The Document Processing Revolution

by

Michael B. Spring[1]
Jeffrey D. Campbell
Department of Information Science
University of Pittsburgh
June 1, 1995

I. Introduction

This paper examines the thesis that processes of creating, publishing, and managing documents are undergoing radical changes. In order to assess the validity of the thesis and to better understand the nature of the revolution, it is necessary to establish a framework for analysis. The paper takes two perspectives to set the stage. First, an historical analysis attempts to posit the current developments in a larger context. Second, the nature and the magnitude of the changes in several aspects of the technology of document processing are explored. Following these analyses, the paper attempts to outline the document processing revolution. This is done in two parts. First, relevant changes in enabling technologies and information infrastructure are explored. Second, the changes in roles, processes, and products of modern document processing are reviewed.

The paper is organized as follows:

1. Historical Context
   a. A narrow view: reprographics
   b. The big picture: communications
2. The Nature and Magnitude of the Change
   a. Trends in technology
   b. Information Earthquakes
3. Enabling Technologies and Infrastructure
   a. Bit pipes and set top boxes
   b. Electronic currency
   c. Network security and reliability
   d. Public policies

1 Michael B. Spring is an Associate Professor of Information Science at the University of Pittsburgh. His research involves the application of technology to the workplace with particular attention to interactive system design, document processing, and standards. Prior to his appointment in Information Science he was Director of the University External Studies Program. He received his BA from the College of the Holy Cross, Worcester, MA, and his Ph.D. from the University of Pittsburgh. Dr. Spring is the author of numerous articles in the areas of text and document processing, information technology standardization, virtual reality, and human-computer interaction. He is the author of two books—**Electronic Printing and Publishing: The Document Processing Revolution** by Marcel Dekker, 1991, and **Hands on Postscript** by Hayden, 1992. He serves on the editorial boards of Computer Standards and Interfaces and Virtual Reality World, the Long Range Planning Committee of X3, and the NRC Review Board for the Computer Systems and Computer and Applied Mathematics Laboratories of the National Institute of Standards and Technology.

4. The Functional Changes
   a. Documents
      i. The document defined
      ii. New electronic document forms
      iii. Hyperdocuments and infobots
   b. Roles
      i. The author in control
      ii. Reader agents
   c. Processes
      i. Reengineering document processing
      ii. Selected processes in transition
5. Conclusions
   a. Economies of scale
   b. People, technology, and organizations

II. Historical Context

What exactly is happening today? Is this the middle of a revolution or is this much ado about nothing? It has been said that the short term impact of new technology is overestimated and the long term impact is underestimated. This section takes two views. The first, a narrower view, positions the changes in the context of the printing or reprographics process. The second view examines the changes in the context of the history of communications.

A. A Narrow View: Reprographics

It is not uncommon to hear the current developments in the area of documents compared to the emergence of the printing press. Eisentstein's book on the Printing Press as Agent of Change[2] provides background for understanding the impact of the printing press. One may view the development of reprographics as marked by four events, beginning with the Gutenberg printing press.

1. Mechanical Reprographics
2. Optical Reprographics
3. Digital Reprographics
4. Ad Hoc Reprographics

These terms characterize four significant developments in the reprographic process. First, 500 years ago, the Gutenberg press signaled the beginning of the era of mechanical reproduction using modular reusable masters. Mechanical reprographics substituted flexible, moveable type and mechanical reproduction for human efforts in reproductions. The printing press replaced the scribe.

Optical reprographics represents two distinct revolutions. The first was lithography, the use of photographic technology to make printing plates. For all practical purposes this put an end to mechanical typesetting causing a significant change in the printing industry. The second was the use of photographic technology coupled with photosensitive dielectrics to give us photo duplication — or Xeroxing. This made every object in the world a master that could be used for the printing process. The revolution here was dramatic. For the first time, virtually anyone could become a printer. The "master" existing in the copier was infinitely flexible and reusable.

2 Eisentstein, Elizabeth, **The Printing Press as Agent of Change**, Cambridge Mass: Cambridge University Press, 1979.

The third revolution, digital reprographics, is somewhat fragmented, and includes the fax machine, laser printers, and the new digital copiers. With the introduction of digital technology, it was no longer necessary that the original have a physical form. If it did have a physical form, the revolution made it possible for the original and the copy to be in different locations.

The most recent reprographic revolution is "ad hoc reprographics". This term is meant to capture the form of reprographics where both the master and the copy are electronic. It is characterized by the "printing" of an electronic master to an all points addressable screen — a computer or TV screen. Ad hoc reprographics is fundamentally different than the earlier stages. No physical resources are consumed. The "document" display is immediately reusable. Neither the master or the copy may ever exist in a non-electronic form.

It is interesting to note that even with ad hoc reprographics there is pressure to maintain the traditional printing, publishing and distribution roles. Part of the attractiveness of CD-ROM is that it maintains a physical object to produce and sell. The media is simply denser and more flexible.

B. The Big Picture: Communications

While the reprographics perspective is helpful, there seems to be something more fundamental and significant happening. A number of authors have suggested that radio and television, which few would deny have had a significant impact on society, have had a profound impact on communications. This can lead one to think about the current revolution as a revolution not in methods of reprographics but in methods of communication. One might postulate the following developments in the communication process:

1. The language revolution
2. The literary revolution
3. The telepresence revolution
4. The meta-literacy revolution

1. The Oral and Literary Traditions

There is evidence that "humans" have been around for a million years or so.[3] It is difficult to pinpoint when spoken language developed as a critical means of communication. Ong holds that social interaction has been occurring for the last 30,000 to 50,000 years[4]. Although it is impossible to date the origin of spoken language, it was clearly millennia before written language. Many scholars refer to this period as the period of the oral tradition. The oldest deciphered written documents are about 6000 years old.[5] Scholars have identified the development of writing as beginning the literary tradition.

Think about the change in documents — the document processing revolution — moving first from a pre-language tradition to language and then from the oral tradition to the written or literary tradition. First, just imagine what kinds of documents existed prior to spoken language. There is not much that easily comes to mind, but there must have been something. The pantomime of the great hunt, the defeat of the great enemy, etc. These "documents" must have been subject to significant misinterpretation. Little of what constituted the great documents of the day have survived. One might speculate that the creation stories all have some common origin, which seem to be bound to physical objects and dramatic events that could be described without language.

---

3 Fromkin, Victoria and Rodman, Robert **An Introduction to Language**. Fort Worth, Texas: Harcourt Brace College Publishes, 1993, p. 22.

4 Ong, Walter J. **Orality and Literacy**. London: Methuen & Co., 1982, p.2.

5 Fromkin, Victoria and Rodman, Robert **An Introduction to Language**. Fort Worth, Texas: Harcourt Brace College Publishes, 1993, p. 22.

Whether these stories predate spoken language is difficult to say. The concepts are surely fuzzy enough and based enough on physical universals — the sun, the earth and the stars.

The literary revolution occurred when arbitrary symbols where used to represent spoken language. After 20,000-30,000 years of the oral tradition, the development of writing must have caused quite a stir. Echoes of the revolution can be heard to this very day. The revolution must have been cataclysmic. Imagine the court poets arguing over the value of these new scribes. Imagine the ridicule when two scribes used different symbols to represent some utterance. The message in spoken form was easily understood by all. Memory in this frame must have been viewed as a far superior form of communication. The human store could be queried, it was easily understood, etc. It is not hard to imagine the highly valued court poets claiming that anything recorded in this new way would surely be lost because even two experts could not agree on how to record a spoken word. Any individual — king to slave — could speak and listen, but only specially trained technologists could read and write. This picture has a modern counterpart. Technological advances never come automatically. Many years after the invention and wide dissemination and acceptance of writing, the advanced Greek society finds one of its brightest luminaries commenting on the technology

> *This invention of yours will produce forgetfulness in the minds of those who learn to use it, by causing them to neglect their memory, in as much as, from their confidence in writing, they will recollect by the external aid of foreign symbols, and not by the internal use of their own faculties.*

Many people find this quote by Socrates humorous. It was so long ago, and so ludicrous given the current state of wisdom. Even to this day, there are places and situations in which the oral tradition prevails, for example, the notion of eye witness testimony versus affidavits.

> *People had to be persuaded that written documentation was a reliable reflection of concrete, observable events. To the modern mind, the evanescence of the spoken word seems more plastic, quixotic, and undependable than the printed word. To members of a highly oral culture, however, the spoken word was connected to the incontrovertible realities of bodily experience, while the written word was a thin, substanceless scratching whose two-dimensionality seemed highly arbitrary.*[6]

Socrates had reason to level his criticism. The written word did reduce attention to memory. Written language is a technology to supplement and improve the communication process. The revolutionary change in communication occurred with the virtually unlimited storage capacity and permanence represented by writing. While the literary tradition was accelerated with the development of mass produced books in the fifteenth century, movable type printing represents a quantitative change in the technology, not a qualitative change. At the same time, the magnitude of the change caused by the vast reduction in the cost of reproducing materials was significant.

2. Telepresence

The age of the first literacy continued unchallenged until the late nineteenth century. The invention of the telephone and motion pictures with sound were the early skirmishes in this next revolution. The widespread acceptance of broadcast radio and later television accelerated the change. Some suggest that the technology was giving birth to a second orality and that this new orality would cause a decline in the literary tradition. The decline in the number of general circulation newspapers was seen as one indication of a reversion from written to spoken sources for current events information.

Television and radio represent something more. Immediate, global television broadcast of events ranging from the Olympic games to the Persian Gulf war represent something of a return to the pre-oral tradition. They represent a form of telepresence that can be stored. This recording is in the same form as the original. It is no longer necessary to retell the story, it can be replayed exactly as it happened. This can be very helpful in the analysis of

6 Zuboff, Shoshana, **In the Age of the Smart Machine: the Future of Work and Power**. New York: Basic Books, 1988, p. 77

transient events. Whether it is Armstrong stepping onto the moon, the Kiwi's winning the America's Cup, or a human or natural disaster occurring somewhere around the world. We now have the ability to be present to the event — neither the oral or written tradition has to be relied upon. In addition, because signals can be stored, the events can be replayed and analyzed repeatedly. There are limits to what is recorded. For film or video, the camera angles and lighting must be correct or there may be important information that is not on the tape. The increased availability of tools to undetectably alter audio tape, photographs and video raises questions about the reliability of the recordings. Nonetheless, electronic recordings represent a separable tradition that in a sense obviates the need for communication about the event. Within the limits of the recording, it is the event.

3. The Meta literacy Revolution

It could be argued that computer mediated communication is a return to a literary tradition. The proliferation of electronic mail contrasts with the decline in letter writing and increase in telephone usage. Hypertext is a return to knowledge expressed in a written form. The majority of World Wide Web resources are text and image based. This material is highly dynamic. Documents are being changed frequently, new documents are added and old ones deleted.

In some ways, these new documents move beyond simple literacy. In the World Wide Web, the generation and management of links approaches the significance of the documents themselves. This emphasis on linking was virtually non-existent in the first literacy and represents a kind of meta literacy. Not only is the message passed on, but recommendations about how to navigate it. This meta literacy also includes a view of written material that sees it is one alternative for communication. There is increasing use of audio and video documents to supplement text based documents. This meta literacy then is characterized by two intertwined developments. First, the communication is now in all the possible forms, symbolic, oral and telepresence. They may be interwoven at leisure. Beyond this however, the author can now weave processing instructions into the communication. These enable the receiver to understand how the author developed the communication or that allow the receiver to continue to modify and change the document as it is accessed.

III. The Nature and Magnitude

of the Change

Base technologies are what make revolutions possible. In this case, it is both the qualitative and quantitative change in base technologies that are important. This section examines the nature of the underlying changes and the magnitude of those changes.

A. Trends in Technology

It is almost impossible to stay abreast of the of the technological changes. They are just coming too fast. What are the implications of Sonet versus ATM versus Frame Relay. Is it POSIX or NT or Taligent for the future. Where will optical disk technology stabilize, and which desktop look and feel will succeed. Should we invest in CORBA or COM or SOM, in SGML or ODA, etc., etc., etc.

While it is difficult to stay abreast of all the developments, there are some trends that can be discerned. Key areas of change in the basic technology underlying document processing include digitization, accommodation, standardization, objectification and integration. Of these, the first two are in my estimation the critical driving technologies of the revolution.

1. Digitization

Increasingly, data is being kept in a digital instead of analog form. Perhaps the most common form of digital storage and processing in common use today the laser printer.[7] The office laser printer transforms a symbolic set of instructions into 8 million bits which are transferred to the paper going through the printer. Each 0 is repre-

sented as a black dot on the paper and each one is represented as a white dot. The result is the printed page you see in front of you. Another common form of digital information is the audio CD. To create an audio CD, music is sampled about 40,000 times per second, and the magnitude of the sound wave is represented by a 16 bit number — a value between 0 and 64,000. This means that a CD-Audio recording uses 4.8 Megabytes of data for every minute of recording.[8] ISDN offers digital telephone service that provides the user with two lines capable of sending 64,000 bits of information per second — the rate at which all modern phone systems transfer the digital representation of phone conversations. Consumer electronics such as telephone answering machines are available with solid state memory for a digital message instead of a cassette tape. High definition television (HDTV) is the next major digital technology and it is expected to hit the U.S. market in the next five years.

Digital data is important in that error correction techniques can be embedded in the representation to allow the data to be self correcting. This is part of the quality advantage audio CDs (which are digital recordings) have over analog records or tapes. Digital representations of images may be stored at high resolution and displayed at a resolution that is appropriate for the use. There is a trade off between the resolution and the speed of rendering the image, so different applications may demand different decisions..

Multipurpose digital signal processing (DSP) integrated circuits are currently being designed to be closely integrated into personal computers. Some manufactures are developing motherboards with built in DSP. The PC DSP provides digital/analog conversion. They replace special purpose circuitry that is currently used to provide modem, fax, telephone (PC voice mail, fax response, etc.) and sound card functions. Features that were implemented in hardware will be implemented by software, providing for easier maintenance and enhancement.

2. Accommodation

Computing power has been doubling every two years. Much of this increased power has been used to improve ease of use, that is to better accommodate the user. It is interesting to note that the text based word processor of 5 years ago had almost all of the capabilities of today's WYSIWYG (What You See Is What You Get) word processor. Yet the WYSIWYG word processor requires a machine that is four to eight times more powerful. Side by side, the functions of the programs seem to operate in about the same human time. Where has all the increased processing power gone? The WYSIWYG program uses all that power to make the word processing functions easier to use. Graphical user interfaces have absorbed almost all of the recent increases in computing power to provide an improved user interface.

One way to understand this is to understand the change that has occurred over the last 30 years. In 1962, the cost of a computer was near $1,000,000 and the average professional was making on the order of $10,000. When a conflict arose between the human and the computer, the human lost — they were the less expensive cog in the wheel. Today, when all costs are considered, a professional costs an organization closer to $100,000 while a computer of the same power as the million dollar computer of 1960 costs about $10,000. The human is now the winner when there is a conflict.

The technology of accommodation involves two areas of technology development. The first is the actual interface technologies, the eye phones, the data gloves, the mice and trackballs, the speech recognition tools. Each of these technological developments is intended to provide an easier, faster, more natural way for a user to input or access data stored in a computer system. Technology is quickly getting to the point where all of these operate at a level that allows the digital representation to be indistinguishable from the corresponding analog representation, given the normal acuity of the human sensory or motor system in question.

---

7 It may well be that the fax machine is even more ubiquitous and representative of the use of digital representation. The fax machine scans an analog image, compresses the data via a technique that uses run length encoding via a modified Huffman code, transmits the data, decompresses it at the remote end and prints it.

8 In reality, compression allows this to be a smaller number and error correction and format information uses additional space.

The second aspect of the technology of accommodation has to do with understanding the capabilities of the human component of the information system. The growing body of psychological and biological research is continuing to yield information about how humans process information and this is being used to improve overall processing speed and efficiency. Take for example the theory of preattentive stimuli. Put very simply, this theory says that some stimuli are processed by the sensory system in a massively parallel fashion. Consider for example the task of finding a small "t" on a page of newspaper print. This process is not preattentive, and requires that each possibility be processed consciously. On the other hand, finding a red letter on a page of newspaper print can be done in an instant — colors are processed preattentively. The theory is more complex than this but consider the implications of such a theory for coding information in a search/retrieval task. These kinds of phenomena are the basis of some of the research on navigating information spaces.

3. Standardization

Mature standards are so widely accepted that they are almost invisible. Over the last 500 years, standards for traditional document processing have developed. These standards have included many things: the form of business letters, the process of filing, the size of stationery, etc.. Paper and book sizes are good examples. Both come in a variety of sizes but not an endless variety. The effect is that equipment ranging from printing presses to book shelves can be produced efficiently because there is limited variation to accommodate. Information technology is heading toward mature standards. There are a thousand areas in which standards bodies operate around the world. Even though less than 1% of these bodies develop standard for information technology, they are producing about 50% of the new standards pages.

Two types of standards are evolving that will significantly impact document processing productivity. First, technical compatibility standards are emerging which will provide simple and direct machine to machine (OSI, MHS, and FTAM) and program to program (SGML and ODA) transfers. Second, organizational and industry standards are emerging that will allow for these documents to be viewed coherently (EDI, CALS, etc.). There are also scores of less formal standards in this latter category that need to evolve. For example, organizations need to establish consistent naming conventions and file organizations for the filing of electronic documents. As structural copymarking becomes de rigueur, organizations will need to adopt organizational and industry standards for copymarking of commonly exchanged documents in accord with meta standards such as SGML and ODA.

4. Objectification

Object orientation is a paradigm shift in how computers are programmed to manage data and processes. This shift is significant in two ways. First, related to the development of systems, it represents a recognition of the need to move to a new more modular view of software. Object oriented programming is more modular than traditional approaches. This is providing developers with productivity benefits in developing, testing and maintaining the software. As the cost of software development rises dramatically with the increasing complexity of the software, there is increased emphasis on programmer productivity. Object orientation appears to some to be a means by which we can control this process.

The second impact of objectification relates to the user. Traditionally, data (text, graphics, audio, etc.) were stored independently from the programs that processed the data. One consequence of this separation was that the user had to keep track of the programs that processed data files. For example, there are a dozen different formats for storing PC graphics data. The user had to know which program to use. In the object oriented approach, the objects contain both the data and "methods" which are the processes that can act on that data. For example, a graphic object might have a 'print' method that would render the data in a printable form. Another graphic object with a different format would also have a "print" method. Now the user only has to know to invoke the "print" method and no longer cares about the details of the storage format. This provides the user with simpler model of the system.

5. Integration

Integration is largely the result of technology maturation. For example, maturing DSP technology allows the integration of modem, fax, sound and computer telephony into one circuit. Integration is manifested through multifunctionality, connectivity, and interoperability.

Multifunctionality is being able to use one advanced item instead of a series of less advanced objects. Many office desks have four terminals on them — telephone, fax, personal computer and workstation or main frame terminal. The trend is towards replacing these with a single multifunction device. Many computers have scanners and printers. A scanner and printer together form a digital copier. A copier with a modem is a fax machine. There have been several commercial products that combine these functions. None have been particularly successful in the market place. Performance/quality improvements and price reductions make the promise of a combined personal copier, fax printer, scanner with a price tag of a few hundred dollars a reasonable expectation in a few years.

This is also a period of rapid integration at the intermachine level. This comes in two forms, First, connectivity is developing beyond the local area network level. It is increasingly important to be connected at the enterprise level. Connections between organizations exist for electronic data interchange (EDI) of business transactions such as purchase orders and invoicing. Electronic funds transfer is well developed. The Internet and World Wide Web are promoting new levels of interconnection.

Second, this is the beginning of true interoperability, which goes beyond connectivity. Interoperability allows cooperation between programs running on different computers. For example, a user on one computer might invoke a function that causes a program to run on another computer and have the second computer display the results on the first. Global weather forecasts are the result of massively parallel data processing efforts. They are ultimately tied together based on the results developed in an independent yet coordinated form on many different kinds of machines all over the world. Similarly, there is increasing use of the information agents and "worms" that work independently across the network to provide information about what is available.

B. Information Earthquakes

While these trends say a lot about what is happening today, many people find it easy to get lost in all of the data. One other way to think about the changes that are occurring today is to measure the extent of the changes. It is interesting to think about how much change is being encountered. Changes in storage, transmission and processing power have been significant. They are so large that they are best measured in orders of magnitude, just like the Richter scale for earthquake intensity, each increase in the number represents a ten fold increase in impact. A rating of one equates to a ten fold increase, a rating of two equates to a hundred fold increase, a rating of three to a thousand fold increase, etc.

1. Processing

Information storage and retrieval is at the center of one of the earthquakes. Documents that would fill a library can be searched electronically in minutes. The electronic search can locate information within documents that could not be located by humans in a physical library in any reasonable amount of time. For example, any search that could not be answered based on bibliographic data and abstracts would require reading every document and would be impossible for all practical purposes. This human processing limit drives the need for bibliographic information. Electronic search allows for the storage and retrieval of full text, reducing the need for the analytic surrogates. The United States Library of Congress contains about twenty million books. Any word could be located by doing a binary search of an index of these books and such a search could be completed within 42 reads of the index.[9] This means the search would be completed in a few milliseconds.

This does not imply that the catalogers, indexers and librarians are not needed. It does imply a change in their role. Skilled indexers will still be needed to add index terms that to not actually appear within the text. Similarly,

a 1965 concordance of Byron's works took 25 years to develop manually.[10] The 285,000 entries can be generated by a computer in a matter of minutes. While concordance construction will change dramatically, expertise will still be needed in the use of concordances.

Looking at the quake in processing capability, one can make a variety of measurements. Overall, the magnitude of the change in information processing speed via technology between 1950 and the year 2000 is in the range of a magnitude 6-7 infoquake.

2. Storage

Storage media are experiencing similar order of magnitude shocks. The density of memory on silicon chips keeps increasing. The first PCs in the early 1980s contained chips with 16K bits on a chip. PC memory chips now hold 16 Mb. State of the art chips are now being developed that store 1 GB in a cubic inch. This corresponds to an 1,000 to 62,500 times increase. Personal computer hard disks have gone from being nonexistent to having multiple gigabyte disks readily available. Similarly, removable media have gone from 360K floppies to CD-ROM holding about 650 MB, another increase of 1800 times. The current CD-ROMs are the first generation. Enhanced capabilities should be expected. CD video is not yet well implemented, but two hours of CD video is achievable, it is simply now a matter of what standard will be adopted.

Looking at the quake in storage capability, measurements would seem to suggest that from the 1950 to the year 2000 the magnitude of the change in information storage capability via technology will be in the range of a magnitude 5 infoquake.

3. Transmission

Electronic data transmission time is the center of another infoquake. Teletype machines ran at 150 or 300 bits per second. Today, 2400 bps modems are common world wide with 28,800 bps modems becoming increasingly common. That is two orders of magnitude over teletype. Basic ISDN (Integrated Services Digital Network) offers a 4 fold increase over this to 128,000 bps. This is significant when one considers that to transmit a book with 210 pages, ten images and fifty graphics would take 90 minutes by 2400 bps modem but 31 seconds via basic ISDN. [11] At ISDN speed, it could be faster to transmit the book to your computer than to find it on a shelf. It would certainly be faster than going to get it from the library. Asynchronous Transfer Mode (ATM) and Frame Relay offer bandwidth on demand. It is reasonable to expect serious users will all have the 10,000,000-100,000,000 bps network connections many academics currently have on their desktops.

Looking at transmission capability, measurements would seem to suggest that from the 1950 to the year 2000 the magnitude of the change in information transmission capability via technology will be in the range of a magnitude 8-9 infoquake. While the transmission infoquake is significant enough in its own right, the co-occurence of the three infoquakes is truly unprecedented.

IV. Enabling Technologies and Infrastructure

There are a number of technological developments that are essential to a revolution in document processing. There is an effort in the U.S. at the current time to define these enabling technologies. In some ways this is the

---

9 Witten, Ian H., Moffat, Alistair and Bell, Timothy C., **Managing Gigabytes**, New York: Van Nostrand Reinhold, 1994, p. 15

10 Witten, Ian H., Moffat, Alistair and Bell, Timothy C., **Managing Gigabytes**, New York: Van Nostrand Reinhold, 1994, p. 2

11 The calculations are based on a book consisting of 210 text pages each with 9.6 KB of text, 10 120 KB images and 50 16 KB graphics.

major focus of the National Information Infrastructure effort. A few key enabling technologies are described in this section. They include:

1. bit pipes and set top boxes
2. electronic currency
3. network security and reliability
4. public policy

A. Bit Pipes and Set-Top Boxes

The technical trends discussed above are driving both communications channels and processing devices toward integration. Thus, while some Americans now have three information lines coming into there house — one for television, one for voice communication, and one for data, this will change with time. Already, the telephone companies have committed to, and are providing, better grade communications lines that handle voice and data equally well. Video is not far behind. The various groups working on the National Information Infrastructure committees are addressing how to connect homes to the network.

A device is needed that provides access to the various kinds of information that can be moved via a digital bit pipe. Early on in the development of ISDN, many people felt that the cost of digital phones, an order of magnitude greater than the cost of analog phones, would greatly hamper the extension of digital service. It is true that digital signal processing capability adds significantly to the cost of a phone, but it only adds marginally to the cost of a PC and it can be used for many other functions in a PC. Thus, we will see the digital phone emerge as a new service to be provided via our PC, which will also act as an intelligent answering machine, appointment scheduler, video conferencing device, etc.

The set-top device has four primary subcomponents. First is the general processing, storage, and memory capabilities that are expected on a rather ordinary desk-top device. Second, add the digital signal processing capability that allows the conversion between digital and analog information and the result is a phone, scanner, and the guts of a digital television. Third, add an array of connections to the box to allow it the be hooked into the home system. Thus, it will need a fiber or coax connection to the outside world, and internal connections to all the devices in the home that it will feed or that it will integrate. It will need coaxial cable out to the televisions, Ethernet to the other microprocessors in the home, twisted pair out to speakers, and signal processing lines in from security and fire devices.

The fourth component is the “set” that the device will sit on top of. The notion in the U.S. is that whatever is built will be a supplement to an existing output device in the home, in particular the television. The assumption is that the device is probably not the television currently in homes, but a digital high definition television. It is interesting to note that some of the same problems that occur with phones and digitization occur with television and digitization. A large home television (30 inches or 76 centimeters) costs $350 while a medium sized computer monitor at work costs $1000. In order for an integrated precision output device to be placed into every home, there will have to be some breakthroughs in flat panel digital display technology. These breakthroughs have been accomplished, and they may be made public between the time of this writing and the time you read this. In any case, they are not far away.

All of these capabilities exist today—a wired home can be had for a few thousand dollars from off the shelf suppliers to the consumer market. Unfortunately, as of today there is nothing to hook the set-top device up to. Remote and computer control devices for electrical outlets, lights and appliances have been available for years but have yet to achieve significant popularity. This leads to the famous S curve of technology adoption and to the notion of network externality. The S curve of technology adoption indicates that the rate of adoption of a technology is slow in the early and late stages and rapid in the middle.

Network externality is an economic term indicating that the value of a network increases as the number of connections to it increases. An example from cable television illustrates this point nicely. One of Ted Turner’s early cable companies was having difficulty attracting advertisers. When cable television had reached 35% market satu-

ration, advertisers—Proctor and Gamble in particular—decided that it was a viable medium and started using it for advertising. Once this happened, more people wanted to provide cable channels to get at the available marketing revenues. Once 35% saturation was reached, which took about 25 years, an additional 35% saturation came in little more than 6 years. At the same time, the number of cable channels swelled from ten to hundreds.

B. Electronic Currency

Assuming that the pathways are established, many services can begin to be provided over the information infrastructure, providing that the parties can establish business relationships. This involves four subcomponents:

1. Electronic data interchange
2. Secure communications
3. Authentication
4. Electronic currency

The area of electronic data interchange is already well developed through the U.S. committee X12 and the international EDIFACT effort. In some industry groups such as banking, aerospace, and automotive, this development is at a reasonable level of maturity. Purchase orders, invoices, and hundreds of other business documents have been designed and are exchanged regularly. There are some minor problems related to system interoperability which have led to the creation of clearinghouses to do format conversions. These are very minor headaches which are being overcome by additional standardization. The education and library industries are somewhat behind the curve, but this is normal. Progress toward standardized forms for electronic transcripts, proposal formats, overdue notices, etc. will take more time, but they are coming.

In the area of security, it is not clear which technologies will win and gain wide spread acceptance. Security presents an interesting dilemma. At one level, secure communication can be accomplished today via any of a number of proprietary methods. Unfortunately, if two machines use different schemes they cannot share secure information. They must all agree to the same security method in order for communication to occur. Thus within well defined communities such as defense and banking, the goal has already been achieved. However, a broader based public system for general use is needed. Added to this, at least in the U.S., is the demand for a system that can be tapped when a justifiable legal need arises. This debate, known as the “clipper chip debate” has caused significant controversy in the U.S..

Third, a means of providing authentication is needed. Authentication verifies a person’s identity and establishes that he or she is authorized to make a given transaction. This is important glue in the system. Bank and phone card fraud are already a significant problem in the U.S.. As the financial system becomes electronic, there needs to be some simple and yet foolproof mechanism for authenticating the user will be needed. This is a complex but critical component of the enabling technology.

Finally, there needs to be electronic money itself. A number of people are studying this problem, and the discussions of the characteristics of electronic money are wonderfully complicated. At the core level, the discussion has centered on the fact that money must meet the ACID test used for computerized transactions. This means that whatever form it takes, the money must be:

1. Atomic: the payment and the purchase are inextricably linked
2. Consistent: the amount the buyer pays must be consistent with the amount the seller gets
3. Isolated: each transaction should be isolated -no linkage between different transactions
4. Durable: if the transaction fails in the middle, there should be no partial completion and things should roll back.

In addition to these concerns, electronic currency will have to have minimal overhead — the cost of “printing” and processing the money must not add significantly to the cost of the transaction. Electronic money must have reasonable granularity and scalability—that is to say, one must be able to buy cars and candy bars. Many people must be able to purchase things electronically at the same time. Electronic money must be trusted. This means

that should be readily accepted. There should not be multiple mutually exclusive forms of money. It should not be possible for someone to make their own money as counterfeiters do now. Similarly, existing money should not be able to be changed in value. Electronic money must hold its value over time, it cannot expire.

The last and perhaps most fascinating feature, is related to the isolated nature of the transactions. Part of this has to do with the anonymity of money. When a newspaper is purchased with cash, it is not known who bought it. The physical currency carries no history with it. The technologists are working on similar capabilities for electronic cash. At a simple level, a smart card could take money from a bank account with appropriate authentication and put it on the card. On card circuitry can then spend that money, debiting the amount on the card without telling the seller anything other than the fact that this card, which is anonymous to the seller, has enough money on it to buy the item in question.

C. Network Security and Reliability

The network must be secure at a number of levels. The issue of transaction integrity has already been addressed. This involved the use of a digital signature of some sort to guarantee the validity of the transaction. Security involves a number of other concerns as well. These include:

1. Methods to assure that a portal to the network is semi-permeable, that it is resistant to unauthorized access.
2. Transmission assurance that the transaction will be delivered only to the proper destination and will not be intercepted or modified en route.
3. Network reliability, survivability, and scalability that will insure that the message/transaction will be able to get though to the destination without failure. In other words, access will not be affected by catastrophe and that network load will not impact performance.

To achieve the first goal, mechanisms are needed to assure that the information does not include Trojan horses and that there are no backdoors to the system that can not be locked and unlocked by the owner. To achieve the second goal, some forms of encryption and authentication are needed. There are also system policies and policing that are required to insure that other people do not get unauthorized access to the bit pipes.

The last goal makes clear that security also includes assurance that if the transaction is sent that it will in fact arrive. This has to do less with the threat of intrusion than it has to do with the reliability of the system. At a simple level, a system needs to have rerouting capabilities to insure that expected system failures are minimized and localized. This also requires however that the system have survivability. Essential services must survive catastrophes such as fires, earthquakes and typhoons. Finally, these services must be over built to survive extreme use conditions. In the U.S., for the transportation system, this is the Thanksgiving holiday. For the phone system, it is Mother's Day. For the Internet, it might be the demand for copies of the photos of the comet hitting Jupiter, or of a war in Kuwait.

D. Public Policies

The kind of revolution made possible by these kinds of technologies and infrastructure will be aided or constrained by public policy. These policies will have to do with access, privacy, intellectual property, regulation, etc. For example, as Samuelson[12] has pointed out, copyright policies at heart have to do with establishing financial incentives for authors and publishers that promote societal goals of spreading information and increasing learning. In the U.S. access to the NII via libraries and schools is a national priority. This kind of public policy will encourage a revolution in how documents are handled. At the same time, some manifestations of new intel-

12 Samuelson, Pamela, Legally Speaking: Copyright and Digital Libraries, **Communications of the ACM**, April 1995, 38(4), pp. 15-21 and 110.

lectual property regulation seem to ignore the issue of making information more available and appear in some ways to be controlled by special interests looking to keep control of the situation.

V. Functional Changes

The document processing revolution that is underway today will make use of this emerging infrastructure. One way to understand what will change is to ask three questions

1. How will the things called documents change?
2. How will the roles of the key players change?
3. How will the processing of these documents change?

Historically, documents have been linear and, for the most part, atomic. New document forms are emerging that are neither linear nor atomic. More responsibilities are being shifted to the creators and users of documents. The process is being streamlined and the stages of processing are shifting. Each of these areas is examined briefly below.

A. Documents

This section develops a definition for a document, looks at some of the new kinds of documents that are emerging, and takes a particular look at hyperdocuments and active documents.

1. The definition of document
2. The likely new products in the near term
3. Two new document types: Hyperdocuments and Infobots

1. The Document Defined

It is interesting that this entity to which so much effort is devoted and which is held so dear is so difficult to define. From an information science point of view, it is essential to develop some definitional clarity to do research in the area. One early effort yielded the following:

> *A document is an identifiable entity having some durable form, produced by a person or persons toward the goal of communication; it may take a number of forms, but must have a least one symbolic manifestation that used to store or communicate information between people. It is a cohesive entity formed of subcomponents in logical, layout, and content form.*[13]

Documents are an essential product in the information age. There was a day a document was a report or a book that consisted predominantly of text written by a single author. Things are no longer so simple.

1. Documents include text, graphic, images.
2. Documents may be cohesive, a letter or a report, or disjunctive, a medical record.
3. Documents may be authored by individuals, groups, or organizations.
4. Documents may have a limited life span or be archival.

2. New Electronic Document Forms

Documents are becoming ever larger and more complex. The support documents for the Boeing 747 weigh as much as the 747. Two copies of the documentation for a new drug application will fill the largest moving van

13 Spring, Michael, **Electronic Printing and Publishing: The Document Processing Revolution.** New York: Marcel Dekker, 1990, p8.

available. Publishing related expenditures reflect 6% to 10% of the gross revenues of the Fortune 1000. Over two trillion pages were printed in 1986 and nearly 4 trillion were printed in 1990.

What new forms will electronic documents take? There are several clues that may be gleaned from the current situation. Several new forms of documents are emerging. These include:

1. custom documents, only what is needed for that customer
2. on demand documents, documents only when they are needed
3. living documents, the source document is always current
4. multimedia documents, the needed mode is used
5. distributed documents, the document is located close to interested parties
6. interactive documents, the document exhibits some intelligence

These various forms of documents overlap, but they can be treated individually for purposes of illustration. The custom document is emerging because of the significantly reduced cost of digital reprographics. It is still true that the cost of mass produced documents is less than the cost of unique products. It is equally true that the smaller size of custom documents makes them more economical. As an example, consider the cost of the journals on my shelves. The entire journal comes at a lower cost per page than a single copy of the articles I am interested in. However, the overall cost of printing one copy of all the interesting articles is less than the cost of the journal containing all the articles. Similarly, copies of advanced text books that include only the material I need are likely to include less than mass produced versions.

In academia, bulletins and course catalogs are used to specify the rules and regulations that apply to graduate students. At any point in time, the official bulletin fails to reflect the rules and regulations currently in force because the document was printed at some prior date. Living documents are those that have strong version control features that allow one to print the document as it exists today or as it had existed at a given point in time.

Multimedia documents are all the rage today, and having experienced the value of being able to switch modes in working with my children I am convinced that they will become ever more important. We have only begun to come to understand how to use these documents. Consider that in preparing to come to Adelaide, we began with a map of the world, shifted when we found Adelaide to a text description of the city, and then moved to pictures of the city. Future encyclopedias might allow us to move to sound bytes and to smell the flowers in the city park.

Distributed documents are somewhat more difficult to describe, but they are growing exponentially thanks to the World Wide Web. The documents exist as discrete components that exist on machines around the world. They are located near the creators of the documents, but are accessible to others from around the world as appropriate. For many, they are the quintessential stuff of the virtual library. Finally, there are interactive documents — the subject of the next section.

3. Hyperdocuments and Infobots

There are two new forms of documents that merit particular attention. The first is the hyperdocument. Vanevar Bush is generally credited with the earliest proposal for a system that has the characteristics that are normally associated today with hypertext, but significant credit goes to many other individuals such as Ted Nelson and Douglas Engelbart. What is significant about hyperdocuments is their base conception. In Bush's mind, they are consonant with the way human's think:

> *The human mind ... operates by association. Selection by association rather than indexing may yet be mechanized. One cannot hope to equal the speed and flexibility with which the mind follows an associative trail, but it should be possible to beat the mind decisively in regard to the permanence and clarity of the items resurrected from storage.*[14]

His initial conception, remember this is 1945, is of a desk and levers.

Consider a future device for individual use, which is a sort of mechanized private file and library. It needs a name, and, to coin one at random, "memex" will do. A memex is a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory.[15]

Hypertext allows users to associate one piece of text with others easily, quickly, and naturally from any point in a manner consistent with human thought. Hypertext may provide new methods of document use beyond current physical and electronic handling. Hyperdocuments allow not only static branches, but dynamic branches, and this makes them interactive.

The second new document form relates to work in virtual reality. In this scenario, documents have not only passive procedures attached, as with current object technology, but active procedures as well. A couple years ago, I wrote about it as follows:

Finally, returning to the library scenario, imagine a situation where we take the notion of a virtual reality a step further. Borrowing from Ted Nelson, we imagine a hypertext where all human knowledge is stored. This hypertext is situated in a cyberspace. Books and articles are objects in the space.[16] *The document object may be active, like a knowbot (Cerf, 1991, p. 74), but for out purposes more like the converse of a knowbot—an infobot. Rather than Cerf's knowbot looking for data as a human surrogate, the infobot is a document surrogate looking for humans to inform. The attached procedures include processes to duplicate information, maybe to charge for it, and to navigate the cyberspace looking for readers. Sophisticated infobots might be sent to a sample population correlate of the Nielsen families to test their attachment rules and to subject them to modification. After suitable testing the objects could be put out on the net to distribute themselves.*[17]

This, can be the ultimate form of junk mail — documents that relentlessly seek out the people to whom they should present themselves. While there are a number of new social structures that will need to be developed here to maintain privacy, the basic notion will work its way into our experience. Rather than people working to find documents, every more intelligent documents will seek out the people to whom they should be presented.

B. Roles in Processing Documents

Technological advances cannot occur unless they are adopted by people. In a very few years, there has been a major transition in the industrialized world. The pace of technological change has put a tremendous strain on people. In the U.S., in the 1870's agriculture was the dominate occupation — with 50% of the work force. By 1910, manufacturing had surpassed agriculture. By 1960, information workers surpassed manufacturing. Today, only 2% of the work force is in agriculture. Information and services sectors dominate with 80% of the workers and by the turn of the millennia, this latter group will grow to 90%. This is a radical upheaval. Looking more specifically at the management of documents, there are even more dramatic cultural shifts:

1. In 1982, the IBM PC was introduced. In 1980, virtually no one had a workstation. By 1995, 90% of the employees in many organizations have advanced multimedia PCs.

---

14 Vanevar Bush,As We May Think **Atlantic Monthly**, 176, 101108, July, 1945.

15 Vanevar Bush,As We May Think **Atlantic Monthly**, 176, 101108, July, 1945.

16 The scenario assumes a cyberspace infrastructure that is based at the very least on the existence of interconnected and interoperable machines. This particular scenario is based further on the assumption that directory information of the type imagined by the X.400 standard is available under X.400, information about the user is stored as an adjunct to the user's identifier, and there is a high level of standardization in data interchange formats.

17 Spring, Michael and Jennings, Michael, Virtual reality and Abstract Data: Virtualizing Information, **Virtual Reality World**, Volume 1, Number 1, Spring 1993, pp. c-m..

2. Five years ago most people had standalone workstations. Today, most organizations are networked, and enterprises are being networked. By 2000, a global network will exist.
3. The need for speed of information delivery has increased from 5 days for mail delivery to overnight mail to minutes for fax or email.
4. The revision rate for information products, the internal complexity of those documents, and the number of people involved in creating them are all increasing dramatically.

Some people are not reacting well to this pressure. Over the last several years, the water fountain of information has been replaced with a garden hose, and most recently with an information fire hose. This information overload problem is often discussed, but unfortunately, it is most often in terms of how to force more into the human brain. There is a need to be sensitive to the human components of the problem and to work to build tools to assist in this process. There are two significant changes in the roles related to documents

1. The Author in Control

Historically, document publication has required that the author work with a number of people to create a document — editors to develop the content, graphic artists to lay it out, and printers to make copies. For the most part, computer systems are making it possible for the author to fulfill all these roles. The most dramatic manifestation of this control comes via the emergence of structural copymarking.

Early electronic typesetting embedded codes that indicated specific physical properties for the type (typeface, font, size, line length, leading, etc.). Modern systems have shifted to style sheets and structural copymarking. Structural copymarking indicates the role of the associated text instead of the physical properties. Instead of marking text as 'Times Roman bold, 18 point, flush right', the structural copymark might be "Major Heading." Structural copymarking was implemented in early systems like Scribe and now appear in leading PC word processors. Definable "styles" and auto formatting capabilities are all examples. WYSIWYG systems have facilitated this shift.

As a consequence of structural copymarking the author can concentrate on the content and the system will produce aesthetically pleasing output. This can eliminate the intermediaries from copy editors to graphic layout artists to proofreaders. Reducing the intermediaries can reduce the preparation time significantly. In many technical fields, the professional communities are concluding that an acceptable editing job done quickly is better than excellent editing done slowly. When the target audience is small and knowledgeable, there is likely to be a high tolerance for minor errors.

The fact that the author can now self publish has two down sides. First, even with the best system, the author is distracted by these new responsibilities. Even if the author successfully manages to format the document, they are doing less authoring and somewhat more designing. Equally critical, the tools are not perfect and there is a risk that the author will do an inadequate job. It is still the case that the results achieved by a professional can far surpass the quality of computer augmented work.

Interestingly, automation is uncovering some weaknesses in the traditional system. Human copy editors make changes that are not consistent across the entire document. In the editing and review process of a recent book, it became clear that the rule-based component-based document was viewed by the editors as a concrete linear document. Some interesting things were found:

1. automated aspects of the document were edited repeatedly — e.g. header and footer.
2. in text changes were made manually rather than by specifying rules. The editor marked every occurrence instead of specifying that all x should be y.
3. modules that were included multiple times were edited differently at different positions.

Typography is also effected by structural copymarking. Instead of getting a specific typographic design for each document, the documents all have a consistent type. This could be an advantage in creating a consistent image or a disadvantage in settling for a generally acceptable format instead of the best one for each document.

Thus, one aspect of the revolution is that the author's sphere of influence and responsibility will increase. Authors will in essence be self publishers, opening up the question of how the reader will be able to distinguish quality work from garbage.

2. Reader Agents

The electronic document market is just beginning to emerge, but it is already clear that there are several consequences for the consumer. They are not confident about where to look for particular information. There is no way to determine that the desired information is not available. At the present time, Internet searches can be analogous to randomly calling everyone in the phone book. There is pressure to better organize the process.

Most physical products are delivered through a multi-level distribution channel consisting of producers, wholesale distributors and retail stores. Consumers rely on the retail stores to have the products they need. The retail stores provide a structure for the shoppers. The customer knows (at least in general) what store to go to purchase a particular type of product. For example, if you needed car parts, you would go to an automobile parts store, not a food store. Since stores specialize in certain types of products, you would have a reasonable expectation of when to give up looking for an item you could not find. Retail stores in turn rely on the wholesale distributors (and to some extent directly on the producers) to provide the products they will sell. The retailers will select the products they will sell from the full product line offered by the producer through the distributors. The distributors keep the retailers informed of new products in an attempt to get the retailers to sell those products.

We find some of these patterns of supply and beginning on the Internet. There are specialized archives and there are services for archives, etc. Without endeavoring to predict specifics, it is clear that agents will be needed to assist consumers of information. The following are tentative examples of the kinds of direct and indirect agents that will have to find a place on the Web:

1. Information Collectors
2. Authorized and Recognized Repositories
3. Cost effective Gophers
4. Distributors

Collectors

The network phone books address the difficulty of finding people on the network. They are in some ways remarkable, in other ways, quite poor. As the cost of usage, some of the directory services now begin by asking the user to provide information. One has built a directory of a million people this way. This service provides access information on the million who have registered. In a few years, as the database grows, the service will cease to be free, especially as it becomes accepted as the significant place to be listed.

Repositories

There will be both authorized and recognized repositories. An authorized repository might be the official place to leave information of archival value or information with legal implications. For example, to record a property deed, it might have to be deposited with a fee at XYZ location on the web. It will then be guaranteed to be available for 20 years. Similarly, there may be places that are recognized for the quality of their collections—either by virtue of inclusiveness or of selectivity and quality. Whether the supplier or user pays for a piece is hard to determine. But an author may want to submit work to XYZ location because he knows that the site keeps everything on the subject and therefore that is where people look. In addition the author might put it there because he knows that all of their material is first rate, and therefore, if they accept his work it will have the characteristic of being considered first rate.

Gophers and Distributors

Imagine that there is a photo of the America's Cup races from New Zealand in two years that everyone in New York City wants. The cost of moving a million bits from Auckland to New York is one dollar. If the cost of

moving it in New York City is ten cents than it makes sense to open a photo store in NYC and have a selection of photos moved there from around the world. If the NYC store sells them for twenty cents, they will make back there investment of a dollar after 10 sales. The other side of this coin has to do with distribution. A supplier may ask a distributor to move something to the places where it will sell if it is available locally at reduced cost. The distributor pays to move it to local retailers and gets a commission as copies are sold.

C. Processes

The processes applied to documents run from creation and dissemination to storage and retrieval. Document processes occur at a number of different levels of detail. For example, creating a document involves the operations of outlining, writing, editing, validating, designing, illustrating, proofing, and displaying. The processes related to accessing them include classifying and organizing them for storage and to formulating queries to retrieve them. This section looks at document processes with an eye to reengineering. In the discussion, the process changes already underway are reviewed.

1. Reengineering

Geoffrey James in **Document Databases** discusses the fact that while technological innovation is required for a revolution, it is not in itself sufficient.

> *Gutenberg's improvement in technology did not by itself cause a revolution in communications. The full impact of his invention was delayed by limitations in the method by which books were produced. Each book was hand crafted to resemble a hand lettered manuscript... The real revolution took place when early forms of mass manufacturing were applied to publications.*[18]

Indeed, there are many who believe that the real revolution related to the book occurred as a result of the efforts of printers such as Aldus Mantius who introduced new book sizes and typefaces.

Michael Hammer is a work reengineering consultant. In essence, Hammer suggests that cutting fat and automating work processes avoids the critical dimension of what needs to be done. He says that reengineering is the process of:

> *Recognizing and breaking away from outdated rules and fundamental assumptions that underlie operations. Unless we change these rules, we are merely rearranging the deck chairs on the Titanic.*[19]

The document processing field is replete with processes that might be reengineered. It has been known for years that the retention of information is increased if text is not printed with justified margins and yet informational text is still justified. Computer systems can format most simple documents based on simple rules. Paragraph and sentence spacing can be done automatically if simply left to the system. Still users persist in doing the physical formatting themselves. Reengineering is a major focus of one of our current research projects called CASCADE — Computer Augmented Support of Collaborative Authoring and Document Editing. Consider a couple of examples from that effort:

1. The processes of logical document structuring and physical document formatting are loosely coupled. Logical document structuring should be able to control and drive document formatting.
2. Entities are referred to directly rather than indirectly making it necessary to manually change references when locations or numbering are changed by the system.
3. Indices are constructed by hand rather than by the system based on stop words and user removal of inappropriate items.

18 James, Geoffrey, **Document Databases**, New York: Van Nostrand Reinhold Co., 1984, p.9

19 Hammer, Michael, Reengineering Work: Don't Automate, Obliterate. **Harvard Business Review**, 90(4), July August, 1990, pp. 104112

4. Documents are created from scratch as unique entities rather than developed from templates and modified as needed.

Document processing needs to be reengineered. In general, there are three types of benefits to be gained.

1. Significant savings in the time to produce the document.
2. A reduction in the costs incurred in producing documents including:
   a. reduced printing cost because of reduced transportation, waste and storage costs;
   b. decreased production staff costs;
   c. increased publication value in sales or revenue related uses;
   d. decreased costs for graphics.
3. New opportunities to do things such as:
   a. updated publications on demand and just in time publishing.
   b. ability to provide (sell) new services
   c. customized publications

2. Processes in Transition

In this section a few trends are examined which are indicative of the kinds of changes taking place in the processing of documents. These include:

1. the increase in the revision frequency
2. the ability to print documents on demand
3. the shift to finding by association

Revision Frequency

Electronic editing has increased the frequency of document revision. Indeed, one of the changes that can be noted is the process of making changes to existing documents rather than starting from scratch each time. The document you are reading reflects the merger and extension of ideas from two older documents on the changes in document processing. Historically, the cost of revisions has been high. Today, there is virtually no cost. In addition, making changes was very costly when it meant that existing inventory had to be scrapped. If there is no inventory the revision cost is reduced to the work needed to make the logical changes. There are several consequences to this reduced cost:

1. Simple corrections can be easily made.
2. Time dependent information can be updated frequently.
3. A series of incremental changes can be made instead of major version changes.

This third change is a two edged sword. With a series of small revisions, how are the readers supposed to know when there are enough changes to merit obtaining a new version. A publisher could announce every small change as “new and improved” hoping to sell more copies of the document. Alternatively, the publisher could withhold any minor updates and only produce infrequent “new editions.” In some ways this is analogous to the problems faced by software companies in deciding when to release program patches to fix problems and new versions with new features. A document with frequent revisions is constantly evolving. If a document is required for archival purposes, this raises the issue of what version(s) to retain.

Printing on Demand

Printing on demand is the process of the consumer obtaining a document electronically and printing a copy for themselves. This has a profound impact on publishing. It also changes the definitions of archival copies. Traditionally, publishers did the printing of documents, creating the plates, putting them on the press, adjust the press until good copies are produced. This set up time can be significant and is not dependent upon the number of copies to be printed once the press is ready. Therefore, there is a strong incentive to print as many copies as will ever

be needed. This spreads the set up cost over a large number of copies and reduces their average cost. Expressed simply, there are economies of scale in long press runs.

Presses are now available that can take the document via a computer interface, make the plates and adjust the press without significant operator intervention. Set up time is reduced but still prohibits economic short press runs. Ultimately, fully electronic printing involves no set up costs, and allows for any number of copies.

There is a more important aspect of printing on demand which has received remarkably little attention to date. Printing on demand shifts production costs from the producer to the consumer. The consumer downloads the electronic document (often paying the communications charges) and consumes their own supplies (paper, ink or toner) and uses their own equipment (printer). This transfer of costs has already occurred with faxes. The advertiser sending the fax is causing you to use your supplies and capital to produce their message. This contrasts with "junk" mail advertisements where the cost is borne primarily by the advertiser.

One of the remaining advantages of mass production in printing has to do with the speed of production. The fastest laser printer just can not compete with a web press. However, with network environments this process is changing. As a result, the whole process of publishing will be turned inside out. Instead of printing and then disseminating, dissemination will occur first and then printing.

Guided Finding by Association

Since the library at Alexandria, the to find material is by looking in an area where information had been stored based on its main focus. Thus, all the books on libraries were located in the same area. Now there new ways of locating relevant materials. These methods were first alluded to by Vanevar Bush in an article published at the end of World War II — As We May Think. Hypertext links between documents allow users to navigate between related documents. One problem that results is that people get confused about where they are, where they have been and where they are going. They are lost in hyperspace. This is particularly like when the environment is dynamic. Changing connections, updating and growth increase the user's cognitive load. Users need a guide to assist with routing, selection and navigation. Just as a teacher serves as a guide to a text book, the guide would increase the value of the document base.

VI. Conclusions

It is our conclusion that document processing is indeed in a state of revolution. There are order of magnitude changes in many areas that have the impact of earthquakes. Basic technology trends including digitization, standardization, accommodation and integration provide a foundation for the revolution. Enabling technologies including the information infrastructure, electronic currency, security techniques and intellectual property rights provide the materials necessary for the revolution. There are a variety of factors that are driving emerging roles for information wholesalers and retailers.

There are changes in products, roles, and processes that also seem to indicate that a revolution is underway. The authoring process has become a logical task with the computer handling much of the physical design. Being able to download and print documents on demand has led to dissemination followed by printing in contrast to centuries of printing followed by dissemination. Dynamic documents are emerging that change the way we think about these heretofore static artifacts of communication. In concluding, there are two points that merit closer examination—the economics of documents and the organizations that support that economy.

A. Economics of Documents

Economics drives much of what goes on in the world. Fortunately or unfortunately, it is increasingly a fact of life. With the move to an information based economy and a digital world, many things will move from atoms to bits and those bits will increasingly be treated as commodities. Negroponte has an analysis in his book "Being Digital".[20] about the conversion from an economy based on atoms to an economy based on bits. His premise is

that where an atom can be replaced by a bit, there are advantages to the substitution and they will have an impact on how business is done.

Books and documents are the quintessential examples of the atoms to bits shift. Where are the new economies of scale in the document business? What are the instances of network externality that are about to emerge? The new bit based documents will require new hardware and software. Users will have to be attached to new distribution system. However, unlike a physical distribution system, the costs of the electronic system are decreasing. As the start up costs of joining the network go down—by the rules of network externalities, it can be expected that users and providers in this new environment will achieve economies of scale. When selling physical goods, the store must invest in inventory for each item sold and each item uses up shelf space in the store. Therefore, there is a significant marginal cost for each item carried. The document market has a marginal cost of "stocking" a document that is close to the cost of the disk space consumed. The additional cost for each item sold is almost zero. If the document seller receive revenue for every document sold and their marginal cost is very low, there is a tremendous incentive to increase sales volume. Physical retail is driven to larger stores by economies of scale that reduce marginal cost. Document retail has a low marginal cost and will be driven to maximize revenue.

Current information providers are analogous to the small, Mom and Pop retailers. With decreasing start up costs, more and more providers are entering the market. Economics provides an argument for increasing size. Customer service (know where to look and when to stop looking) will further drive the consolidation into stronger wholesalers and retailers of electronic documents.

B. People, Technology and Organizations

There are significant unknowns about the process of introducing technology in the workplace and in a culture. The process can become destructive, or at very least nonproductive, if it is not handled well.

At the Institute for Research on Interactive Systems (IRIS) at the Rand Corporation, Bikson and others have been studying the process of introducing technology in the workplace.

> *We found that site to site variations in the success of implementing new technologies were more fully explained by differences in the implementation process itself than by differences in the systems or in the organizations.*[21]

Among the specific findings were the following:

1. Better implementation is achieved if social as well as technical issues are addressed during implementation.
2. Satisfaction and production increased when learning to use the computer is supported.
3. Implementation is more successful when users feel capable of exerting a substantive, positive influence on the spread of technology.
4. Implementation succeeds when the software is suited to user needs
5. Implementation succeeds when the program was modifiable by the user.

In **In the Age of the Smart Machine: the Future of Work and Power**, Shoshana Zuboff traces the long history of automation. The application of power technology to the production process is automation. She reviews the impact of Frederick Taylor and the industrial engineers who developed the time motion and document flow studies. They are the people that took the holistic joy out of finishing the job you started. In essence, they are the people who created word processing pools and piece work. They are the people who said anyone can type anything for anybody — it is simply an I/O process.

---

20 Negroponte, Nicholas, **Being Digital**. New York: Knopf, 1995.

21 Bikson,T., Gutek, B., and Mankin, D., Implementing Computerized Procedures in Office Settings. Rand Institute for Research on Interactive Systems, R3077NSF/IRIS, p. v.

Zuboff says that the introduction of the computer has caused something new to occur. Where the steam engine caused automation, the computer causes information. Rather than automating, we are informating. More importantly, while the industrial revolution was characterized by piece work and the fragmentation of work, the information revolution may well be characterized by informated control of the overall process once again. As an example, the use of computers has made it possible for a single individual to control the paper making process in a very large plant.

The task then is to make use of this new technology in a way to help people help organizations. What is needed is a new way to think about people and organizations. In the informated organization:

> *Members can be thought of as being arrayed in concentric circles about a central core, which is the electronic data base. . . On the innermost ring, nearest to the core, are those who interact with information on a real-time basis. They have responsibility for daily operations.*[22]

22 Zuboff, Shoshana, **In the Age of the Smart Machine: the Future of Work and Power**. New York: Basic Books, 1988, p. 396