is needed in electing to bring them to the attention of the director or other Office member and whether to acknowledge receipt to the sender.

9a

1. If the document is addressed to an individual, attach the address label to the document. No number is to be assigned at this stage to a document received at the Office.

9b

2. Send a card of acknowledgement if this seems appropriate. Forward to the Cataloger, with note to route to addressee if it is assumed addressee would wish to see it.

9c

DATA ELEMENTS FOR CORRESPONDENCE LOG

10

These data elements are a subset of those designed for cataloging all information items (10937,) including Journal items and external documents.

For outgoing letters, the elements will be picked up by cataloging programs from the online files. For incoming hardcopy letters, a coder will select the relevant information from the copy used in coding.

Number: Number should be input in the form (Cxxxxx). This functions as the statement name.

From: *a1 Author
 #1 job title
 #2 corporate affiliation
 #3 suborganization, if any
 #4 street address
 #5 city, state, zip

To: *b5 Addressee
 #1 job title
 #2 corporate affiliation
 etc.

Re: *c1 Title [or supplied title in brackets].
 All incoming letters or memos not carrying a title
 will have a bracketed title supplied by the coder.
 All outgoing letters and memos should contain one
 supplied by the writer.

Rec'd: *d1 Date item was received

Dated: *d2 Date carried on item

Encl: *n6 Documents accompanying this letter or memo

Abstract: *y1 Not essential, but a place for noting information of
 interest from or about the communication. For
 outgoing memos this field will be picked up from the
 Comments field.

Keywords: *y3 Keywords, preferably selected from an authority list.

EXPERIMENTAL DEVELOPMENT OF A
SMALL COMPUTER-AUGMENTED INFORMATION SYSTEM

SRI-ARC 31 DEC 73 21453
Annual Report to ONR

Appendix C

Scenario for Producing Subject Bibliographies

6 August 1973

SCENARIO FOR PRODUCING SUBJECT BIBLIOGRAPHIES
Jeanne North with the aid of Dean Meyer

Introduction

A very useful capability of NLS is its application to the production of listings of references on a given subject. Using various NLS processes, a list of citations on one subject can be extracted from a large data base of citations on various subjects. An ordered list can be produced from a list in some other order or without order. Print directives can be added which will format the output for improved legibility.

This is a rigid scenario giving only an indication of the rich capabilities of NLS in the production of information tools. Its intent is to direct the novice to a useful product without letting him be overwhelmed by the options actually available. The interested and more experienced user is referred to the Content Analyzer Primer (userguides, L10-contentanalyzer,1), the TNLS User Guide (userguides,tnls-beginners,0) and the Output Processor User Guide (userguides,op-intro.NLS,).

Preparation of the Data Base

The data base is any list of citations prepared by the user. These citations may be in any format. Examples of formats are:

Schroeder, Michael D.; and Jerome H. Saltzer. A Hardware Architecture for Implementing Protection Rings. Communications of the ACM, vol. 15, no. 3, March 1972. p.157-170.

Keywords: Computer Access; Privacy;

Reddy, D. Raj; Lee D. Erman; and Richard B. Neely. A Model and a System for Machine Recognition of Speech. ARPA Network SUR Group, SUR Note 44. 5 September 1972. 26p.

Keywords: HEARSAY;

Erickson, Garwood Elliott. An Audio Response User Information System. Dartmouth College, Thayer School of Engineering, Hanover, N.H. June 1970. 87p.

The appearance of the finished bibliography will be improved by the adoption of a uniform bibliographic format for input, because the individual citations are not restructured by the processes. In the examples above, the keyword information is inserted at the end of the statement by a carriage return and 9 spaces, to facilitate reading.

These citations are to be input as top-level statements in a null file created for the purpose. New citations may be added at any time by loading the file and entering statements at the beginning of the file.

SCENARIO FOR PRODUCING SUBJECT BIBLIOGRAPHIES

6 August 1973

Jeanne North with the aid of Dean Meyer

Although files can be of much larger size, it is good practice to limit file size to 1000 statements or less because of the space consumed by partial copies and updated copies, and because all stages of processing are simpler with smaller files. A data base may consist of a number of files which may be searched in separate operations and the results from each accumulated in a new file as described in this scenario.

Selecting a Search Pattern

The pattern to be searched for will usually be a single term and often be a single word. For the examples above, search patterns might be:

["Reddy"]	["Dartmouth"]
["Audio Response"]	["Computer Access"]
["SUR Note"]	["SUR"] AND ["Reddy"]

Note that the brackets and quotes are parts of the pattern and are to be typed in.

Creating a File for the Resulting Bibliography

Assuming that the product of the search is to be a bibliography all of whose items match the search requirement, a file should be created to contain this bibliography. Once created, a file may be used to accumulate the contents of later searches if this is desired. For purposes of this scenario assume the creation of a file named SPEECH. (CR indicates a carriage return. [] indicates system response, with the exception of use in patterns, as above):

```
[*] N[ull File F:] SPEECH CR
```

Recovering From a Mistake

In the steps that follow, if an incorrect command or other error makes it desirable to abort the effort, the following command will restore the file presently loaded to its condition at the last update.

```
[*] E[ecute] U[nlock File <USER>SPEECH.NLS;2] CR [Really?] CR  
[Yes]
```

SCENARIO FOR PRODUCING SUBJECT BIBLIOGRAPHIES
 Jeanne North with the aid of Dean Meyer

6 August 1973

Executing a Search and Creating the Bibliography

Load the source file of references, called DATABASE in the example below. Type the material outside the brackets (except in the case of the pattern, in which the brackets ARE typed in). Substitute the desired pattern for ["Audio Response"] and substitute the name of the empty null file for SPEECH, inserting it with the comma and parentheses which make it a link.

```
[*] L[oad] F[ile F:] DATABASE CR
[*] G[oto] P[rograms] C[ontent Analyzer type in?] Y[es]
[T:] ["Audio Response"] CR      (Type the brackets)
[Compiling User Program]
[*] E[xecute] A[ssimilate Structure] P[lex]
[to A:] (SPEECH,) CR
[from A:] CR
[L:] CR
[viewspecs V:] 1 CR
[*] v[iew specs] C[hange]
[V:] J CR
[*] U[pdate] F[ile] CR
```

To demonstrate that the items are selected, it is possible, but not necessary, to print out the results. In the command below, viewspec x will print only the first lines of items; viewspec w will print the whole file:

```
[*] P[rint] B[ranch A:] CR
[V:] x CR
```

6 August 1973

SCENARIO FOR PRODUCING SUBJECT BIBLIOGRAPHIES

Jeanne North with the aid of Dean Meyer

Sorting the File

If the source file is in bibliographic form with the author's last name as the first word of each item, the following procedure will sort the output file alphabetically by author:

```
[*] G[oto] S[ort] P[lex at A:] .1 CR
```

```
[ok?] CR
```

To demonstrate that the items have been sorted, it is possible, but not necessary, to print out the results. In the command below, `viewspec x` will print only the first lines of items; `viewspec w` will print the whole file:

```
[*] P[rint] B[ranch A:] .0 CR
```

```
[V:] x CR
```

Printing a Bibliography

A bibliography in a standard form with hanging indention can be produced by adding Output Processor Directives. In statement 0, after the semicolon, insert print directions:

```
[*] I[nsert] T[ext after A:] .0> CR
```

```
[T:] .H1="(any title you like) .Split; (name and/or date)";  
.YBS=1; .PN=0; .IRest=5; .PES; CR
```

```
[*] U[pdate] F[ile] CR
```

When ready to print, give command:

```
[*] L[oad] F[ile] SPEECH CR
```

```
[*] O[utput] D[evice] T[eletype Go?] CR
```

EXPERIMENTAL DEVELOPMENT OF A
SMALL COMPUTER-AUGMENTED INFORMATION SYSTEM

SRI-APC 31 DEC 73 21453
Annual Report to ONR

Appendix D

Scenario for Using the Network Journal

SCENARIO FOR USING THE NETWORK JOURNAL
Jeanne North
Network Information Center

p.1

Network users may send mail to individuals and to groups, input as messages or entire files, through the Network Journal, using SNDMSG or their site's mail system. The mail is converted at NIC into NLS files, Journalized, and sent to specified recipients. Short messages may be received as messages, longer ones as citations to files which may be retrieved immediately, and also later, by using FTP. Mail sent to NIC with a "/" in the user-name field invokes the Net Journal.

SENDING THE MESSAGE OR FILE BY TENEX SNDMSG

Construct user field with slash and NIC idents:

[Users:] sender ident/addressee ident(s)@NIC

e.g. "JEW/DHC MAP@NIC". To send to a group, use group ident, e.g. "JEW/NLG@NIC". To combine Journal SNDMSG with SNDMSG to others, add others' Network addresses after commas, e.g. "JEW/DHC@NIC, PRATT@ISI".

See ARPANET Directory for NIC idents and Network addresses.

[@] SNDMSG <CR>

[Type ? for help]

[Users:] JEW/NGG <SP> DHC@NIC, PRATT@ISI <CR>

[Subject:] Title of Message <CR>

[Message (? for help):] Text of message ...

Note: ↑B allows insertion of a sequential file as the message or at any point in the text of the message:

↑B [(Insert file:)] <directory>file <ALT> [ext ...EOF)]

<↑Z>

[jew/ngg dhc@nic -- ok] (or [-- queued -- timed-out])

[pratt@isi -- ok]

When using SNDMSG through TELNET, change TELNET escape character, to ↑Q for example, allowing ↑Z to be used for end-of-message.

RETRIEVING THE FILE FROM NETWORK JOURNAL

Substitute the citation received, for example "<GJOURNAL>21695", for "<journal>number" and supply a filename in the following:

[@] FTP <CR>

[HOST FTP User process x.xx.x]

[*] CONN <SP> NIC <CR>

[Connection opened]

[*< OFFICE-1 FTP Server x.xx.x - at DAY DATE TIME]

[*] LOG <SP> ANONYMOUS <SP> NIC <CR>

[*] GET <SP> <journal>number.NLS;xnls <CR>

[to local file/ filename <CR> [New file/ <CR>

[< IMAGE retrieve of <journal>number.NLS; started]

[< transfer completed]

[*] DISC <CR>

[*] QUIT <CR>

[@] COP <ALT> [<File list>] file <ALT> [<TO>] LPT: <CR> [OK] <CR>

SENDING A MESSAGE BY TELNET, FTP, OR OTHER MAIL SYSTEM

TELNET by TENEX Users:

```
[@] TELNET <CR>
[User Telnet x.x DATE ....]
[#] NIC <SP> FTP <CR> [is complete.#]
[300 OFFICE-1 FTP Server x.xx.x - at DAY DATE TIME]
MAIL <SP> JEW/RWW <SP> DHC <CR> (pause)
[350 Type mail, ended by a line with only a "."]
Re: Title of Message <CR>
First line of message <CR>
second line of message <CR> ...etc.
. <CR> (pause)
[256 Mail completed successfully]
<↑Z>
[#] DISC <CR>
[#] QUIT <CR>
```

FTP by TENEX Users:

```
[@] FTP <CR>
[HOST FTP User process x.xx.x]
[*] CONN <SP> NIC <CR>
[ Connection opened]
[*< OFFICE-1 FTP Server x.xx.x - at DAY DATE TIME]
[*] QUO <ALT> MAIL <SP> JEW/DHC RWW <CR>
[*] (pause) [Type mail, ended by a line with only a "."]
[*] QUO <ALT> Re: Title of Message <CR>
[*] QUO <ALT> First line of message <CR>
[*] QUO <ALT> second line of message <CR> ...etc.
[*] QUO <ALT> . <CR> (pause)
[*< Mail completed successfully]
[*] DISC <CR>
[*] QUIT <CR>
```

Using Other Mail Systems:

It is not possible to give a generalized scenario for use with all local mail systems.

The general procedure, to be applied to the local mail system, is to supply:

"NIC" as the host name, and

Sender's NIC ident / Addressee's NIC ident as user name

See RFC 543, NIC 17777, for more detail on Network Journal.