# INFORMATION STORAGE AND RETRIEVAL OF COMPOSITE DOCUMENTS: A USER ORIENTED MODEL OF COMPUTER MESSAGE SYSTEMS

by

Dr. Edward A. Fox

TR-85-34

August 1985

# INFORMATION STORAGE AND RETRIEVAL OF COMPOSITE DOCUMENTS: A USER ORIENTED MODEL OF COMPUTER MESSAGE SYSTEMS

Edward A. Fox, Department of Computer Science
Virginia Polytechnic Institute and State University

## ABSTRACT

This paper outlines a new model of computer message systems based on a user-oriented rather than a communications perspective. Most network users spend their time working with various types of 'composite documents' such as mail messages, news, reports, or entries in directories of names. They perform operations that are aided by a logically centralized but physically distributed world-wide information storage and retrieval system. Their concerns should be with creating, filing, submitting, searching, and retrieving documents, not with routing or transmission matters.

Computer networks are rapidly proliferating. Their number and size are increasing, and interconnection is commonplace. High volume applications like mail and news transmission are well established in offices, large corporations, and governmentally supported internets. Computer conferencing and directory assistance services are continuing to evolve. Standards for message formats, transmission protocols, and name/address directories have been recommended by DARPA, NBS, and CCITT.

Research in document modeling, distributing databases, and retrieving text objects is being adapted to improve the functionality of computer message systems. Advanced techniques for text analysis, automatic indexing, query construction, description of comprehensive interest profiles, and retrospective or current awareness retrieval can simplify and improve the effectiveness of user agent software. Initial experimentation has shown the value of recognizing the structure of composite documents.

Current work with the CSNET name server database, and with a collection of messages extracted from AIList digests distributed over the DARPA Internet, indicates that the proposed model accomodates mail, news, and directory assistance. Other work with book passage retrieval suggests that large files can be processed too. It is hoped that integration of these services, along with conferencing and database handling, will occur, as a new network for the Virginia Center for Innovative Technology is proposed and eventually implemented.

**CR Categories and Subject Descriptors:** C.2.3 [**Computer-Communication Networks**]: Network Operations; E.2 [**Data Storage Representations**]: Composite Structures; H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing; H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval; H.4.3 [**Information System Applications**]: Communications Applications-*Electronic Mail*

**General Terms:** Design, Experimentation, Standardization

**Additional Key Words and Phrases:** composite documents, computer based message systems

---

Project funded in part by grants from the National Science Foundation (IST-8418877), the Virginia Center for Innovative Technology (INF-85-016), and VPI&SU College of Arts and Sciences.

## 1. INTRODUCTION

As telecommunication services become more and more available to users of computer systems, improved techniques are needed to provide access to the growing logical store of global information. This paper outlines a new user-oriented model aimed at describing and satisfying these requirements. The model deals with the creation, analysis, storage, and retrieval of composite documents in the world network.

### 1.1. PROLIFERATION OF NETWORKS

A rapid proliferation of computer networks is now taking place around the world. Individuals with terminals or microcomputers are frequently allowed to connect to other computers besides those locally available. Packet switched data networks, ideal for transferring mail and news messages, began with the ARPANET in the late sixties and have now led to global internetworking including North America, Western Europe, and Southeast Asia [10]. Sophisticated systems using distributed processing have simplified communication within some large corporations, where basic issues of naming, authentication, and locating have been resolved [6]. Lower bandwidth, less formal networks have evolved from common software packages; more than 2000 machines intercommunicate using a program called UUCP and more than 500 of those share news and messages through the loosely coupled USENET [25]. Many North American universities have joined BITNET in an effort to simplify exchange of messages, text files, and programs [34], and that network is being extended to Europe as well. CSNET, the computer science network, is a logical network spanning a variety of physical networks such as the DARPA Internet, Telenet, Uninet, and Phonenet [13]. Many other networks are being developed to meet the needs of industry, educational institutions, and governments.

### 1.2. UTILIZATION OF NETWORKS

In the office environment, where small local networks abound, users are generally pleased with new capabilities for handling electronic transmission of mail and other items [4]. Use of computers and communications is becoming a key element of office operations [59]. The creation, transmission, receipt, reading, and filing of mail requires a good deal of the time of modern day office workers, and so their experiences must be considered in the development of new electronic aids [9].

In addition to mail handling, networks are useful for less private communications. Some computer users prefer to receive news electronically [33]. Public bulletin boards are also available for sharing information and comments among participants in various interest groups [39]. Computer conferencing systems provide even better organization and reduce the amount and complexity of work required by each user [58]. Some systems are for special purposes [20]. Others are for general use and are becoming more readily available [55].

To facilitate network utilization, various aids are provided to help identify remote systems or users of those systems. For example, during the evolution of the ARPANET, a name server had to be developed to keep track of the large number of computers attached [56]. For BITNET and CSNET, name servers handle the names of individuals as well as of computers.

## 1.3. STANDARDIZATION

The transmission of mail and other electronic messages frequently requires cooperation between a number of computer and telecommunication systems. Standardization of message formats, transmission protocols, and of names and addresses has become essential.

In the DARPA Internet, a text based memo format is utilized [17]. The U.S. National Bureau of Standards called for the design of a message format standard with a somewhat different orientation [24] which has led to a Federal Standard [52]. More recently, NBS proposed standards for message transfer [53].

For global communications, the need for standardization is also urgent [49]. Development by the CCITT has let to draft recommendations for message handling systems [11] that are appropriate for the various PTTs around the world [50]. In Canada, for example, the EAN system has been implemented along CCITT guidelines [54]. ISO standardization has followed the blueprint laid out by CCITT.

The related problem of global directory assistance is also being considered by standards organizations. It is important to separate out the separate notions of naming, addressing, and routing [44]. A consistent model and conceptual framework for these efforts is essential [57]. The DARPA Internet has developed a domain naming scheme for its host computers [47], but that has been rather difficult to develop and implement. The National Bureau of Standards has broadened the issue to handle names and addresses of computer users in another proposed standard [7]. IFIP Working Group 6.5 has considered the related problems of naming and directory assistance, and has distributed papers dealing with required services [40] and naming conventions [41]. CCITT has likewise inaugurated a study of this important problem.

## 1.4. RELATED INVESTIGATIONS

A variety of research investigations have been undertaken to deal with the many important problems relating to messaging in the global network. Efforts in the areas of office information systems, information storage and retrieval, document modeling, database management, and computer messaging are of particular interest.

In offices, handling of mail and other correspondence is crucial. Croft has developed several experimental test collections for manipulating text in offices [18]. He has considered how information storage and retrieval methods can be applied to office tasks [19]. Integration of these methods with those of database management systems is clearly needed to handle the enormous variety of objects found in offices.

In 1979 Dattola described one of the first systems combining information retrieval and database methods [21]. Naffah advanced a scheme for modeling and manipulating complex objects with text, graphics, and voice components [51]. Various proposals have been made to extend the relational model to support text and retrieval applications [26]. The entity-relationship model could likewise be extended to handle documents [12]. Special architectures for an integrated retrieval and database system have also been proposed [3,5]. Distributed systems of this type are certainly needed for resolving global addressing and directory assistance requirements [35].

Modeling of large, complex text objects has been considered. As manuscript preparation and submission using electronic means becomes more popular, standards for tagging of the variety of types of information are essential [43]. Kimura has suggested a hierarchical model for such objects, where the structure and content are

distinguished and separately manipulable [45].

A variety of tools have been provided for handling mail. SENDMAIL is a comprehensive tool for resolving addressing and distribution problems [1]. MS [16] and MH [8] are two popular packages that serve as user agents, connecting individuals with the mail transport system. More comprehensive systems such as HERMES provide users with powerful editing and searching capabilities [48].

Experimental systems for mail handling are based on more general models than that of memo-based systems. EDMAS utilizes capabilities for mailboxes of originators and recipients and maintains chains to describe interrelationships among messages [2]. DRAGONMAIL [14] has been developed to manipulate 'conversations' [15] rather than mail or conferences. In this model, graph theory [37] can be employed to describe not only the connectivity of computers and their users, but also to describe the relationship among mail items that are physically distributed but could logically be considered as part of a global information store.

## 2. COMPOSITE DOCUMENT MODEL

In order to model the variety of information manipulated by users of computer message systems, it is necessary to describe the items that users are interested in, at the proper level. Access to these information items, which will be called composite documents, can be viewed as the responsibility of an information storage and retrieval system that deals with a global information store. Issues relating to the design and implementation of such a system can be understood in the light of efforts to devise some of the parts essential to the overall scheme. Experimental studies are necessary for validating these ideas.

### 2.1. COMPOSITE DOCUMENTS

A composite document will be defined as an entity containing at least some textual information, that has a particular structure. Creation, storage, location, and manipulation of composite documents therefore requires the ability to deal with both structures and arbitrary amounts of text.

There are many examples of composite documents found in computer message systems. Mail, memorandum, articles, reports, manuscripts, news releases, entries in conferences or conversations, listings in directories or name servers, and parts of databases are all composite documents.

What is important in considering composite documents is the set of operations defined on those objects. In addition to the operations characteristic of database management systems (eg., creation, update, deletion), it is necessary to be able to (recursively) access all components of documents. Furthermore, inexact matching between composite documents must be well defined, so that a similarity in the range of zero to one is produced. Whereas inexact matching of strings has been extensively studied [36], and there are a variety of techniques for matching possibly misspelled names [22], matching of more complex items is problematic. Database systems, in their normal processing of queries, sometimes allow matching of entities that include both data and text, but similarity values are always binary. Some experimental retrieval systems can compute a real-valued similarity, but cannot easily specify separate handling of different types of data or components of a composite object.

This author has proposed extending the vector space model of information retrieval so that an entity can be viewed as made up of various types of data [29]. A special query notation is employed to allow access to each type of data, using an extended form of Boolean logic [29], where the logical operators are given a 'soft' interpretation. Several collections have been gathered to allow experimentation with this model [31]. An experimental retrieval system has been devised to allow search of these collections. Real valued similarities are produced, based on matches among the various types of data [30].

## 2.2. INFORMATION STORAGE AND RETRIEVAL

One can represent the various objects that are handled by computer message systems as composite documents. From the perspective of users of such systems, the global computer network can be considered as a large, logical information storage and retrieval system. Users of information retrieval systems can add to collections, search collections, or be notified of interesting items when the system carries our current awareness or selective dissemination of information (SDI) operations based on predefined profiles. Matching can be based on factual information (eg., recipient name, date of distribution), subject information (eg., topic, words in message body), or interrelationship information (eg., citation of previous message).

Mail handling is easily modeled by these operations. Message submission is equivalent to addition of a composite document to the global store. Receipt of mail is the same as examining the results of SDI operations where the profile selects items with the current user as recipient. Replying is similar to submission, where the new item includes interrelationship information (much like a citation in a paper) connecting it to an earlier message.

Bulletin boards and news services can be modeled using either factual or subject information components of the composite documents. Those components of a document would identify the news group or groups which should receive that document. Members of groups would have a component in their SDI profiles to select any item connected with the group. The same scheme can be employed for computer conferences, where messages and profiles could both refer to conferences. Since composite documents can have an arbitrary number of components, there is no problem with dynamically creating or terminating conferences. Since interrelationship information of arbitrary complexity is allowed in profiles and documents, conversation-based mail systems can also be implemented using the proposed model.

The global directory that describes computer systems and users of those systems is also easily modeled. Entries in that directory are composite documents, which may include descriptive elements, and can be searched using exact match techniques or other approaches. Groups or distribution lists are simply retrievable sets defined over the global store. Finding a user is equivalent to carrying out a distributed search. Added flexibility arises due to the handling of inexact matching of queries and directory entries.

It is clear that this information storage and retrieval perspective allows modeling of many of the uses of computer message systems. What is unique about this model, however, is that in addition to handling the typical messaging situations, there are many new possibilities. Profiles can be made up using any combination of factual, subject, and interrelationship information. Messages from particular users or groups

can be automatically selected or rejected. Messages that are urgent can be ranked as more similar to a profile and hence retrieved more rapidly. A variety of profiles can be created by a user, and can operate disjunctively, allowing ready separation into categories. Searches for information can be initiated at will, so that old messages can be viewed in addition to new ones.

While this model is sufficiently general to encompass other proposed message handling schemes, convenient implementations are not presently available. The following sections describe initial efforts to design, implement, and test various aspects of this model.

## 2.3. DESIGN, IMPLEMENTATION, TESTING

The proposed model calls for a distributed system combining features of information storage and retrieval, database management, and computer messaging systems. While current work addresses the development of such a facility, previous work has dealt with various components of the overall scheme. Since the information storage and retrieval model for computer messaging has evolved out of the need to accomodate bibliographic searching of composite documents, it is important to consider the approach and results obtained in that domain.

The SMART experimental information retrieval system has been redesigned and implemented to allow manipulation of composite documents [30]. The vector space model incorporated in earlier versions of SMART allowed documents to be represented by a list of attributes, along with importance weights for each attribute. In the new version of SMART, documents are viewed as having attributes of various types. Similarity computations are based upon the effects of matches of attributes distributed among each of the several types.

Two collections of bibliographic items, one dealing with computer science and the other with information science, have been gathered for testing of the composite document model [31]. The first collection has author, journal, volume, number, title, abstract, keywords assigned, classification categories, citations to other articles, citations from other articles, and measures of both cocitation and bibliographic coupling strengths with other articles. The second has similar entries, but lacks the direct citations and the bibliographic coupling data. Tests on both collections have been run using a variety of questions. Search results are best when a combined similarity measure is used, counting all types of information in proper proportion [28].

In addition to searching collections of distinct composite documents, it is important to find component composite documents embedded in large text objects. Several books published by the Baha'i Publishing Trust have been obtained in machine readable form. One book, *Gleanings from the Writings of Baha'u'llah,* with 165 sections and 345 pages, has 707 paragraphs. A set of 52 queries have been organized in several different forms. The original query terms are present in one set, those terms and related terms (eg., from thesaurus categories) in another, and terms plus pages (obtained from the back of book index) in yet another. Comparison of the effectiveness of searches for paragraphs, pages, or sections, using each of the query forms constructed, is now in process. Insight into handling of large text objects, with various types of information present, should result and help validate proposed models for composite documents with multiple levels of structure.

Based on the initial success with composite documents containing bibliographic information, and with retrieval of passages in large text objects, adaptation of those methods to computer networks has begun. Two separate activities, one with the CSNET name server database, and another with the archives from the ARPANET digest, AIList, are underway.

The CSNET name server is implemented as a centralized database manager with distributed local agents. Entries from the name server database are being analyzed to investigate the usefulness of clustering, based on various similarity measures, for the purpose of identifying distributed 'invisible colleges.' Refinement of the storage and retrieval component of the name server is also being considered, to facilitate more flexible searching of this network directory.

AIList is mailed a few times each week as a digest including a number of messages of widely different types dealing with current concerns of people studying artificial intelligence. Digests are one type of composite document, as are messages. Manual analysis of several years of archives is underway as a prelude to development of automatic text analysis methods. Using the Prolog language, an analyzer similar in some ways to FRUMP [23] will be devised to classify and separate out key information items in preparation for automatic indexing. Since unrestricted text in a variety of formats is involved, very flexible pattern based methods are needed [38]. The output of such an analysis will include frames linked together to represent the content and structure of messages [46]. Eventually, students at Virginia Tech will use microcomputer based software [32] to search old archives or filter incoming messages. Further refinement, especially of the user interface for handling messages and profiles, will also draw on previous related work [42].

As these various efforts continue, it is hoped that the information storage and retrieval model of handling composite documents will be further developed and find clear application in the domain of handling computer messaging in the evolving global network.

## 3. SUMMARY AND FUTURE

In light of the rapid proliferation of computer networks, the success of many different standards organizations, and the interest of research workers in various fields, a new model for handling global message requirements has been proposed. Based on manipulation of composite documents, which represent any one of the various entities handled by such systems, a user oriented perspective has been followed. Viewing the global information store as a logical whole that is physically distributed, access to that store involves information storage and retrieval operations. Computer mail, bulletin boards, conferences, conversations, directory assistance, and other types of systems can be modeled by such an approach.

Research to date has suggested that proper handling of composite documents requires a more complete model than that commonly considered. Studies of passage retrieval from books, clustering of CSNET name server entries, and analysis and retrieval of AIList digest messages are now underway. A study is beginning of the needs for the Virginia Center for Innovative Technology, where it is essential to integrate computer mail, news, conferencing, and database access in a new logical network to serve educational, industrial, and governmental institutions in the

Commonwealth. It is hoped that as these investigations continue, the role of composite documents in networking will become better defined and serve as a conceptual and practical tool for implementing services according to the user-oriented information storage and retrieval model.

## REFERENCES


(1) Allman, Eric. SENDMAIL - An Internetwork Mail Router. *UNIX Programmer's Manual.* Berkeley Release 4.2; 1983.

(2) Almes, Guy; Black, Andrew; Bunje, Carl; Wiebe, Douglas. Ed. as: A Locally Distributed Mail System. Technical Report 83-07-01. Dept. of Comp. Sci., Univ. of Wash.; July 1983.

(3) Bartschi, M.; Frei, H. P. Adapting a Data Organization to the Structure of Stored Information. In Gerard Salton and Hans-Jochen Schneider, Ed. . *Research and Development in Information Retrieval, Proc., Berlin, May 18-20, 1982.* Berlin: Springer-Verlag; 1983: 62-79.

(4) Bikson, T. K.; Gutek, B. A. Advanced Office Systems: An Empirical Look at Utilization and Satisfaction. Rand Report N-1970-NSF. Santa Monica, CA: The Rand Corp.; Feb. 1983.

(5) Biller, Horst. On the Architecture of a System Integrating Data Base Management and Information Retrieval. In Gerard Salton and Hans-Jochen Schneider, Ed. . *Research and Development in Information Retrieval, Proc., Berlin, May 18-20, 1982.* Berlin: Springer-Verlag; 1983: 80-97.

(6) Birrell, Andrew D.; Levin, Roy; Needham, Roger M.; Schroeder, Michael D. Grapevine: An Exercise in Distributed Computing. *Commun. ACM.* 25(4): 260-274; April 1982.

(7) Bolt Beranek and Newman, Inc. Naming and Addressing in Computer Based Message Systems. Draft Report No. ICST/CBOS-82-4. Dept. of Commerce, National Bureau of Standards; Aug. 1982.

(8) Borden, Bruce. MH: A Mail Handling System for UNIX. *UNIX Programmer's Manual.* Berkeley Release 4.2; 1983.

(9) Bruder, J.; Moy, M.; Mueller, A.; Danielson, R. User Experience and Evolving Design in a Local Electronic Mail System. In Ronald P. Uhlig, Ed. *Computer Message Systems.* Amsterdam: North-Holland; 1982: 69-78.

(10) Casey, M. Packet Switched Data Networks: An International Review. *Inf. Tech.: Res. & Dev..* 1(3): 217-244; July 1982.

(11) CCITT. Message Handling Systems. Draft Recommendations X.400, 401, 408, 409, 410, 411, 420, 430. Int'l Telephone and Telegraph Consultative Comm.; Fall 1983.

(12) Chupin, J. C.; Joloboff, V. A Data Model for Office Systems. In N. Naffah, Ed. *Office Information Systems.* Amsterdam: North-Holland; 1982: 39-56.

(13) Comer, Douglas. The Computer Science Research Network CSNET: A History and Status Report. *Commun. ACM.* 26(10): 747-753; Oct. 1983.

(14) Comer, Douglas E.; Peterson, Larry L. DRAGONMAIL: A Prototype Conversation-Based Mail System. *Proceedings of the USENIX 1984 Summer Conference;* June 1984: 42-51.

(15) Comer, Douglas E.; Peterson, Larry L. Conversations: An Alternative to Memos and Conferences. *Proceedings of the Guelph Workshop on Computer Conferencing and Electronic Messaging;* January 1985: 13-21.

(16) Crocker, David H. Framework and Functions of the 'MS' Personal Message System. R-2134-ARPA. Santa Monica, CA: The Rand Corp.; Dec. 1977.

(17) Crocker, David H. Standard for the Format of ARPA Internet Text Messages. RFC 822. ARPANET Networking Group; Aug. 1982.

(18) Croft, W. Bruce. Experiments with Automatic Text Filing and Retrieval in the Office Environment. *ACM SIGIR Forum.* 16(4): 2-9; Spring 1982.

(19) Croft, W. Bruce. Applications for Information Retrieval Techniques in the Office. *ACM SIGIR Forum and Proc. 6th Annual Int. ACM SIGIR Conf. on R&D in IR.* 17(4): 18-23; Summer 1983.

(20) Daney, Charles. The VMSHARE Computer Conferencing Facility. In Ronald P. Uhlig, Ed. *Computer Message Systems.* Amsterdam: North-Holland; 1982: 115-127.

(21) Dattola, Robert T. FIRST: Flexible Information Retrieval System for Text. *J. Am. Soc. Inf. Sci..* 30(1): 9-14; 1979.

(22) Davidson, Leon. Retrieval of Misspelled Names in an Airlines Passenger Record System. *Commun. ACM.* 5(5): 169-171; May 1962.

(23) DeJong, Gerald. An Overview of the FRUMP System. In Wendy G. Lehnert and Martin H. Ringle, Ed. . *Strategies for Natural Language Processing.* Hillsdale, NJ: Lawrence Erlbaum Assoc.; 1982: 149-176.

(24) Deutsch, Debra. Design of a Message Format Standard. In Ronald P. Uhlig, Ed. *Computer Message Systems.* Amsterdam: North-Holland; 1982: 199-220.

(25) Emerson, Sandra L. Usenet: A Bulletin Board for UNIX Users. *Byte Magazine.* 99(8): 219-236; Oct. 1983.

(26) Fox, Edward A. Implementing SMART for Minicomputers Via Relational Processing with Abstract Data Types. *Joint Proceedings of SIGSMALL Symposium on Small Systems and SIGMOD Workshop on Small Data Base Systems, ACM SIGSMALL Newsletter.* 7(2): 119-129; Oct. 1981.

(27) Fox, Edward A. Automatic Document and Passage Retrieval Methods: Aids to Searching the Baha'i Writings. *Proceedings Annual Meeting Assoc. for Baha'i Studies;* Quebec; April 1981.

(28) Fox, Edward A. Combining Information in an Extended Automatic Information Retrieval System for Agriculture. *Infrastructure of an Information Society (Proc. 1st Int. Information Conf. Egypt, 13-16 Dec. 1982);* 1983.

(29) Fox, Edward A. Extending the Boolean and Vector Space Models of Information Retrieval with P-Norm Queries and Multiple Concept Types. Dissertation. Ithaca, NY: Cornell Univ.; Aug. 1983.

(30) Fox, Edward A. Some Considerations for Implementing the SMART Information Retrieval System under UNIX. Tech. Report 83-560. Ithaca, NY: Cornell Univ., Dept. of Comp. Sci.; Sept. 1983.

(31) Fox, Edward A. Characterization of Two New Experimental Collections in Computer and Information Science Containing Textual and Bibliographic Concepts. Tech. Report 83-561. Ithaca, NY: Cornell Univ., Dept. of Comp. Sci.; Sept. 1983.

(32) Fox, Edward A. Information Retrieval with Undergraduate Microcomputers. *Coll. papers Amer. Soc. Inf. Sci. 19th ASIS Mid-Year Meeting;* 20-23 May 1984: 253-264.

(33) Frost, Martin. The News Service System. Stanford Art. Intell. Lab. Operating Note 72.2. Palo Alto, CA: Stanford University; Feb. 1975.

(34) Fuchs, Ira H. BITNET - Because It's Time. *(IBM Corp.) Perspectives in Computing.* 3(1): 16-27; March 1983.

(35) Garcia-Luna, Jose J.; Kuo, Franklin F. Addressing and Directory Systems for Large Computer Mail Systems. In Ronald P. Uhlig, Ed. *Computer Message Systems.* Amsterdam: North-Holland; 1982: 297-314.

(36) Hall, Patrick A.V.; Dowling, Geoff R. Approximate String Matching. *ACM Computing Surveys.* 12(4): 381-402; 1980.

(37) Harary, Frank. *Graph Theory.* Reading MA: Addison Wesley; 1969.

(38) Hayes, Philip; Mouradian, George V. Flexible Parsing. *Amer. J. Comp. Ling.*. 7(4): 232-242; Oct.-Dec. 1981.

(39) Horton, M. How to Read Network News. *UNIX Programmer's Manual, 4.1 Berkeley Software Distribution, Vol. 2.* 1979.

(40) IFIP WG 6.5. Naming and Directory Services for Message Handling Systems. Working Paper, Version 4. IFIP WG 6.5; July 1983.

(41) IFIP WG 6.5. A User-friendly Naming Convention for Use in Communication Networks. Working Paper, Version 3. IFIP WG 6.5; March 1984.

(42) Jacob, Robert J. K. Using Formal Specifications in the Design of a Human-Computer Interface. *Commun. ACM.* 26(4): 259-264; April 1983.

(43) Jennings, Margaret. The Electronic Manuscript Project. *Bulletin of the Am. Soc. Inf. Sci.*. 10(3): 11-13; Feb. 1984.

(44) Kerr, Ian H. Interconnection of Electronic Mail Systems - A Proposal on Naming, Addressing and Routing. In Ronald P. Uhlig, Ed. *Computer Message Systems.* Amsterdam: North-Holland; 1982: 315-326.

(45) Kimura, Gary D. A Structure Ed. tor and Model for Abstract Document Objects. Tech. Report No. 84-07-04. Seattle, Washington: Univ. of Washington, Dept. of Comp. Sci.; July 1984.

(46) Minsky, Marvin. A Framework for Representing Knowledge. In P. Winston, Ed. *The Psychology of Computer Vision.* New York: McGraw-Hill; 1975.

(47) Mockapetris, P. Domain Names - Concepts and Facilities. RFC 882. Univ. of S. Calif, Information Sciences Inst.; Nov. 1983.

(48) Mooers, Charlotte D. The Hermes Guide. Report No. 4995. Cambridge, MA: BBN, Inc.; Aug. 1982.

(49) Myer, Theodore H. Global Messaging ... Issues and Approaches. *J. Telecommunication Networks.* 1(2): 173-187; Summer 1982.

(50) Myer, Theodore H. Standards for Global Messaging: A Progress Report. *J. Telecommunication Networks.* 2(4); Winter 1983.

(51) Naffah, Najah. Design Issues of Presentation Protocols in Office Systems. ARC 2.509. Rocquencourt FRANCE: INRIA; 1980.

(52) National Bureau of Standards. Message Format for Computer-Based Message Systems. Federal Information Processing Standards Publication (FIPS PUB) 98. NTIS; March 1983.

(53) National Bureau of Standards. Specification of the Message Transfer Protocol. Draft Report. Inst. for Comp. Sciences and Tech., NBS; July 1983.

(54) Neufeld, G.; Sample, R.; Hilpert, B.; Derring, S.; Demco, J. The EAN distributed message system. User Manual Version 1.0. Vancouver, B.C. CANADA: Dept. of Comp. Sci., Univ. of British Columbia; Nov. 1983.

(55) Palme, J. COM/PortaCOM Conference System Design Goals and Principles. Report. Stockholm University Computing Centre; March 1984.

(56) Pickens, John R.; Feinler, E. J.; Mathis, J. E. The NIC Name Server - A Datagram Based Information Utility. *Proc. 4th Berkeley Workshop on Distributed Data Management and Computer Networks*; August 1979.

(57) Sirbu, Jr, Marvin A.; Sutherland, Juliet B. Naming and Directory Issues in Message Transfer Systems. *Proc. IFIP 6.5 Working Conf., Nottingham, England*; May 1984.

(58) Strom, B. Ivan. Computer Conferencing - Past, Present, and Future. In N. Naffah, Ed. *Office Information Systems.* Amsterdam: North-Holland; 1982: 287-316.

(59) Uhlig, Ronald P.; Farber, David J.; Bair, James H. *The Office of the Future: Communication and Computers.* Amsterdam: North-Holland; 1979.