COMMUNICATIONS
FORUM

MAKING ELECTRONIC MAIL MORE INTELLIGENT

Thomas Malone,
M.I.T.
Kenneth E. Mayers,
Digital Equipment Corporation

October 31, 1985

Genga Arulampalam, Rapporteur

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Massachusetts Institute of Technology
Communications Forum

MAKING ELECTRONIC MAIL MORE INTELLIGENT

Thomas Malone,
M.I.T.
Kenneth E. Mayers,
Digital Equipment Corporation

October 31, 1985

Genga Arulampalam, Rapporteur

MAKING ELECTRONIC MAIL MORE INTELLIGENT

I. DIGITAL'S ELECTRONIC MAIL SYSTEM (EMS)

Dr. Kenneth E. Mayers, Digital Equipment Corporation

EMS History

Mayers began by reviewing the history of Digital's Electronic Mail System (EMS). Electronic mail started as a single node system about six years ago and has grown to over 300 nodes and 40,000 mail boxes worldwide today. Digital implemented EMS to keep up with the rapid movement of employees between offices and facilities and to replace some of the paper communication which was growing at 40% per year. According to Mayers, it was obvious that Digital had to begin to move from the "paper page" to the "electronic page."

Digital initiated a pilot study in January 1978 which ran for 18 months. Usage grew rapidly-- from 40 users at the beginning of the pilot to 750 users by the end. The company utilized various forms of feedback, including questionnaires and face-to-face interviews, to understand the popularity of the system and study its impacts. The researchers' conclusions were that EMS had the potential of becoming as pervasive and important as the telephone and that within ten years everyone at Digital (with the exception of production line workers) would have an EMS account. In fact, the spread of EMS has proceeded much faster than any of the pilot study researchers expected. Today, only 6 1/2 years after cut-over to a full-scale electronic mail system, virtually every professional at Digital has an EMS account.

In August 1979, Digital initiated a full-scale, three-node EMS system. It soon grew to a multi-node system with gateways to other internal Digital communication networks and to other public communication networks. Today the system is capable of delivering a hard copy message to anyone within the Digital corporation and, through an interface, can transmit a telex anywhere in the world. Mayers believes that EMS was probably the first office-oriented mail system in heavy daily use in the United States.

Results of EMS Operation

During the last 6 1/2 years of operation, EMS has been studied closely. One interesting findings was the shift in primary user over the 1981-1982 time period. In 1981, 63% of users accessed the system themselves while 37% accessed the system through an intermediary. Two years later, these figures had shifted to 91.6% direct access and 8.4% via an intermediary. Mayers explained this change as managers becoming comfortable enough with the system to use it themselves rather than relying on their secretaries. He stated that the ability to use either a menu mode and a command mode contributed to the system's high ease of learning. A majority of users (72%) found that they were comfortable with the system in less than one month. Another measure of the integration of EMS into everyday work was the substitution of single-purpose, shared EMS terminals with multi-purpose, single-user terminals. In this way, EMS access became more convenient, both in the office and at home. Because three out of five Digital professionals have terminals at home, access to EMS over standard terminals is important.

Mayers summarized the strengths and weaknesses of EMS as follows:

STRENGTHS:
- -saves time
- -easy to use
- -reduces phone traffic
- -reduces paper flow
- -accessible from any location

WEAKNESSES:
- -access problems (both largely eliminated by increased)
- -insufficient accounts (both largely eliminated by increased number of accounts)
- -fear of junk mail (less of a problem than expected)

Challenges to Expansion

Mayers expressed the challenges to expansion as the problem of supporting additional users and a larger absolute volume of mail while limiting the number of users serviced by a given node. Support for remote nodes and directory control become especially difficult as an electronic mail network grows. System accessibility can suffer if node and network capacity is strained.

Increased intensity of use can pose as much of a problem as the growth in the number of users. Mayers explained that EMS usage is similar to telephone usage, with a few very heavy users and many moderate to light users. He defined intensity of use as follows:

Average Use	Heavy Use
1-3 times daily	3 times daily
15 minutes total	45 minutes total
12 in, 2 out	50 in, 12 out

He also noted that EMS usage peaks in the early morning and late afternoon. About 22% of subscribers log-on between 8 and 9 a.m. and about 20% log-on between 3 and 5 p.m. The slowest time period is the 12 to 1 lunch hour, when just over 10% of users log on. Digital applies usage intensive billing to the heaviest users.

Economics of EMS

Digital estimated the cost savings resulting from the usage of the EMS. Although it is difficult to measure all benefits in such a situation, it was decided to estimate cost savings from the reduction in telephone calls and paper mail. The resulting cost avoidance was estimated to be $2 million in 1981 (assuming a base of 3000 users) and $8 million in 1983 (assuming 8000 users). Mayers added that EMS' versitility in carrying graphics as well as text increases its utility to users.

Evolution of EMS

According to Mayers, EMS is two systems-- a user interface provided by Digital's office systems (e.g. All-in-One) and a message transport service which delivers the messages to the relevant user(s). He noted that Digital has multiple mail systems, all of which "talk" to one another. EMS was

designed to meet the messaging needs of corporate professionals. At present, Digital's computer network has 6000 registered nodes.

Mayers traced the development of EMS from its beginning in 1979 as a collection of dedicated mail machines serving dumb terminals to its evolution by 1983 to a layered system providing multiple services (e.g. mail, text editing, data base access) through a single terminal running on a general purpose transport facility (i.e. DECnet). This layered design separates the user interface from the transport function and allows each to develop independently in a manner best suited to achieving its goals. Their was a parallel evolution in the investment per user and number of users per distributed office machine:

	1979	1983
Investment per user	$3000	$4000
-per terminal	2000	2000
-per host-share	1000	2000
Subscribers per machine	500	250

He projected forward to 1987 when he expected an integrated office network to provide integrated text, voice and image manipulation and transport to multiple interconnected local area networks (LANs). These LANs would serve about 1500 subscribers each and would connect PCs to both local host-based servers (providing access to printers, data bases, communication services, and local mail systems) and remote networks (including a worldwide EMS). He expected the investment per user to jump to $5000, most of which ($4500) would be for advanced local work-stations. The exact location of a specific service or user would be tansparent to the user. The network would be smart enough to know where and how to get the requested service and provide it to the user.

Mayers summarized his requirements for a message delivery system by stating that it should:

-co-exist with data processing applications
-support all common mail packages
-use a standard mail-interchange protocol
-permit decentralized control of mail nodes
-take an innovative approach to the directory management issue (i.e. how to track frequent employee relocations)
-emphasize management information

He noted that current network design delivered messages to local cluster servers. These servers had detailed directories for all cluster users and accomplished the final message delivery task. At present, the sender needs to know what cluster (e.g. Los Angeles) to send a message to. In the future, centralized directories will eliminate the need to identify the target cluster. All the user will need to do is specify the recipient's full name.

Mayers stressed the complementary nature of voice mail and text mail. He did not believe that one would displace the other. He also noted that image mail will not be cost effective until Group IV FAX is available and image and text can be mixed on a single page and scanned by a single FAX device.

Mayers closed by summarizing the evolution of the concept of "messaging" and increasing messaging integration as follows:

	Messaging	Mediator
Increasing integration	Person to person	None
↓	Machine to person	Message generators
↓	Person to machine	Parsers
↓	Machine to machine	Software bus

II. MAKING ELECTRONIC MAIL MORE INTELLIGENT

Dr. Thomas W. Malone, Sloan School of Management, MIT

Overview

Malone identified one of the key problems "lurking behind" any electronic mail systems as the difficulty of promoting effective information sharing within organizations. As the geographic scope of electronic mail systems expands and it becomes possible to broadcast messages to a very wide audience, the dual problems of identifying the appropriate recipient(s) for a message and avoiding information overload become important. Malone suggested that Artificial Intelligence (AI) techniques could help to improve the message routing problem and provide, in essence, "expert support systems" for users. He noted that improved group communications and information sharing could support entirely different organizational forms in the future.

Sharing Information in Organizations

According to Malone, the challenge of better information sharing within groups is composed of two related problems:

> -information overload (too much unwanted information given to the individual), and
>
> -relevant information scarcity (not enough of the information really needed for decision-making available to the individual).

Previous solutions to this problem include electronic mail (i.e. making it easier to get information to other members of the organization), distribution lists (in which users add or subtract their names from centralized mailing lists) and computer conferencing (in which each mailing list is focused on a particular topic of common interest). The latter two approaches operate by limiting the number of messages routed to a given user but still require the user to scan an often large number of messages to determine actual relevance. Malone stressed that AI goes beyond these previous solutions by allowing the user to formalize the "rules" he or she uses for sorting and prioritizing incoming mail and relying on an "intelligent mail sorter" to help with this mail processing task. A similar procedures could be used for handling outgoing mail. The goals of an AI-based mail routing system would include:

-finding interesting messages in a pool of "public" information,
-sorting incoming messages into categories,
-prioritizing messages based on importance and urgency, and
-finding appropriate receivers for an outgoing message.

Mechanics of Mail System

Malone illustrated the mechanics of an AI-based electronic mail system by describing a system under development at MIT. He focused on the system's ability to aid the user in composing messages, sorting incoming messages into convenient categories and searching a public access information pool for interesting messages.

The MIT mail system provides users with screen-based message "templates" which permit the user to compose common types of messages quickly and also facilitates the system's scanning and processing of messages. These templates outline a skeleton message and leave blanks for the user to complete. For example, the meeting announcement template requests that the user specify the message sender and receiver, the meeting time, place, subject and sponsor. The template provides default values (e.g. sender = person who logged on) and a list of common alternatives (e.g. sponsor = Sloan School or Center for Information Research) which are selectable from pop-up menus on the screen. Users can develop a unlimited number of such templates.

The system also provides a "production rule" editor which uses the same message templates to allow the user to specify the rules to be used to route, sort or delete different messages. For example, if a user is never in the office on Tuesdays, he or she can specify that any message whose date field indicates a date falling on a Tuesday should be deleted (or stored in a special "hold" file). Malone pointed out that this procedure improves on previous keyword scanning techniques which can not differentiate between a date which falls on a Tuesday and the word "Tuesday" used in the text (e.g. "The Tuesday Massacre"). Another, very interesting example cited by Malone was the ability to sort messages based on the "social" standing of the sender. For example, if the sender were Silk or Siegel (deans of the Sloan School) most (Sloan School) users would want to file that message in a "priority" folder.

Malone also described the "Information Lens," a public pool of messages that performed the routing of messages between senders and recipients that did not know each other or were not aware of each other's interests. If a user sends a message to the Lens, the system will scan the message for critical descriptive elements (e.g. subject, sender) and send it to other individuals who have expressed interest in receiving such messages. In future versions of the system, a user will be able, at any time, to scan the Lens for interesting information that has been sent to it in the past. A user who sends a message to the Lens is implicitly agreeing that anyone in the organization may view the entire contents of the message. Malone noted that the Lens concept thus raises many important legal and organizational issues about how to use such technology.

He also stressed that the system has the important but subtle quality of providing an incremental adoption path. There are many levels of usage, each of which provides additional benefits. However, there are no "cliffs" that the user has the climb before the user can begin gaining value from the system.

Criteria for Information Sharing

Malone suggested that there were three criteria that the Information Lens could use to share "public" information within an organization. These included a cognitive approach, social approach and economic approach.

The cognitive approach is based on the assumption that users can specify their interest in terms that can be applied to characteristics identified for each message. For example, a message requesting information about AI or Lisp would use a keyword search on the subject field of a meeting message type or on the text field of other message types to extract possibly relevant communications. The general process involves giving users access to simplified knowledge representation, using "inheritance lattices" for common message types, senders using templates to construct messages and receivers using rules to describe interests, sort categories and set priorities.

The social approach filters messages, on part, based of the characteristics of the sender. Malone explained a more elaborate use of this approach by noting that no researchers he knew read all the periodicals in their field. Instead, they relied on recommendations and references to important articles from colleagues. This implicit form of communication "endorsement" could also be used by the Lens mail filter.

Finally, the economic approach assumes both that a person's time and information has value. This method would require that a sender "buy" a person's time as reimbursement for reading a message or that a recipient "pay" for especially desirable or valuable information. Malone said that the proper method for valuing a person's time and the value of information is unclear. However, Malone felt that the concept may provide a powerful means for both reducing certain types of junk mail and encouraging the flow of more useful information.

Summary

Malone summarized the key elements of the proposed AI-based electronic mail system as:

1. Knowledge representation
   -semi-structure messages represented as frames
   -production rule sets for processing messages
   -inheritance lattice for message types

2. "Direct manipulation" editors for rules, messages and message types

3. Incremental adoption path for individuals and groups

Organizational Impact

Malone hypothesized that the more effective and efficient sharing of information may support the development of entirely new organizational forms. He first traced the changes in dominant organizational forms in American business since the mid 19th century, moving from decentralized markets to functional hierarchies to product hierarchies. He then suggested a path that organizational structure might take in the future given the significant

changes in information sharing technologies. He predicted a trend toward "decentralized market" organizational structures, also termed "adhocracies" by Toffler and Mintzberg and "networks" by Naisbitt. The key structural characteristics of these new organizations would include:

-decentralized decision-making
-many-to-many communication
-shifting networks rather than rigid hierarchies

III. SPEAKERS' COMMENTS AND RESPONSES TO QUESTIONS

A question was raised regarding the uselfulness of voice mail. Mayers mentioned that voice mail is not especially convenient in its present state of evolution. There are varying degrees of linkage between voice mail systems and text mail systems. Commenting on the ability to move between the two systems, Mayers said that in his experience those people who had experience in text mail found voice mail a convenient addition. However, the same adoption path did not apply in reverse-- those with voice mail experience did not necessarily know about text mail and were not motivated to learn about such systems.

A question was raised about the potential for natural language foreign language translation. Malone said that such translation was not presently possible for unrestricted message texts; however, that translation of limited, clearly defined messages elements using templates may be feasible.

Regarding the existence of something similar to registered mail in an EMS, Mayers stated that there were basically two types of such services-- return receipt mail (which merely indicates that the recipient got the message, not that that person read or understood the message) and read/put away (which indicates that the person opened the message). Mayers said that Digital had an average of 6.2 addressees per message. They are considering the use of logging devices to track various message usage patterns.

In response to a query about the compatability of various Digital systems with other systems, Mayers replied that they had adopted international standards in principal and were gradually moving to complete usage.