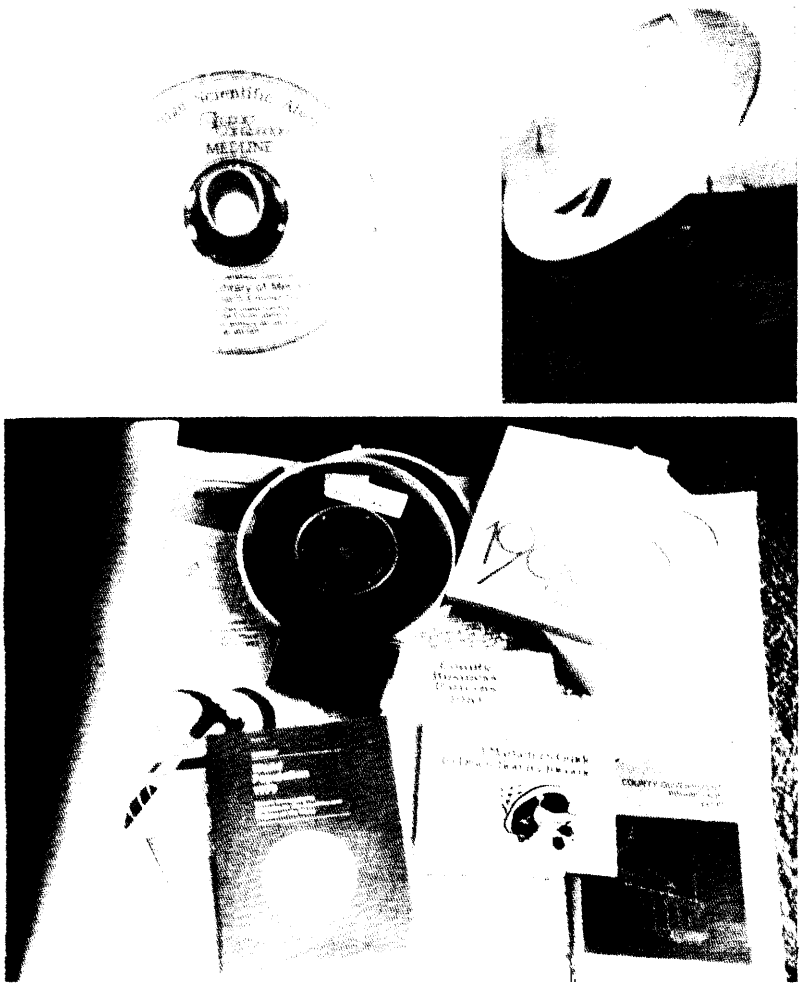
Chapter 3

Key Technology Trends Relevant to Federal Information Dissemination



Clockwise from top left: National Library of Medicine CD-ROM Disk (photo credit: Doug Jones, National Library of Medicine); satellite (photo credit USA Today, all rights reserved); and gallery of Bureau of the Census data products (photo credit: Neil Tillman, Bureau of the Census).

CONTENTS

Chapter 3

Key Technology Trends Relevant to Federal Information Dissemination

SUMMARY

The results of the General Accounting Office (GAO) surveys of Federal information users (see chapter 4) and prior studies on the future of paper and paper-based media (such as books) indicate that paper is expected to decline only marginally as a preferred format in the next few years, although this decline could become significant for specific types of information (e.g., bibliographic, reference, statistical, scientific, and technical) that are highly suited to electronic access and manipulation. The outlook for microform is less favorable. Microfilm is currently used very little for Federal information dissemination; microfiche, while used extensively, is expected to decline significantly as a preferred format, except for records storage and archival purposes.

In sharp contrast, the GAO surveys of Federal agencies (see ch. 2) and Federal information users (ch. 4) indicate that plans and preferences for dissemination in electronic formats (e.g., electronic mail and bulletin boards, optical disks) are projected to increase dramatically.

This chapter surveys a number of major technologies and key technical trends relevant to Federal information dissemination. Several key technical trends are expected to continue conservatively for 3 to 5 years and in many cases for at least 10 years, and are combining in such a way that most of these plans and preferences are likely to become reality. These trends include:

- continued steady improvement in the price/performance of microcomputers, which already bring the power of mainframe computers to the desktop at the cost of a stereo set; microcomputers provide the technological underpinning for numerous information collection, processing, and dissemination activities;
- continued, rapid proliferation of desktop publishing systems, comprised of a microcomputer, nonimpact printer, and page composition software (and sometimes a scanner for paper input) in the most basic configuration;
- continued, rapid improvement in the power of desktop publishing software to handle more complex documents, formats, fonts, and the like;
- continued, steady improvement in the price/performance of nonimpact printers, with low-cost desktop printers offering output quality acceptable for most documents, and high-end printers offering quality comparable to some phototypesetters;
- similar improvement in the price/performance of scanners, with the capability of high-end scanners (to handle a wide range of type styles and sizes) migrating to desktop scanners;
- as a combined result of the above trends, overall continued improvement in the ability of desktop systems to produce higher quality, more complex documents, thus further reducing the gap between desktop and high-end electronic publishing and phototypesetting systems;
- for complex, large-volume, and/or large institutional applications, continued improvement in high-end electronic publishing characterized by:
  - –declining cost of software and workstations;
  - —increasingly heavy competition between desktop and high-end systems;

—rapidly growing networking of desktop and high-end systems;
—rapidly growing networking of workstations with high-end nonimpact printers and phototypesetters;
—increasingly heavy competition between and among software, workstation, phototypesetter, and computer equipment vendors, as well as systems integrators and service bureaus and;
—continued migration of electronic publishing to other applications such as forms management and multi-format output;

- continued, rapid increase in the number and use of computerized online information services, especially for information search and retrieval, electronic display, and remote printing-on-demand when needed;
- continued, steady increase in the number of online information gateways that provide the channels for electronic information exchange (such as electronic data transfer, mail, facsimile, and bulletin boards), but not the information itself; these gateways include common carriers (interexchange and bell operating companies), value-added companies, and nonprofit and governmental systems;
- continued advances in the telecommunication technologies that underlie online information services and gateways, including packet switching, fiber optics, satellite networking, FM subcarrier transmission, and integrated switched digital systems;
- rapid advances in optical disk technologies and applications, especially for purposes of information storage and dissemination; advances include:

—accelerating penetration of Compact Disk-Read Only Memory (CD-ROM) as remaining standards issues are resolved;
—maturation of Write Once Read Many-times (WORM) and erasable optical disks (compact and full size) as technology stabilizes and standards are established;
—emergence of Compact Disk-Interactive (CD-I) and other optical disk applications;

- rapid advances in development of expert systems applicable to many aspects of information dissemination-including technical writing, indexing, information retrieval, and printing management; and
- continued, steady progress in development and acceptance of standards for telecommunication, data transfer, optical disks, and page description and text markup.

The net, cumulative effect of these technical trends is to afford the Federal Government the opportunity to realize the kind of significant performance improvements and cost reductions that have been demonstrated in the private sector. Also, the convergence of these technical trends, along with progress in standards-setting, makes information systems integration a real possibility for the Federal Government and other users. Systems integration permits the coupling of input, storage, processing, and output technologies in ways that permit multi-media (e.g., paper, microform, online electronic, and stored electronic) dissemination from the same electronic database. In essence, the key technologies and technical trends highlighted above are central to the emerging movement towards systems integration.

INTRODUCTION AND OVERVIEW

A major objective of this study is to identify and discuss new or evolving ways in which information technology can or might be applied by the U.S. Government Printing Office (GPO), National Technical Information Service (NTIS), and other Federal agencies to the dissemination of Federal information. An important step in this process is the identification of key technology trends that are relevant to information dissemination.

OTA has surveyed a representative cross-section of major technologies relevant to information dissemination. The presentation in this chapter emphasizes electronic technologies, although paper and microform are discussed briefly. Conventional ink-on-paper printing technologies, including prepress and binding, are considered in chapter 4 in the context of alternative futures for GPO. As discussed in chapter 4, GPO has already upgraded its conventional printing technology to a level comparable to most of the private printing industry, However, GPO has much less experience with online information systems, expert systems, optical disks, and high-end electronic publishing. NTIS is in a similar situation (see ch. 5), as is the Depository Library Program (DLP). Libraries in general, especially the larger research libraries, have more experience with electronic systems (see ch. 6). Thus, this chapter is highly relevant to general consideration of future technological directions for GPO, NTIS, and DLP, as well as the Federal mission agencies.

This chapter emphasizes significant current or emerging technical trends that are expected to persist for at least 3 to 5 years into the future. In many instances, the key trends are likely to continue even longer–perhaps for 10 years or more. In the aggregate, the trends provide what OTA believes is a reliable overall technology planning framework for Federal information dissemination. However, the presentation in this chapter is not intended to be used in the evaluation and selection of specific equipment and systems. Some examples of equipment, vendors, and/or applications are provided, but for illustrative purposes only. Some cost and performance data also are included for illustrative purposes only. These data change rapidly, and should be checked with vendors if product or service-specific evaluation is contemplated. Also, the presentation is geared to the level of the informed lay person, not to the technical specialist. For discussion of specialized information technologies not included here (e.g., digital data tapes, digital cartographic systems), the reader should consult a forthcoming OTA staff paper on *Federal Scientific and Technical Information Dissemination.* Finally, for the discussion of telecommunication technologies not covered here (e.g., digital facsimile, videotext, cable television), the reader should consult OTA'S forthcoming report on *Commumnications Systems for An information Age (1989).*

Information Systems Integration

One important finding is that a combination of technological advances, cost reductions, and current or pending standards has opened up for the first time a real possibility of information systems integration in the Federal Government. The importance of this development cannot be overstated. Until recently, the Federal Government, along with other major information system users, had no choice but to obtain an essentially incompatible complement of information technology hardware and software, such that system integration across the government and major agencies was very difficult, if not impossible. Integration of specific systems within agencies was possible. But even here, major integration efforts, for example in the Department of Defense, still led to numerous incompatible systems.

The relevance to Federal information dissemination is immediate and direct. It is now possible to conceive of, plan for, design, and implement an integrated information dissemination system (or more likely a network or hierarchy of systems) for the Federal Government. This is possible because of advances in a whole range of relevant technologies-including input, storage, processing, and output technologies-that can deal with the entire range of media, including paper, microform, magnetic disk, optical disk, and direct electronic. And the cost/performance trends in these technologies are likely to make a wide range of applications cost-effective when compared to conventional methods.

Two other related trends are equally important. One is the trend toward standards for systems interconnection at the hardware, software, and applications levels. There is strong movement among the vendor and user communities and in the various national and international standards bodies towards a hierarchy of standards that will make it possible for a wide range of information systems to talk with and exchange information with each other.

Another trend is the rapid penetration of computerized information systems in all sectors of society, but especially in the business, educational, and research communities. This means that many of those who provide information to the government and use information from the government can now or soon be electronically connected, and can, where appropriate, send and receive information in a variety of electronic formats. This in no way suggests an end to paper-based information products—but only that paper can be used where it is really needed and in a more efficient and cost-effective manner.

Realizing this potential for information systems integration requires, of course, more than just the technology and standards. A variety of institutional and policy changes maybe necessary, and various alternatives will be discussed in later chapters.

Nonetheless, it appears that the technology, the industry, the standards, and the government are all moving towards systems integration. It is now possible to envision, in the relatively short term, a Federal information dissemination environment that includes the following illustrative elements:

- document/data entry (e.g., scanning, word processing, facsimile);
- document revision/composition (e.g., electronic publishing-desktop and high-end, computer graphics);
- document storage (e.g., electronic database, optical disk);
- document output (e.g., electronic publishing, laser printing, photo offset); and
- document distribution (e.g., optical disk, electronic mail, computer diskette, online electronic, paper copies, microform).

Indeed, electronic publishing can be viewed as a key integrative technology because it can serve to integrate the various formats (paper and microform as well as electronic) of information input, processing, storage, and output within a common technical framework. Electronic publishing can also serve to connect the various so-called islands of automation in an organization-office automation, publishing systems, database systems, records management, document storage systems, and the like. Standards on information exchange are critical, as is the need to find ways for the people who work in various areas of automation to work more effectively together.

The Microcomputer Revolution

Most Federal Government information is either collected from the private sector, State/local government, and the general public, or is created by Federal employees and contractors as the result of studies, analyses, research, and the like. Even information collected from outside the government is frequently subject to analysis by Federal employees, and in that sense has a creative or value-added aspect.

The dominant technology relevant to the collection and creation of Federal information is the microcomputer. Over the last 5 years or so, the United States has witnessed a revolution in computer technology that has brought the power of the mainframe computer to the desk of millions of public and private sector employees and citizens. And this revolution is expected to continue for at least another 5 years.

The sheer magnitude of this microcomputer revolution can be measured in many different ways. For example, the Federal Government itself has gone from only a few thousand micros in 1980 to roughly 200,000 in 1986 to 500,000 in 1988 to a projected 1 million by 1990. The percentage of school districts with computers had already increased from about 18 percent in 1981 to over 90 percent by 1985, according to the National Center for Educational Statistics.[1] In the business community, microcomputers drew even in total computing power with mainframes and superminis as of 1985 and are projected to dominate by 1990, according to Dataquest.[2] Indeed, various projections show microcomputers growing at an average 10 to 15 percent through 1990, compared to about 5 percent for mainframes and superminis. The logic of this trend is understandable when one considers that the IBM personal computer systems, selling for less than $10,000 are equivalent in computing power (measured in millions of instructions per second) to the IBM 370-168 mainframe computers that sold for several million dollars when introduced in the mid-1970s.

Even the home market has had significant microcomputer penetration, with about 19 million households buying a microcomputer since 1981 (about 14 percent of all households).[3] Link Resources projects an ultimate home penetration of about 35 percent, although this may be conservative if full function microcomputers drop to the $300-500 price range in the next 3-5 years.[4] This would place the microcomputer in the same price range as a good quality 19-inch color television.

The continuous improvement in price/performance of microcomputers is driven in part by advances in semiconductor chip technology, which shows no signs of slowing down. The 32-bit chip family (such as the Intel 80386 or the Motorola 68020) made possible the latest personal computer systems that are more powerful, more user friendly, and more compatible with each other. Price/performance is expected to continue to improve as the 32-bit chips are further assimilated in microcomputer product offerings and as next generation microcomputers are developed and introduced.

The important impact on Federal information collection and creation is that an already large and increasing amount of information is generated in electronic form, that is, by capturing electronic keystrokes with a microcomputer or word processor. Today, much of this electronic information is submitted to or provided by the government in paper form. But the potential exists to substantially reduce the amount of rekeyboarding, and presumably the cost of such information, by maintaining the information in electronic form as long as possible.

The Continuing Role of Paper and Microform

A note of caution with respect to the role of paper is in order. Despite the dramatic increase in computer technology and electronic information, paper documents are expected to have a continuing, major role for several reasons. First and foremost, for documents of significant length, research has found that reading from a computer screen is much more difficult than reading from paper, despite improvements in the design and resolution of

[1]Cited in J. Bloomdecker, *Computer Crime, Computer Security, Computer Ethics* (Los Angeles, Calif. National Center for Computer Crime Data, 1986).

[2]Cited in G. Lewis, Zoom! Here Come the New Micros, " *Business Week*, Nov. 1, 1986 pp. 82-92.

[3]E. Roth, "Power Surge in Personal Computers, " *Editorial Research Reports*, vol. 1, No. 1, Jan. 9, 1987, p. 4.

[4]Ibid, p. 6.

screens and terminals. Even extensive practice at electronic reading does not appear to make a significant difference. Second, paper continues to be a more convenient and portable medium for many purposes, and accommodates a wide range of reading styles and locations. Third, for many documents, paper is still a bargain, although this is changing with the advent of optical disk storage technology. And of course, electronic publishing can significantly increase the efficiency of paper use, even when the final product is still in paper format. Fourth, the paper format (especially for lengthy reports and books) permits the reader to browse through material and use a variety of conscious or subconscious search patterns that may be difficult if not impossible to replicate even with today's computer-based search and retrieval software. Reading paper formats can lead to greater comprehension.

Overall, most studies on the future of paper and paper-based media (such as books) have concluded that the paper format will play a major role as a medium of information storage, exchange, and dissemination for the foreseeable future.[5] The results of the GAO surveys of Federal information users (summarized in ch. 4) indicate that paper is expected to hold steady or decline only marginally as a preferred format in the next few years, although this decline could become significant for specific types of information (e.g., bibliographic, reference, and statistical) that are highly suited to electronic access and manipulation. At the same time, the preference for electronic formats (e.g., electronic mail and bulletin boards, floppy disks, and compact optical disks) is expected to increase dramatically.

The outlook for microform (microfilm and microfiche) is not as favorable as for paper or electronic formats, but there is likely to be continuing significant use of microforms for records storage and archival purposes for the foreseeable future, or at least until electronic alternatives have been fully established and stood the test of time. Microform is well suited for archival purposes because it requires less storage space (compared to paper), has a longer shelf life (compared to paper and electronic, although this may change), is a stable access technology (compared to electronic), and is lower in cost (compared to paper and some electronic).' Microform offers a lifetime of 100+ years, whereas the lifetime of acidic paper is perhaps several decades, and magnetic media (tape and disks) a few years to a decade or two. The main competitive threats to microform for archival purposes are from acid-free paper (which can last 100 + years, but would still require more storage space and be more costly than microform) and optical disks. Optical disks do not as yet have proven archival capability (although manufacturers are claiming 40+ years), require less storage space, and can be less expensive than microform. Microform is likely to continue as a major archival medium at least until optical disks (or some related electronic-storage technology) are well established.

However, for many nonarchival purposes, microform is not the preferred medium even today. For reading lengthy written materials, users find microform to be inconvenient, uncomfortable, and inefficient compared to paper. For information search and retrieval, users frequently prefer electronic formats, including online database systems as well as, increasingly, offline media such as CD-ROMs. The results of the GAO surveys of Federal information users (summarized inch. 4) indicate that microfilm is little used today for Federal in-

[5]See, for example, Priscilla Oakeskott and Clive Bradley (eds.), *The Future of the Book: Part I -- The Impact of New Technologies* (Paris Unesco, 1982); U.S. Congress, Joint Committee on the Library, *Books in Our Future*, A Report From the Librarian of Congress to the Congress, S. Print 98-231, U.S. Government Printing Office, Washington, DC, 1984; John Y. Cole, *Books in Our Future: Perspectives and Proposals* (Library of Congress, Washington, D.C. 1987).

[6]See, for example, Kenneth E. Dowlin, *The Electronic Library: The Promise and the Process*, (New York, Neal-Schuman, 1984; F.W. Lancaster, Libraries and Libraries in An Age of Electronics (Arlington, VA: Information Resources Press, Arlington, VA, 1982); Edward Gray, "The Rise and Fall of Technological Applications: Considerations on Microforms and Their Possible Successor, " *International Journal of Micrographics and Video Technology*, vol. 15, No. 1, 1986, pp. 31-38; National Research Council, Committee on Preservation of Historical Records, *Preservation of Historical Records* (Washington, DC, National Academy Press, 1986).

formation dissemination, and that microfiche, while used extensively, is expected to decline significantly as the desired format for dissemination of many types of Federal information. It should be noted, however, that the micrographics industry is itself using electronic technology to continuously upgrade microform access technologies, such as computer-assisted microfiche retrieval systems and computer-output microfilm systems.[7] AIso, the tecology for microfiche to paper conversion continues to advance. For example, the Defense Technical Information Center recently funded the development and installation of duplex (two-sided) microfiche copier machines.

In sum, however, the current and future use of paper, microform, and electronic formats will depend largely on the type of information and the type of information user. The kinds of criteria that will be relevant in selecting format include:

- amount of information to be stored, accessed, and/or distributed;
- amount of storage space available;
- frequency of information access or retrieval;
- length of time information is to be stored;
- desired speed of access or retrieval;
- costs of storage, access, and retrieval;
- number of users; and
- technical expertise of users.

The rest of this chapter focuses on several key electronic technologies relevant to Federal information dissemination. The price/performance characteristics of these technologies make them highly competitive with paper and microform for those types of information well suited for electronic formats.

[7] See, for example, Coopers and Lybrand, *Information and Image Management: The Industry and the Technologies*, study conducted for Association for Information and Image Management, Silver Spring, MD, 1987.

ELECTRONIC PUBLISHING

Desktop Publishing

One of the microcomputer applications most relevant to this study is desktop publishing. Desktop publishing combines elements of advanced word-processing and computerized page layout and composition systems. Desktop publishing can be defined as a set of hardware and software, including a multifunction personal computer, which has the ability to produce near-typeset quality output, and utilizing multiple type fonts, sizes, and styles and multiple page layouts. The characteristics of desktop publishing are:

- low cost (about $10,000 for an entire system),
- user-friendly software (frequently employing icons and a mouse and a 'what you see is what you get or WYSIWYG screen display) that requires minimal training,
- near-typeset quality output (but still considerably less than high-quality book and magazine printing, for example), and
- relatively simple and straightforward documents (although the desktop publishing software is much more sophisticated than typical word-processing software).

A typical desktop publishing configuration includes:

- a microcomputer with mouse or digitizing tablet, keyboard, and screen (roughly $2,000 to $4,000 inclusive);
- page composition software (about $500 to $1,000);
- a low-end laser printer (about $1,500 to $3,000); and
- a low-end scanner for paper input (about $2,000 to $4,000).

Desktop publishing is expected to become a standard part of personal computing, and to grow significantly over the next several years. Growth in desktop publishing reflects the substantial potential savings for those types of documents that do not require higher levels of quality and complexity. For simple reports, newsletters, pamphlets, and the like, desktop publishing can cut composition costs from the $50 or more per page range (for commercial composition and typesetting) to the $1 to $5 per page range. Compared to straight word-processing text, desktop publishing can reduce the page length by perhaps 40 percent on average, and this translates into substantial savings from reduced paper and mailing costs. There are also major savings from a streamlined revision process, minimal rekeyboarding, and the ability to store text and graphics for future use and revision.

Two significant limitations of the low-end desktop publishing are limited ability to deal with complex documents (e.g., complicated layouts using text and graphics) and limited print quality (due to the typical 240 dots per inch (dpi) or 300 dpi resolution of low~nd laser printer output). The first limitation is being mitigated rapidly by ever more powerful desktop publishing software releases. Also, users can invest in more sophisticated software and, if necessary, obtain software that supports phototypesetters as well as laser printers. Finally, low-end laser printers are improving output resolution, thus reducing the print quality differential between laser printers and phototypesetters.

Desktop publishing has made dramatic inroads in the newspaper and newsletter industries. An estimated 80 percent of newspapers with a circulation of over 100,000 use Macintosh-based desktop publishing, including the *Wall Street Journal* and *USA Today,* as do an estimated 75 percent of newspapers with a circulation over 50,000.[H] Knight Ridder and the Gannett Corp. are using desktop publishing systems to create and distribute graphic designs nationwide. While major newspapers generally use photocomposition equipment for typesetting in order to get higher print quality, newsletter and some small newspaper publishers frequently find that laser printer typesetting quality is good enough.

The potential implications of desktop publishing for Federal information dissemination seem just as significant. An increasing percentage of Federal information collected, created, and disseminated would appear to be well suited for desktop publishing.

High-End Electronic Publishing

The distinction between desktop publishing and so-called "high-end" electronic publishing is somewhat arbitrary, since microcomputer-based desktop systems can be connected or networked to high-end work stations, typesetters, and printers. Electronic publishing is considered to be the electronic preparation of material at all pre-press stages of the publishing process, including text and graphics preparation, page layout, and composition, with the actual printing in any of a variet y of formats—paper, microform, magnetic tape or diskette, optical disk, or direct electronic. In general, high-end electronic publishing is distinguished by:

- high volume (in number of pages and copies),
- high quality (of the final product),
- high complexity (of the page layout and composition), and
- high cost (compared to desktop systems).

High-end systems typically cost $30,000 to $150,000 depending on the configuration, compared to $5,000 to $10,000 for desktop systems. For the software alone, high-end publishing systems typically cost $15,000 to $30,000 compared to $500 to $1,000 for desktop software.

The demand for high-end electronic publishing (and to a lesser extent desktop publishing) is driven by a powerful combination of advantages that translate into significant cost savings and productivity improvements. For example, electronically published materials are generally found to be:

- more attractive,
- easier to read,

"F. Seghers, "In News Graphics, Macintosh Makes the Front Page," *Business Week,* Jan. 19, 1987, p. 87.

- more timely (publishing time can be anywhere from 25 to 90 percent faster), and
- much less expensive.

Cost savings can be realized in several ways. For example, electronic publishing generally reduces the total number of document pages by 40 to 50 percent, but occasionally up to 80 percent, since typeset pages contain more text than typewritten pages. This can dramatically reduce paper costs for hard copy print runs. For documents with limited demand and low volume, electronic publishing makes printing-on-demand a realistic option. Electronic publishing also facilitates the revision process by minimizing rekey boarding and graphics redesign.

Various market surveys project a strong demand for electronic publishing over the next 5 years, based on a perceived need for electronic publishing by major corporations and government agencies.

Electronic publishing systems have made rapid technical advances in just a few years. This trend is expected to continue due, in part, to heavy competition among graphics workstations, publishing software, traditional photocomposition services, and computer equipment companies, as well as systems integrators that combine hardware and software from numerous vendors. At the heart of electronic publishing systems is the 32-bit workstation that permits complex manipulation of text, graphics, and, increasingly, halftones. These are the same types of workstations used for computer-assisted design (CAD) and sophisticated graphics applications. This workstation is now an established technology, with a substantial track record. According to Dataquest, 32-bit workstation sales ($15,000 to $50,000 per workstation price range) grew from about 100,000 units in 1983 to 1 million units in 1985, an estimated 2 million in 1987, and a projected 4 million in 1989.[9]

The technical power, sophistication, and flexibility of electronic publishing systems are illustrated by a typical system which uses 32-bit workstations (such as Sun, DEC, or Apollo), a 19-inch monochrome display with a high resolution screen, and a local area network. The typical system can accept input from CAD workstations, scanners, graphics (raster and vector, line art and halftones), spreadsheets, and text (in standard formats compatible with almost any mainframe, mini, microcomputer, or word processor). The system provides output to various laser printers (such as Xerox, Kodak, Imagen, Apple, and Sun) and phototypesetters (such as Linotronic and Compugraphic). Advanced software capabilities typically include:

- integration of text and graphics in non-structured pages;
- free-form drawing with a mouse;
- tracing tablet to copy drawings:
- editing of digitized line art;
- pixel-by-pixel editing of halftone photographs; and
- simultaneous editing of different portions of the same document.

Overall trends in electronic publishing include the following:

- movement from a fragmented market to an integrated market;
- aggressive competition from electronic publishing systems offered by traditional phototypesetters and by electronic publishing service bureaus;
- standardization of information exchange among different types of hardware and soft ware;
- declining price/performance ratios; narrowing of the technical differences between desktop and high-end publishing systems;
- increasing integration of direct-to-plate printing technologies; and
- increasing speed and quality of performance (including higher resolution, color, and multiple languages).

In the corporate community, investment in electronic publishing is generally claimed to have a rate of return of 50 to 60 percent and a payback period of 2 years or less. Also, companies typically claim to have cut overall pub-

[9]Cited in G. Lewis, "New Micros,' op. cit.

lications turnaround time by 50 to 75 percent. While similar data are not yet available from government users, Interleaf Corp. indicates that the following Federal agencies are using Interleaf electronic publishing systems: Defense Advanced Research Projects Agency (for technical reports); Office of Naval Research (for research studies); intelligence agencies (various applications); U.S. Coast Guard (for technical manuals); U.S. Army (for technical manuals); Department of State (for regulations); Department of Agriculture (for statistical documents); Bureau of the Census (for statistical reports); and Federal Reserve Board (for financial analyses). Xyvision reports sales of electronic publishing systems to, among others: the National Center for Health Statistics, National Center for Disease Control, U.S. Geological Survey, Bureau of the Census, and Central Intelligence Agency.

Electronic Forms Management

Another growing application of systems related to electronic publishing is electronic forms management. Several companies specialize in this applications area. The typical stand-alone workstation, including a processor and hard disk along with software and a high resolution display, costs in the range of $25,000 to $60,000, depending on memory size. The typical system has many of the capabilities of electronic publishing systems discussed earlier, and can be used for designing newsletters, manuals, and technical documents as well as forms.

However, it is not necessary to have full-capability electronic publishing systems for many forms-management applications. For example, among Federal agencies, the Air Force, Army, Navy, Internal Revenue Service, Social Security Administration, Federal Reserve Board, Federal Deposit Insurance Corporation, and National Aeronautics and Space Administration are all using personal computers and laser printers to manage forms. Microcomputers are used to enter the data, and laser printers are used to merge the data with standard forms for printing. The agencies indicate direct savings on the order of 40 percent over preprinted forms.

For larger print runs, even greater savings may be possible where offset printing can be used to reduce the per-page printing cost of about two cents assumed for laser printers. Further savings seem likely since the forms can be stored and edited electronically, minimizing rekey boarding and redesign. As another example, the combination of microcomputers (or mainframe terminals) and laser printers can be used to permit direct electronic input of data collected by agencies into standard reporting forms stored on the laser printer. The completed forms can be transmitted electronically to a regional office or to Washington, DC, eliminating both cost and potential errors associated with rekey boarding and the time delays associated with mail delivery. Paper copies can be printed out for archival purposes.

Computer Graphics

Advances in computer graphics are central to the recent breakthroughs in desktop and high-end electronic publishing. Indeed, computer graphics capabilities are key aspects of most electronic publishing systems. And publishing applications have themselves become one of the driving forces for further advances in and broader use of computer graphics. Other driving forces include:

- graphics needs of the scientific community;
- military applications of computer graphics, most recently stimulated by the Strategic Defense Initiative's requirements for very sophisticated, three-dimensional, dynamic computer graphics and modeling;
- continued movement toward graphics standards; and
- continued breakthroughs in price/performance ratios.

Major technical trends in computer graphics include:

- the continuing transition from film-based techniques to digital processing;

- development of relatively low-cost (under $15,000) desktop color scanners and printers;
- further improvements in high resolution graphics (up to 2,000 x 2,000 pixels);
- further development of full color, interactive, three-dimensional graphics workstations at relatively modest prices (e.g., $30,000);
- continued migration of high-end workstation capabilities to low-end workstations;
- and progress in developing standards for exchanging graphics data between workstations, such as the Digital Data Exchange Standard.

Further technical progress in computer graphics seems assured as various companies continually develop new products for top secret military applications. Advanced digitized mapping techniques are used by the Defense Mapping Agency and by various Federal civilian agencies such as the Fish and Wildlife Service, National Oceanic and Atmospheric Administration, and Bureau of the Census. Computerized graphics products can interpret infrared aerial imagery and produce maps. In general, computerized mapping offers advantages similar to computerized printing in that:

- The original map preparation is much faster.
- Maps can be stored electronically to facilitate relatively easy updating.
- The original map and an-y revisions can be displayed on a video screen.
- Hard-copy output can be obtained relatively quickly with a plotter or laser printer.

Scanners and Printers

Almost all desktop and high-end electronic publishing systems are configured to include one or more printers, and may include one or more scanners. The price/performance of scanners and printers has dropped dramatically in the last few years. Scanners are used primarily to digitize text and images that are initially in paper formats. Scanners are not as efficient as direct electronic input, but are much more cost-effective than rekey boarding or redrawing those materials not in electronic format.

The cost of scanners has dropped to the point where low-end scanners are available in the $2,000-4,000 price range with a speed of up to several pages per minute and a scanning resolution of 200-300 dots per inch. While satisfactory for many desktop applications, higher speed and resolution are generally needed for high-end publishing purposes. High-end scanners are available in the $15,000 to $40,000 price range with speeds of 1 or 2 pages per second and resolution levels up to 400-dp~. Thus, the high-end scanners achieve speeds and resolutions similar to the high-end laser printers discussed later.

A major advantage of high-end scanners is the capability to approximate graphics-quality halftone pictures. This is accomplished by scanning the image at up to about 120 scan lines per inch and recording multiple bits for each pixel, rather than the one bit commonly used for scanning text and line art. Instead of recording black or white with one bit per pixel, multiple bits permit the recording of the degree of blackness for each pixel, known as gray-scale scarming. Also, many high-end scanners can scan a wide range of type styles and sizes, and some scanners can be programmed to learn new (to the scanner) type styles. These capabilities are expected to migrate to the desktop scanners.

The technical status of printers is more complicated because printers are now used for functions other than printing, such as typesetting, graphics input, and forms management.

For printing of straight textual material, electromechanical line printers (known as impact printers) are efficient for low-copy runs (i.e., one or a few copies per original). Medium performance impact printers can print at about 1,000 lines per minute (20 pages per minute at 50 lines per page) and cost $10,000 to $15,000. High-end impact printers can reach output speeds of about 3,600 lines (72 pages) per minute. Low-end desktop line printers print at a few pages per minute.

The role of impact printers is expected to continue to decline, because of the need for printing graphics and complex page layouts, the use of variable type styles and fonts, and the integration of forms and data at the point of printing.

Non-impact printers (using laser, light emitting diode array, ion deposition, and other technical processes) provide better quality, greater flexibility and diversity, and faster speed (at the high end). While on a per-page cost basis, non-impact printers may be more expensive than impact printers, this is not an appropriate comparison in most cases. When serving as a typesetter or graphics printer or proof printer, the non-impact printer can be an order of magnitude cheaper than conventional methods. For example, for low-volume applications where 300 dpi output resolution is satisfactory, a desktop laser printer at $1,500 to $3,000 may be perfectly adequate for producing camera-ready copy, compared to a phototypesetter at $35,000 to $70,000.

The high-end, non-impact printers are still quite expensive, typically in the $100,000 to $200,000 range although these prices are expected to come down. A typical high-end, non-impact laser printer prints at 90 to 120 pages per minute. By comparison, a desktop laser printer prints a few pages (e.g., 3 to 6) per minute and costs as low as $1,500. At the next level up, a typical mid-range laser printer might print at 12 to 20 pages per minute and cost $10,000 to $15,000. Again, price/performance ratios continue to fall.

Non-impact printers are not well suited for jobs requiring high output quality and/or print volume. With respect to quality, most non-impact printers can achieve an output resolution of 300 dpi (assuming that the input resolution is at least at that level). This output quality is adequate for a wide range of purposes, but not for high-quality publications. By comparison, photocomposition equipment can produce typeset output at resolutions of 1,200 or more dpi. Technical advances are reducing this quality differential. Indeed, 400 to 1,000 dpi laser printers are now on the market. High resolution non-impact printers are adequate to meet many electronic publishing needs, either for demand printing or as camera-ready copy to be used in subsequent plate-making and photo-offset printing. Continued technical advances and market forces are likely to push the typical output resolution of laser and other non-impact printers into the 600 to 800 dpi range over the next few years, thus further closing the quality differential.

With respect to print-volume requirements, it is still far cheaper to use conventional photo-offset printers for high-volume print runs than laser and other non-impact printers. One can debate the various break-even points as a function of the length, format, number of copies, and desired turnaround time for specific documents. In offset printing, plate-making (preparation of masters or negatives of the original images by which ink is transferred onto paper to make copies of the original) costs anywhere from a few dollars to $50 and up per page. This cost indicates that non-impact printing is frequently less expensive for short print runs of under a few hundred copies per original. For larger print runs, the printing cost is likely to be cheaper with photo offset rather than laser printing.

It appears that the cost of non-impact printing (including xerographic) is rarely below 2 cents per page. Thus, assuming $2 per page for plate-making (this is for desktop relatively low volume applications), and assuming all other costs are equal (purchase or lease, maintenance, supplies, and labor), the break-even point would be about 100 copies. In this hypothetical and oversimplified case, print runs under 100 copies per original would use a non-impact printer and print runs over 100 per original would use a photo offset printer. Other elements besides cost may enter into the printing decision, such as quality, speed, turnaround time, and control. In the future, the break-even point between non-impact and photo-offset printing will depend in part on their relative technical advances and cost reductions.

ONLINE INFORMATION DISSEMINATION

Online Information Retrieval

Previous discussions have focused on a number of electronic information technologies–microcomputers, page composition and publishing software, computer graphics, scanners, printers—with the information maintained in electronic form through many or all stages of the publishing process. The primary final output has been in paper format. Advances in technology make it possible to disseminate the output information in a variety of electronic formats as well as paper. For some purposes and some kinds of information, electronic formats may be preferable to paper. This is especially so for bibliographic, reference, statistical, and bulletin board information where the user may not want to see the whole document, but is only interested in locating specific pieces of information. The private sector information industry has given high priority to computerizing access to these types of information, whether the original source of the information is the government, academic, research, or commercial sectors.

This section discusses the technology and application of online information retrieval systems in the context of the private sector, since this is where much of the online activity is occuring. From a technical viewpoint, these private sector applications are directly relevant to the Federal Government.

The technology of online information retrieval is well established. Customer access is typically via a microcomputer or terminal connected to a modem. Residential customers normally tie into the local telephone company network (e.g., Bell Operating Companies, independent telephone companies) and, if accessing a database from long distance, then connect to an interexchange carrier network (e.g., AT&T, MCI, U.S. Sprint) or a value-added network (e.g., Tymnet, Telenet). Business customers can sometimes bypass the local telephone company and connect directly to an interexchange or value-added network. At the other end, access to the desired database is frequently via an online database services company (e.g., Lockheed DIALOG, Pergamon Infoline) or a database gateway company (such as is available from Western Union Easy Net). Gateway companies serve as intermediaries between the customer and the database source and do not maintain the database itself. Online database service companies actually maintain copies of the databases online, so that referral to the database source is not necessary. Some database source companies do provide for direct customer electronic access to the database, without going through a gateway or online services company. Companies that maintain online databases need:

- a host computer and memory necessary for handling the volume of data and frequency of use, and
- the necessary front-end processor and communications equipment for handling remote inquiries and transmitting responses.

The growth of the online information industry has been phenomenal. From less than $500 million in annual revenues in 1978, the industry has grown to about $2 billion total revenues in 1986, $3 billion in 1987, and is projected to reach about over $4 billion by 1990.

A typical commercial online database service charges about $40 to $80 per hour, of which about 40 to 45 percent is for acquiring and preparing the actual data, and another 40 to 45 percent is for sales, marketing, and administration. About 6 to 9 percent is for communications (including the cost of customer premises equipment, e.g., computer terminal and modem, local exchange access, and interexchange link if applicable), and about 6 percent is for data processing (including the cost of hardware and software for database storage and data communicaticm).[19]

[19] Studies by Cuadra Associates and Elsevier Science publishing cited in P. W'. Huber, *The Geodesic Network: 1987 Report on Competition in the Telephone Industry,* prepared for U.S. Department of Justice, Washington, DC, January 1987, p. 7.13.

One implication of the above cost structure is that substantial savings can result to the extent that the data are already in the appropriate electronic format. If as a result of electronic publishing government statistical or reference reports were produced in electronic form as a matter of course (even if the ultimate product is in paper format), then the electronically formatted information could, at least theoretically, also be made available as an online database. This could significantly reduce the cost of data acquisition and preparation. This is a major cost element, regardless of whether the government and/or commercial firms disseminate the data. Sales, marketing, and administrative costs may not be as amenable to reduction for commercial firms, unless they are working under contract to the government in such a way that the market was, in effect, guaranteed. For the government, distribution to information intermediaries (e.g., libraries in the depository program) might help reduce marketing and other costs. Technology is only a small part (perhaps less than 15 percent) of the cost of online databases.

Telecommunications

Online information retrieval services and several other kinds of electronic information dissemination (e.g., electronic mail and facsimile) are dependent on telecommunication technology and systems. A number of developments are converging to facilitate and most likely reduce the relative cost of data communication. One key trend is the transition from analog to digital telecommunication networks that are designed to transfer information much more efficiently than the conventional analog telephone networks. A second trend is the rapid movement towards national and international standards for data networks of all kinds. A third trend is the maturation of Ku-band satellite, fiber optic, and FM subcarrier technologies for data transmission.

The implementation of FTS 2000, the upgraded Federal Telecommunication System, is intended to make state-of-the-art data communication capability available to all major Federal agencies. As currently planned, FTS 2000 will include:

- switched voice (up to 4.8 kilobits/second transmission capacity);
- switched digital integrated service;
- packet-switched services;
- video transmission (including graphics, facsimile, limited and full motion video); and
- dedicated voice or data transmission circuits.

The switched digital integrated service and packet-switched service should be especially useful for online database retrieval or electronic document transmission. The switched digital integrated service is designed to be the equivalent of the Integrated Services Digital Network (ISDN) for digitally integrating voice, data, images, and video over the same transmission medium. As planned, the FI'S 2000 version will be consistent with ISDN international standards and will have a 1.544 megabit/second transmission capacity subdivided into 24 individual channels of 64 kilobits/second each. For illustrative purposes, one 64 kilobit/second channel can transmit about four pages of text per second (at 250 words per page x 8 characters per word x 8 bits per character). The planned IWS 2000 packet-switched service will be consistent with international standards for open systems interconnection and interoperability with public data networks and public electronic mail services. The FTS 2000 packet-switched service is planned as a 24 hours a day, 7 days a week operation with 99.5 percent uptime and 98 percent availability.

The basic concept of packet-switchingis that data can be transmitted most efficiently when assembled into packets (or bunches) of bits of information. The U.S. packet-switching volume for 1985 has been estimated by International Resource Development at about 47 million kilopackets, of which 7 million kilopackets were for database access, 3 million for electronic mail, and 0.3 million for electronic data interchange. 11 Typical commercial rates for

[11] Cited in P.W. Huber, Telephone *Industry,* op. cit., table PA.1.

packet-switching have been estimated at about $0.50 per page of text for local packet-switching and roughly 3 times that for national packet switching. ” While these rates compare favorably with electronic mail and may be acceptable for very short documents, the cost of packet-switching long documents would be quite high. Whether FTS 2000 will significantly reduce packet-switching costs in unknown at this time.

Data transmission networks of all kinds are expected to incorporate both satellite and fiber optic technology wherever appropriate. For example, a high speed (56 kilobit/second) packet-switched data transmission network can incorporate both a fiber optic terrestrial component and a Ku-band (12-14 gigahertz) satellite component. The Ku band permits use of lower-cost, very small aperture (VSAT) earth stations with receiving disks that are 1.2 or 1.8 meters in diameter. Such a system could be used for such functions as transmitting data collected from remote locations.

Over the next few years, a balanced network of satellite and fiber optic transmission links is likely to evolve. Fiber optic links are likely to be used primarily for heavy volume, point-to-point transmissions, while satellite links are expected to dominate for point-to-multipoint transmissions. Experimental tests of fiber optics have attained transmission rates of 4 billion bits/second over relatively short distances. By comparison, this is more than a 1,000 times the 1.544 megabits/second transmission capacity specified in the ISDN standard, and is equivalent to transmitting an entire 30-volume encyclopedia in 1 second.

The integration of fiber optics with satellite, microwave, and copper wire circuits will be facilitated by the continuing development of teleports, with respect to traffic between major U.S. metropolitan areas and overseas traffic. Teleports are essentially buildings and facilities that serve as a platform or bridge for interconnecting different modes of telecommunication all at one location.

“I bid., table PA.2 and accompanying text.

The trend towards so-called intelligent buildings will facilitate integration across different telecommunication technologies and services. Intelligent buildings are prewired during construction with local area networks (LANs) capable of handling digital data communication. LANs can carry information much faster and more efficiently than the conventional telephone and PBX (private branch exchange) analog circuit and switching systems. The cost of LAN installation is much reduced if completed during building construction rather than retrofitted. The trend toward intelligent buildings is expected to accelerate in response to the rapid increase in networking of microcomputers, mainframe terminals, peripheral equipment (including scanners, printers, and graphics workstations), and the like in the office environment.

A final telecommunication technology to be discussed in this section is FM (frequency modulated) radio subcarrier transmission. The FM subcarrier is an excess portion of the bandwidth assigned to FM radio stations, and was deregulated by the Federal Communications Commission in 1983. The FM subcarrier appears to be cost-effective for point-tOmultipoint transmission of time-sensitive digital data traffic, such as news and public affairs information. For example, MultiComm Telecommunications Corp. (Arlington, VA) is using Western Union’s Westar IV satellite to transmit information to 90 participating FM radio stations, where the information is in turn retransmitted on the FM subcarrier to receiving sites equipped with a special, low-cost FM receiver. The information can be stored on a microcomputer or printed out. MultiComm sells the receiver/printer for $500 or leases the equipment for a nominal fee of $25/month. The costs of the service per receive site range from 20 cents per page of information transmitted for immediate delivery (e.g., within 19 seconds), 10 cents per page for delivery within 2 hours, and 5 cents per page for overnight delivery. This is far cheaper than courier service, especially for shorter documents. The 90 participating FM stations broadcast to an estimated 85 percent of the U.S. population. Ku-band small satellite earth stations could be used to reach rural and remote areas. MultiComm offers a Federal News Service that transmits transcripts of White House briefings, congressional testimony, and the like to hundreds of newspapers and trade associations, and an Infowire service for low-volume users who need time-sensitive information on, for example, White House and agency press releases, advance schedules of upcoming hearings, and the like. Other private firms are using the FM subcarrier to distribute such information as stock market quotes.

Electronic Mail

Another technical option for online information retrieval and two-way information transfer is electronic mail. As discussed previously, electronic mail capability is planned as part of the FTS 2000 system. Electronic mail has grown more slowly than initial expectations, but appears to breaching a critical threshold of viability.

The outlook for electronic mail is being enhanced by several key trends:

- Electronic mail is increasingly included as a basic capability of office automation systems, such as those offered by Data General, DEC, IBM, Wang, and NBI;
- Vendors are providing much improved capacity for interconnections or gateways between electronic mail systems (e.g., Wang and IBM, DEC and MCI Mail, IBM and Western Union Easy Link, MCI Mail and CompuServe Easy Plex);
- Enhanced electronic mail capabilities are being developed that can handle graphics and spreadsheets besides American Standard Code for Information Interchange (ASCII) text; and
- There is growing acceptance of the CCITT (Consultative Committee on International Telephone and Telegraph) X.400 standard for electronic mail and messagehandling service. X.400 is based on the 0S1 (Open Systems Interconnection) model and will permit interconnection among various electronic mail services.

Many electronic mail systems require a modem (modulator/demodulator) at each end of the circuit, to convert the digital signals from the sending computer into analog signals for transmission over the telephone lines (at least in the local exchange) and back again from analog to digital at the receiving computer. However, modems are likely to be less of a constraint in the future for at least two reasons. First, the cost of modems continues to drop–a 300 bits per second modem now costs $100 to $200, a 1,200 bps modem (the de facto standard for remote computer networking including electronic mail and bulletin boards) $200, and the higher speed 2,400 bps modem about $300 to $400. Second, in the future, all-digital data communication and telephone networks will eliminate the need for modems almost entirely. Modems will be necessary only to the extent analog phone systems are still used.

The cost of electronic mail varies according to the length and the volume of the messages and the type of electronic mail system used. For an inhouse personal computer or office automation-based electronic mail system, the cost range has been estimated at roughly $1 to $2 per 3-page message (7,500 characters) at a monthly volume of 1,000 messages, and is estimated to drop to about $0.10 to $0.20 per 3-page message at a monthly volume of 10,000 messages. By comparison, electronic mail service bureaus typically charge in the range $1 to 3 per 3-page message regardless of volume.

Other alternatives for transmission and receipt of electronic mail include: electronic bulle tin boards, digital facsimile services, and videotext services. For discussion of these and other related telecommunication technologies, see *Communication Systems for An Electronic Age* (OTA, forthcoming, 1989).

OPTICAL DISKS

For information that neither changes frequently nor requires immediate, online remote access, optical disk technology is a viable technical option for purposes of information storage and dissemination and as an important component of electronic publishing systems. (Other optical technologies not discussed here, such as optical or laser cards on strips, could provide storage and dissemination of smaller amounts of information.) While some standards issues still need to resolved, the significant technical advantages of optical disks are becoming more and more evident as a result of numerous development applications, prototype tests, and, commercial offerings.

Optical disk technology uses a laser beam to record data on plastic disks by engraving pits in the surface. The disks can then be subsequently read by a low-power laser beam to retrieve the data. There are several different types of optical disk, and some are further along in terms of technology and standards than others. Standards are essential for optical technology to ensure compatibility among different types of disks and disk readers, and to minimize the possible need for future rerecording of data due to incompatible equipment.

The major advantage of optical disk technology is the ability to store and disseminate large amounts of information at very low cost. For example, a 4.72 inch (12 centimeter) CD-ROM (Compact Disk-Read Only Memory) can store up to roughly 540 megabytes (millions of bytes) of data. Assuming that one typewritten text page averages 250 words or about 2,000 bytes per page, one CD-ROM can store up to 270,000 pages of typewritten text. Grollier has recorded its entire 20-volume Academic American Encyclopedia on about one-fifth of one disk. One floppy diskette (single-sided, single density) can store about 360 kilobytes of data, which is equivalent to about 180 pages of double-spaced typewritten text. Thus, 1 CD-ROM can store the equivalent of about 1,500 floppy diskettes, about 54 of the 10-megabyte hard disks, or about 10 of the 1,600 bits-per-inch magnetic computer tapes. A 12-inch (30 cm) WORM (Write Once Read Many times) optical disk can store up to 1 gigabyte (billion bytes), which is roughly double the capacity of a CD-ROM. All of these storage capacities are per single side, and would be doubled for two-sided disks.

The total and per bit or byte manufacturing costs of both 4.72-inch CD-ROMs and 12-inch WORM optical disks are already quite low. CD-ROMS can be mastered for $4,000 to $5,000 and can be reproduced in quantities ranging from $30 per disk for 100 copies to $6 per disk at volumes of several thousand. Some estimates suggest per disk costs as low as $3 for volume runs. The 12-inch WORM disks are more expensive to produce, at about $150 a copy, but are still far cheaper per byte than floppy diskettes or hard disks. These costs do not include the cost of data acquisition and preparation, which apply to any storage medium, and the cost of equipment needed to read the disks. All that is necessary to read CD-ROMs is a CD-ROM reader, available from several vendors in the $500 to $1,000 price range, and a personal computer and screen. Thus, for users already owning a microcomputer system, the incremental cost of CD-ROM equipment is in the same range as the medium to high-end consumer-oriented compact digital audio disk players. WORM readers are considerably more expensive—several to tens of thousands of dollars range-although this can be modest for the institutional (corporate and government) users who are the likely clients for 12-inch WORM disks.

Optical disks also offer other advantages:

- rapid access to stored data (i.e., in one second),
- the ability to use a microcomputer for data access and retrieval,
- high levels of data integrity,
- very minimal disk or equipment wear,
- convenience and portability, and
- relatively long media life.

The latter point is somewhat controversial as initial manufacturer estimates of 10 to 20 years have now been extended to 40 to 50 years. Some suggest that 100 years is possible under ideal conditions. Disks could be recopied at periodic intervals if necessary.

The high level of commercial and governmental activity is indicative of the potential for CD-ROM and WORM disks. Vendors (such as Lockheed DIALOG, Cambridge Scientific Abstracts, Aide Publishing, and VLS, Inc.) are offering many new optical disk-based products and services. Many of these include databases that originate in whole or in part from the Federal Government.

Federal agencies are actively pursuing a wide range of development and prototype projects. For example:

- The Nuclear Regulatory Commission (in cooperation with the Smithsonian Institution) is developing a WORM optical disk system to keep track of submissions regarding nuclear waste disposal under the Nuclear Waste Policy Act of 1982. The system uses personal computers, scanners, and 12-inch WORM optical disk drives.
- The Library of Congress is prototyping optical disk technologies for general research, archival, and information retrieval purposes, including the use of a 100-disk optical jukebox for 12-inch WORM disks. The jukebox has a potential storage capacity of 200 gigabytes.
- The National Library of Medicine is prototyping various optical disk technologies for medical applications, research, archival, and instructional purposes.
- The Bureau of the Census is prototyping the use of CD-ROMs for storing and distributing maps that will result from the 1990 census. The Census Bureau is also examining the potential of CD-ROM for a broad range of geographic and topographic maps as part of the Topographic Integrated Geographic Retrieval (TIGER) project being conducted jointly with the U.S. Geological Survey.
- USGS is prototyping the use of CD-ROMs for the possible goal of providing all (or a large part) of USGS earth science information in CD-ROM format such as seismic data from the National Earthquake Information Center. USGS officials believe that CD-ROM offers the potential to make earth science data much more accessible at lower cost.
- The U.S. Navy's Printing and Publications Service is implementing a print-on-demand system for 1.2 million pages of military specifications and standards, including text and graphics images. The Navy is using a 12-inch WORM optical disk unit to record the disks, which are then placed on two 32-disk juke boxes. More frequently requested documents are concentrated on a few disks, and output is printed with Xerox 9700 laser printers. The system is intended to:

  —reduce warehouse space and printing costs,
  —improve response time,
  —eliminate dissemination of out-of-date documents, and
  —serve as a prototype for many other applications-for example, technical manuals, training materials, and handbooks.

  The Navy intends to develop interactive applications for document updating, alternative storage media (such as CD-ROM and 5.25-inch WORM optical disks), and document search capability.
- Other Federal agencies actively pursuing optical disk technology include the: Internal Revenue Service, Patent and Trade mark Office, National Archives and Records Administration, National Aeronautics and Space Administration, Central Intelligence Agency, and numerous components of the Department of Defense including the National Security Agency.

Another popular optical disk technology is the analog videodisk. This is heavily used for educational and training purposes, and can store up to 54,000 images per disk. Videodisks are roughly similar to CD-ROMs in cost—about $2,000 to master, $18 per disk for the first 100 copies, and under $10 per disk for runs of several thousand. Videodisk readers cost in the range of a few hundred to a few thousand dollars.

Beyond the CD-ROM, WORM, and videodisk, there are several other optical disk technologies under active development and application. The most noteworthy are: the CD-I (Compact Disk-Interactive) that combines text/data, video, audio, and software storage, editing, and retrieval on one disk; and the erasable 12-inch or 5.25-inch optical disk. CD-I is of particular interest because it will make possible such compact disk applications as: talking CD books, "smart" CD books (using expert systems), CD book (or library) of the month, and interactive audio, video, and database software. CD- I will be a disk with powerful capabilities including:

- graphics (e.g., digital video still frames, limited motion video, encoded colors, full screen animation);
- audio (e.g., digital audio, hi-fi, mid-fi, speech quality); and
- text (e.g., bit-map text storage for display only, character-encoded text for editing/processing).

Optical disk technologies and applications continue to advance at a rapid pace. Double sided 12-inch WORM disks are now available with up to 4 gigabytes storage capacity per disk. The initial commercial 5.25-inch erasable disks already have been introduced, and CD-I disks now are in the prototype stage, with commercial introduction expected in 1989 or 1990. Some vendors have expanded the capacity of CD-ROM disks up to 750 megabytes, and others offer CD-ROM juke boxes that can access up to 240 disks. Personal computer compact disk readers are entering the market, as are specialized PC-CD/ROM applications (e.g., using hypertext or hypergraphics software). A fledgling CD-ROM service bureau industry is developing, not far behind and perhaps eventually to be integrated with the electronic publishing service bureau industry.

A major critical path item for optical disk technology is the development of standards. The current status and outlook are briefly discussed in a later section, along with consideration of other standards issues.

EXPERT SYSTEMS

Expert systems, sometimes known as knowledge-based or rulebased systems, are typically computer soft ware packages that permit users to have the benefit of expert knowledge in specified subject areas. The 'expert inexpert system means that both the knowledge and rules (decision paths and criteria) built into the software come from relevant subject matter experts. Expert systems have advanced to the point where widespread application to many aspects of information dissemination is likely-ranging from technical writing to information access and retrieval to the management of electronic publishing.

The expert systems applicable to information dissemination are no different in principle from the systems that have been successfully applied to various scientific, industrial, and educational areas. For example, expert systems have been used to:

- help make agricultural management decisions regarding pest control as a function of the type of crop, landscape, weather conditions, season, other vegetation, infestation history, and the like;
- help students explore and master a subject or skill and even monitor the learning progress of the students (known as "electronic or intelligent tutors"); and
- help technicians interpret technical data from computer-assisted manufacturing systems.

Expert systems can be tied into both online bibliographic and full text information retrieval and to electronic publishing. For example, prototype expert systems with sophisticated search strategies are being used to retrieve and deliver full text information via electronic publishing systems. These kinds of information retrieval expert systems could eventually work hand-in-hand with expert systems designed to efficiently manage electronic publishing. One can easily envision the day when expert systems will help optimize the electronic publishing and dissemination (paper and electronic) of information products (or packages of products), given the specific profile of the product (number of pages, composition, type style, use of graphics, etc.), anticipated user needs (e.g., size of demand by format), and the mix of dissemination channels (initial press run of paper copies, provisions for demand printing, online database access, optical disk distribution, etc.).

Numerous expert system applications for information search and retrieval are under development. For example:

- The National Records and Archives Administration (NARA) developed a prototype expert system to assist with routine inquiries from researchers. The objectives of the project were to evaluate the capability of an expert system to capture the expertise of experienced archivists and to relieve them of the significant expenditure of time needed to answer routine inquiries. Test results indicated that if the prototype system were expanded to full scale, the system could be expected to agree with its human counterpart more than 90 percent of the time. NARA plans to extend testing of expert systems to other areas of records management.
- The Defense Technical Information Center has established an Artificial Intelligence/Decision Support Laboratory that is working to apply the full range of expert systems and even more powerful artificial intelligence technologies to information access and retrieval. The ultimate objective is to facilitate the capture and transfer of knowledge from the experts to the users of DTIC (and other DoD) information systems, utilizing innovative information display techniques and full integration with the DoD Gateway Information System that is interfaced with hundreds of online databases.
- The National Agricultural Library (NAL) has developed prototype expert systems that query users on their information needs and route them to the appropriate bibliographic sources. The prototype was provided to over 700 librarians in a floppy disk format that runs on a microcomputer. NAL hopes to create a critical mass of expert system users, and believes that expert systems could help free librarians from the more routine ready reference and directional questions. NAL is also exploring linking expert systems to other government and commercial online databases and CD-ROM players. Expert systems could be used to query the user on his or her information needs, help sharpen the request, and then route the request to an online bibliography, a disk-based bibliography, or a full text document on videodisk or CD-ROM with electronic printing on demand. The possibilities are almost endless.

TECHNICAL STANDARDS

The pace of development and application of several of the technologies discussed earlier is dependent on the development of and agreement on national and international standards. Standards-setting efforts are underway in all critical areas, although the intensity of activity varies. The major standards organizations include:

- the International Committee on Consultative Telephone and Telegraph (CCITT), which is a unit of the International Telecommunications Union and whose formal members are 160 governments;
- the International Standards Organization (1S0), whose members are the national standards bodies of 89 countries;

- the American National Standards Institute (ANSI) that represents the United States in the IS0 and coordinates voluntary standards activities in the United States;
- the National Bureau of Standards (NBS), which is the lead U.S. government agency in many standards areas; and
- the Federal Communications Commission with respect to certain telecommunication standards.

A new organization, the Corporation for Open Systems (COS), was established in 1986 to promote open systems interconnection standards. COS members are primarily telecommunication and information equipment and services companies.

The following discussion highlights standards-setting for optical disks, page description languages, and test markup languages. For discussion of other standards areas (e.g. electronic data interchange, integrated digital services), see *Communication Systems for an Information Age* (OTA, forthcoming, 1989).

With respect to optical disk standards, the two leading manufacturers of optical disks—Sony and Phillips-took the lead and developed a set of proposed standards for CD digital audio, CD-ROM, and CD-I, known as the Red Book, Yellow Book, and Green Book standards, respectively. The proposed CD-ROM standards (Yellow Book) included detailed technical specifications for CD encoding, mastering, replication, decoding, and reading, such that any CD-ROM disk can be read by any CD-ROM disk drive, and have become de facto industry standards. In addition, standards for the logical formatting of CD-ROMs were initially proposed by the so-called High Sierra Group and subsequently adopted by 1S0 as an international standard. Data preparation is the one area not fully specified by the proposed standards. While data must be logically organized, formatted, and prepared prior to conversion into opticaI disk format, the standard permits use of a wide range of computer operating systems. Although the other specifications insure that disks are physically readable by any disk drive, the data may not be accessible except through proprietary software. However, this approach is consistent with the usual industry practice for disk drive standards. Standards for WORM, Erase-able, and CD-I disks are in earlier stages of development.

Another very important standards area involves the page description and the text markup languages used to code the format, style, and composition of documents. If the text markup language used to prepare a document is not compatible with the language used by the composition and/or output devices, then significant additional work is required to strip the markup commands from the document and reinsert the commands in a compatible language. Sometimes it is easier just to rekeyboard and recode the entire document, at significant additional cost. Alternatively, a page description language can be used to make the conversion automatically, if there is page description software compatible with the particular text markup language and output devices in use.

One possible page description language is the PostScript language (by Adobe Systems) that is becoming a defacto standard at least for desktop and WYSIWYG publishing systems due to the fact that both Apple and IBM, among others, use PostScript. This possibility is under consideration by NBS, ANSI, and 1S0. A related effort involves the development of a Standard Page Description Language (SPDL). These approaches are intended to match the applications software (e. g., for editing and composition) to the output devices and eliminate the need for the so-called device driver which is a separate set of instructions needed to make the applications software communicate with the output device.

Text markup standards are particularly important to realize the full benefits of electronic information dissemination. If government documents (whether reports, pamphlets, manuals, other text, or text plus tabular and graphics material) are not prepared in a standardized electronic format using standardized codes and descriptors, substantial recoding and rekeyboarding may be necessary at later stages of the dissemination process. Any significant recoding and rekey boarding is costly and can offset some or all of the cost advantages of electronic formats.

Text markup standards are intended to establish a consistent set of codes for labeling key elements of a document–such as chapter titles, paragraph indentations, tabular presentations, and the like. Such standards establish a logical structure for the elements of a document, in a hierarchical order—such as chapter, paragraph, line, word, and character. The elements are assigned codes (which can be a letter, number, symbol, or combination thereof) that are keystroked along with the text, tables, and graphics included in the document. If these electronic codes are widely agreed upon and used (i.e., standardized), then the documents can be electronically transferred from one stage in the dissemination process to another with little or no additional effort.

A wide range of information dissemination functions would be facilitated by text markup standards, including:

- authoring
  - —creating the document
  - —editing
  - —revising
- archiving
  - — short-term
  - –long-term
- disseminating in multiple formats
  - —conventional printing
  - —electronic printing-on-demand
  - —online electronic
  - —offline electronic (e.g., magnetic tape, floppy disk, CD-ROM)
  - —microform
  - —specialized outputs (e.g., braille, foreign languages, voice)
- disseminating through multiple channels
  - —agency clearinghouses and information centers
  - —governmentwide clearinghouses and sales programs
  - —press, libraries, and commercial vendors

For example, text markup standards would help ensure that NTIS and/or GPO are able to efficiently reproduce and disseminate agency electronic documents. This would also facilitate private vendor repackaging or enhancing of agency documents, if the vendors utilized the same standards.

Three major approaches to text markup standards are: 1) GPO's logically structured full text database standard; 2) the Standard Generalized Markup Language (SGML) that has been adopted by the Department of Defense; 3) and the Office Document Architecture standard.

The GPO standard is used almost exclusively by GPO, congressional committees and offices, and Federal agencies—primarily those agencies that submit magnetic tapes to GPO for typesetting and printing. Full text database standard or specification is the application of a logical coding structure to the full text or content of the document, including tables as well as text. GPO staff recently completed training on how to write software programs that can translate from SGML to GPO's full text database standard. GPO indicates that it is prepared to write such software at customer request. The GPO standard is designed primarily to meet the needs of publishing professions. SGML, on the other hand, while also meeting publishing needs, is recognized as an international standard, endorsed by DoD and some vendors, and is being issued by NBS as a Federal Information Processing Standard (FIPS). SGML is a set of rules for developing the element codes for a document, whereas the GPO standard includes both the rules and the specific codes themselves. Both SGML and GPO use a logical structure, so in principle SGML codes should be convertible to GPO's codes, and perhaps vice versa, although some of these applications are still under development.

Office Document Architecture (ODA), a related protocol, is directed primarily to meeting office, not publishing, needs, and thus the document complexity is reduced (due to fewer fonts, formats, etc.). ODA is a method of encoding software that essentially converts documents to a common code compatible with a wide range of office automation systems. ODA is a protocol for converting the codes used to format individual documents into a common format for the interchange of the documents among different systems. ODA was initially defined by the European Computer Manufacturers Association (ECMA) to be consistent with the Open Systems Interconnection standard developed by the 1S0 and has been issued as an international standard. Officials at NBS believe that there may be a need for both ODA and SGML standards within the Federal Government.

Finally, there is intensive work by all major standards organizations to refine and implement the open systems interconnection (0S1) concept. An 0S1 reference model has been developed under 1S0 auspices. The model services as a master standard for an integrated telecommunications-information systems environment. It also incorporates already established standards such as those for packet-switched data networks and electronic mail. Many vendors and users have recognized the need for rapid 0S1 implementation. In the United States, NBS is coordinating an 0S1 prototype system known as OSINET that is intended to be a test of the 0S1 reference model. The results are being made available to the standards-setting organizations.

The Federal Government commitment to 0S1 is already significant, with a growing consensus that 0S1 is necessary to move to interoperability of the now confusing and largely incompatible range of equipment and software in the government inventory. Indeed a Federal interagency committee has recommended that 0S1 standards be mandatory for new Federal computer and telecommunication procurements and be a first option for retrofits of existing systems. The suggested 0S1 procurement standard would be consistent with the 1S0 reference model. This 0S1 procurement standard is being issued by NBS as a Federal Information Processing Standard (FIPS).