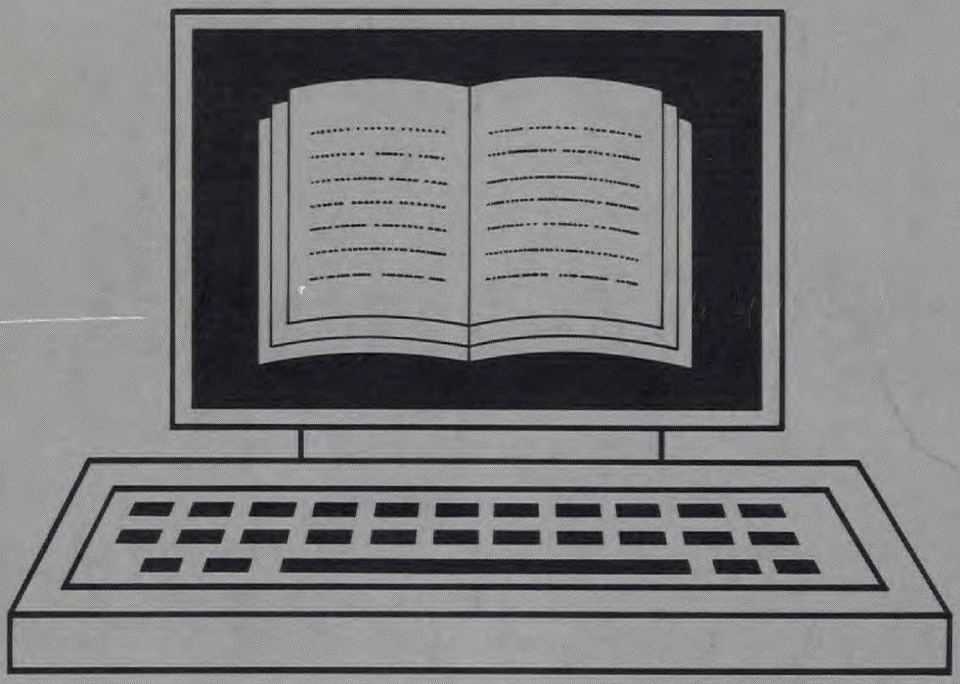
AN ASSESSMENT OF ELECTRONIC PUBLISHING PRODUCTS & INDUSTRY

1·9·8·8



Prepared For
The Department of Communications
Government of Canada

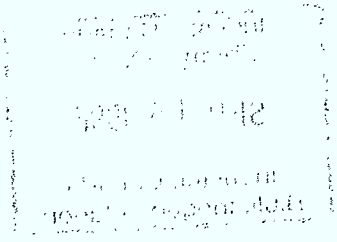
Derek Murray Consulting Associates Inc.
In Association With:
Evans Research Corporation, and
Stevenson Kellogg Ernst and Whinney

STANDALONE DATABASES: OPTICAL PUBLISHING

November, 1987.

Submitted By:

Derek Murray Consulting Associates Inc.



STANDALONE DATABASES: OPTICAL PUBLISHING

November, 1987.

Submitted By:

Derek Murray Consulting Associates Inc.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

1.0 INTRODUCTION

Optical media are one of the newest and most exciting avenues through which publishers of standalone databases may distribute their products in the marketplace. Because this media is so new, this chapter has a fairly heavy emphasis on providing an overview of the technology, as well as comparing it to existing methods of vending standalone data. As well, existing and potential application areas for the media are discussed.

If an emphasis on the technology exists, then there is a de-emphasis on forecasting market size and revenue, simply because there is not yet enough history behind the media to contemplate future trends.

When interview program participants were polled as to the major factors driving the growth in the market for online databases, some fairly fundamental issues emerged, including:

- the formation of an information-driven society; and
- the gradual assimilation by society of computers and computerization as a simple tool, no different than an automobile.

Fundamental factors such as this will also further drive the demand for standalone data in a variety of application areas. As greater and greater amounts of data are produced and distributed, two fundamental issues will emerge - storage and accessibility. Optical media addresses both of these issues in an extremely efficient manner, and thus, our discussion of standalone database publishing will focus mainly on developments occurring in this area.

An overview of the existing technology follows beginning with an examination of current trends and systems.

2.0 TECHNOLOGY OVERVIEW

2.1 Current Trends and Systems

2.1.1 -- Videodiscs

Videodiscs were, for the most part, developed by consumer electronics companies to compete with video cassette recorders (VCRs) in the consumer market for video recording and playback systems. The competition was analogous to that existing between gramophone records and audio cassette systems. Many advantages were claimed for the videodisc systems including:

- low disc replication costs and, thus, low disc prices;
- cheaper players than VCRs;
- superior image quality to that offered by the videocassettes;
- far less wear during playback; and
- a far more flexible system, offering facilities such as freeze-frame, random access, and slow speed forward and reverse.

The advantages of the videodisc system, as well as its major disadvantage, the lack of a recording facility, is reflective of many of the attributes of CD-ROM systems and other optical storage media. Time will tell how far the analogy will carry, however, but for the videodisc, the lack of a recording facility prevented the technology taking off as a consumer item. The many extra features available with videodisc systems made them very attractive to the institutional market where they are very widely used for interactive training and point of sale systems.

One of the more successful products in this market was the brand name of Laservision, developed by Phillips. This system established a defacto standard for videodisc mastering which is now widely observed. The program information, which can be stored on film or tape, is read by a helical scan videotape unit and used to modulate the intensity of a thin film of positive photoresist. Once recording is complete the exposed area of photoresists are etched away to leave a master disc containing video information encoded in its periodicity and length as pits. The master's recording surface is then metallized using an evaporated metal coating and becomes the source for generating identical submasters.

Two main processes exist for the mass production of replicated discs. Phillips and 3M use a cold photopolymerization process referred to as 2P while Pioneer, Sony and others make use of a compression injection molding technique. Both replication materials involve the pressing or stamping of discs and thus are closely allied to gramophone record production techniques and the same economies apply -- the more copies produced, the lower the marginal cost.

While replication techniques vary slightly, all laser optical videodisc systems currently on the market make use of the same basic playback technique which involves using a low power laser as a stylus and a photo-detector to collect light reflected by a disc. Such systems are called reflective systems. The players may range in price from $800 for serial playback consumer products up to over $4,000 for units with their own internal microprocessor which can be programmed to retrieve and display frames in a predetermined order and which can be linked to and controlled by external computer systems.

The non-contact optical videodisc offers a number of significant advantages in terms of playback facilities and durability over serial player videocassette systems which make it especially suitable for training systems or interactive point of sale display systems. These are:

(1) The laser pickup device can be moved at random across the disc and will operate at a range of different speeds or in still-frame mode.

(2) There is no stylus contact with the disc so there is no likelihood of wear either in normal usage or when freeze frame facility is required.

(3) Since a laser is utilized to read the data, it can be focused in an extremely accurate manner, ignoring the protective coating on the outside of the disc and focusing straight on the recorded information. This means that dust or scratches on the protective coating do not result in a loss of picture quality.

The Laservision videodiscs have a diameter of twelve inches. The information is encoded on the disc as a series of pits which vary in length and spacing, arranged on a spiral track. A single track of pits and flats contains all the information necessary for a colour video program with stereo sound or two separate sound channels, plus control data for playback operation. There are 54,000 tracks per disc side and the disc rotates at 1800 rpm. With one frame recorded on each track, a playing time of 30 minutes per disc side is provided. The information on the disc is read by pointing a laser beam onto the underside of the disc and focusing it directly into the bottom of the pits. The light is then defracted and reflected back through source-tracking mirrors to a photodisc.

During playback of a disc, tracking of the read head proceeds from the innermost to the outermost track. The discs can be configured to offer a choice of two playback modes designed to cater to the needs of two different markets.

With Constant Angular Velocity (CAV) or active play mode, the frame is recorded on each track and the disc rotates at constant velocity so the linear tracking velocity increases from the innermost track where it is least to the outermost track where it is greatest. Because each track holds one frame, a single frame can be frozen on the screen by directing the read head to scan one track continuously. Further, the speed and direction of the pick-up device can be exactly controlled and each frame on the disc, up to 54,000, can be numbered and the read head directed to go to and read any one in a fraction of a second.

CAV is therefore the ideal mode for use in interactive videodisc systems where individual frames need to be called up and displayed to answer specific enquiries or as part of a computer-based training system.

In the consumer market, however, where the main requirement is simply for serial replay of feature films, CAV mode, with one frame per track, represented a poor use of space and restricted the playing time available on each side of the disc to 30 minutes of video. For the consumer market, then, Phillips abandoned the principle of one frame per revolution of the disc and varied the velocity of the disc to provide Constant Linear Velocity (CLV) mode or long play discs. Using CLV mode, the playing time of each side of the disc can be expanded as on the outer tracks of the disc with a large circumference and two or three frames can be recorded. Using this recording technique, and varying the speed of rotation of the disc at the playback stage, the playing time can be expanded from 30 minutes to 60 minutes per side. The trade off, however, is that individual frames cannot easily be accessed on a track by track basis and the freeze-frame facility is forgone.

From the standpoint of database publishing, CAV discs with the fast random access facilities show the most promise. Various companies have devised techniques for encoding digital data within the video signal of the videodisc so the vast storage capacity and random access facilities provided by the videodisc can be used to distribute digital databases in a number of electronic publishing applications. This use of videodisc technology is called hybrid videodiscs and will be discussed later in the chapter.

2.1.2 -- Compact Discs

Compact discs (CDs) appeared later on the market than optical videodiscs. While the videodisc was first shown in 1973 and made available commercially in 1978, CDs were shown in 1980, and released commercially in 1983. In spite of this lag, however, sales of CD players to the public have already surpassed sales of videodisc players. The CD player has been described by a number of sources as one of the most successful launches of a consumer electronic product ever, with several million players sold within two years of its introduction.

Part of the reason for this is that compact disc systems are all standardized on the format and digital recording techniques designed by Phillips and error detection and correction systems jointly developed by Phillips and Sony. The main aim of the standard is to achieve total compatibility so all CDs produced to the standard can be played on all CD players designed to meet the standard. The standard is popularly known as the 'Red Book'.

The physical processes involved in producing a compact disc are virtually identical to those described above for a videodisc except that the material recorded on the disc is in a different format and the master and replicate discs are single sided with a 12 cm or 4.72 inch diameter rather than a 12 inch diameter.

The key difference between CDs and videodiscs, other than their size, is the format in which the information is stored. Audio information is stored in the CDs in digital form while videodiscs are essentially an analogue storage medium. With CDs the audio information is digitally encoded prior to storage on the discs and is read off the discs as a bit stream at rates in excess of four million bits per second. At the output stage, the digital information on the CDs is passed through a digital to analogue converter for output as a series of audio signals which can then be replaced on standard consumer high fidelity systems making use of the same amplifier and loudspeaker systems as would be used to replay the signals picked up from a record or cassette deck.

For compact disc replication, a polycarbonate called macrolon is used. This raw material is liquidized and injected under high pressure and heat into the injection moulding machines. After moulding, the single disc surface which carries the pits is coated with a thin layer of aluminum to produce a reflecting surface for the laser beam. The surface itself is coated with a protective lacquer which is dripped on the rotating disc. Before labeling and packaging, the disc is centred and the centre hole punched in.

As is the case with analogue records, CDs are produced with a run-in section, a music section, and a run-out section, all of which contain information in digital coded form. The run-in section of CDs contains a table of contents consisting of the number of recorded tracks and the time from the start of the disc to each track. Data is recorded onto each CD master in a spiral of tracks starting at the inside of the disc and working outwards. A provision exists for dividing each CD into a maximum of 90 tracks. The standard 12 cm disc

contains some 20,000 spirals in the program area and in these are recorded a series of laser cut pits. Using this recording technique one CD can store over 60 minutes of high quality stereo, which, when error correction and detection codes are accounted for, represents a storage capacity of seven billion bits.

Playback of discs in a CD Player involves the use of a laser optical system which must first be focused over the relevant track on the disc and then accurately focused on the disc's recorded surface. CD players use the same reflective systems as the videodisc system to read data from the discs. A low power laser beam is shone onto the disc's recorded surface. The intensity of the reflected laser beam is less when the laser reads a pit then when it reads a band (the area of the disc which has not been etched away). The different intensities of reflected light are detected by photocells and constitute the raw digital signal.

The discs are played back at a Constant Linear Velocity (CLV) of 1.25 m/sec. from the inside to the outside. The rotational speed then varies from 430 to 200 rpm. CLV mode is employed to maximize the storage capacity and in CD-audio terms, the playing time. However, this translates into an access time slower than if CAV mode was utilized.

In practice, a piece of music is not accessed via its track number (1-99) but by its position on the disc spiral measured in terms of the time it takes to reach it in normal playback from the beginning of the disc. This is inconsequential in terms of the storage of audio data but it does limit CDs as a computer storage peripheral. Last, since CDs are a digital medium, they are supplied with a sophisticated error detection and correction system devised by Phillips and Sony which serves as the basis for the error detection and correction system used on CD-ROMs.

2.1.3 -- Compact Disc Read Only Memories (CD-ROMs)

In the case of videodiscs, an analogue consumer product had been adopted and used by a number of companies to develop information delivery systems based on hybrid videodiscs. Correspondingly, when a digital consumer product (the compact disc) was introduced, computer hardware manufacturers, software suppliers, publishers and database providers also took a considerable interest in the potential of the new medium as a computer storage peripheral and an electronic publishing medium. Within two years of the launch of the first commercial CD disc player, Phillips showed a CD-ROM (Read Only Memory) player which was a direct adaptation of the CD system for publishing and data processing applications.

Both CD-ROM and CD-audio discs are produced using the same physical mastering and replication processes. Plants designed to master and replicate CDs can be easily adopted to master and replicate CD-ROM. The discs have the same physical dimensions - diameter, thickness, and chemical composition. CD-ROM drives rotate at the same speed as the CD players and both CD and CD-ROM discs are recorded and read in Constant Linear Velocity (CLV) mode. Both applications of the media use the same modulation systems and error detection and correction systems which are defined in the industry standard Red and Yellow Books.

However, CD-ROMs are designed to store digital data whereas CDs were designed to store digitally encoded audio information. The result is that the CD-ROM "Yellow Book" standard has made provision for more accurate addressing of data and for an extra level of error detection.

Both CDs and CD-ROMs use the same basic control and display (subcoding) systems for locating/addressing information. As we have seen above, the subcoding system for CDs and CD-ROMs divides the disc into three parts: the lead-in area, the program area and the lead-out area. With CDs and CD-ROMs the program area can be divided into a maximum of 99 tracks or programs. For CDs storing only digitally encoded audio information, this is sufficient to enable any section of audio to be addressed quickly.

With CD-ROM, as defined in the Yellow Book standard, such a division is not sufficient and there is a need to add in a facility to access specific data more precisely and to provide for extra error correction. Hence, provision is made for every one of the 99 tracks to contain either digitally encoded audio information or pure computer data. The two types of information are handled in different ways and are referred to as two different modes. In practice most CD-ROMs currently being produced contain just one data program but the standard will support mixed mode discs holding both pure computer data and digitally encoded audio and in future a range of other types of information, each defined as another mode.

The error detection and correction (EDAC) specified in the Red Book standard for audio information storage is adequate for storing audio information. Here, one lost bit will not affect the sound reproduction and data. Audio data "tends to degrade gracefully". However, it would not be adequate for storing pure digital data where "one incorrect bit error could mean the difference between a plus or a minus sign in your bank balance". The result is that CD-ROM drives require a further level of error detection and correction than compact disc players.

To achieve this higher precision, the CD-ROM Yellow Book standard divides the data programs on a CD-ROM into logical units -- two thousand byte blocks - one of which is stored on each physical sector on the disc. Each sector on the disc contains 98 CD frames, each with 24 bytes, to give a total sector capacity of 2,352 bytes of data of which 2,048 bytes are always designated for user data and the remaining 304 bytes are used for data which the drives need to locate data on the disc, additional error correction systems, or additional user data.

A mode byte describes the nature of the user data. At present "Mode 1" is defined for computer data storage where the additional error correction is required. However, the CD-ROM standard also provides for other information to be stored on the disc - that which "degrades gracefully", and where existing EDAC systems are adequate. "Mode 2", then, is defined where an auxiliary data field is used to store more information rather than the data necessary for extra levels of error correction.

In the future, the same basic CD-ROM drive could be used for both audio playback and data retrieval and already a number of combined audio/data CD-ROM players have been shown, targeting the domestic market where they could be linked to the personal computer and rack hi-fi systems. The combined player would have two outputs. The audio CD output would flow through a digital to analogue converter and from there to an amplifier and loudspeaker system. The digital CD-ROM output would flow to a computer or output device such as a monitor or printer.

CD-ROMs, like CDs, are single-sided 12 cm diameter discs with a total user capacity, after formating of 550 million bytes.

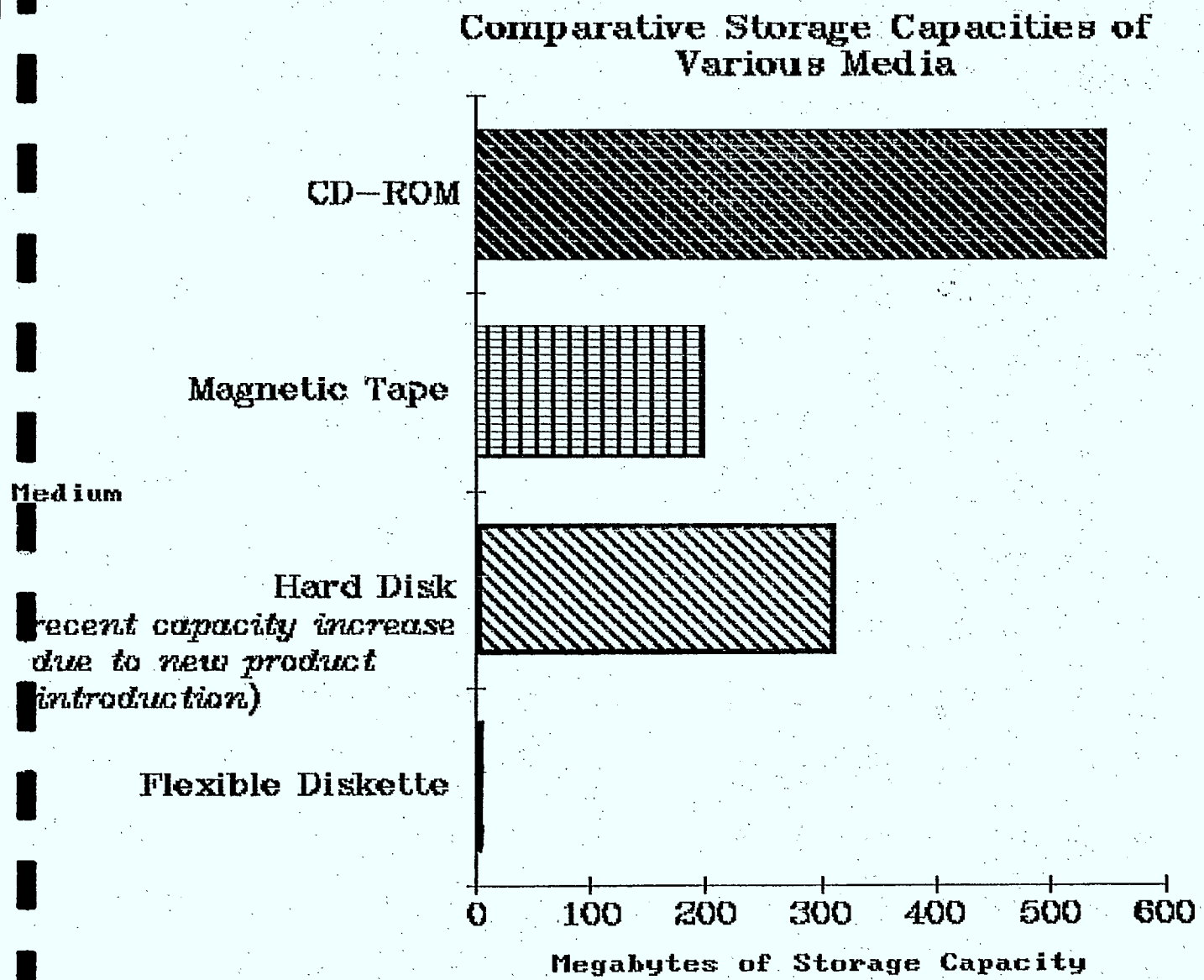
This is an overwhelming storage capacity compared with flexible diskettes, hard discs and magnetic tape, exhibiting storage capacities of 0.5 to 1.5 megabytes, 70 megabytes and 200 megabytes respectively.

CD-ROM also offers random access. Since the entire 550 megabytes are in one continuous stream it is possible to select a particular byte number to start and a following number of bytes. CD-ROM data, being organized sequentially along a three-mile long spiral, and read with Constant Linear Velocity, can be addressed in absolute time. The total time is that required to read the data sequentially. This time is represented in terms of 0 to 60 minutes, 0 to 59 seconds, and 0 to 74 fractions of a second. A microcomputer can then interpret these instructions via a simple time/position algorithm to effect random access.

Although the issue of standards is dealt with completely in Chapter 3, a portion of the question must be addressed in the sphere of a technological overview.

In spite of the fact that any CD-ROM produced to the Yellow Book standard will fit in and can be read by only CD-ROM drives, compatibility problems still arise due to the fact that the drives are computer peripherals and users need a computer to access the data.

The first problem is the lack of standardization with respect to the volume and file structures, or the manner in which the data is logically organized on a CD-ROM disc. This issue is not addressed because it is directly linked to the computer and operating system environment where the CD-ROM is likely to be used. For CD-ROM to be readable on a range of computer


Comparative Storage Capacities of Various Media
CD–ROM
Magnetic Tape
Medium
Hard Disk
recent capacity increase due to new product introduction)
Flexible Diskette
0
100
200
300
400
500
600
Megabytes of Storage Capacity

systems and by a range of operating systems, a volume and file structure must be standardized. The High Sierra Group in the U.S. are attempting to address this issue by producing a "Working Paper for Information Processing - Volume and File Structure of Compact Read Only Optical Discs for Information Interchange". The implications for the publishing industry of this standard, if adopted, are discussed in the next chapter.

The second problem is related to CD-ROM drives. Drives have different controllers and the commands which they accept and the format in which they must be given are not standardized across manufacturers. Defacto standards must also be worked out between drive manufacturers and computer manufacturers/ operating system suppliers for the physical/logical connection between the host microcomputer and the controller.

The third problem is related to retrieval software. If a supplier wishes to provide free-text retrieval facilities - an essential requirement when trying to access 550 million bytes of data with a personal computer - he must choose a software supplier and organize his data according to the requirements of that supplier.

Additionally he must supply versions of that retrieval software package with the database, either on the CD-ROM itself or on a separate floppy disc and that fact will immediately limit the range of computer/operating systems that the CD-ROM database can be used on. Limiting factors will be how many versions of the software package there are, which operating systems they run on, and the minimum RAM and magnetic disc storage required on the PC to support the software.

At the time of writing, the cost of mastering one CD-ROM (assuming that the data is already available on tape in the required format) is approximately $4,500 and the cost of replicating the discs is $25 to $40 per disc, depending on whether a run of hundreds or thousands is required. As capacity increases in the future, predictions of $2,000 mastering costs and high volume replication costs of $2 per disc exist. Current turnaround times from delivery of final formated tape to dispatch of replicated discs is 2 to 3 weeks but again as capacity grows, turnaround times of one week or less are anticipated.

2.1.4 -- Compact Disc-Interactive (CD-I)

On February 17, 1986, Phillips and Sony announced plans for the CD-Interactive (CD-I) media specification. CD-I is an application-specific branch of CD-ROM and builds on the compact disc data format standards surrounding CD-ROM.

CD-I, as defined by Phillips and Sony, is the consumer entertainment and education branch and attempts to meet all the requirements of that marketplace.

The CD-I recording format is identical to CD-ROM. CD-I simply gives the user some standard ways in which to represent common forms of information that they can use when the forms are appropriate for their application.

Because it is aimed at the consumer market, CD-I attempts to define a complete system as the requirement there is seen to be for an appliance which provides a particular function or set of functions which is as simple to use as a record or CD

player. CD-I has to be a simple system and hence, while CD-I is logically an extension of the CD-ROM standard, conceptually, CD-I also represents an extension of the CD digital audio concept of consumer product offering total compatibility of discs and players by defining both the way in which information is recorded on the disc and the equipment needed to read it.

With CD-I, Phillips and Sony have defined not only the medium, but also the equipment that will be needed to read it, the audio processes, the video processes, the choice of microprocessor and the operating system so they can achieve their goal of being able to play all CD-I discs on all CD-I players.

In addition, because CD-I players will be used in the home, it is a requirement that CD audio discs can be played on CD-I players too. It is also expected that data recorded on CD-ROM, according to High Sierra Group provisions, would be potentially readable on a CD-I player.

For a database or CD-ROM to be usable on CD-I, then, the software vendor would need to have developed his software specifically to run on the CD-I player and use the CD-I player's operating system. Although this would potentially be done in the future, very little standard retrieval software would be able to run on the CD-I player.

While CD-I is still a computer peripheral and a computer interface to link a CD-I player to other computer systems will be available, a computer itself will be incorporated into the CD-I player so that there will be no need to link

the player to a standard television or monitor and control it via a keypad or mouse device. The CD-I player will have no magnetic disc drives and thus, all information including the application program, the additional information needed for the operating system, and all the data must be on the CD-I disc itself as that will be the only medium used to run the complete application.

The operating system and microprocessor specified for CD-I needed to take into account the requirements of real-time applications such as entertainment. Such real-time applications require executable object code and system calls, a specific microprocessor and operating system are both defined for CD-I. Specifying the microprocessor family and operating system means it is possible to produce discs carrying audio, video text, binary data and application programs which will work on all CD-I drives from all manufacturers. CD-I users can either code text according to the conventions of the operating system, according to the convention defined by the application software or, if text manipulation is not important, by using bit-mapped images of text on discs.

To ensure universal disc/drive compatibility, dedicated hardware and interfaces are required, including specialized chips for video display processing, audio processing and real-time data via the CD-ROM interface.

If CD audio tracks are also present on a CD-I disc then the first CD-I track will contain only essential information such as the Super TOC (Table of Contents) and possibly the bootstrap and file directories. This information which is typically no more than a few seconds long, is muted if a CD-I

disc is played on a CD audio player. If CD audio tracks are present on a CD-I disc, they will be located after this first CD-I track. Any other CD-I information is contained in a track after the CD audio tracks.

Since audio blocks on a CD-I can be one of three types -- video data, audio data or computer data, synchronization is necessary to ensure that the information which is being read from the disc is directed correctly into the audio, video or computer processors; all the time ensuring that the right pictures and text are viewed at the right moment and in step with each other and accompanying audio.

Before it can be stored on any CD, information must first be prepared or authored in a particular way. For the production of CDs, recording studios are now well equipped to produce the master tapes necessary to make the discs.

For CD-ROM, text and data must be written in computer readable form. Specialist software houses are already active in transferring existing printed or database information into the correct form on magnetic tape, together with the indexes or inverted directories needed for effective searching.

For CD-I, while use can be made of the above services for text and PCM audio, they need to be supplemented further by techniques for ADPCM audio and video processing as well as the integration of the video, audio and computer data. The The CD-I specification will lay down the guidelines that need to be followed by specialist software houses in doing the work.

As a system, CD-I, like CD audio, will be software driven and Phillips expect CD-I software to emerge from a variety of market sectors, including the entertainment industry, the computer industry, the publishing industry and the computer games industry. Given the unprecedented nature of the CD-I medium, moreover, much of the software will have to be created from scratch. In theory a publisher could simply transfer a book onto a CD-I disc but the real opportunities presented by CD-I involve interactivity.

Prototype CD-I players will be introduced in the U.S. by late 1987/early 1988 and the first commercial products could be available by mid to late 1988. According to Phillips/Sony some 70 companies have already taken a licence for CD-I but it remains to be seen how many decide to pursue it through to the stage of producing commercial products.

2.1.5 -- Recordable Digital Optical Discs

Recordable digital optical discs allow the user to record data on the discs himself in-house. These discs are often called WORM (Write Once Read Many Times) discs. The name, however, is somewhat misleading. Data may be recorded more than once. In fact, storage could be carried out over a number of years until the disc is full. It is not, however, possible to overwrite data onto a sector that has already been used, or to erase and re-record data on the same sector. This characteristic makes WORMs an ideal archival medium, likely to replace magnetic tape in a number of applications as a data archiving medium and, in some applications, micro-

film, as a way of storing facsimile scanned images of documents.

A number of key differences exist between the production of WORMs and read-only media which restrict the potential of recordable disc systems for publishing applications:

1. With read-only media all the information is recorded onto the master disc in the mastering stage together with formating and indexing data and in some cases, retrieval software. With recordable and erasable/recordable media only the formating and indexing data is recorded at this stage, leaving the rest of the disc blank for users to record their own data onto it at a later stage.

2. After mastering, the blank formated discs are replicated in much the same way but instead of simply being coated with a reflective layer and a protective coating - as is the case with videodiscs, CD-ROMs and CD-Is - the recordable discs are firstly coated with a sensitive recording layer and then a number of different protective treatments are applied and the replicates are usually loaded into plastic cartridges to provide further protection during handling.

3. At the playback stage, users of recordable optical disc systems require a combined recorder/player rather than simply a player. In the recorder/player or drive the optical head comprises either two separate lasers - a high power one for recording and a lower power one for reading - or one laser which can operate at different levels of intensity; and

4. While there is one physical standard for CDs and CD-ROMs and one defacto standard for optical laser videodiscs, there are currently no standards relating to recordable digital optical discs. They are available in a range of diameters, including 5.25 inch, 8 inch, 12 inch and 14 inch and they make use of a wide range of subtrates or base materials including glass, aluminum and plastic. A range of different recording mechanisms exist as well, including pit forming, bubble forming, phase change, and dye ablate. Further, single, dual, trilayer and quadrilayer recording surfaces exist together with grooved or non-grooved tracking and pre-formatted, non-formatted, or post-formatted discs.

   The result is that, with a few exceptions, each commercially available recordable digital optical disc is designed to play in one particular device and is not recordable on drives from other suppliers.

   Considerable work is being done on standardizing a 5.25 inch WORM but they will be discussed in more detail in Chapter 3.

5. The last difference is pricing. Read only media are publishing media and hence, while there is relatively high mastering cost, replicate discs can be produced very cheaply and the players are relatively inexpensive. With recordable digital optical disc systems, there is, of course, no commercial mastering cost as the data is recorded on the discs by users in-house but the discs are relatively expensive at present because the coating pro-

cess is critical and they have to be supplied with cartridges. Current prices for 12-inch recordable discs range from $400 to $1,000. A typical cost of a 12-inch single disc drive at present is $20,000.

These differences will help determine each of the markets which CD-ROM and WORM will serve. CD-ROM is well suited to the distribution of digital databases to a large number of clients but is not suited to the in-house creation and storage of databases as users would have to have a magnetic tape drive on their computer system, send the tape to an intermediary and have them reformat it, produce a master CD-ROM and one or two replicates and then send them back to the user company - an expensive and time consuming exercise which could not be justified unless multiple copies of the database were to be distributed.

A recordable digital optical disc, however, would be ideal for archiving in-house databases as the user would simply download data from their computer system onto the recordable disc and it would be stored and instantly available. A number of companies, such as Aquidneck Data Systems, Data General and others offer optical disc data storage subsystems which are designed to emulate magnetic tape drives and link with most host processors. A number of other specific examples will be discussed in a later chapter.

The technology is not currently suitable, however for making multiple copies of these archived discs. The only way to copy digital optical discs at present is serially by reading the data from one disc and recording it onto another. This is a slow process requiring the availability of at least two digital optical disc drives. One full disk would take

approximately 60 minutes to copy this way. The next problem is the cost of the discs ($400 to $1,000 each) and the fact that at present, most only play on proprietary drives which, in turn, will only interface with certain computer systems.

If the physical and logical structures of 5.25 inch WORMs are standardized resulting in the ability to play different discs on different drives, then it will be possible to agree on an interchange standard, similar to that which exists for magnetic tapes, to allow 5.25 inch WORMs to be used for the exchange of database information. They would not compete with CD-ROMs for distribution to a mass market but could complement them. Users of on-line databases, for example, can expect to be offered the option of purchasing all or a portion of the databases on CD-ROMs for a fixed yearly subscription, enabling them to search the databases locally on their PCs. Such systems will tend to be single user systems with limited post processing capabilities.

The majority of first generation recordable optical disc drives use 12 inch diameter discs, based on the popular videodisc format. The discs offer storage capacities of 1 to 1.6 billion bytes per side, and can be single or double sided. For example, the OSI Laserdrive 1200 intelligent optical disc drive can accept a single sided, one gigabyte, twelve-inch preformatted glass disc using Phillip's air sandwich construction. There are 32,000 tracks per disc with 32 sectors of 1,024 user bytes per track. The drive can be supplied with an ISI or SCSI interface and offer a sustained transfer rate to the host of 250,000 bytes/second. The average access time is 150 ms and the disc rotates at 480 rotations per minute on the drive. Corrected error rates

from the drive are less than 1 in 10^{12} and the drive's mean time between failure is rated at 12,000 hours.

A new generation of 5.25 inch recordable digital optical discs are also now being developed and marketed by companies such as Optotech and Information Storage Inc. These first generation 5.25 inch discs offer storage capacities of 100-300 megabytes per side.

Finally, we may see within two years a recordable CD-ROM which has already been dubbed by Phillips as CD-PROM (Compact Disc Programmable Read Only Memory). It would make use of irreversible phase change recording techniques and would offer the same storage capacity access speeds and data transfer rates as current CD-ROM drives. However, the CD-PROM would be provided in its own cartridge for media protection and would not be readable on a standard CD-ROM drive. Phillips will, however, launch a new CD-PROM drive which will be able to play both CD-PROMs and CD-ROMs and potentially, CD audio discs.

2.1.6 -- Erasable Digital Optical Discs

Erasable digital optical discs can best be viewed as a third generation of optical discs which will give users the ability to not just record data on the discs but the ability to erase and re-record data many thousands or even millions of times on the same track sector, just as they can with magnetic media today.

The latter facility is essential for many active computer applications whose files must be retrieved, modified and re-written on a regular basis and the majority of operations and applications software packages assume the use of erasable media. Hence erasable optical disc suppliers will be aiming to compete directly with both flexible and rigid magnet discs for a share of the computer storage peripheral market.

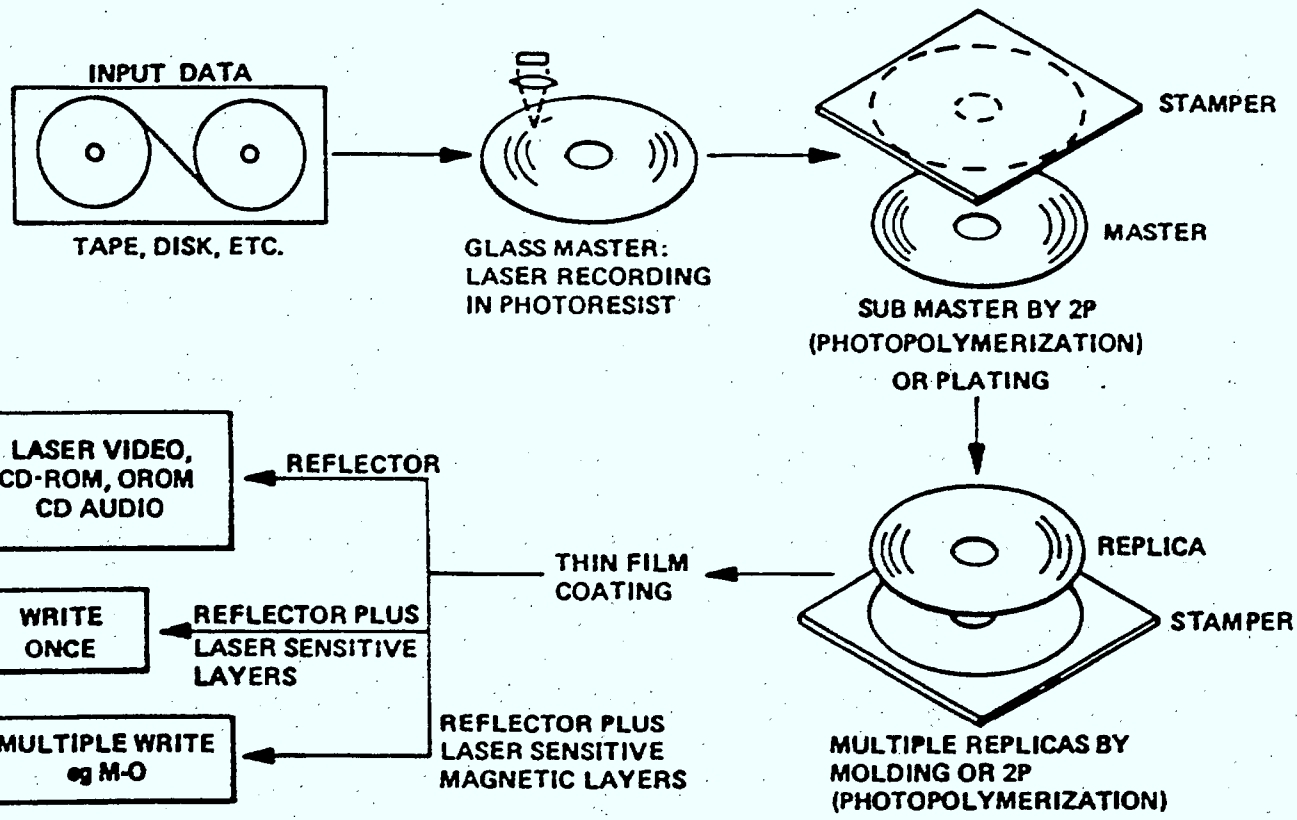
Four main advantages over magnetic media will be extolled by optical media suppliers:

1. High Storage Capacity - in excess of 200 - 300 megabytes per 5.25 inch disc.
2. Low Cost Drives
3. Removable discs so the off-line capacity is unlimited and jukebox type devices could offer automatic disc handling facilities.
4. Durability, so the use of non-contact laser record/read heads could avoid the risk of headcrashes.

As is illustrated in Figure 1, the processes involved in mastering and replicating erasable discs are almost identical to those for recordable discs. The mastering stages are the same and much of the replication, but, at the coating stage, the chemical structure of the recording layer is different so that the recording/erasure process can be repeated on a cyclical basis rather than being a single, inalterable process.

Although working prototypes of erasable optical disc systems have been shown, no erasable systems are commercially

FIGURE 1


GENERIC OPTICAL MEDIA OUTLINE
INPUT DATA
TAPE, DISK, ETC.
GLASS MASTER:
LASER RECORDING
IN PHOTORESIST
STAMPER
MASTER
SUB MASTER BY 2P
(PHOTOPOLYMERIZATION)
OR PLATING
REPLICA
STAMPER
MULTIPLE REPLICAS BY
MOLDING OR 2P
(PHOTOPOLYMERIZATION)
THIN FILM
COATING
LASER VIDEO,
CD-ROM, OROM
CD AUDIO
REFLECTOR
WRITE
ONCE
REFLECTOR PLUS
LASER SENSITIVE
LAYERS
MULTIPLE WRITE
eg M-O
REFLECTOR PLUS
LASER SENSITIVE
MAGNETIC LAYERS

available at present and hence exact specifications are hard to give. 3M is one of the leading companies in this area together with KDD, Sharp and other Japanese companies. Two basic techniques are employed - phase change and magneto optic - and latest estimates are that erasable disc subsystems using these two techniques will begin to become available from 1988 onward.

Phase change systems are based around the observable fact that certain tellurium-based alloys can exist at room temperature in either the crystalline or amorphous form and can be switched reversibly between each state. The optical properties of the film, its reflectivity, are different for the optical read-out of stored data. Matsushita are one of the leading proponents of this recording technique.

The recording of data in magneto-optic systems involves the use of a laser to provide heat and a magnetic coil to provide the magnetic field. Prior to recording, the entire recording layer (the magnetic film) is perpendicularly magnetized. Then, at the recording stage, a micron-sized region on the disc is heated up by the write laser to a point above the so-called Curie point. As the material cools in the presence of an external magnetic field oriented anti-parallel to the initial direction of magnetization, a small region (one micron) is formed that is reverse polarized. This region is the equivalent of a pit or bubble in recordable discs and playback is accomplished with a low power laser read beam.

The first generation of erasable digital optical discs, when they become widely available, are likely to be 5.25 inches in diameter and will be recorded in Constant Angular Velocity

mode to provide fast access times and high data transfer rates. 3M have recently formed an agreement with Optimem in the U.S. to produce and build a universal drive that will accept read only, write once, and erasable, rewritable 5.25 inch discs recorded in CAV mode. The read only medium used in such a drive would not be CD-ROM but would be a double-sided, 5.25 inch diameter CAV mode Optical Read Only Memory.

2.2 Advantages and Disadvantages of CD-ROM as a Publishing Medium

CD-ROM technology exhibits a number of advantages and disadvantages when viewed in the context of applicability as a publishing medium. Currently, the advantages of CD-ROM when viewed in this light are:

2.2.1 -- Advantages

High Storage Capacity

One CD-ROM disc possesses the capability to hold 550 megabytes of data which is the equivalent of 250,000 pages of coded text, or equal to what could be captured in a stack of 1,500 floppy discs.

Low Cost

CD-ROM technology is based on a mass-produced audio medium and hence, disc and media costs can be kept low. Further, the technology is already proven and can be produced in existing CD-audio plants with minimum retooling. As well, economies of scale exist for CD-ROM publishing. Thus, the

more that can be produced and sold, the lower the per unit production costs. It follows, therefore, that the most popular material can be sold very cheaply and still be very profitable to the publishers.

Standardized Formats

The computer and information industries have cooperated to develop a file format which should enable CD-ROM designs adhering to that standard to be read by a range of different operating systems running on different computers which are in turn linked to different CD-ROM drives.

2.2.2 -- Disadvantages

The technology is not without its drawbacks, however. These drawbacks currently include:

Extended Turnaround Time

The data preparation, premastering, replication, checking, packaging and distribution cycle is a complex one and turnaround times are inevitably long. No publisher, at present, is considering a more frequent than monthly update cycle. The implication of this is that in many cases, material published on CD-ROM could be eight weeks old by the time it is available to the user.

Technology Change is Rapid

Even in context of the fact that there exists an industry-wide standard for CD-ROM production, the CD-ROM is a computer peripheral and the PC market is extremely volatile

and competitive. While this results in cost reductions in hardware, widening the market for CD-ROM products, it can also pose the publisher certain problems as the computer and software companies continue to update operating systems and hardware more rapidly in order to gain competitive advantage. For example, CD-ROM may be challenged in the future by other read-only memory formats such as OROM. OROM, or Data ROM is a multifunction 5.25 inch optical disc drive capable of accepting read only, write once, and erasable optical discs. The technology also utilizes the fast access Constant Angular Velocity format.

Relatively Slow Access

Since CD-ROM is based on the CD-audio Constant Linear Velocity format, it exhibits slow access time and hence, databases and retrieval software must be designed to overcome this limitation.

Poor Development of Graphics

CD-ROM technology lacks development in the direction of graphical memory representation. Most current CD-ROM systems deal only with coded textual material.

Clearly, as the industry develops, graphics of all kinds will be recorded on CD-ROM. However, the High Sierra Group Working Standard does not specify how graphics should be recorded on CD-ROM. Much work needs to be done in investigating the areas of resolution standards, workstation design, storage requirements, compression and decompression, and the cost implications of handling graphics on CD-ROM.

2.3 CD-ROM as Compared to Competing and Complementary Technologies Utilized for the Production and Distribution of Databases

CD-ROM can be compared to a number of traditional competing and complementary technologies currently used to produce and distribute databases and information in general.

The comparison which follows was made against a standard set of criteria which includes:

- Storage Capacity
- Type of Material - Graphics, Colour, Moving Video, Resolution Quality
- Ease of Production
- Cost of Production/Replication
- Cost/Ease of Distribution
- Turnaround Times
- Ease of Updating
- Retrieval Hardware Required
- Portability of the Medium
- Accessibility of the Information
- Ease of Use and User Acceptance
- Ease of Integration with Existing Systems

2.3.1 -- Storage Capacity

While there is a limit to the amount of information that may be stored on a single page of paper, there are, theoretically, no bounds to the number of pages we may package together in a volume or series of volumes. There is, however, a direct relationship between storage capacity, costs and space consumed. For each page of information, there are additional production, consumable and distribution costs and

an increase in the bulk of the medium. Thus, the realistic limit which is imposed on the storage capacity of paper is the amount which can be handled conveniently. If storage capacity was to be measured by the number of pages that could be stored in a particular area, then clearly paper would rate far behind microfilm or optical discs.

To illustrate, one-100 foot roll of 16 mm microfilm can store over 2,500 frames, each containing the image of an A3 or smaller sized page. However, where each frame contains the image of an A5 or smaller document (i.e. a cheque), the roll can contain as many as 4,000 to 5,000 frames. In terms of fiche, one A6 size source document microfiche stores 98 frames and one Computer Output Microfilm (COM) can store some 270 - 400 frames. Therefore, in pure storage terms, microfilm is a very compact medium. In fact, studies have shown that in large scale systems, microfilm can save some 95% to 98% of the space occupied by equivalent paper filing systems.

In terms of versatility, however, microfilm rates lower than paper. With microfilm, if only one page needs to be copied and distributed, an entire microfiche has to be produced or a strip of film has to be loaded into a jacket.

As is the case with paper, there is theoretically no limit to the potential storage capacity of an on-line system. Realistically, again, it is limited to the number of magnetic disc drives which can be controlled by any host while still offering realistic response times. At present, most textual databases can be accommodated on on-line systems. However, if suppliers wish to build up large graphic databases in the

future, then optical storage systems and higher bandwidth networks will be needed.

There is a considerable cost involved in adding large numbers of high capacity magnetic disc drives to the system but this must be weighed, as in most other situations, with the benefit provided by the system. In the case of in-house systems, the database must be valuable enough to the company to justify the expense or, in the case of commercial databases, is of sufficient interest to enough people to cover the costs of storing the data and making a reasonable profit.

Lastly, and certainly not least, each CD-ROM has a storage capacity of 550 megabytes which is equivalent to the memory required to contain 250,000 pages of coded text or 10,000 scanned images. This makes CD-ROM a far more compact medium than paper or COM and approximately twice as compact as roll microfilm for the storage of images. However, the dense nature of CD-ROM storage means that it does neither compete with paper as a way of distributing a few pages of information nor microfiche as a way of distributing thirty or forty pages of information. The mastering costs associated with CD-ROM and the hardware costs associated with accessing the data stored on CD-ROM mean that it will not be cost effective for distributing less than 1,000 to 2,000 pages of information or approximately five microfiche or 10-20 megabytes and will not really come into its own until one is distributing 10,000 pages of information. CD-ROM, then, like microfilm, is an incremental medium.

2.3.2 — Type of Material

Virtually any type of material can be recorded on paper. The only limitation is that the data is static as, obviously, no capability for portraying animation or motion video exists. Very high resolution images may be recorded on paper provided that high quality paper is used.

Source document microfilm, like paper photocopying systems, can be used to capture all types of material - text, graphics, additional material. For colour representation, however, a special type of film is required. Also, as with paper, microfilm is a static medium so there is no provision for reproducing motion video.

The quality of reproduction which can be achieved with microfilm can be extremely high but typically is closer to that obtained through photocopying rather than in graphic art applications.

Most of the currently available on-line databases are limited to text and numeric data but many also incorporate computer graphics and, via videotex interfaces, some companies are now experimenting with low resolution digitized photographic images. A limitation with image representation on-line exists, however, due to the fact that most external systems rely on telephone lines and data can not be transmitted down them fast enough to support images. Further, most on-line users work with standard computer terminals and the monitors do not have the necessary resolution to support image display. Lastly, with on-line systems, as with microfilm and paper, motion video can not be supported.

CD-ROM is, at the present time, being primarily used to store text and computer graphics. However, CD-ROM and other CD systems can be used to store audio information, roster scanned images, video and full motion video. Inherently, it is the most versatile of standalone database mediums in terms of the type of material it can contain.

2.3.3 -- Ease of Production

Although a wide range of techniques for producing and replicating material or paper exist, ranging from photocopiers to laser printers, the original material can come from one of two sources. In the first case, the material is already on paper and copies are made via a photocopier or lithographic process. In the second case, the material exists in digital form which may then be transferred onto paper with a line, daisy wheel, or non-impact printer.

Ease of production with impact print, non-impact print and offset lithography systems depends on the volume of material being produced and the frequency of updates required. For example, the photocopier or offset system is ideal for a 10 page report. However, for voluminous technical manuals which need updating every month, the publisher should capture the entire manual in digital form and design a sophisticated database structure through which text and graphics can be recalled, displayed, updated, merged and output in the required sequence on a regular basis.

In terms of microfilm, for in-house use where the originals are in paper form, the production process can be a fast and

simple operation. High speed rotary cameras film thousands of documents per hour, and even with awkward material such as bound volumes and journals, rates of several hundred pages per hour can be achieved.

Two hurdles exist in microfilm publishing. First, source documents must be sorted into a desired order before they are filmed and that order is then fixed for the life of the microfilm. The second is that once the documents have been filmed, the film must be processed before it can be used. This results in a delay and necessitates staff trained in the use of chemicals.

Setting up an on-line database is a complex exercise involving data capture database design, indexing and loading. Compiling an on-line database can be compared to using a laser printer to output a database onto paper prior to distribution; to computer output microfilm, where a database is output onto microfiche; or to CD-ROM where the entire database or a subset of it is output onto CD-ROM and distributed for local access. In the context of these comparisons, the ease of production would be approximately the same for all four media and indeed, the same database can be output in all four ways if required.

In some cases, however, production of CD-ROM based databases is more complex than that of on-line as developers work out their own proprietary file structures. In the future, as the High Sierra Group Standard is adopted, the process will be simplified. But to make full use of CD-ROM as with on-line distribution, it is essential to go through the expensive one-time process of creating a database. This limits CD-ROM

to applications requiring the provision of large amounts of information to large numbers of end users with a regular update cycle and where the information is valuable enough for the end users to pay a premium for the service.

2.3.4 -- Cost of Production and Replication

A fixed cost is involved in capturing and organizing information before placing it on paper, microfilm, CD-ROM or making it available on-line.

In terms of paper, when dealing with low volume, simple reports, all that is required is to copy the original using a photocopier or create a few plates on a lithograph machine. In this degree of applications, the cost can be considerably lower than that for placing the information on-line or CD-ROM. If one is dealing with very large volumes of material, however, and wanting to update it regularly and print it out via a high speed laser printer, then the steps which have to be gone through in capturing, writing and organizing the material into a fully computerized database are almost exactly the same whether the final output is paper, microfiche, CD-ROM, or on-line distribution.

Basic cost elements involved with paper are toner/ink, machine time, and staff time for supervising production. While with paper production and replication, costs are related directly to the number of pages produced, this is not the case with microfilm or CD-ROM.

Source document microfilming is a relatively low cost operation. If 16 mm roll film is used, the commercial bureaux

charge rates of $22.20 to $44.40 per one-thousand frames to film the documents and process the film. With source document microfiche, typical costs are $17.72 to $33.30 for a ninety-eight frame microfiche. In terms of replication, one-hundred foot reels of 16 mm roll film holding approximately 2,500 frames can be copied onto diazo film for as little as 25 cents. Thus, for in-house use where large numbers of documents have to be filmed and only one or two copies are required, roll film is ideal. For publishing purposes where hundreds of copies of a report of substantial length must be distributed microfiche is the ideal medium and, combined production and replication costs are far lower than with paper.

Computer Output Microfilm (COM) fiche are produced at even higher reduction ratios than source document fiche so as many as 400 frames can be put on one fiche. Production costs for a COM fiche range from $3.30 to $8.90 and duplication costs are the same as for source document fiche. Costs of producing and replicating COM fiche are far lower than the cost of producing the same amount of material on paper via a laser printer.

In terms of designing and creating an on-line database, development is a complex task and represents a considerable overhead to the in-house or commercial publisher. However, once the database has been created, it can be updated on-line and maintenance is not excessive.

Not only can potential on-line distributors receive help and guidance in designing and creating the database from system integrators and software suppliers, but also they can, if

they lack the resources, the market, the database and the hardware to maintain it, enlist the services of a host who will take the database, install it on their mainframe system, and market it to their client base.

The cost for creating a database on CD-ROM currently ranges from $20,000 to $220,000 depending on the amount of data capture needed, the size of the database, and the complexity of the database structure. Fixed costs include mastering of $4,000 to $6,500 per ROM and replication which, depending on the volume of the replication, ranges from $10 to $60 per disc.

The major difference between CD-ROM technology and the other mediums is that not even the largest companies would master or replicate discs in-house as the cost of such a facility falls in the region of 100 to 200 million dollars. This will have implications on turnaround times and could limit the usage of CD-ROM where security is an issue.

2.3.5 -- Cost/Ease of Distribution

In terms of paper, the more pages that exist, the more expensive it will be to distribute them, both with respect to staff handling costs and postal and transport rates.

Microfilm, and in particular, microfiche, has proven to be a very attractive distribution medium in in-house and technical publishing applications due to its low production and replication costs and its compact nature. Ten COM fiche, or the equivalent of 2,700 pages, can be distributed for the price of a first class stamp. Microfiche are ideal for on-demand

print systems. One master microfiche can be kept and used in the production of replicates as orders are received.

In terms of on-line system distribution, the growth of host companies is testament to the fact that the sale and marketing of on-line systems has proved difficult for many information providers. By using a host, the information provider does not effectively have to worry about distribution at all. He simply produces one or two copies of the database and gives them to the host to put up on their mainframe system.

In the case of CD-ROM, with database creation and mastering costs being one-time costs, the economics of CD-ROM become more attractive the more potential users there are. CD-ROM, then, competes less with photocopying or microfilm in distributing small numbers of a report as it is more geared to distributing large numbers of a medium containing large size databases to end users. In terms of ease of distribution, CD-ROM is very compact and hence, postage is quite low. If the disc is full, consumable costs are low at $22 to $44 for 200,000 pages.

2.3.6 -- Turnaround Times

The flexibility of paper is again apparent in terms of turnaround times. If a simple copy of a document is required, the original may be photocopied in seconds or printed out on demand via laser printer. However, as the volume of material increases and the number of copies required goes up, so the process of printing out, collating, and dispatching the material lengthens. However, large volumes with twenty-four hour turnaround times are still possible (i.e. newspapers).

Turnaround times do suffer when large reports must be updated regularly and be distributed to a large number of users via traditional postal routes.

With source document microfilming, turnaround is again a function of the size of the original and the number of copies of that original required. Replication can occur at much higher rates than any printing process with duplicators capable of producing 2,000 to 3,000 copy microfiche per hour, with each fiche containing 270 frames of information.

Turnaround times can be even quicker with COM as modern recorders can produce one 270 frame fiche per minute as well as having integral processors which produce master microfiche ready for replication.

With an in-house, on-line system, the turnaround time is defined by how long it takes to capture the new data and process it so that it is in a form in which it can be added to the database. In terms of commercial systems, it is the time it takes for the host to load the new data up onto the database. Clearly, there is a lengthy turnaround between the decision to set up a large on-line database system and the actual implementation, but once the system is up then it can be updated relatively easily.

CD-ROM can not compete with any of the other media in terms of turnaround times. This is due to the production cycle which entails updating the database, sending the tapes to the mastering facility, producing a check disc, stamping replicates, and then sending them to the end user or publisher.

It is unlikely that anyone will have a turnaround time shorter than one month. The only possible exceptions are for small-scale, in-house applications where, apparently, companies such as 3M, with their replication techniques, can offer two or three day turnaround with limited numbers of replicates. This may, for the aforementioned limited application, bring the turnaround time down to one week.

However, even this weekly turnaround is not competitive with COM which can provide twenty-four hour, or even six hour turnaround. Even in large scale applications, such as banks, COM fiche are distributed to thousands of branches every day.

2.3.7 -- Ease of Updating

Although updating a small existing volume of data contained on paper is an easy task, as the volume increases, so does the necessity of storing the data in some sort of database format where pages can be brought up in any desired sequence and formating and repagination can occur automatically.

Microfilm, as with paper, is not the best way of disseminating highly volatile information. Once produced, microfilms cannot be physically annotated. With source document publishing, the update process consists of refilming all the material, both old and new. With COM, it is possible to update a large database by printing and indexing the new material.

Ease of update is one of the main advantages of on-line distribution. Once the database has been created, data can be updated in seconds, which explains why all providers of

volatile information, such as airline seat booking systems and financial databases use it.

In terms of CD-ROM, the only way to update is to reissue. The only means by which the long turnaround characteristics of CD-ROM can be offset is through combined CD-ROM and on-line service where backfiles of static information are made available on the disc, while volatile information is accessed on-line.

2.3.8 -- Retrieval Hardware

One of the main advantages of paper as a publishing medium is that it requires no retrieval hardware so there is no expense involved and no compatibility problems.

Users of microfilm, however, need at least an optical reader, if not a reader-printer. This is a major barrier to acceptance as a consumer medium and for books and newspapers. Current readers, however, offer large improvements over early models, being microprocessor-controlled, providing fast and accurate retrieval.

Use of on-line systems is also impeded by the required retrieval hardware. A terminal and a modem to provide access to a telephone line are the bare essentials, preventing use in remote locations without access to telecommunications.

The main benefit that CD-ROM possesses in terms of retrieval hardware is that personal computers are selling large quantities so that for many users, CD-ROM will just be another peripheral. The price of CD-ROM workstations will not likely

decline for a year, however, at which time standards which have been put forth should be accepted and economies of scale taken effect in the production process.

2.3.9 -- Portability of the Medium

Depending on the volume, paper is an extremely portable medium as no hardware is required for retrieval of the information stored on-line. As volumes of information increase, other mediums discussed become more beneficial in the context of portability.

As microfilm has to be viewed using specific hardware, however, it is not nearly as portable as paper. One and two decades ago, attempts were made to develop book size readers powered by batteries or the sun, but no one has resolved the problem of optically enlarging and displaying the images stored on microfilm.

With respect to on-line databases, the increasing penetration of portable, compact computers into the personal computer market, coupled with the advance of in-car telephone technology will continue to improve the "portability" of this medium.

Lastly, personal computers with CD-ROM drives possess a low degree of portability. However, when half-height 5.25 inch drives are available, knee-top CD-ROM workstations will likely closely follow.

2.3.10 -- Accessibility of Information

The fact that paper is basically a serial medium makes it unsuitable when the information contained on it is aimed at multiple users who each want to access it in different ways. The only way to adapt paper to a wide range of retrieval requirements is to print the same list out in a number of different sequences.

However, as the problems such as cost and space overhead begin to become a detriment, the need for an electronic database format becomes apparent.

Traditional source document microfilm systems also force the selection of one or two sequences in which to order information. Even though microfilm is more compact and cheaper than paper, reprinting in a number of sequences still incurs large costs with extensive databases.

Accessibility is another major advantage of on-line databases. The entire database can be searched using structured or unstructured full text searches. The only hurdles to accessibility are the log-in and various security procedures necessary in the control of such a resource.

Accessibility is also foreseen as one of the main advantages that CD-ROM will possess given it will combine the ease of searching on-line databases with the advantages of local access such as graphics and motion video, unlimited access time, and the ability to offer extensive help facilitles to the user.

Multi-level software has already been developed for use in concert with CD-ROM systems so that experts and novices can interact with the reader through the same interface.

2.3.11 -- Ease of Use

In terms of utility publishing, paper as a medium has some major limitations as large multi-volume directories, for example are intimidating and physically and logically difficult to use. Again, information is laid out in only one sequence.

With respect to microform systems, some resistance to the technology in many application areas exists. Information contained on this medium is not searchable in full text form and as personal computers become more of an established presence, users will demand that data be fully searchable.

A definite learning curve exists in the use of on-line systems. End users need training in the search software and command systems that a particular database employs. This may be a deterrent if many different databases and systems must be utilized. Once search strategies are mastered, however, data is far more accessible than with paper or microfilm-based systems.

Ease of use is purported to be a major advantage of CD-ROM systems once standards are agreed to. The log-on process will be simpler and it will be considerably easier to install more help facilities on CD-ROM than on an on-line system, providing the opportunity to include tutorials on the same disc on which the data is resident.

2.3.12 -- Integration with Existing Systems

Information held in printed form cannot easily be entered into a computer system. With the growth of personal computers in many application contexts, there is a growing requirement to take data from published sources and process it further into a form suitable for the subsequent user.

Microfilm is also not easily machine-readable. Even COM, once data is output, cannot be read back into a computer. As with paper, users can not access selected segments of the data and then download them into their personal computer for subsequent processing or modeling.

On-line systems, however, can easily be integrated into existing systems by using a mode as the interface between the telecommunications lines and the personal computer. The data found on the on-line system can be downloaded into the user's personal computer for subsequent processing as it is already in coded form.

CD-ROM offers an attractive alternative to re-keying information from published sources for further processing. For example, bibliographic data distributed on CD-ROM could form the basis of an on-line ordering system or an on-line catalogue card. Further, if directories are distributed on CD-ROM, sets of company addresses can be selected, downloaded onto magnetic media, and used for mailing lists.

2.4 An Examination of Some Other Storage Mediums

2.4.1 -- OROM - Optical Read Only Memory

We have already seen the key advantages of CD-ROM as a publishing medium as being standardization, high storage capacity, the availability of mastering and replication facilities, and relatively low cost hardware.

However, two drawbacks associated with CD-ROM technology exist which may leave it vulnerable to OROM for certain applications. One is the type of Constant Linear Velocity made recording which results in relatively slow access times of as much as 1 to 2 seconds in current CD-ROM drives. A second drawback is that the CD-ROM is 4.72 inches in diameter and hence it is not compatible with the emerging 5.25 inch WORM discs and prototype 5.25 inch erasable optical discs which tend to use the Constant Angular Velocity format.

A number of companies, the first of whom being 3M, floated the concept of a 5.25 inch Constant Angular Velocity mode read only optical disc, called OROM or Data ROM. This new medium could operate in the 5.25 inch optical disc drive environment which is capable of accepting read only, write once, or erasable optical discs. Thus, not only would OROM be compatible with future optical media standards, but it would possess a faster access time and data transfer rate. Such a medium would be easily adaptable to higher storage densities as they become feasible and would serve the high-end computer marketplaces where multi-user access was a requirement. The low-end publishing marketplace would likely be kept to CD-ROM. The prototype system, a collaborative

venture between 3M and Optimem, provides storage capacity of 200 megabytes per side, access times of 100 msecs, and data transfer rates of 5 mbits per second.

Sony also announced, in 1985, that a read only optical disc called Data ROM, would be introduced which would be a compatible format to their write once and erasable optical discs. Not only would this permit write once, erasable, and read only data on the same discs, but faster access and data transfer rates than is possible with CD-ROM would result.

A last key player in the OROM segment of the market is Phillips. Like 3M and Sony, Phillips intends to commit itself to CD-ROM and CD-I for the domestic consumer and low end electronic publishing marketplace. However, a Phillips and Control Data Corporation subsidiary called Laser Magnetic Storage will be fielding a write once 5.25 inch drive which will also accept OROM discs prepared to similar formats to the 3M and Sony media.

These and other product introductions point to the fact that there is a market for a faster, higher performance read only optical disc which will be compatible with planned 5.25 inch WORM and erasable disc formats. It is also believed that IBM is interested in supporting this type of optical product - an endorsement which would likely ensure the OROM format a firm position in the marketplace.

OROM is expected to have very little impact on CD-ROM-based products in the short term, apart from creating further confusion on the part of the consumer, and thus, a delay in purchasing decisions.

In the long term, CD-ROM and OROM will settle into their own niches in the optical storage media marketplace. In commercial publishing applications, the emphasis will be on low cost workstations, standards and the widespread availability of mastering and replicating facilities and data penetration facilities. CD-ROM is already meeting these requirements for technical publishing applications. As well, from many technical publishing applications where CD-ROM is in competition with existing paper and microfiche-based systems, the key requirements will be low cost workstations and low replication costs for the media. In other words, the CD-ROM media will not likely be challenged by distributing commercial databases, reference works and images to PC users.

However, OROM will be a high end product meeting the data distribution needs of mainstream computer users. For software distribution and the distribution of raw data and graphics libraries for use in particular computer systems and for a number of in-house data distribution applications, OROM looks potentially attractive. If the media does, in fact receive the backing from IBM speculated on above, it will be widely adopted by mainstream computer users, due to its advantages of multi-user access, compatibility with other formats, faster access and faster data transfer rates. Even where OROM is used for distributing software, documentation, and graphics libraries, possible shortage of mastering facilities and even the lack of standards will not prove a significant drawback. 3M can master OROM media already and if OROM receives the backing of major computer companies, then many current mastering facilities can be adapted to produce and replicate OROM discs.

2.4.2 -- Hybrid Videodiscs

When videodiscs were first launched as an analogue storage medium for the domestic consumer marketplace, companies such as Reference Technology, Laserdata and TMS saw the potential of videodiscs for data distribution or optical publishing.

Techniques were developed for encoding digital data within the video signal of the videodisc. Proprietary optical publishing systems were also based around videodiscs. In these systems the publisher sent his database in coded form to one of the above companies and they encoded the data, recorded it on a videodisc and produced replicate discs containing the digital data. For playback the user then needed an industrial videodisc player linked to a decoder/control unit which in turn was linked to a microcomputer. The data was then read off the videodisc, converted in real time from analogue to digital format, underwent error detection and correction routine and displayed on the computer screen.

These systems are called hybrid videodisc systems because they recorded digital data in the video format. Thus, it is theoretically possible to combine both analogue and digital data on the same videodisc, providing a hybrid system which can support still frame and motion video, text, data, graphics and audio.

The major problem haunting suppliers of hybrid systems was the lack of standards in the field. Each supplier used different techniques to encode the data, different hardware and different error detection and correction systems. The lack

of standardization resulted in two major difficulties. First, there was no interchangeability of discs and players and the large commercial publishers did not feel confident that any one supplier would triumph. Second, due to the proprietary nature of the hardware and software used in the systems, prices were too high to make hybrid systems attractive to the generic publishing market. This experience is yet another testament to the importance of standards to the electronic publishing and computer industry as a whole.

Hybrid systems, then, appeal only to a very narrow segment of the market. With CD-ROM available, the main application areas for hybrid videodiscs will be where high storage densities are required (one video disc can store more than one billion bytes of data) or where a mixture of high quality motion video and digital data is required.

A recent hybrid videodisc product introduction is the Laservision ROM, which differs markedly from most such systems. Phillips has not attempted to store data in the 55,000 video frames on each side of the disc - they have reserved this for still-frame or motion video. However, the company has devised a technique for recording data in the twin audio tracks on a disc and claim that if no audio is required, they can store 324 megabytes of data on <u>each side</u> of the twelve inch videodisc plus 55,000 video frames.

2.4.3 -- Optical Cards

The best known product falling under the optical card type storage medium is the Drexon lasercard developed by Drexler

Corporation. The Drexon lasercards are basically plastic credit cards with the optical recording medium applied in a strip as a replacement to the magnetic strip used in other smart cards. Like optical discs, the cards may be read only or write once.

The Drexon recording strip measures 35 x 80 mm and has a storage capacity of two megabytes, although versions possessing a capacity of ten megabytes have been demonstrated. These cards are wallet-sized, like a credit card, and are fully encapsulated in polycarbonate plastic for security and usability. The Drexon cards can be prerecorded in the factory for read only memory applications in cases where many copies of permanent, identical data are requested. Mass production of the ROM cards is achieved through high speed photolithography, utilizing specialized equipment in cleanroom conditions. The data bits recorded on the cards are read using photodetector arrays. In high volume the read only laser cards could cost as little as $5 per piece.

Like the read only memory card, the recordable laser cards have a storage capacity of two megabytes, or approximately 800 pages of standard text. In addition to cardholder data and transaction information, potential applications include storage of digitized photographs or fingerprints, encryption codes and a range of other deterrents against counterfeiting and other fraudulent uses of lost or stolen credit cards. The cards will be laser recordable at a secure point of issue or at point of use for cumulative recordkeeping or debiting and a standard blank recordable card could sell for roughly two dollars in high volume.

A key element in the success of the cards, as with other optical media, will be the cost of the reading/writing equipment and its reliability. A number of Japanese companies, namely Toshiba, Nipponcoinco, Canon and Olympus, have demonstrated both card readers and card reader writers. The Toshiba reader measures 150 x 210 x 290 mm and offers a transfer speed of 32 - 64 kilobits per second. It uses a CCD array to read the data. The reader/writer unit is capable of recording up to 2 megabytes onto a card at a write speed of ten kilobits per second. The read speed is 64 - 128 kilobits per second.

Laser cards will be useful for distributing simple software packages to PC users or possibly for updating material distributed on CD-ROM. The prime applications, however, will be in the transaction handling environment.

2.4.4 -- WORM Discs

Two key application areas exist for WORM discs - as a data archiving medium in competition with magnetic tape and for the storage of facsimile scanned images of documents in competition with microfilm. Current twelve inch WORM discs and forthcoming high storage capacity 14 inch discs from Kodak and others will be used in large centralized systems. Clearly, there will be little overlap here between WORM discs and CD-ROM systems as in many applications, textual and image databases will be captured and maintained on large WORM discs and then sub-sets of the databases will be downloaded onto a CD-ROM for distribution and sale.

There are a growing number of 5.25 inch WORM drives coming onto the marketplace too and although these will be used to provide PC users with data archiving facilities and as part of PC-based personal filing systems, one of the key application areas for such WORM discs could be the distribution of selected in-house databases and here WORM discs could compete with CD-ROM systems, depending on the number of end users and the need to update the information.

The previous section compared CD-ROM with microfilm and specifically with Computer Output Microfilm for the distribution of in-house databases and technical documentation. Many of the in-house applications of COM involved distributing from 1 to 10 copies of a database to specific workers within an organization. The reason for doing it is that it is quicker and simpler for those users to access the data from COM fiche structured to meet their requirements than it is for them to access the database on-line in the form in which it is held to the mainframe and where they are in contention with other users.

With such a small number of end users CD-ROM could not be cost-effective and could not offer the immediate turnaround of COM systems. WORM discs can, however, provide the solution. The particular sub-set of the database required by the end users would be downloaded onto a 5.25 inch WORM disc and the required number of copies of the disc could be produced by serially copying disc to disc and the discs then distributed to the end users for access on their PCs. Currently it takes about 40 minutes to serially copy one side of a 200 megabyte WORM disc but this production time will decrease as the raw error rates of the medium improve and the data trans-

fer rates are increased. The obvious benefit of course, is that the database is still in coded form and hence would be more easily searchable and processable.

In applications, then where either the number of end users does not justify the cost of mastering a CD-ROM, or where the turnaround times offered by CD-ROM are not acceptable, or where users do not wish to use external mastering and replication facilities for security reasons, WORM discs could, and are, being used to replicate and distribute databases.

Two companies who have already demonstrated such systems include MARCIVE in the U.S. who distribute bibliographic databases to libraries on 5.25 inch WORM discs and a company called Teletrak who have developed a complete Computer Optical Disc System designed to challenge current in-house COM systems.

With the Teletrak system companies produce a print tape formated exactly as if they were going to load it on their COM recorder and download the data onto Computer Output microfiche. The tapes are loaded but instead of being displayed on a CRT screen and recorded onto film, the print image data is recorded onto 5.25 inch WORM discs using a software package developed by Teletrak. Indexing data is created concurrently and stored on a hard disc. The double-sided 5.25 inch WORM discs hold 400 megabytes of data, or the equivalent of 70,000 computer report pages, 1,100 floppy diskettes, or 300 COM microfiche. In addition, as mentioned above, they can be copied and distributed for access by specified user departments.

3.0 THE ROLE OF STANDARDS IN THE STANDALONE DATABASE PUBLISHING INDUSTRY

CD-ROM is the most advanced of the optical publishing technology in terms of commercial development and market introduction and acceptance. It is logical, then, to focus our attention in this chapter, on this medium, tracing the evolution of its related standards as well as assessing the future implications of this degree of standardization.

CD-ROM borrows its advanced degree of standardization from CD-audio technology. Compact discs were first shown in 1980 and were launched commercially in 1983. Prior to the product introduction, however, much work was done by Phillips and Sony Corp. in the direction of creating a standard format for production of the playback device as well as the storage medium. The CD-based standard, referred to as the "Red Book", ensures that all CDs could be played on all CD-players. The fact that the world's major electronic manufacturers rallied around one standard format was one of the key factors in making the compact disc the most successful introduction of a consumer electronic product in history. Without this standard and the confidence it instills on the part of the consumer, market penetration would have fallen far short of what it is today.

CD-ROM is a computer peripheral. Thus, given its application, it must meet the dual requirements of the computing and publishing industry to ensure its long term success and to attract the sort of investment that will be needed to take it from its present position as a prototype distribution system to a mainstream publishing system. One of the main requirements is for internationally accepted standards governing all aspects of CD-ROMs that could affect system compatibility.

The Phillips/Sony Yellow Book Standard, which is based on the Red Book, defines the physical dimensions of the disc, the recording techniques, the data format, the main aspects of drive performance, data modulation and the error correction and detection system. This gives CD-ROM a major advantage over the hybrid videodisc systems where each supplier developed proprietary data encoded systems, proprietary error detection and correction systems, and proprietary hardware - a situation militating greatly against compatibility.

The need also exists for a logical format standard - an agreed logical format dictating the manner in which data is organized on CD-ROM and how the CD-ROMs themselves are identified (the volume, directory, and file structures). This is necessary if different operating systems running on different computer hardware are to be able to access the information stored on one CD-ROM. This logical volume and file structure standard must be built on top of the Yellow Book physical standard.

In order to set a context for this discussion, it is necessary to point out that the logical format is one of three key components of a CD-ROM file management system. The file management system stands between the application software running on the computer and the controller which controls the disc drive. A file management system basically comprises software and data structures designed to convert the physical, sector-oriented view of the disc held by its controller into a logical view that the application program can use. For example, in the case of CD-ROM drives, a CD-ROM drive controller views the CD-ROM disc as a sequence of 2 kilobyte sectors spread over sixty minutes of recording time with 75 sectors in each second of the recording - a total of 270,000 sectors. Taken from a logical perspective however, a CD-ROM disc is a collection of files, each with its own unique name, and the application program needs to call for a file by that name. The application program must be able to open the file and read it, regardless of the fact that the information in each file must be read in 2 kilobyte sectors or that each sector has a precise physical address.

Besides the structure or logical format of the data, the other two components of which a file management system consists are the software that writes the data in that format (the origination software) and the software that reads and translates the logical format for our use (the destination software).

The logical format of the CD-ROM disc will determine where to put identifying data on the disc, where to find the directory or directories of files on the disc, how the directory is structured, whether the subdirectories are supported, how many files can be stored on a CD-ROM, the performance cost of storing large numbers of files, how large a file can be, whether files can span multiple volumes and whether files must consist of sequential consecutive sectors.

The logical format is clearly different from the physical format of the disc (defined in the Phillips/Sony Yellow Book) which is considered as given by the file management system. The logical format is crucial to the file management system - it defines the system's structures and operating characteristics and hence must be standardized if CD-ROM compatibility is to be achieved.

The CD-ROM 'Yellow Book' standard was made available in 1985. The National Information Standards Organization (NISO) organized a meeting of information industry representatives to discuss the need for CD-ROM standards. Consensus was reached that a minimal standard was needed which allowed interchangeability and supported optimal performance without limiting applications. It took six months to name this committee and draw up the goals, during which time, numerous CD-ROM products were announced.

The committee was charged with developing a standard which at base established a standard enabling:

(1) Developers of software to have a uniform file environment in which to design and implement applications.

(2) Publishers and distributors of information to master a single compact disc and know that replicated copies will be readable throughout the world on most combinations of CD players (drives), computer hardware and computer operating systems.

(3) Manufacturers of compact disc systems to write a single version of file server software as most for most computer operating systems.

(4) Mastering services to validate replicated copies of compact discs using the directory and other standard information elements to verify accurate placement of data on the disc.

While NISO were naming their committee and drawing up the goals of the standard, Digital Equipment Corporation proposed their own defacto file structure standard for use on read only media. There was genuine concern that if some working standard for CD-ROM file formats was not agreed upon quickly, then the number of different CD-ROM products would proliferate and suppliers would become entrenched with their own file structures with the result of nil compatibility.

Another industry group known as the "High Sierra Group" met in the autumn of 1985. This ad hoc committee included among its members representatives from:

- Apple
- Digital Equipment Corporation
- Hitachi
- Microware/Phillips/Sony
- Microsoft
- Reference Technology
- 3M
- TMS
- Videotools
- Xebec
- Zellnick Inc.

The aim of this mix of large and small companies was to come up with a working interim standard for CD-ROM volume and file structure which the industry could use prior to NISO making a ruling. In fact, given the influence of the companies represented on the HSG it was widely believed that, provided a genuine consensus of opinion was reached, their proposal would certainly be adopted as a working standard and would in all probability be accepted as the official standard, with only minor amendments deemed necessary.

In drawing up the functional requirements for the proposed CD-ROM volume and file structure standard, the HSG contracted and received representations from a number of industry associations and committees including the American Library Association's technical standards committee and the U.S. Information Industry Association's Standards Committee. These two groups set up a CD-ROM sub-committee and put forward three main requirements:

(1) A primary functional requirement of a CD-ROM standard is the interchangeability of CD-ROM discs with CD-ROM drives interfaced with various common computers running various common operating systems.

(2) Since CD-ROM has certain inherent performance limitations in seek times and transfer rates, the CD-ROM standard should support optimal performance of the media/device combination.

(3) Development of applications for CD-ROM should not in any way be restricted or precluded by the CD-ROM standard. It is understood that the CD-ROM standard cannot ensure the success of the applications, rather it should (not) restrict their freedom nor inhibit their implementation.

The standard specifies three noted levels of interchange:

- Level 1 is restricted so MS-DOS can read it - there is stipulation that a file name shall not contain more than 8 characters and a filename extension not contain more than 3 characters;

- Level 2 is set up to be a proper subset of what the HSG believes the CD-I group will do;
- Level 3 has everything.

Since its release, the HSG working standard has been studied and commented on by a number of industry groups to determine whether it meets the functional requirements set out by the NISO and IIA. Examining the four NISO goals:

(1) provided it is adopted the HSG working paper would enable developers of software to have a uniform file environment in which to design and implement applications;

(2) provided it is adopted, the HSG working paper should enable manufacturers of CD-ROM drives to write a single version of file server software for most computer operating systems;

(3) The paper would also enable mastering services to validate replicated copies of compact discs using the directory and other standard information elements to verify accurate placement on the disc;

(4) The fourth goal of the NISO was that the standard should enable publishers and distributors of information to master a single compact disc and know that replicated copies will be readable throughout the world on most combinations of CD players (drives), computer hardware and computer operating systems. The High Sierra Group working paper, on its own, will not entirely fulfill this goal and it is doubtful whether it will ever be fully realized at the publishing environment but the paper appears to be a major step in the right direction. For this goal to be fully achievable one would also need standard hardware and software interface supports from the operating systems to the range of CD-ROM drives on

the market. A standard hardware interface from the CD-ROM drives to the computer system, as well as the requirement that publishers would have to use standard indexing and retrieval software packages that could run on all the main operating systems on all main computer hardware would also be necessary. Such software, if it were ever developed, would undoubtedly suffer from severe limitations but it would perhaps be used in certain applications where publishers simply distributed a set of files on CD-ROM and PC users simply run a utility piece of software to search the directory and download specific files from CD-ROM onto their magnetic disc for subsequent processing.

Thus, the fourth requirement should be put forth as a goal to which progress continues within the realm of realism and applicability. Looking at the database publishing application, however, where free text retrieval will be required, it appears certain that, for the short to medium term at least, publishers will have to form close arrangements with the suppliers of retrieval software. There will be a fixed relationship between the structuring of at least the index file if not the full text file on the CD-ROM and the retrieval software that has been selected for use. Inevitably, therefore, if a publisher chooses to use a software package that runs only on one operating system he will limit the potential market for his CD-ROM database.

The HSG paper's fulfillment of the IIA requirements can also be evaluated:

(1) The interchangeability of CD-ROM discs with CD-ROM drives interfaced with various common computers running various common operating systems simply reiterates the fourth goal of NISO;

From the standpoint of database publishing, CAV discs with the fast random access facilities show the most promise. Various companies have devised techniques for encoding digital data within the video signal of the videodisc so the vast storage capacity and random access facilities provided by the videodisc can be used to distribute digital databases in a number of electronic publishing applications. This use of videodisc technology is called hybrid videodiscs and will be discussed later in the chapter.

2.1.2 -- Compact Discs

Compact discs (CDs) appeared later on the market than optical videodiscs. While the videodisc was first shown in 1973 and made available commercially in 1978, CDs were shown in 1980, and released commercially in 1983. In spite of this lag, however, sales of CD players to the public have already surpassed sales of videodisc players. The CD player has been described by a number of sources as one of the most successful launches of a consumer electronic product ever, with several million players sold within two years of its introduction.

Part of the reason for this is that compact disc systems are all standardized on the format and digital recording techniques designed by Phillips and error detection and correction systems jointly developed by Phillips and Sony. The main aim of the standard is to achieve total compatibility so all CDs produced to the standard can be played on all CD players designed to meet the standard. The standard is popularly known as the 'Red Book'.

The physical processes involved in producing a compact disc are virtually identical to those described above for a videodisc except that the material recorded on the disc is in a different format and the master and replicate discs are single sided with a 12 cm or 4.72 inch diameter rather than a 12 inch diameter.

time and effort needs to be spent in planning, organizing and indexing a CD-ROM database. Any CD-ROM logical format standard must support the efforts of the publishers/data preparers and system integrators to produce effective CD-ROM publishing systems that optimize the potential of CD-ROM as a read only publishing medium. This point has been clearly taken into consideration by the High Sierra Group in preparing the working paper.

(3) It is commonly agreed by the industry that the Paper meets this requirement, of ensuring that the development of applications for CD-ROM should not in any way be restricted or precluded by the CD-ROM standard. Delegates to the European Optical Disc Forum in Luxembourg, representing the European information industry, were asked to complete a questionnaire regarding the Working Paper. The results were very positive, with no one saying they would not adopt the standard and most planning to adopt it within six months. Most also felt it would be a significant factor in opening up the market for CD-ROM publications but by no means the only factor. Other issues, including the hardware/software interface question, as well as economic and marketing issues must be resolved before the CD-ROM market reaches its full potential.

As was mentioned previously, CD-ROM is the farthest along in the standardization process. When interview program participants were polled as to whether there was hesitance or concern with compatibility on the part of the consumer because only defacto standards exist with respect to logical formatting of data on optical disks, 40% felt that such concern existed. However, considerable optimism was expressed as to the long-term effectiveness and acceptability of the standards as 100% of respondents were satisfied with the standards in their current state.

In terms of CD-interactive, Phillips and Sony have now announced plans for a third standard - The Green Book. The CD-ROM standard is the basis for CD-I. As CD-I is application specific, it attempts to define a complete system. Because it is aimed at the consumer market, CD-I has to be a simple system.

While CD-I is logically an extension of the CD-ROM standard, CD-I also represents an extension of the CD digital audio concept of a consumer product offering total compatibility of discs and players by defining both the way in which information is recorded on the disc and the equipment needed to read it.

With CD-I, therefore, as with CD, Phillips and Sony have defined not only the medium but also the equipment that will be needed to read it, the audio processes, the video process, the choice of microprocessor and the operating system so that they can achieve their goal of being able to play all CD-I discs on all CD-I players.

Work on standardizing the Write Once Read Many (WORM) medium is also proceeding. Distribution of this medium began with the introduction of 12 inch discs. Suppliers were so far advanced with commercialization of incompatible 12 inch WORMs that it was impossible to try to provide for a standard before the discs were launched. There is now some work on physical standards on the 12 inch discs but it is at an early stage.

However, with 5.25 inch discs, the bulk of the companies developing drives and media (estimated to be in excess of 50 major companies) are delaying production of drives and media because they are awaiting the outcome of the considerable efforts which are currently being made at both national and international levels to introduce physical standards for 5.25 inch WORMs.

The organizations involved included the British Standards Institution, American National Standards Institution (ANSI) and the Japanese Industrial Standards Committee. As well, the International Standards Organization (ISO), the European Computer Manufacturers Association (ECMA) and the International Electro Technical Commission are contributing to the standardization effort.

There are four main aspects to the 5.25 inch standard. The first is definitions and environment and this is agreed; the second is mechanical, physical and dimensions and this is almost complete; the third and fourth sectors on optimal characteristics and format still need considerable work.

There is some way to go before an equivalent to the Phillips/Sony Yellow Book Standard could be agreed for the 5.25 inch write once digital optical discs and drives. In the interim the lack of such a standard and the expectation among the industry that such a standard may become available in 1987, is delaying the introduction of such products.

The lack of standardization in terms of the logical format of the discs as well as the retrieval software has limited transportablility and increased the capital costs of hardware for accessing databases from different publishers.

Customers who faced expensive microcomputer purchases several years ago only to find that they had a limited selection of applications software and could not interface with other systems hesitate to invest in a new product which may have the same problems. Until standards are generally accepted, consumer confidence will suffer due to the uncertainty of the market and customer support.

Indeed, the main problem for all suppliers of hybrid videodisc systems was the lack of standards in the field. All the suppliers used different techniques to encode their data, different hardware and different error detection and correction systems with the result that there was no interchangeability of discs and players and the large commercial publishers did not feel confident that any one supplier would triumph. In addition due to the proprietary nature of the hardware and software used in the systems, prices were too high to make hybrid systems attractive to the generic publishing market. The result is that hybrid systems serve only a narrow, specialized segment of the market.

Apparently, CD-ROM is well on its way to surpassing the spectre of this problem. According to Patrick Gibbins, Managing Director of Archetype Systems Limited, a successful British systems integrating company, "... the position concerning both the physical and the logical file standards is satisfactory from our point of view. In all the discussion over CD-ROM standards, it is easy to exaggerate; the position already is far happier than it ever has been over floppy discs, or even magnetic tape."

The logical format standard will certainly facilitate the interchange of data on CD-ROM. However, it will not on its own meet all the requirements of the publishing industry, some of which may not prove obtainable by any means.

The Yellow Book is a de facto standard governing the physical aspects of CD-ROM. The High Sierra Group working paper, which assumes the existence of the Yellow Book standard, standardizes the logical format. It remains for these de facto standards to be adopted by ANSI and ISO but it is expected that they will be, in due course. Meanwhile, the industry has working standards to build on.

The next desirable development would be to see industry standard hardware and software interface supports from the operating systems to the range of CD-ROM drives on the market. As well, a standard hardware interface for the CD-ROM drives to the computer systems is desired. Until that happens, retrieval software vendors and CD-ROM data preparation companies will have to develop and supply separate software drives for each combination of CD-ROM drives and operating systems.

	Portability of Medium	Accessibility of Information	Ease of Use and User Acceptance	Ease of Integration with Existing Systems
PAPER	Paper is extremely portable for low volume applications. Further, no retrieval hardware is needed.	Paper is basically a serial medium. This makes it unsuitable for multiple users who each want to access it in different ways.	In terms of utility publishing, paper has some major limitations. For example, large, multi-volume directories are intimidating and physically and logically difficult to use.	Information held in printed form is not easily entered into computer systems. However, with the growth of personal computers in many application contexts, there is a growing requirement to take data from published sources and place it further into a form suitable for the subsequent user.
MICRO-FILM	Microform is not nearly as portable as paper as it must be viewed using specific hardware.	Traditional source document microfilm systems also force the selection of one or two sequences in which to order information.	Information contained on this medium is not searchable in full text form. As personal computers become more of an established presence, users will demand that data be fully searchable.	Microform is not easily machine-readable. Thus, users can not access selected segments of data and download them into their personal computer for subsequent processing.
ON-LINE	The increasing penetration of portable, compact computers into the personal computer market, coupled with the advance of in-car telephone technology will continue to improve the portability of this medium.	The entire on-line database can be searched using structured or unstructured full text searches.	End users need training in the search software and command systems that a particular database employs. This may be a deterrent if many different databases must be utilized.	On-line can easily be integrated into existing systems by using a modem as the interface between the telecommunications lines and the personal computer.
CD-ROM	Personal computers with CD-ROM drives possess a low degree of portability. However, when half-height 5.25 inch drives are available, knee-top workstations will likely closely follow.	CD-ROM will combine the ease of searching on-line databases and the advantages of local access such as graphics and motion video, unlimited access time, and the ability to offer extensive help facilities to the user.	It is considerably easier to install help facilities on CD-ROM than on an on-line system. The opportunity exists, as well, to include tutorials on the same disc on which data is resident.	CD-ROM offers several attractive alternatives to rekeying information from published sources for further processing.

	Ease of Updating	Retrieval Hardware	Profitability of Medium	Accessibility of Information
PAPER	Storage capacity is unlimited, but extremely bulky. Approximately 200-250 words per page.	Any type of static material can be stored on paper. Obvious limitations are animation or motion video, but capacity for high resolution images exists if good quality paper is used.	Ease of production, whether utilizing impact print, non-impact print or offset lithography, depends on the volume of material being produced and the frequency of updates required. If voluminous, with frequent updates required, entire product should be captured on database where specific segments of text and graphics can be recalled.	The cost for creating low-volume paper-based reports can be considerably more than that for placing the information on-line or on CD-ROM. The cost elements involved with paper-based reproduction are toner/ink, machine time, and staff time. These paper, production and replication costs are directly related to the number of pages produced.
MICRO-FILM	Large scale systems may save 95%-98% of the space occupied by equivalent paper filing systems.	Ability to store text, graphics and tonal material. Again limited to static material. Special film required for colour representation. Reproduction comparable to that achieved through photocopying.	Production can be fast in in-house uses where originals are in paper form. High speed rotary cameras can film thousands of documents per hour. However, documents must be sorted into a desired order before filming. Also, prior to filming, film must be processed before it can be used.	Roll film is ideal for in-house use where large numbers of documents have to be filmed and only one or two copies is required. For publishing purposes, where hundreds of copies of larger reports are needed, microfiche is the ideal medium. Combined production and replication costs are far lower than with paper.
ON-LINE	Theoretically, no limit to storage capacity. Realisticaily, limited to the number of magnetic disc drives which may be controlled by one host, while still providing realistic response times.	Most on-line databases are currently limited to text and numeric data. Some incorporate low-resolution, digitized photographs, but these are not effective unless using high-speed transmission and high resolution terminal screens.	Producing an on-line database is a complex exercise, involving data capture, database design, indexing and loading.	Designing and creating an on-line database is a complex task and represents a considerable overhead cost to the in-house or commercial publisher. However, system integrators, software suppliers, and 'hosts' can be retained to aid in the development and distribution process.
CR-ROM	Capacity to store 550 megabytes of information. Equivalent to 250,000 pages of coded text or 10,000 scanned images. Twice as compact as roll microfilm for the storage of images.	At present, CD-ROM is primarily used to store text and computer graphics. CD-ROM and other CD systems can also store audio information, raster-scanned images, video and full motion video. It is the most versatile of standaione database media in terms of material it can contain.	Production of CD-ROM-based databases is more complex than that of on-line as developers work out their own proprietary file structures. The adoption of the High Sierra Group Standard will simplify the process somewhat. Not suitable at the current time for frequent updates.	Cost for creating a database on CD-ROM ranges from $20,000 to $220,000 depending on amount of data capture necessary, size of the database, and complexity of its structure. Fixed costs include mastering of $4,000 - $6,500 per ROM and replication of $10 to $60 per disc, depending on volume.

	Cost of Distribution	Turnaround Times	Ease of Updating	Retrieval Hardware
PAPER	The greater the number of pages that exist, the more expensive it will be to distribute them, respecting both staff handling costs and postal and transport rates.	Paper is extremely flexible with respect to low volume applications. However, as volume increases and the numbers of copies required goes up, the process of printing, collating and dispatching the material lengthens.	The greater the volume of data contained on paper, the greater the necessity of storing the data in some sort of database format in order to make the process of formatting and repagination a more efficient one.	No retrieval hardware is necessary in order to access data contained on paper. Thus, no compatibility issues arise.
MICRO-FILM	Ten COM fiche, or the equivalent of 2,700 typed pages, can be distributed for the price of a first class stamp.	Turnaround is again a function of the size of the original and the number of copies required. Replication can occur much more rapidly than print, with production rates of 2,000 to 3,000 microfiche per hour, each fiche containing 270 frames of information.	Microfilm is not an efficient means of disseminating highly volatile information. Once produced, microfilm can not be physically annotated. All material, both new and old, must be refilmed. With COM, it is possible to update a large database by printing and indexing the raw material.	Microform users need at least an optical reader. Current models are microprocessor controlled, providing fast and accurate retrieval.
ON-LINE	By using a host, the information provider may bypass distribution concerns. The host puts up the database on his mainframe for access by all of his existing and potential clients.	A lengthy turnaround between the decision to set up a large on-line database system and the actual implementation. However, once the system is up, then it can be updated relatively easily.	Ease of updating is one of the main advantages of on-line storage. Once the database is created, data can be updated in seconds - even while the database is in use.	A terminal and a modem to provide access to a telephone line are the bare essentials.
CD-ROM	As database creation and mastering are one-time costs, the economics of CD-ROM become more attractive the more potential users there are. CD-ROM is geared to distributing large numbers of a medium containing large size database to end users.	CD-ROM can not compete with any of the other media in terms of turnaround times. This is due to a production cycle entailing the updating of the database, sending the tapes to the mastering facility, producing a check disc, stamping replicates and then sending them to the end user or publisher.	The only means by which to update CD-ROM is to reissue. Current turnaround intervals sit at one month.	CD-ROM may become just another add-on peripheral to the thousands of personal computers existing in the home and in the workplace today.

4.0 TYPES OF APPLICATIONS AND SERVICES

The potential market segments which may be addressed by optical publishing technology appear to be quite diverse. This situation was illustrated in our interviews with a number of players in the North American industry. When queried generally on their perception of the greatest opportunities offered the optical publishing industry and more specifically on the markets or applications in North America which hold the most potential for optical publishers, a number of different perspectives emerged including:

- the field is still wide open - it is not possible to limit markets at this time;
- general reference, education;
- entertainment;
- medical, financial, insurance;
- consumer products/educational institutes - encyclopedias, translation dictionaries, medial applications;
- business-oriented applications - building codes, maps, accounting, legal and marketing data;
- images tied with data, i.e. parts catalogues, reference manuals;
- anyone with a critical need for information.

Keeping in mind that this broad range of opinions has emerged from members of the same industry, optical publishing is definitely showing its age - infancy at best.

In an interview published in the "Electronic and Optical Publishing Review" in December, 1986, Patrick Gibbins, the Managing Director of Archetype Systems Ltd., said that "In terms of market, however, (CD-ROM and optical publishing in general) is a somewhat amorphous product -- it's a bit like trying to work out 'What is the market for pencils?' The marketplace for which we are aiming is essentially that of publishing - by which we encompass both commercial publishing and the broad range of 'in-house' publishing."

Tony Hendley, however, in his book CD-ROM and Optical Publishing Systems is somewhat more definitive in his view of the potential application of CD-ROM and optical-based publishing. His segmentation of the market is as follows:

- The Professional Market
  - In-House Publishing
  - Technical Publishing
  - Commercial Publishing
    - Software Distributing
    - Graphic Databases
    - Database Publishing
    - Specialized Information Publishing
    - Reference Works
    - Backfile/Collection Distribution
- The Library Market
  - Library Automation Services
  - Database and Reference Materials
- The Educational Market
- The Domestic Consumer Market

This framework will be suitable for organizing the balance of this chapter.

4.1 The Professional Market

4.1.1 -- In-House Publishing

In-house or corporate publishing systems are systems for capturing, organizing, storing and disseminating information within a company or public sector body. Typically, these tasks are accomplished through the use of word processing systems and/or in-house reprographic units or commercial print shops to produce and print reports, telephone directories, sets of statistics, minutes of meetings, technical reports, and sales forecasts. Paper is currently the most popular medium for distributing this in-house information; many organizations also use computer output microfiche and on-line electronic distribution via PC networks and private viewdata systems. More and more of this information is being produced with word processors or microcomputers.

In addition, with desktop publishing systems being introduced which can handle text and graphics and enable users to carry out sophisticated page make-up functions, there will be increasing pressure to store and disseminate the resulting internal publications electronically than print them out and incur heavy consumable costs.

What has in fact happened, over the years, in the commercial publishing and library area, is that electronic information services have become centralized. The use of CD-ROM will back this trend attempting to convince users to return to the

local delivery of physical packages of information. With in-house publishing systems, there will be conflict between those who advocate putting all corporate information up on the mainframe and making it available to users via PCs linked on a network and those who advocate that individual departments should download their information onto CD-ROM and distribute it to those users who really need it. The advantages of using CD-ROM is that there will be no contention and users can spend as long as they like accessing the CD-ROM database. As well, the user may have access to the entire 550 megabytes instead of just the portion which they have called for and stored locally in the magnetic disc buffer. The retrieval software used for specific CD-ROMs can be geared specifically to the needs of the users and the data provided on CD-ROMs can be packaged according to the requirements of the specific targeted users.

There are, however, some major disadvantages of using CD-ROM with in-house applications such as this:

(1) The main disadvantage of CD-ROM in this application is the turnaround time. Corporate users will have to send data outside to get discs mastered. It is possible to have a turnaround time of as small as three days but this service carries high premium as normal turnaround times for small runs is two weeks. In addition, there is currently a chronic shortage of CD-ROM mastering and replication facilities.

This disadvantage has been partly overcome with a system developed by a company called Tallgrass. The system, called Lightfile, will be discussed below.

(2) CD-ROM-based in-house publishing systems also have implications for corporate staff responsible for handling this type of data. Certain skills and resources are needed to convert a database into a form where it can be easily updated and reviewed on a quarterly or monthly cycle on CD-ROM. The deadlines imposed by the CD-ROM production cycle will impose stricter discipline on the staff responsible for creating, amending and updating the corporate databases than was previously the case with printed or microform publications. There will be a need for some re-assessment of the ways in which databases are structured in order to make them as flexible and as easily updatable as possible. Most users today still have their information in page or image form. The investment required to convert it into a versatile database format is proving a major disincentive to their adoption of CD-ROM publishing solutions.

(3) There is a high one-time cost of producing the master CD-ROM which would have to be absorbed over what could be a relatively low number of copies of the discs in many applications. Again, this figure will drop as volumes increase and more mastering and replication facilities come on stream.

(4) There is a high one-time cost of CD-ROM drives. Presently drives from DEC and Reference Technology are priced at approximately $500 to $2,000. If high resolution terminals are required for image display then the cost of the workstation rises significantly.

(5) Lastly, there is the timescale and cost involved in implementing an in-house CD-ROM publishing system. As the

> major companies move into this area, procedures may be simplified and costs will come down. At present, however, according to Pete Rudnisky of Digital Equipment Corporation, the installation of a medium-size in-house CD-ROM distribution system could cost close to $500,000.

In spite of the various disadvantages of CD-ROM in this application, a number of major office automation, computer and reprographic companies are evaluating where CD-ROM can fit into future in-house electronic publishing systems.

Apparently, Digital Equipment Corporation is taking the lead in this market, offering their existing customers the opportunity to experiment with CD-ROM-based publishing systems in-house.

As well, Reference Technology in the U.S. have recently announced a teaming agreement with Xerox Corporation to market CD-ROM systems for in-house information distribution. According to Stephen Snyder of Reference Technology "CD-ROM applications are rapidly growing in popularity among government agencies, especially with the U.S. Department of Defence. Distribution of technical documentation and print-on-demand systems are examples of personal computer-based applications where these agencies can accomplish their data management missions at significant cost savings."

A number of interesting products are in the prototype or market introduction stage. For example, Laserdata has developed a prototype IBM PC compatible electronic page imaging system called Laserview. There are two main elements of the system: a filing system and a retrieval workstation.

The filing system consists of:

- an IBM PC XT, AT or compatible
- a 300 dot per inch scanner
- digital compression and decompression hardware
- a high resolution MS-DOS compatible monitor
- a laser printer which supports standard printer output in addition to hardcopy of the page images
- write-once optical discs for storing the page image databases
- a complete package of indexing and retrieval software; used for data pre-mastering if the database being created on the filing system is to be distributed on CD-ROM or network interface hardware if the page images are to be networked.

In short the filing system is a combination production and retrieval station which allows the user to develop databases of tens of thousands of page images.

The second part, or the retrieval workstation, is intended for applications which require distribution, viewing and printing of electronic page images. It can accept networked data from the filing system or the page images can be accessed from the write-once optical discs or CD-ROMs. The workstation comprises an IBC PC bitted with decompression monitor control and printer control hardware, a high resolution monitor capable of displaying an entire 8.5 x 11 inch page at once, a table top laser printer and a CD-ROM drive or networking hardware. The Laserview filing system is priced at under $40,000 in the U.S. and the retrieval workstation at under $15,000.

Another WORM-based system which is actually on the market is a product called Lightfile made by Tallgrass. Although not as integrated a system as Laserview, Lightfile connects to a

card in an open slot in an IBM PC or compatible. The drive has a capacity ranging from 200 MB to 800 MB.

The Lightfile system is targetted at anyone who collects data which they do not want erased. In other words, doctors, lawyers, engineers, or anyone with more than 100 floppy discs. The lightfile system, with 800 MB in storage capacity sells for nearly $22,000.

Two examples of agencies actually adopting this technology follow. The Nuclear Regulatory Commission, Division of Waste Management, has developed an on-line records management system providing for full text search of ASCII text as well as retrieval of the original image of the document on a single high resolution monitor. The system utilizes a WORM drive as well as an optical character reader.

Additionally, the Government of British Columbia is experimenting with the idea of distributing sets of government data such as statistics, for example, to various departments on CD-ROM.

These systems aim to offer the ability to create mixed mode image and coded text documents, store them on write-once optical discs locally and then make them available to remote users either via a network system or by serially copying the write once optical discs and distributing them to users with retrieval workstations, or if there are enough users, mastering a CD-ROM and distributing copies to end users. Most systems in existence today operate on a small scale but serve to demonstrate the potential options open to users of in-house electronic publishing systems in the future.

4.1.2 -- Technical Publishing

Technical publishing systems and in-house publishing systems are fundamentally similar. They both involve the capture, manipulation, storage, and dissemination of information. However, with respect to technical publishing systems, the information is typically technical documentation such as parts lists, instruction manuals, specifications, book order lists, standards, pricing information, maintenance schedules, timetables, and statistics. The most significant difference between technical publishing and in-house publishing systems, however, is that usually the information has to be distributed outside the originating organization to a controlled set of users such as clients, suppliers, dealers, distributors, retailers or branch offices.

One specific example of this application is where a car manufacturer produces and sends out parts catalogues and price lists on microfiche on a monthly or quarterly basis to their many thousands of dealers. Therefore, technical publishing is closer to traditional commercial publishing than in-house publishing in that the technical publisher is usually aiming at an external audience. However, that audience, in the case of the example of dealers, is usually a captive one and can be given an instruction to purchase the necessary hardware and software needed to access the published information.

A prototype of a CD-ROM-based repair manual is now in existence. This manual is intended for the use of airline

mechanics who must closely follow FAA guidelines in any repair procedure. The multitude of parts that make up a modern jet, coupled with the complicated procedures for installing some of them make optical media ideal for this application. The vast textual information can be efficiently indexed and the numerous graphic data can be compressed to maximize disk usage and minimize access time. The primary input device, as it now stands, is a keyboard with a number of programmable keys or 'macros'. Future modifications might include a pointing device which would prove useful in stepping through illustrated parts breakdowns. A voice recognition system that could respond to commands such as "zoom", "scroll", or "next" would also be desirable.

CD-ROM, however, can not be justified for this type of application with cost considerations as the only point of comparison. However, an optically based parts catalogue, for example, would be much easier to use than the existing microfiche based publications. Optically based manuals can be searched on a number of keyword types (i.e. model numbers, part number, part name) while microfiche documents tend to be more static in their accessibility. The CD-ROM system could also easily be used in conjunction with other software which would allow inventory tracking as well as actual parts ordering. Lastly, CD-ROM based publications and the reader needed to access them can be integrated into existing computer based systems.

Another technical publishing application whose CD-ROM systems may complement or replace computer output microfiche is in the distribution of book ordering lists.

For example, the Ingram Book Company offers an electronic book identification and ordering system called Lasersearch.

This system allows the user to search over a million English language book titles by author, title, ISBN, publisher, LC number, or keyword. An order can then be transmitted electronically to Ingram's in-house computer. The reader and CD-ROM disc are offered at the subscription price of $795 per year.

4.1.3 -- Commercial Publishing

Commercial publishing can be defined as an application in which a publisher identifies the market for a particular set of information, organizes that information, and then makes it available for a fee. Commercial publishing can be broken into a series of smaller categories which include:

- Software Distribution
- Graphic Data Bases
- Data Base Publishing
- Specialized Information Services
- Reference Works
- Backfiles of Serials

4.1.3.1 -- Software Distribution

One of the most attractive applications for CD-ROM will be as a software distribution medium. As today's software becomes more and more complex, the manuals and other documentation are becoming tremendously unwieldy. CD-ROM could be used to distribute the operating system, the application software, the user manuals, tutorials and all other information related.

A number of companies are offering software libraries on CD-ROM for PC users. Reference technology has issued a CD-ROM disc containing a library of software for the IBM PC. In this manner, an organization's entire software base could

be held in the palm of the hand. Since CD-ROM drives are interfaceable with micro-computers, the software desired at a certain time can be downloaded onto the user's hard disk, thus overcoming CD-ROM's characteristically slow access time. At the microcomputer end of the market, however, it is not likely that optical media will be ideal for distributing pure applications software as most packages can be held on floppy discs. However, when several popular packages are combined on a disc, together with the documentation and an audio training program, the concept becomes much more attractive.

4.1.3.2 -- Graphic Data Bases

Although there are still problems relating to the storage of graphics on CD-ROM, their resolution will allow a number of interesting applications.

One potential use of optical media in the graphics arena is in the distribution of font libraries to printers and publishers with their own desktop publishing systems. As thousands of typefaces exist, they may be scanned and stored in memory so that they may be retrieved for use at the laser printing stage. The amount of memory consumed by one high resolution piece of type is considerable and only CD-ROM would allow large type font libraries to be offered in electronic publishing systems. Large in-house uses of electronic publishing systems could build up their own libraries of unique scanned graphics, including logos, scientific symbols, chemical structures, architectural symbols, and computer flowchart symbols through a system utilizing a scanner, a piece of facilitating software, and a WORM disc.

Another very interesting graphics application area for optical media would be map information. A map data base could be

used for computer assisted navigational systems in planes, military vehicles and cars, and for utility companies who could then overlay their pipes or telephone lines and issue their service engineers with CD-ROM drives and monitors so that they could locate broken water mains, etc.

Phillips has developed a computer assisted retrieval and navigation system using CD-ROM based digitized maps which they expect to be utilized in luxury cars by 1990. The system consists of a combined CD-ROM /CD-Audio player, an on-board microcomputer, and a monitor. The goal of the system is to reduce fuel consumption, decrease travel time, and link car functions with road and weather conditions. The driver enters his intended destination into the computer and then road information on CD-ROM is compared with the car's actual speed and route. Route information is given to the driver through a speech synthesis unit which will also give instructions for correction if a wrong turn is made. Further, the European Radio Data System broadcasts digital traffic information. This data would be captured by the onboard system to revise routes in the event of traffic jams.

As well, the first digital world atlas on a CD-ROM has been developed by DeLorme Mapping Systems. The first edition of the world atlas contains political boundaries, roads, cities, rivers, lakes, islands, land elevations and worldwide ocean depths. The monitor map scale is about twenty miles to the inch for most of the world, although the U.S.A. and Europe have been mapped to a scale of one inch equals 400 feet.

There are currently sixteen levels of zoom controlled by the database. The user simply places the cursor over the area of

interest and presses the mouse button. The location of the last mouse click becomes the new map centre. Another method of data base access is to type the latitude and longitude along with the scale desired and the location of these coordinates will appear centred on the screen.

A last example in this exciting application area is the Conquest System developed by Donnelley Marketing, a Division of Dun and Bradstreet. This system superimposes geo-demographic market research data over maps of varying scale and detail, allowing users to pinpoint their existing and potential markets from a geogrpahic perspective.

4.1.3.3 -- Database Publishing

Online database publishing is a full-fledged commercial publishing activity. It is also one of the first areas of the publishing industry to take the potential of optical publishing seriously.

There are four possible impacts which CD-ROM can have in the online database publishing industry. These are:

(1) CD-ROM can be used as an alternative to an existing medium in an established application, i.e. a database is published on CD-ROM rather than made available online;

(2) CD-ROM can be used alongside an existing system in an established application, i.e. a data base is published on CD-ROM and made available online. At best, this significantly expands the market for the data base and at worst, it will fragment the market and simply increase the overhead of the online publishing industry;

(3) CD-ROM can be used to create new applications combining sections of existing databases or distributing new data bases; or

(4) It may not be used at all.

When the concept of data storage on CD-Audio disks was introduced, and thoughts centred around its application in the database publishing area, numerous advantages which the optical media might have over online distribution emerged.

The first was that the cost of online services kept many potential users from ever using the database. Optical storage would allow the distribution of the database on a subscription basis with unlimited use after that point.

Second, since there is no need to log-on or use modems to access a CD-ROM based database, and because CD-ROMs offered vast storage capabilities, database providers had an opportunity to make CD-ROM versions of their data base extremely easy to use with help facilities and multi-experience level search software.

Lastly, data bases could be packaged to meet the specific needs of users. For example, instead of one large legal database, there could be numerous sub-sets of the legal database aimed at specific types of practises.

Fueled by the initial optimism, numerous companies hurried into production of optically-based versions of their online databases without really harnessing the true power of the media. The necessary standards had not yet been agreed to and little marketing research had been done to see what users wanted, or were prepared to pay for.

Database providers found themselves in a real dilemma when it came to pricing CD-ROM products. In order to create a wider market among end users and smaller companies, they desired to

keep the price low. On the other hand, if the entire database is available on CD-ROM at low cost, online income could be eroded substantially.

Following the initial enthusiasm, more and more firms are taking a more realistic view of the potential for these products. Some companies, like Digital Equipment Corporation, have decided to concentrate on data preparation services rather than to actively market data bases. Seventy-five percent of the representatives contacted during our interviews with the online industry did not have plans to introduce any database services on CD-ROM in the short-term, 37% felt that they might consider using it in the future. Those who had no interest in CD-ROM often stated that they did not feel that the technology had yet proven its value.

Several success stories emerged, however. More and more information providers are experimenting with CD-ROM and are announcing systems which are approaching what users are prepared to pay for. There are still some key issues which will have to be addressed before CD-ROM finds its real niche in the online industry.

First, CD-ROM is ideal where telecommunications do not exist. CD-ROM can be used to distribute selected databases in the third world and other remote locations where there are no telecommunications links. Many of the users will now be able to access large bibliographic databases via a computer system rather than relying entirely on microfilm or printed bibliographies. The Commonwealth Agricultural Bureau is a major provider of agricultural databases to a worldwide audience who will likely be a prime example of this application.

Second, combined CD-ROM/online terminals are required. A number of database providers have shown terminals where users can use the same software to access databases online or locally via CD-ROM drives. In many applications, it is desirable to not only hold large static databases locally, but

also to access the current database online to ensure that the most up-to-date information is accessed. H. W. Wilson has developed such a PC-based workstation. Such a combined terminal would offer database providers, hosts and users with the potential to develop or use hybrid systems which make the best use of the two delivery alternatives. Such terminals could also begin to solve the pricing dilemma faced by these organizations.

Third, if CD-ROM is to open up new markets for databases, then CD-ROM must be aimed at end users rather than intermediaries like librarians.

To achieve this, CD-ROM databases must have good help facilities, be menu driven, and, if possible, artificial intelligence must be used to guide the user around the data base.

As well, it will not be sufficient to simply make one year of an existing database available on one CD-ROM. Subsets from a number of data bases can be published together on one CD-ROM with common search software. These packaged database products would be targeted at the needs of a particular professional.

4.1.3.4 -- Specialized Information Services

Specialized information services can be seen as an attempt to target specific groups of professionals, analyze their information requirements and produce specialized libraries comprising all the reference type information that these professionals need to carry out their work efficiently.

Some of the most promising application areas for specialized information services would appear to be medical, legal, financial, scientific and technical areas where workers all need access to vast stores of recorded information. In the future these and other new publishers will begin to transfer firstly the index databases to CD-ROM and eventually, much of the full text information itself so that users will be able to search both the index and full text information on CD-ROM via low cost computer terminals.

If the installed base of CD-ROM drives increases and as costs drop due to the influence of standards and volume sales, the number of specialized information services will increase. They will become increasingly more sophisticated and specialized until it becomes economically viable to produce specialized information services on CD-ROM that sell one hundred or less copies.

4.1.3.5 -- Reference Works

Reference works are simply more widely used information services such as existing published directories, general encyclopedias, dictionaries, telephone books, travel guides and government statistics.

Although this market overlaps the library and consumer marketplace, the emphasis for the professional market will be increasing the functionality of the directory.

For example, Silver Platter has placed the British Post Office's entire file of 23.5 million postal coded addresses on to a single CD-ROM which is selling for one-fifth of the

price for the previous microfiche and magnetic tape-based products. The address file can now be used to find any private or business address in the United Kingdom within segments. A Canadian company has developed a CD-ROM package which contains the addresses of 500,000 Canadian businesses on one disk and seven million household addresses. The file can be accessed via name, telephone number, or postal code, with the ability to print mailing labels from the file. This product will have a range of applications in the direct-mail market.

In the Government sector, many economic statistics, census data, or consumer survey data may be made available at a premium on CD-ROM so that marketing departments and researchers could process the data and incorporate it into modeling systems. The Government of British Columbia is currently pursuing this avenue.

4.1.3.6 -- Backfile/Collection Distribution

The former five application areas within the commercial publishing sphere, relate to cases where the information distributed on CD-ROM would be actively used and processed by the end users and hence the CD-ROM product would have potentially considerably greater value than the equivalent printed product.

This last application area refers to a specific area which consists of publishing of backfiles of journals and newspapers and the publication of collections of government reports.

Examples include the popular backfile copies of U.S. Government reports distributed on microfiche by the National Technical Information Service, and the Educational Resource Information Center (ERIC).

Microfiche is often used as a means of distributing facsimile images of the pages of services and reports usually purchased

by company information centres to service requests for specific articles and reports. Typically, the indexes to the material are available online so users can conduct a search, identify the specific articles they are in need of, and request copies of those articles. To satisfy the request the library retrieves the microform, prints out the article or report using a reader-printer or copies the microform if it is a large report and sends the printout or replicate microform to the user. In large information centres, and in centralized document supply centres where demand is high the use of two or three systems is cumbersome and time-consuming. Hence publishers and document supply centres are experimenting with new techniques for distributing images of articles and reports electronically either on CD-ROM or via facsimile networks.

Although many plans are in place and a few prototype systems exist, there are several technical problems to be overcome before large collections of images can be distributed commercially on CD-ROM but it appears that the publishers and document supply centres see sufficient advantages in such a system to invest considerable amounts in their time development.

For the short to medium term, at least, professional applications will constitute by far the largest market for CD-ROM systems because these are the applications where productivity gains can be easily measured and paid for and where the necessary base of computer expertise lies.

4.1.4 -- The Library Market

The library market was one of the first potential market areas to be targeted by the optical publishing industry.

A number of CD-ROM-based products which are aimed at libraries currently exist.

4.1.4.1 -- Library Automation Systems

One of the first and most obvious applications for CD-ROM systems in the library market was the distribution of bibliographic databases. The Library Corporation in the U.S. have produced a catalogue production system called 'Bibliofile'. This is based around a CD-ROM system connected to an IBM PC. Over 1.4 million MARC records are stored on two CD-ROMs containing all of the Library of Congress's English language catalogue entries since 1964 plus popular titles since 1900. Libraries can use the system for retrospective catalogue conversion, the creation of MARC tapes and the printing of catalogue cards and labels.

In terms of automated book ordering systems, the Ingram Book Company were one of the first to show the previously described Lasersearch System. The system was aimed primarily at bookstores but the concept is equally applicable to libraries. Other systems which have been recently introduced or are soon to come out are Bowker's Books in Print as well as Whitakers British Books in Print.

4.1.4.2 -- Data Base and Reference Material on CD-ROM

Traditionally, there are three categories of libraries - public, academic and special. Special libraries refer to company libraries and information centres, libraries in public sector organizations including government departments, libraries in research establishments and professional associations and practices, and lastly, large national libraries.

Cimtech feels that special libraries are likely to represent the most lucrative market for CD-ROM databases, just as they are the heaviest users of online databases. These will be followed by academic libraries and finally by public libraries.

There are, however, a number of concerns which must be addressed before the CD-ROM technology is embraced by the library community.

The first of those concerns is that of standardization. Standards issues include: the machine or hardware connections at the physical level; data formats for location and retrieval of information and software from the disc at the logical level; and consistency in search procedures at the applications format (CD-ROM, CD-I, digital videodisc, write-once, and erasable) are at different stages of development or are non-existent. Librarians want the assurance that within a specific optional technology, any disk will play on any manufacturer's player and moreover that the content of any disk is retrievable on any manufacturer's player. Librarians are not concerned with the internal workings of applications software, but rather the ease of use and consistency across systems to aid the user of multiple systems and databases.

Reliability and ongoing maintenance of equipment must be addressed by vendors for librarians to have confidence in these new products since access to the information is hardware-dependent. The ability to upgrade workstations as systems require more storage and processing power is also of concern, as rapid obsolescence of equipment will discourage large capital investments in systems.

Another issue relates to the type of material which will be required by the various library markets. Clearly, many of the special interest libraries will have interests aligned with the professional association or research body or company which they are designed to serve. Such libraries typically use two or three core data bases online and a number of others on an infrequent basis. In such libraries, a trend might be towards purchasing all or sub-sets of the core data-bases on CD-ROM in future, and use them on combined online, CD-ROM terminals so that, if required, they can also go online to other data bases or to check the latest version of the core databases.

If many of the predictions made about the impact of optical publishing over the next five to ten years are true, then the impact of special libraries is likely to diminish. This is due to the fact that publishers are tending away from selling their product to intermediaries and toward marketing it to end users by making them specialized and easy to use. In certain applications such as the financial and legal areas and certain of the engineering disciplines where the data is structured and requests can be easily formulated, the trend toward individual knowledge support is irreversible.

4.1.5 -- The Educational Market

There are a number of different application areas for optical publishing which fall within the educational sphere.

Applications for CD-ROM publication in the administrative area of education will include catalogues of audio visual aids and educational products for distribution to education authorities, distribution of educational databases, distri-bution of educational statistics, databases on courses at university, and careers information databases.

CD-ROM also has an application in generating test questions. Tescor Inc. have announced a CD-ROM system called the national item bank and Test Development System which aims to enable teachers and educational administrators to create individualized tests, print them out and score and analyze the results. The disc includes 40,000 test questions that have been used and validated by teachers in the past. The teacher specifies a particular test and the level of the pupils. The software searches the disc and selects test items that match the request.

Another company, Quantum Access Inc., are developing a CD-ROM database product containing all the legislative and regulatory material that a Texas school administrator is likely to need.

One of the major barriers for optical based educational courseware for use in primary and junior schools is the problem that CD-ROM and CD-I-based programs will largely be suited for highly interactive individual instruction rather than classroom teaching. The type of courseware that could prove attractive would include educational games whose responses from the class or a small group would lead to visual responses on the screen and, as an audio visual aid, slide sequences with audio commentary and questions and answers which again lead to different slides being shown and different audio commentary.

The secondary school environment is more conducive to individualized instruction or programs geared to small groups. These systems could be installed in libraries, audio-visual laboratories, and in computing departments. Educational games and modeling systems may be attractive in this area while, in the library environment, encyclopedias and phonetic dictionaries will likely be feasible.

Acceptance of optically-based technology in the educational marketplace will likely depend on the success that CD-ROM and CD-I have in the consumer marketplace. With the low cost of PCs, it may come to the point where pupils can receive tuition at school based around an optically-based product and then go home and use the same product to help them with homework.

In the higher educational sector, the potential for optical products grows even further, particularly in areas where training courses are expensive and involve the use of expensive hardware and consumables, i.e. medicine, chemistry, electronic and mechanical engineering. In these environments, it will prove cost effective to use simulated systems where the students are motivated to use self-instructional learning materials. The visual content in many training programs is crucial so that for any headway to be made in this area, combination CD-ROM and videodisc systems will likely have to be used.

Further, trainee accountants, lawyers, doctors and engineers will have their own personal computers at home. Coursewares, textbooks and specialized libraries will be made available on CD-ROM so that studying can be carried out at home without the need to visit libraries and be tied by their opening hours.

4.1.6 -- The Domestic Market

Since their launch, compact disc players have sold five million units worldwide according to Cimtech. This still represents only one-sixtieth of the three hundred million record

players currently installed in the world and an even smaller percentage of the number of audio cassette recorder/players in the world today. In the U.K., CD players are said to have achieved the penetration of the marketplace, 1% short of the 5% needed for a product to be recorded as established.

As discussed earlier, a growing trend exists for students and executives to take work home and use a PC at home to do that work. This work may entail the use of CD-ROM-based databases, reference works, and specialized information products. Domestic PC users will also want to utilize their CD-ROM players for more than work, however. The extra storage capacity offered by CDs will enable games designers to produce high quality graphics, stereo sound and animated sequences which will bolster sagging demand for games products.

The self-instructional area is also a promising one, with low cost audio-based training systems and more expensive encyclopedias and domestic dictionaries offering the greatest potential. At this point, the only encyclopedia available on CD-ROM is the Grolier Academic American Encyclopedia. The version currently available is restricted to textual information only. Instant access to every occurrence of a word, topic or phrase in the entire encyclopedia can be provided in an average search time of 5 to 15 seconds.

Cimtech feels that there is much less of a market for CD-ROM and CD-I in the home than in the professional marketplace. The most promising products for CD-ROM in the home mirror the professional uses so that the home market will likely be mainly restricted to students and professionals.

4.2 A View to the Future

It is still, at this point in time, extremely difficult to predict the full impact that optical publishing will have on the standalone database industry and the publishing industry in general. A number of perceived threats to optical media emerged in our discussions with the industry:

- the growth of CD-ROM may be threatened by WORM and ever more flexible media which will allow users to carry out their own publishing in-house;
- the absence of user interface standards as well as the slow development of standards in all areas may mitigate the success of the media;
- the initial product pricing of various CD-ROM products by publishers may stymie growth;
- increased hard fixed disk capacity may lessen the attractiveness of optical media in storage applications;
- the rapid evolution of new technologies (i.e. CD-Interactive, WORM) and hardware advances encourages a 'wait and see' attitude among consumers.

All of the above issues will affect, to a certain extent, the progress which optical publishing media will achieve in each of the previously mentioned application areas. Cimtech has ventured a number of predictions as to the penetration optical media will eventually achieve:

- With respect to software distribution and graphic library distribution, CD-ROM will be targetted at single user PC systems while OROM, for example will better serve the needs of more powerful computer systems.
- In terms of commercial database publishing as well as the provision of specialized information services, two key issues must be addressed as a prerequisite for success. With respect to database publishing, provision of information on CD-ROM must be positioned so as not to erode online services. With respect to specialized information services, the issue is ease of use and accessibility.

- No short term success is seen for CD-ROM in the distribution of backfiles or collections due to the cost-effectiveness of microfilm;

- In terms of the library market, a likely successful product will be bibliographic databases for staff use. For end users, or library patrons, however, several issues remain unresolved:

  - the need for the investment in large numbers of workstations

  - the necessity for supervision of hardware and software, and

  - the high subscription cost to CD-ROM products.

- With respect to the educational and domestic consumer segments, the market is too fragmented to even guess at likely outcomes. One prediction that can be made with a fair degree of certainty will be that the media will have to prove itself in the professional marketplace before making significant in-roads in these segments.

Cimtech poses this interesting question - How quickly will the momentum for optical publishing media build up? Will the entry cost and the cost of hardware seriously restrict the market for the media, or will success in various application areas increase steadily over the years until it becomes a generic publishing medium and a generally accepted alternative to paper or microfilm?

The best guess is that CD-ROM will prove itself to be a long term publishing medium for PC users but it will be bounded by the demand from mainstream computer users for higher performance multi-user systems on one side, and on the other, the demand from schools, libraries, and domestic users for lower cost hardware and software. The first demand may be met in the long-term by OROM and the latter by dedicated systems such a CD-I.

Various market research firms have predicted that by 1990, we could see anywhere from one million to twenty-two million CD-ROM/CD-I drives installed. Freeman Associates predicts that by 1991, 440,000 CD-ROM drives will sell each year at a market value of $154 million. The market for WORM disk drives is, however, predicted at $1.488 billion in 1991.

In terms of the CD-ROM media, the market is predicted to reach half of that of CD-audio, which is predicted to reach $800 million by the early 1900's.

In 1984, the market for data storage was estimated at fifteen billion dollars with optical disc systems capturing $105 million or 0.7%. One source put the value of the overall market at $31 billion for 1988, with optical discs capturing 7.4% or $2.204 billion. When polled on the solidity of this prediction, our interview program participants unanimously disagreed. Their readjustments of the prediction ranged from: at most pessimistic, optical products capturing only 1% of the market; at middle-of-the-road, saying that the projection is twelve to eighteen months ahead of actual development; to most optimistic, putting the revised market capture at fifteen percent.

5.0 CANADIAN AND INTERNATIONAL INDUSTRIAL PLAYERS

This section outlines the major Canadian and international players in the sphere of optical publishing. Companies classified as European-based may have subsidiaries or branch offices in North America, and more specifically, Canada, and thus their products are available on this continent as well.

5.1 Optical-Drive Suppliers

A variety of drives capable of reading optical media are currently on the market. The manufacturers, the model number of their product, and an address where they may be contacted are listed below:

- Denon American
  DRD 550 CD-ROM Drive
  27 Law Drive, Fairfield, N.D. 07006 U.S.A.

- Hitachi Ltd.
  Hitachi New Media Products - CDR 150ZS, 250C, 2500S
  Hitachi House, Hayes, Middlesex, UB3 4DR U.K.

- JVC Corporation
  CD-ROM Drives currently under development - model names not available
  41 Slater Drive, Elmwood Park, N.D. 07407 U.S.A.

- Panasonic
  Panasonic (Matsushita) - SQD-100
  1 Panasonic Way, Secaucus, N.J. 07094 U.S.A

- Phillips Electronics
  Phillips - CDX-1000, CM100, CM110
  City House, 4201430 London Road, Croyden, Surrey, CR9 3QR

- Laser Magnetic Storage (Subsidiary of Phillips and Control Data Corp.)
  Model names not available
  P. O. Box 218, Building SA11-5, Eidenhoven, Netherlands.

- Sanyo Electric
  CD-ROM Drives currently under development - model names not available
  1200 West Artesia Blvd., Compton, CA 90220 U.S.A.

- Sony Corporation
  Sony - CDV100, CDV200, CDV5002
  Sony House, South Street, Staines, Middlesex, U.K.

- Toshiba
  Toshiba - XM2000

- Corel Systems Corp.
  Ottawa, Ontario
  Model Names Not Available

- Tallgrass
  Tallgrass Lightfile

5.2 Data Preparers/System Integrators

These firms are primarily involved in organizing the data according to standard formats as well as providing the access software a directory, delineating exactly where each data item may be located.

5.2.1 -- International Firms

- Archetype Systems Ltd.
  91/93 Charterhouse Street
  London, England
  EC1M 6LN

- Battelle Institut E. V. Abt Software Products
  Postback 900160
  Frankfurt, West Germany

- Battelle Institute Ltd.
  15 Hannover Square
  London, England
  W1R 9AJ

- Bertelsmann AG
  Carl Bertelsmann Street 161
  Guerterslote 05241801
  West Germany

- BRS Europe
  73-75 Mortimer Street
  London, England
  W1N 7TB

- CEDROM Technologies
  68 Quai de la Seine
  75109 Paris, France

- Digital Equipment Corporation
  CD-ROM Marketing, Engineering Division
  P. O. Box 121, Imperial Way
  Reading, Berks

- Elanders
  P. O. Box 10238
  S-4301 Kungsbodsa Sweden

- Harwell (STATUS)
  Marketing and Sales Department, A.E.R.E.
  Harwell Soloratory
  Oxfordshire, U.K.
  OX11 ORA

- Laser Magnetic Storage
  P. O. Box 218
  Building SA11-5
  Eindhoven, Netherlands

- Logica Communications and Electronic Systems Ltd.
  64 Newnon Street
  London, England
  W14 4S3

- Office Workstations Ltd. (OWL)
  2 Easter Road
  Edinburgh, Scotland

- Pergamon Infoline (Compact Solution)
  12 Vandy Street
  London, England
  EC24 2DE

- Phillips Electronics
  City House, 4201430 London Road
  Croydon, Surrey
  CR9 3QR

- Scientific Consulting
  Vollshovener Weg 172-176
  500 Koln 71
  West Germany

- Silver Platter Information Ltd.
  10 Barley Mow Passage
  Chiswick, London
  W4 4PH

- STET
  00198 Romor
  C So D'Italia
  858430, Italy

5.2.2 -- North American Firms

- Access Innovations
  P. O. Box 40130
  Albequerque, New Mexico, USA
  87196

- AMTEC Information Services
  3700 Industry Avenue
  Lakewood, California, USA
  90714-6050

- Battelle Software Products Center
  505 King Avenue
  Columbus, Ohio, USA
  43201-2693

- BRS
  120 Route 7
  Latham, New York, USA
  12110

- Computer Access Corporation
  Suite 324, 26 Brighton Street
  Belmont, MA, USA
  02178-4008

- Corel Systems Corp.
  1600 Carling Avenue
  Ottawa, Ontario
  K1Z 7M4

- Cuneiform Inc.
  11 Murphy Drive
  Nashua, NH, USA
  03062

- Digital Equipment Corporation
  2 Mt. Royal Avenue
  Marlborough, MA, USA
  01752

- Group L
  481 Carlisle Drive
  Herndon, VA, USA
  22070

- Knowledgeset Corporation
  2511 Garden Road
  Building C
  Monterey, CA, USA
  93940

- Laserdata
  One Kendall Square
  Building 200
  Cambridge, MA, USA
  02139

- Microsoft
  16011 N. E. 36th Way
  Box 97017
  Redmont, WA, USA
  98073-9717

- Microtrends Inc.
  650 Woodfield Drive
  Ste. 730
  Schaumburg, IL, USA
  60195

- NCR Corporation
  292 Madison Avenue
  New York, NY, USA
  10017

- Online Computer Systems
  20251 Century Blvd.
  Gernontown, MD, USA
  20874

- Reference Technology Inc.
  5700 Flatiron Parkway
  Boulder, CO, USA
  80301

- Reteaco Inc.
  716 Gordon Baker Road
  WILLOWDALE, Ontario, Canada
  M2H 3B4

- TMS Inc.
  110 West 3rd Street
  P. O. Box 1358
  Stillwater, Oklahoma, USA
  74076

6.0 DATABASES CURRENTLY AVAILABLE ON OPTICAL MEDIA

- **A-V Online**

A-V Online is produced by the National Information Center for Educational Media (NICEM) and distributed by Silver Platter Information Inc. A-V Online is an audio-visual material database, covering numerous non-book materials, videotapes, films and other media on any subject. The database is intended for the library market and for anyone who needs such information.

- **Blaise**

Blaise is produced and distributed by the British Library R and D Department. The database contains bibliographic records of books published in Britain, and published conference proceedings. The product is intended for the library market.

- **Bibliofile**

Bibliofile is produced and distributed by the Library Corporation. The database represents the Library of Congress MARC database. It can be used for creating catalogue cards, shelf lists, order lists, and various other applications relevant to library technical services.

- **BiblioMed**

BiblioMed is produced and distributed by Digital Diagnostics of Sacramento, California. The BiblioMed Citation Series covers 45 journals over a span of three to five years. The BiblioMed Specialist Series contains full text articles of key journals.

- **Books in Print Plus**

  Books in Print Plus is produced and distributed by Bowker Electronic Publishing. The database is a tool for library acquisition departments used for the tasks of pre-order searching, ordering, filing, and accounting. It is also possible, through the system, to order directly from Ingram.

- **Compact Cambridge - Aquatic Sciences and Fisheries Abstracts**

  This database is produced and distributed by Cambridge Scientific Abstracts. The information contained in this database comes from literature published worldwide and supplied by agencies such as UN Department of International Economics and Social Affairs, Food and Agricultural Organization, and the International Oceanographic Commission. The product is intended for agricultural libraries and anyone requiring access to information in the fields of Aquatic Sciences and Fisheries Abstracts.

- **Compact Cambridge - Life Sciences Collection**

  The above database is produced and distributed by Compact Scientific Abstracts. It contains eighteen subject-oriented subfiles containing abstracts from more than 5,000 core journals, books, serial monographs, conference reports, international patents and statistical publications. The product is intended for science libraries.

- **Compact Cambridge - Medline**

  Compact Cambridge - Medline is also produced and distributed by Cambridge Scientific Abstracts. The database covers virtually

every subject in the broad field of biomedicine. It is an international database of indexes to articles and proceedings from 3,000 leading medical and biomedical journals. The target market for this product is the health sciences libraries.

- **Compact Disclosure**

Compact Disclosure is produced and distributed by Disclosure of Bethesda, Maryland. The database contains financial and textual information on over 10,000 public US and foreign companies. Two hundred and fifty data items have been extracted from documents filed with the US Securities Exchange Commission. The database is intended for anyone who requires information on US and foreign companies.

- **CDI Corporate**

CDI Corporate is produced by Corporate Technology Services Inc. and distributed by Datext Inc. The database offers comprehensive business and financial information on 10,000 public companies, 900 lines of business, 50 industries and thousands of key executives. The database is intended for anyone who requires corporate and financial information for immediate use.

- **CD/Corp. Tech**

Corporate Technology Services Inc. and Datext Inc. also combined efforts to produce and distribute CD/Corp. Tech. This product offers information on more than 12,000 US high tech companies both public and private covering 40,000 products. This database is intended for those who need factual information on high tech companies.

- **DisCat (Public Access Catalogue)**

  DisCat is produced and distributed by Utlas International US Inc. While the databases mentioned previously in this section have been contained on CD-ROM, DisCat utilizes a twelve inch videodisc. The system provides patrons online access to the holdings of a particular library or group of libraries by searching a database stored on optical discs. The database is intended for the library market or similar agencies.

- **DisCon**

  DisCon is produced and distributed by Utlas International US Inc. and also utilizes the twelve inch videodisc. This product provides the user with approximately six million Library of Congress cataloguing records, accessible by L.C. Card Number, ISBN, and various title fields. The system is targeted at the library market for retrospective conversion.

- **Dissertation Abstracts Ondisc**

  This product is produced and distributed by University Microfilms International. This database is the only central source for accessing almost every doctoral dissertation accepted in North America. The database is targeted at the library market as well as any researcher who requires similar information.

- **Drugdex**

  Drugdex is produced and distributed by Micromedex Inc. of Englewood, Colorado. The database is a drug information system, intended to be used for the identification and treatment of abuses of toxic substances and prescription drugs. The database is intended for hospitals and medical facilities.

- **Ebsco (Serials Directory)**

  This database is produced and distributed by Ebsco Industries, Inc. and is, in essence, a directory containing references to serials publications. The database is intended for library serials departments.

- **Educational Resource Information Center (ERIC)**

  ERIC is produced by ORI Inc. and distributed by Silver Platter Information Inc. The database is actually compiled from two sources: Resources in Education (RIE) and Current Index to Journals in Education (CIJE). The system is intended for anyone interested in the fields covered by the ERIC System.

- **The Electronic Encyclopedia**

  The Electronic Encyclopedia is produced and distributed by Grolier Electronic Publishing Inc. of Danbury, Connecticut. The entire set of twenty volumes of the Academic American Encyclopedia (containing nine million words and more than 30,000 articles) is contained on one compact disc. It is possible to search for any word or word combination that appears anywhere in the encyclopedia and retrieve the results in records. The database is intended as a general reference tool for libraries or anyone interested in such information.

- **Emergindex**

  This database is produced and distributed by Micromedex Inc. The product is intended to be a reference tool in hospital emergency rooms for emergency treatment information on a variety of problems and is suitable for hospitals and medical facilities.

- **Federal Government Publication Catalogue**

  This database is produced and distributed by the Brodart Co. Library Automation Division. The product covers U.S. government publications dealing with finance, business, demographics, agriculture, medicare, and public health and is aimed at the library market.

- **GPO LaserFile**

  This product is produced and distributed by Library Systems and Services, Inc. of Gormantown, Maryland. The database encompasses all records from the U.S. Government Printing Office, including books, serials, and other forms of material and is targeted at the library market.

- **Government Publications Index**

  This product is produced and distributed by Information Access Company and serves as an index to the monthly catalogue of United States Government Publications. The database is targeted at the library market.

- **Indentidex**

  Micromedix Inc. is responsible for the production and distribution of the Indentidex database. The product contains information on the tablet and capsule identification system which provides manufacturer imprint codes as the basic identification source. Information is also provided on U.S. and foreign ethical/prescription drugs, OTC preparations, and generic drugs. The database is intended for hospitals and medical facilities.

- **Infomark**

  Infomark is produced and distributed by National Decision Systems, Encinitas, California. The database deals with marketing and market research. Data has been gathered from the US Census Bureau and other sources. This product is aimed at commercial marketing departments.

- **InfoTrac**

  This database is produced and distributed by Information Access Company of Belmont, California. Articles are drawn from 900 business, technical and general interest publications and are integrated into a single comprehensive database. The product is intended for library reference services.

- **Laserquest System**

  The Laserquest System is produced and distributed on 12 inch laserdisc by the General Research Corporation of Santa Barbara, California. The product is targeted at library technical services.

- **Laser Search**

  This product is produced and distributed by the Ingram Book Company. Laser Search is an integrated Book Identification and Acquisition system holding 1,275,000 titles in English that may be searched from a variety of perspectives. This product is also targeted at the library technical services market.

- **Lasertrak Pathfinder**

  This product is produced and distributed by the Lasertrak Corporation of Boulder, Colorado. Navigation charts and map data are provided for filing of alternate flight plans and related tasks. The product attempts to provide a complete information service for pilots and navigators.

- **LC-LAW**

  This product is produced and distributed by Horizon Information Services of Los Angeles, California. The product is a reference tool specially designed for legal librarians, attorney and legal specialists.

- **Library of Congress - Science and Technology**

  Horizon Information Services is also the producer and distributor of this database for librarians and specialists in Science, Technology and Medicine.

- **LegalTrac**

  LegalTrac is produced and distributed by the Information Access Company of Belmont, California. The database is contained on a twelve inch video disc and is an index of over 720 legal publications including law reviews, bar journals, and seven legal newspapers. This product is targeted at law offices and law libraries.

- **Library Information Science Abstracts**

  This product was produced and distributed by the Library Association of London, England. This database for librarians and information specialists covers all aspects of the profession.

- **MARC Laserfile**

  This product is produced and distributed by Library Systems and Services of Gormantown, Maryland. The database contains 2.5 million MARC records of the Library of Congress. The database is designed for retrospective conversion and cataloguing purposes, allowing libraries to print cards, labels, and store subsets of the database on hard disk.

- **MARC-English**

  MARC-English is another product of Horizon Information Services. The database contains English language legal material catalogued and distributed by the Library of Congress and specially designed for the legal librarian to simplify the cataloguing task in the law library.

- **MARC-STM**

  Horizon Information Services also produces and distributes MARC-STM which is a cataloguing tool designed for the scientific, technical and medical library community.

- **Medline-CD**

  This database, produced and distributed by Horizon Information Services, contains the full English language section of the National Library of Medicine's Medline database. The system provides access to over 1,000 of the world's best known biomedical journals used in research and direct patient care. The database is targeted at health-sciences as well as anyone interested in the bio-medical field.

- **MicroLinx**

  MicroLinx is produced and distributed by Linx Services, The Saxon Co. The database is essentially a serials management system and is designed for libraries.

- **NewsBank Electronic Index**

  This database is produced and distributed by NewsBank Inc. The product contains index entries or references to articles on current issues. Access to over 500,000 newspaper articles taken from the newspapers of over 300 U.S. cities are provided. The database is intended for those interested in current newspaper information.

- **NICEM (AV) Laserfile**

  This database, designed for the library market, is produced by Access Innovations of Albequerque, New Mexico and distributed by Library Systems and Services Inc. The database contains information taken from the National Information Centre for Educational Media database of audio visual, non-book material, and other similar items. The product is contained on a twelve inch videodisc.

- **OCLC Compact Disc - Cataloguing System**

  This database is produced and distributed by the Online Computer Library Centre. This system enables local production of cataloguing products such as catalogue cards, spine labels, and a number of other functions.

- **OCLC Compact Disk-Reference Package**

  This database is produced and distributed by the Online Computer Library Center. This reference package provides on-site patron and library staff access to a specially prepared database on CD-ROM.

- **PC Laser Library**

  The PC Laser Library is produced and distributed by The Library Corporation of Washington, D.C. The database contains nearly 9,000 public domain software programs for IBM PC and compatible machines.

- **Poisindex**

  Poisindex is produced and distributed by Micromedix Inc. This is a toxicology database designed to identify and provide ingredient information for over 300,000 substances, as well as to provide management/treatment protocols for toxicology problems. The product is designed for hospitals and medical facilities.

- **PsychLit**

  PsychLit is produced by Psychological Abstracts Information Services of Arlington, Virginia and distributed by Silver Platter Information Inc. The database is a major resource for information in psychology and the behavioural sciences. The database covers more than 1,300 journals and monographic series from approximately forty-five countries in more than two dozen languages.

- **Public Affairs Information Service**

  This database is produced by Public Affairs Information Services Inc. (PAIS) and distributed by Silver Platter Information Inc. PAIS is the primary reference database for the public policy aspects of social sciences, business, law and economics. The items indexed are selected from more than 1,200 leading journals plus 9,000 monographs, including books, pamphlets, directories, and government publications. The product is suitable for library reference services and for those interested in the fields covered by PAIS.

- **RE-MARC Laserfile**

  This product is produced and distributed by Library Systems and Services Inc. The database is contained on a twelve-inch videodisc and is meant for retrospective conversion of a library catalogue. It is also capable of transferring records directly to various existing library software.

- **Science Citation Index (Permuterm Subject Index)**

  This database was produced and distributed by the Institute of Scientific Information of Philadelphia, Pennsylvania. This database contains the Science Citation Index and allows users to access a full bibliographic record while continuing to brouse the index. The product is aimed at the library's reference department.

- **Socio-File**

  This database is produced by Sociological Abstracts Inc. and distributed by Silver Platter Information Inc. SocioFile offers users access to detailed abstracts from more than 1,200 social science serials published worldwide since 1963. The product is intended for library reference services and for anyone interested in sociological literature.

- **Software Library Dataplate**

  This product is produced and distributed by Reference Technology Inc. and includes software for the IBM PC and compatible machines. Business/financial programs, word-processing software, Database Management Systems, spreadsheets, and games are included. The database is intended for the home and business market.

- **Title 10 USC/CFR Disk**

  This database is produced and distributed by Alde Publishing of Minneapolis, Minnesota. The product is contained on both CD-ROM and videodisc and includes U.S. Code and Code of Federal Regulations material related to the social security field.

- **Title 26 USC/CFR Disk**

  This database is also produced by Alde Publishing. The product is an internal revenue code and regulations database including U.S. Code and Federal Regulations material.

- **Title 42 USC/CFR Disk**

  Alde Publishing also is responsible for the production and distribution of this product which is a Public Health Code and Health and Human Services regulations database including U.S. code and code of federal regulations material.

- **Ulrich's Plus**

  Ulrich's Plus is produced and distributed by Bowker Electronic Publishing of New York, New York. This database provides all the accessing ease and possibilities for periodical references from pre-order searching to actual generation of an order. The product is targeted at library services departments.

- **Wall Street Journal Database**

  This product is produced and distributed by the Information Access Company. This database provides full text coverage of the editorial content of the Wall Street Journal indexed in the InfoTrac database (mentioned earlier). Coverage includes up to twelve months of data, but is only available to those in possession of the InfoTrac database. The product is aimed at those interested in financial and business information.

- **Word's Business Directory**

  The Word's Business Directory is produced and distributed by the Information Access Company. The database, which is contained on a twelve inch videodisc, holds the profiles of 100,000 public and private companies worldwide.

- **Who's Who in Electronics**

  This version of Who's Who is produced and distributed by Knowledge Access of Mountain View, California. This database represents the Harris Selectory of electronics firms allowing users to create and code mailing labels or lists chosen by company, product, geography, size and/or title. This database is designed for those interested in personnel working in the electronics industry.

- **WilsonDisc**

  WilsonDisc is produced and distributed by the H. W. Wilson Co. of Bronx, New York. This database provides all the capabilities needed to access seven major indexes including: Applied Science and Technology Index, Book Review Digest; Business Periodical Index; General Science Index; Index to Legal Periodicals; Reader's Guide to Periodical Literature; and the Social Science Index. The product is intended for library reference service departments.

. **Zip + 4 Directory**

The Zip + 4 Directory is produced and distributed by Alde Publishing. The product is available in both CD-ROM and videodisc formats and includes the complete U.S. national Zip + 4 Directory provided by the Post Office. The database is intended for anyone who requires detailed and random access to this type of information.

The annual subscription prices for the above material databases range from $79.95 to $4,995. with the average price falling at around $1,918. The variability of the product's pricing reflects more than differences in database coverage -- it is symptomatic of the immaturity of the industry.

A range of update schedules for the databases are available, depending on the importance of information currency. For the database products described above, the update frequency ranges from annually to every twenty-eight days.

Just over 35% of the optically-based database products mentioned above have corresponding print products. Thus, there has already been a fairly large degree of innovation in terms of database content.

7.0 THE CANADIAN SITUATION

The optical publishing industry is not well developed in Canada. While there are a number of firms possessing the potential to manufacture optical media, very few firms produce the hardware necessary to read that media in this country.

The major players and potential players in the Canadian industry are:

- Praxis Technologies Corp.
- Reteaco Inc.
- Corel Systems Corp.
- Laser Disc Technologies
- Cinrom
- Americ Disc

7.1 Praxis Technologies Corp.

Praxis Technologies Corp. is located in Mississauga, Ontario. The $30 M plant became the first Canadian operation to reach commercial production of prerecorded compact discs in September of 1985. Praxis manufactures discs for CBC, CBS, Capitol, Polygram, Alert, and Anthem Records and is now one of the largest CD manufacturers in North America. By 1989, the company expects to be producing more than 25 M units a year, fast approaching the plant's annual capacity of thirty million discs per year. According to Michael Sifton, the president of Praxis Technologies Corp., the window of opportunity has already closed on the CD market. The demand for discs still exceeds supply, but most major manufacturers have staked out their territories in the marketplace and start up costs are extremely high.

Praxis reported a loss of $2.8 M on revenues of $3.3 M for the year ended June 30, 1987.

7.2 Cinram

Cinram accounts for almost fifty percent of all cassette and album production in Canada. The company expects to produce 3.7 million pre-recorded CDs at its $15 M plant in 1988. Within two years of production, the plant, which is just outside Toronto, is expected to double capacity to twelve million CDs annually from six million CDs. Capacity will eventually reach 24 million units. Cinram feels that "CD audio is only the tip of the iceberg" and that "the bulk of the market will be in computer applications like CD-ROM." Cinram sported a profit of $2 million for the nine months ended September 30, 1987 on revenue of $19.8 million over the same period.

7.3 Americ Disc

Americ Disc, whose CD manufacturing plant is located in Drummondville, Quebec, is a joint venture of SNC Group Inc. of Montreal, L'Enterprise Francaise Moulage Plastique de l'Ouest (MPO) of France, and the Quebec Government. The plant cost $28 million to construct and capacity will be eight million units by the end of 1987. In contrast to Cinram, Americ Disc does not plan any other applications for its plant besides the production of audio CDs.

7.4 Laser Disc Technologies

Construction of the Laser Disc Technologies plant near Grande Prairie, Alberta was initiated in 1986. The cost of the

facility was $60 M and a production capacity of 24 million discs per year.

On a global level, CD sales are expected to reach 250 million units in 1987. Last year, four million CDs were purchased by Canadians. With most of the production facilities in the country struggling to keep up with the demand for pre-recorded CDs, their attention is not likely to shift to the production of CD-ROM discs until that market softens somewhat.

That market-softening is predicted for the future however. The optimistic perspective is that since CDs are so durable, withstanding wear from normal playing as well as spills and other accidents, they won't need to be replaced, which means they will reach a plateau in the marketplace. The pessimistic viewpoint is that digital audiotape, which offers all the clarity of compact disc audio with the added feature of recording capability, will erode the CD market. It is at this point where CD manufacturers will effect the minimum retooling necessary for the production of the CD-ROM media.

7.5 Reteaco Inc.

Reteaco Inc., which is a Toronto-based company, is the only Canadian company to initiate production of CD-ROM Platters. One of Reteaco's most recent projects was to process all the 1985 issues of the Globe on to a CD-ROM database. Another project which the firm has undertaken was to put on disc the Canadian Telephone Book, which includes seven million Canadian phone listings, complete with postal codes. Other projects include a directory of North American real estate data, automotive parts catalogues, the past 10 years contents of the Harvard Business Review, and a French-English translation list of 860,000 technical terms for the Federal Government.

Starting with magnetic computer tapes, Reteaco analyzes a customer's data in Toronto and has a master disc made by a sub-contractor in the U.S. Copies are then manufactured in Canada. Reteaco's price for creation of a master disc, of course depending on the scope and complexity of the data, is about $15,000. Thus, in volumes of 600 or more the cost per disc is $25 or less.

7.6 Corel Systems Corp.

Corel Systems Corp., located in Kanata, Ontario, is offering a system which includes a personal computer and WORM drive, suitable for a variety of applications previously described in Chapter 4. Corel's engineers have altered the operating program language driving the PC's memory units so that it may utilize the WORM drive. The company is now in the process of negotiating a venture with optical disc equipment makers in the U.S. As well, the company has been approached by several unnamed optical disk makers.

Corel has landed several contracts, including the sale of the systems to the Supreme Court of Canada for $200,000 in March of 1987.

The firm has 35 employees and plans to expand its marketing efforts throughout North America. For the first quarter ended February 28, 1987, the company had a profit of $121,000 on revenue of $1.8 million. For the start-up year ended November 30, 1986, Corel had a loss of $1.8 M on revenue of $1.8 M.

8.0 POTENTIAL BUSINESS OPPORTUNITIES FOR CANADIAN INDUSTRY

As revealed in the previous chapter, Canadian firms are not well positioned for a large measure of success in the hardware manufacturing area. Various firms in the United States, the United Kingdom and Continental Europe and Asia have a large edge over Canada in this aspect of the industry.

There remains, however, one opportunity of the "hard" side of the optical publishing industry. This opportunity relies on the fact that there exists a fairly large established CD-audio disc manufacturing capacity in the country. With a minimum of retooling all, or portions of, these plants may be converted to production of CD-ROM media. This is not likely to happen in the short term, however, as most disc manufacturing plants in the country are at or near capacity. As demand for CD-audio levels, however, or as supply in various other producing countries increases, any loss of production may be offset by manufacture of CD-ROM media.

On the "soft", or service side of the industry, Canada possesses a great deal of potential. Reteaco Inc. is a good example of what may be achieved in the Canadian market in terms of data preparation services. With the vast amount of government and private sector data bases in existence, there will be a large demand for data preparation services as CD-ROM becomes a more and more accepted information storage medium.

Another area in which Canadian firms could likely succeed is the development of access software. Again, as the CD-ROM medium becomes increasingly acceptable and the installed based of CD-ROM drives reaches a comfortable level, competitive attention will

turn to making these systems easier and easier to use. User-friendly software access systems will be in high demand assuming the technology considers to grow.

The ultimate extension of user-friendly access software will be front-ends driven by artificial intelligence. Canada possesses a responsive and growing artificial intelligence and natural language processing community which should be positioned to take up the challenge of linking the concept of artificial intelligence with CD-ROM based data access.

The first step toward forging the alliances necessary to catalyze the various elements of Canada's industry into becoming a force in the world optical publishing market, would be a conference similar to the Online Meeting regularly held in New York City. Only nineteen representatives of Canadian companies were in attendance by thousands of delegates in total. A conference such as this allows various interested parties not only to become aware of the status of the industry in Canada and the rest of the world, but also to form allegiances which may result in successful joint venture production or service partnerships in the future.

STANDALONE DATABASES : OPTICAL PUBLISHING