Electronic Mail and Message Systems
Proceedings of the AFIPS Workshop on
Technical and Policy Perspectives
Washington, DC December, 1980.

1. The Information Marketplace

Michael L. Dertouzos
Professor and Director

Laboratory for Computer Science
Massachusetts Institute of Technology
Cambridge, Massachusetts

Introduction

By *Information Marketplace* I mean the collection of people, computers, communications, software and services that will be engaged in the intra-organizational and inter-personal informational transactions of the future. These transactions will involve the processing and communication of information under the same economic motives that drive today's traditional marketplace for material goods and services. The Information Marketplace already exists in embryonic form. I expect it to grow at a rapid rate and to affect us as importantly as have the products and processes of the industrial revolution.

To sharpen up these abstractions, let us try to imagine the makeup of the Information Marketplace from a point of view that is 20 years ahead:

> *Large organizations of the year 2000 have been using computers and communications since the late 1980's to communicate business data, electronic memos and still images among their own plants. Automated inter-organizational transactions have grown substantially in the early 1990's, and the toy personal computers of the early 1980's have become useful and powerful machines owned by small businesses and by many individuals. Office automation has come of age and has led to increased productivity, and to reductions in the use of paper and travel for certain routine activities. A wealth of private and public networks interconnect all of the machines which number in the ten millions. Entrepreneurs and a new breed of information companies offer a variety of legal, financial, medical, recreational, educational and governmental information services for a fee. Many traditional ways of doing business have changed. For example, advertising is done in reverse, by a service that responds to consumer inquires with products and services that match. An informational labor force supplies, and many people and organizations consume, all of these services from remote rural, or inner city locations.*

This paper summarizes the reasons that will lead us to the Information Marketplace, the underlying technology that makes it possible, some of the future services that the Information Marketplace will support,

and certain legal and socio-economic consequences that are likely to follow. Electronic mail and message systems will play a central role in their own right, and perhaps a major supporting role, in a host of other services that can be envisioned. The approach taken in this paper, while based on a recent book[1] and on current trends, is ultimately the product of the author's imagination and personal bias.

The Driving Force

The main force driving us toward the Information Marketplace is the ongoing relentless improvement of some 30% per year in performance/cost and size/cost of primary solid-state memories and processor components. This improvement which has been going on for some twelve years is expected to continue well into the 1980's and early 1990's. By the end of this century a 256-fold improvement is likely relative to today, leading to a cost of perhaps $50 (1980 dollars) for storing one million characters in a computer memory. This means, for example, that a personal information base equivalent to 100 books may be stored for the price of an automobile. The far more ambitious undertaking of storing the world's written knowledge would still be very expensive but not prohibitive at about one half billion dollars per LOC.[2] These expected hardware improvements are so huge that were they to happen in the field of personal transportation, they would promise by analogy a future price of $10 for today's cars or a future fuel efficiency of 5,000 miles/gallon at today's car prices. The incredible socio-economic impact of such an absurd transportation revolution is, by the same analogy, indicative of the socio-economic consequences of the Information Marketplace.

The second important technological force behind the Information Marketplace is the ongoing evolution of communications technology: Satellite communication makes possible the transmission of data between any two points on earth at affordable and progressively decreasing real costs. Local-network technology makes possible the communication of information among a few tens or hundreds of interconnected machines in the same building. Glass fibers promise substantial increases in speed and decreases in the cost of land-line communications. Mixed-media packet network techniques have already evolved to a level that makes possible the transmission of voice, computer data and images in digital form, hence with increased noise immunity and with the ability to easily mix and process such data by machine.

The above two major trends in *computer hardware* and *communications* have already caused a marked growth in the creation of new companies and a substantial change in the strategies adopted by existing organizations. The picture, however, is not all that rosy! Unlike the hardware with its spectacular and predictable trends, the software needed to make the hardware useful continues to be very expensive. Two reasons can be identified for this imbalance between hardware and software cost trends. The first is generic to the computer field where there are no "natural laws" and few formal ways to design effectively new programs. This weakness is evident in teaching programming -- there is little that a master programmer can explicitly teach to an aspiring youth. Apprenticeship, talent and other intangibles seem to distinguish a programming athlete from a plodding programmer, with staggering productivity differences of 10 to 1 or more. The second reason for the high cost of software stems from the historic desire of individual users to tailor programs to their varying needs rather than to conform their needs to a standard product. Improvement in this area is already visible with the advent of standard application programs for the small user who can now afford the

---

[1] M.L. Dertouzos and J. Moses, *The Computer Age: A Twenty-Year View*, MIT Press, 1979.

[2] The LOC (for Library of Congress) unit of memory was established half jokingly, half seriously by the author to represent large amounts of information. It is 100 trillion characters.

hardware but not the older tailor-fitting software. Software costs, however, are likely to continue to be high and will undoubtedly be the main tempering factor in the growth of the Information Marketplace.

Geographically Distributed Systems

Historically, the computer has been a *centralized* resource. Because of its very high cost in the early days of its development, the computer was shared either through different programs on a *batch* basis (waiting for your turn to put in your cards and get your answers), or on a *time-shared* basis (spreading computer power in round robin fashion to several people so fast that each thinks he has the computer to himself). Both of these approaches which form over 90% of today's computer systems involve a centralized hardware and software structure which "knows" all that it must about different users, different programs, and different data stores. Such a centralized structure, because of inherent complexity limitations has an upper bound on the number of people that it can simultaneously serve. Today, this bound is somewhere in the vicinity of 50 to 100 users. This limitation is comparable to a human being's inability to cope simultaneously with more than a few tasks.

With the continuously decreasing hardware costs and communications advances discussed in the preceding section, it is now possible to interconnect many different machines so that in effect data is sent from one centralized installation to the other over a satellite or local network. This trend toward decentralization can extend quite rapidly to the extreme where each computer serves one user and all such computers are interconnected. The Xerox Palo Alto Research Center has demonstrated this principle with some 700 interconnected single-user computers.

The Information Marketplace that we envision will be a mixture of predominantly single-user, as well as multiple-user computers, all interconnected through a variety of paths, as in today's world-wide telephone network. The reasons leading to such interconnections are economic. First, large corporations with distributed plants and offices will seek such interconnections among their own locations in order to make their business more efficient -- this we discuss further in the section below on Office Automation. Later, perhaps in the early 1990's, different businesses will seek connections with each other to improve the effectiveness of automatically handling communications and transactions on an inter-organizational basis. Somewhere in this time frame, individuals may seek interconnections to certain data banks and to each other for a variety of reasons discussed in the section on Home computers. However, while the future evolution of inter-and intra-company interconnections appears fairly certain today, the interpersonal development is not as clear at this time. Indeed it may exhibit an avalanche effect, as was the case in CB radio -- if enough people and organizations are interconnected then the resultant Information Marketplace will be more useful, leading to more people seeking interconnection.

In conclusion, the technological infrastructure of the Information Marketplace lies in these geographically distributed and decentralized systems which are needed simply because people, hence the collection, processing, and use of information, are geographically distributed in the first place.

Intelligent Programs and Service Automation

Intelligent Programs are expected to play an important role in the Information Marketplace. Today's research programs that are characterized as "intelligent" exhibit expertise in such diverse fields as clinical decision making, mathematics and circuit design. Take for example a recent program developed at the MIT

Laboratory for Computer Science[3] that tries to behave like an expert physician in the administration of the drug Digitalis. Given the patient's history and symptoms, this program recommends appropriate dosage amounts. The program contains within it a good deal of knowledge about Digitalis, much like a book on that subject. Unlike a book, however, the program can respond to a non-specialist physician's questions. Matching human queries to machine explanations is one of several features that distinguish such intelligent programs from specialist texts.

Features such as the above, along with other capabilities and on-going improvements suggest the future use of intelligent programs for the automation of services. Indeed it is not far fetched to extend the behavior of early research experiments to recreational, financial, governmental and business service applications. The automation of certain services by computer further suggests as a major potential advantage the tailoring of services to individual needs.

Even if the development of intelligent services is slow, we will still see a progressively greater automation of services, albeit with more traditional schemes. We have several examples today of companies that provide library or legal services by searching a data bank to find cases that match a pattern specified by their clients. Such a recently established company grew from zero to some sixty million dollars in sales over a three-year period, suggesting an existing pent up demand for the automation of informational services.

Office Automation

The automation of certain office functions is likely to be the first major and rapidly growing application of the Information Marketplace. The reason lies in the coincidence of strong supply and demand forces for a new office technology. On the demand side, office workers continue to be under-capitalized and to rely on minor improvements in capital equipment. At the same time, they are confronted with increasingly complex information management requirements and rising productivity expectations. On the supply side, the technical trends discussed in the preceding sections form a cost effective technological foundation for growth of a new office technology.

This technology is likely to lead (in order of increasing difficulty) from (1) word processing and text formatting; to (2) low-cost transmission of mail and messages (data, voice and still images); (3) automated intra-company office procedures; (4) automated inter-company business transactions; and (5) sophisticated filing and retrieval of information.

The first two of the above developments are self explanatory. *Intra-company* office automation includes electronic mail within the organization, calendar management, forms management, the automatic processing of messages and the formalization of certain office procedures. *Inter-company* office automation is organizationally and technically more difficult than its intra-company equivalent because it involves interaction among a number of *different* and autonomous organizations: The organizational problems involve such issues as the need for common inter-company communication conventions and business transaction standards. The technical problems are associated with the interconnection of thousands, and later perhaps hundreds of thousands if not millions, of cooperating computer ports such as terminals, small personal computers, and larger communal machines. Moreover these aggregates must have at minimum the

[3] By Professor Peter Szolovits and his research group.

functionality of today's centralized systems even though they will be necessarily decentralized. Accordingly, we do not yet know how technically difficult or easy will be their effective inter-connection.

To get a better idea of the applications that are possible in an office environment, consider as an example some of the ad hoc and casually developed programs that I use for my own office transactions:

1. A *calendar* program that keeps track of my time commitments, prints a daily card of appointments and a daily lookahead calendar for the next ten weeks. A *history* file where calendar appointments end up automatically after they have happened. I have used this history file often along with a retrieval system to cost account for my time. For example, I can rapidly assess how much time was spent during the year for any recurring task.

2. A *people* file that lists addresses, phone numbers and other data of some 2,000 people that are my business contacts. This file is updated and printed as my personal directory every quarter. It is accessed on-line several times a day for making phone calls. This file can also be searched for patterns, *e.g.* for all people I know in city X or company Y.

3. A *mail and message* system through which I can send, receive, copy, annotate, and file messages to people within my organization and outside (via the ARPANET).

4. A *travel* program that tells me what to pack for an n-day trip and keeps track of itinerary and related information.

5. A *to do* list that keeps track of the things I must do, as well as of the tasks that I have delegated to others (with deadlines and interim comments). A recent elaboration of this program permits me to manage tasks with an "accounting" approach. That is, I enter new work in a journal and then "post" entries to individual "accounts", i.e. to the people to whom work is delegated, with appropriate comments. These people in turn receive such assignments on their terminals by electronic mail. At any time I can request a "balance sheet" which lists my "work assets" and "liabilities", i.e. work owed me, as well as work owed by me to others -- together with related comments and deadlines.

6. Means for accessing communal data bases such as the organizational calendar of events, the organizational directory, budget, and expenditures.

7. A text-editor (rarely used by me in the office, but frequently used from my home).

I have been using most of the above programs and files for several years and find them indispensable to the running of my own office.

We turn next to the prospects for more *sophisticated filing and retrieval* of information. In spite of much recent fanfare surrounding data bases and data base languages, the problems of filing and retrieving information effectively continue to be substantial. One common problem is the aging of data and the progressive inconsistencies that plague current data bases. A far more serious problem, however, is the limitation of today's data bases to answering only fixed-format keyword-based queries that were anticipated by the data base designers. The far more useful prospect of organizing incoming data in such a way that it can be retrieved when a relevant, but unknown-at-filing-time query arrives is still largely unsolved -- for example,

asking for footwear manufacturers and getting a list not only of those manufacturers that have the word *footwear* in their product description, but also *shoe, stocking, sock, nylons* etc. The prospects for more sophisticated filing and retrieval systems are not clear at this time. The problem is technically difficult because it necessitates automated "understanding" of the query -- something that is currently in the forefront of current research. Such sophisticated filing, however, is needed and will certainly revolutionize office technology if and when it arrives.

Home Computers

We can buy today toy computers, generally without intercommunication capabilities that are somewhat beyond the level of a desk calculator and typically offer entertaining games and a few simple applications. It is likely that these machines are the awkward predecessors of tomorrow's widely used home computers.

Such home computers are expected to provide educational, recreational, medical, financial and other services primarily through their interconnection with the Information Marketplace, and secondarily through packaged locally contained programs. These machines are also expected to maintain a wealth of personal information, make possible electronic mail, and link the office to the home.

In my case, I find the home computer (which is connected to my office) most helpful for checking my office calendar; sending a last minute memo to someone or to my secretary on Sunday evening, when I know that I will be on travel during the week; reaching other people on the system with quick message exchanges; and using it exactly as I use it in the office when I work at home. In addition, both the children and I enjoy playing games, most of which are addictive and some of which are quite educational.

One such deceptively entertaining yet cryptically educational word game involves several players in remote locations who are confronted through their terminals with 10 randomly chosen English letters. The goal is to make, within 3 minutes, the largest number of longest possible valid English words, using part or all of the given letters. As a player composes a successful word, that word appears simultaneously (under his name) on all the player's screens and belongs to him -- that is, no one else can make and get credit for that word. If a player tries to form a nonsense word, the computer rejects it on the basis of a 40,000 word built-in dictionary. While this game is wildly competitive and is accompanied by fast keystrokes and nerves on edge, it has a beneficial educational side effect. Specifically, players will experiment by trying to make what they think are valid words using known words and a trial and error approach. If such a proposed "word" is accepted by the computer, then the player automatically learns that this is indeed a legitimate English word that may be used to win future games. Another program gives the meaning of any requested word, thereby closing the learning cycle -- otherwise, a player may learn and remember new English words without knowing their meaning. I have personally observed my daughter increase her vocabulary substantially and painlessly by playing this game since age 11.

Some Socio-economic Consequences

First, *human displacement* by automated services is likely, yet at a slow rate and over a period of several generations. To start with, there are some human tasks for which machines are better suited, such as highly repetitive jobs that are certainly not contributing to our humanization -- these are likely to be gradually taken over by machines. Regardless, however, of the types of jobs that will be replaced by machines, we should feel no more and no less threatened by such events than by the earlier displacement of people from certain jobs as

a consequence of the Industrial Revolution. If such a transition is slow; if it replaces lower-level functions and if it spans several human generations, as seems to be the case here, then at least a segment of the human labor force is redistributed over more challenging and less mundane activities.

Offsetting this negative effect on employment are the increased opportunities for employing the handicapped, the rural population and people that are confined at home because of children or other reasons. This potentially productive force can be utilized through the Information Marketplace -- terminals or single-user computers at remote sites can offer informational employment opportunities such as *managing accounts, editing text, processing forms, writing manuals and reports,* even participating in *joint informational activities* with workers at other remote sites. In effect, the information Marketplace can be used as a vehicle that distributes geographically the demand for and supply of informational work.

A second consequence concerns the prospect of our eventual *dehumanization* by excessive use of machines. We often hear about such a potential dehumanization, without stopping to consider that we have been already considerably dehumanized through the Industrial Revolution. Gone are the artisans and craftsmen of the pre-industrial era with their tailor-fitting hand-signed products and services. The low-cost, mass-produced goods and services of today have reduced us to affordable uniformity and impersonal numerical identities. To my thinking, the much feared computerization of our society may indeed, if not reverse, at least balance some of these de-humanizing trends -- in particular, the Information Marketplace, may make possible through service- and office-automation the tailoring of services at affordable cost to the most variable of demand centers, ourselves! Emergence of this *mass individualized* service industry may turn out be one of the most important consequences of the Information Revolution.

A third consequence of the Information Marketplace may involve some *mental atrophy* as, and if, intelligent programs become more capable and effective. Such atrophy has already started in arithmetic with the advent of the inexpensive calculator. This is clearly an area which we should watch with caution and try to anticipate by monitoring our educational system.

Another socio-economic consequence involves the possibility of *undue trust* placed on machines by people who are either unaware of a machine's capabilities or who purposefully wish to influence the opinions of others. While such cases will undoubtedly arise, it is my belief that they will not be frequent, since people will seek to comprehend and question the results of computing machines as they have done for other complex systems in the past.

Beyond the above consequences that cast a negative shadow, we can look forward to some more positive consequences such as (1) *increased productivity* through the benefits of automation; (2) *reduced energy dependence* through selective replacement of energy consuming travel by relatively inexpensive data communications; (3) *information filtering, i.e.* selection by machine of information important to us and the screening away of the informational junk that bombards us at an increasing rate; and (4) *improving our way of life* through increased convenience, and through the availability of useful services.

Some Legal Consequences

One of the first such issues that comes to mind concerns the placement of *responsibility and liability* for programs that have been written by many programmers over a long time and which may be incomprehensible by a single person. I cannot conceive of a physician who will install and use a program on digitalis therapy

without comprehending how that program works and without identifying a human organization that is accountable for the program's actions. This situation is no different than the issue of liability in the context of other complex systems such as a jumbo aircraft where many human designers are involved and where it is equally difficult to argue that any one person understands the entire system in detail. In short, the traditional issues of liability that focus on a responsible *individual* should apply without change in the Information Marketplace.

The most important concern that I have about the legal consequences of computers involves the prospects of *reduced privacy* and related *computer crimes*. This issue, in turn, is linked to the future decentralization of the management and control of computer resources. Where such resources are centralized, it is inevitable that information pertaining to us will be aggregated, correlated and ultimately misused. Even if a benign organization has centralized control of such information, the information may eventually come into the wrong hands at the wrong moment. Consider for example the imaginary case of a political leader who asks for a program that "wakes up" whenever 10 or more left-wingers meet in any U.S. city. Such a task cannot be easily pursued today, simply because the information needed by the program is either unknown, or distributed among many independent organizations in the form of airline manifests, credit cards, hotel registration forms and so forth. If all of these data bases, however, were controlled by one central authority, then they could be easily searched by computer. It is this last factor of *easily reachable information* by machine that makes privacy and computer crime such dominant issues in the Information Marketplace.

While in the U.S. and other democratic societies, centralization of information is unlikely, the opposite holds true for *autocratic* political systems, which, by their very nature, are likely to make sizable investments in centralized installations and control techniques. Fortunately, the pluralistic and heterarchical Information Marketplace that we forecast is no more centrally controlled than the marketplace for goods and services -- a bright prospect for the future of privacy in democratic societies. In this area, it is the obligation of the citizenry and of the government to maintain active vigilance toward potential privacy violations and to provide safeguards for avoiding such violations in the first place.

In spite of the reassuring decentralized aspects of our envisioned Information Marketplace, a major problem remains: It concerns the possibility for *surreptitious* and *unauthorized* explorations over the interconnected machines that make up this marketplace. For example, a "malicious" program may be sent to several machines, where after invading them it searches for sensitive information, finds it, copies it, and removes all traces of its presence. Auxiliary "safe-cracker" programs may be used to patiently try a huge number of possibilities in order to break down the (usually cryptographic) defenses of a given installation. Such activities may be pursued for financial gain, for terrorizing, and for industrial or political espionage.

To avoid such potential problems, we must insure that *no data bank or user can join the Information Marketplace unless they meet certain safeguards*, which naturally increase with the sensitivity of the information that they control. Such an approach, although technically possible, may be difficult in a climate where governmental regulation is viewed with suspicion and business reasons dictate whether two data bases should become interconnected. To my thinking, this is a very important area that will eventually require *new laws* and *new regulatory means*. If, in fact, we do not tackle these questions at an early stage, we may find ourselves in trouble after a good part of the Information Marketplace is established, at which time there may be little back-tracking that we can do.

The safeguarding of data involves the technology of insuring (1) *privacy*, i.e. the transmission of information from A to B, or the storing of information in a file without it being detected by others; and (2) *authentication*, i.e. the verification that the purported signatory of a message in the Information Marketplace is not an impostor and that no one has tampered with the message. Several techniques, based on cryptography, are currently available for insuring the privacy and authentication of data. However, since such techniques are critical to our governmental security and to the current charter of the National Security Agency (NSA), there is a built-in conflict between secret governmental and open civilian uses. It is unlikely that this conflict will be resolved by NSA, which has obligations toward the governmental uses, and it may indeed be necessary for Congress to consider either changing the charter of NSA or involving new, or other existing, agencies such as the FCC to worry about the civilian side of the question. Ultimately, the Congress and the President will have to become knowledgeable and concerned about this subject so as to pursue the necessary checks and balances for the overall national good.

Going beyond privacy, the Information Marketplace will bring into prominence the notion of *information* in a new context which will undoubtedly require re-thinking of traditional values and meanings, and the formation of related new laws and institutions: Consider for example the extent to which computer programs should or should not be *protected from unauthorized copying*. Programs have the unique characteristic that they are not clearly "material embodiments" hence patentable, nor are they clearly "writings", hence copyrightable. Instead, they seem to be somewhere in between. To be sure there can be as much creativity and novel art involved in forming a new program as there is in creating a new device. Yet, copying a program is considerably easier than copying a device, and somewhat easier than copying a music record. If the Information Marketplace is to be effective, new means must be provided for protecting the entrepreneurs who invest effort and funds in creating unique data bases and informational services.

Another legal issue concerns the extent to which information should be treated *like or unlike tangible goods* and products. We have a tendency to think of information as a "second-class citizen" to goods. For example, we are frequently asked by government and others to fill lengthy forms at zero cost, while no one in his right mind sends us a furniture kit, asks us to assemble it for free and send it back! Yet respectable work is involved in both cases.

Yet another legal issue associated with the Information Marketplace concerns the development of criteria for the kind of information that *should or should not be stored in machines* and if stored *for how long*. Should there be cases of mandatory machine forgetfulness, or should machines be constrained to remember everything forever?

Then there is the issue of *computer crime* -- its discovery, and its judicial treatment. Should there be, for example, *mandatory audit trails* whenever anybody or any program tampers with a particular sensitive computer in the Information Marketplace. Also, are there computer crimes and fraud that may parade now or in the future as *computer "mistakes"*?

The above sample consequences are by no means a comprehensive summary of the problems that arise when information processing and communications become a dominant factor of our socio-economic infrastructure. They are instead indicative of the potential for trouble if we do not take the time now to identify and confront the nature of information and its relationship to us in the forthcoming Information Marketplace.