

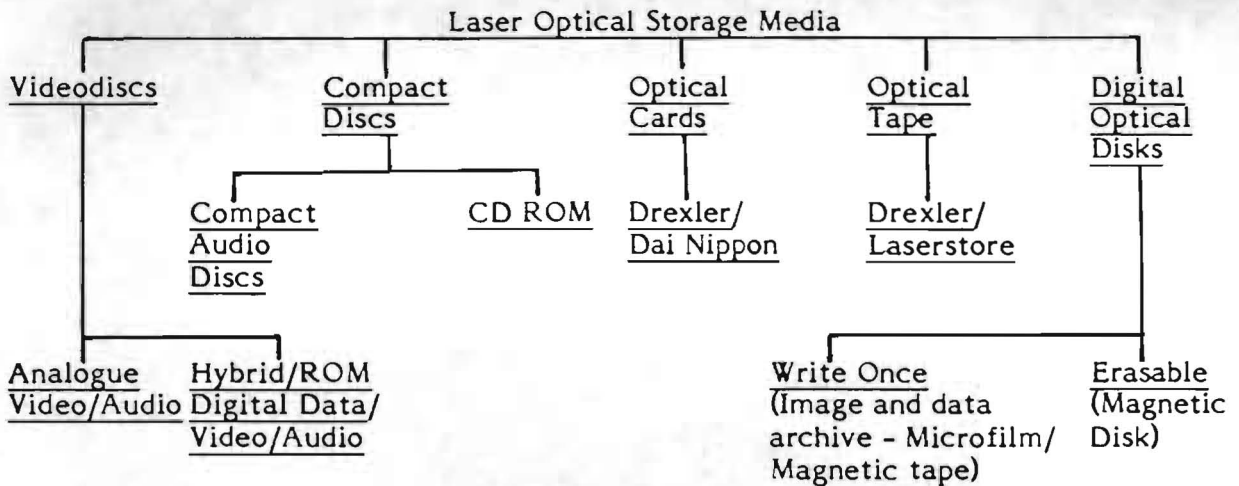
ReprintsCONTENTS

1.	BACKGROUND	3
2.	CD ROM	4
2.1	Production of a CD ROM	4
2.2	Ties with Compact Audio Discs (CADs)	6
2.3	CD versus other optical read only media	7
2.3.1	Hybrid videodiscs	7
2.3.2	OROM (Optical Read Only Memory)	8
2.4	How far can CD ROM standardisation go and how far has it got?	8
2.4.1	Logical file structures	9
2.4.2	Search software	11
2.4.3	Hardware and software interfaces	12
2.5	Marketing issues	13
3.	CONCLUSION	16
	Appendix 1 Executive summary	17
	Appendix 2 Comments on European CD ROM Standards Forum	19

AN INTRODUCTION TO CD ROM (COMPACT DISC READ ONLY MEMORY)

1. BACKGROUND

CD ROM is one of a large number of laser optical storage devices which have been leaving the laboratories since the late 1970's and moving into the marketplace. The following table lists the main types of optical storage devices currently available or under development.



Though they all employ the same basic technology (laser recording/reading head, sensitive thin film recording surface) the above media are aimed at different markets and have different characteristics which suit them to specific application areas. We choose to divide them up into three categories:-

1. **Read Only** - where data is recorded onto the medium by a publisher or information provider and all the user needs is a player to read the information on a pre-recorded optical disc, card or tape (e.g. records on compact discs, feature films on videodiscs, databases on CD ROM).
2. **Write Once Read Many Times (WORM)** - where blank recordable disks are distributed to users who can then use combined recorders/players or drives to both record information onto the discs in-house and read the information back in-house (e.g. digital optical disks which will replace microfilm for the storage of document images and magnetic tape for the archival storage of computer data).
3. **Erasable/Rewritable** - where blank disks are distributed to users who can then use combined recorders/players (drives) to record information onto the disks and read it back and who can erase information from the disks and re-record data onto the same section of the disks many times, as with current magnetic disk and tape media.

Discs are the preferred format because the use of a disc provides high storage capacity and relatively fast random access capabilities. Optical tape systems are intended for archival applications where speed of retrieval is not essential and cost is an important factor while optical cards are aimed at applications where storage density is not important but the low cost of the medium, the reader and the convenience of being able to carry it in a pocket are significant (e.g. the first application for Drexler cards - carrying the medical details of medical insurance company clients).

There are a wide variety of disc sizes in use:

4.72 inch diameter (12cm) Compact Audio Discs
CD ROM

5.25 inch diameter (13cm) Optical Read Only Memory (OROM)
Write Once Disks for small systems
Erasable Optical Disks
(Popular format as modelled on Winchester disk format).

8 inch diameter Music videodiscs for consumer applications (sponsored by Japanese companies)

12 inch diameter Consumer videodiscs
First generation recordable digital optical disks aimed at data archiving and document image storage market (magnetic tape and microfilm replacement)

14 inch diameter Based on mainframe Winchester format and aimed at mainframe data storage market

2. CD ROM

The above is designed to put CD ROM into perspective. CD ROM systems are based very closely on the consumer Compact Audio Disc systems. Indeed there are already dual purpose players that will accept both Compact Audio Discs (CADs) and CD ROMs. At the risk of oversimplifying, one can say that with CADs audio information is coded digitally, stored on the disc in digital form and at the playback stage the digital data is converted back into analogue form and the audio signals are played back via amplifiers in domestic hi-fi systems. With CD ROM digital data is recorded on the disk in the same way but at the playback stage it is not converted into analogue form it is read off in digital form, block by block, checked for errors (which are corrected) and then sent to the memory or the magnetic storage device of a computer.

CD ROMs therefore, as their name implies, are 4.72 inch diameter discs which hold digital data (coded text, digitised images, vector graphics, software programs) and have potential for the distribution of full text databases, software packages or any other form of information likely to be required by users of a range of computer systems from personal computers to mainframes.

2.1 Production of a CD ROM

Like all the optical discs listed above, CD ROMs start life as glass master disks. Data is recorded onto the glass master via a laser beam and the master is then used to produce submasters from which stampers are made which are then used to stamp out as many replicate CD ROMs as are required.

The raw material used for compact disc replication is a polycarbonate called Macrolon which is relatively inexpensive and hence in volume the cost of replicate CD ROMs can be very low - making CD ROMs potentially very attractive as a publishing medium. As CD ROM is a read only medium all the information is recorded at the mastering stage and the replicate discs are then sealed with a protective coating prior to shipment to the customer.

The mastering process is almost identical for videodiscs, write once digital optical disks and erasable disks but in the case of write once and erasable disks the only data that is recorded on the disks at the mastering stage is formatting data and the replicate disks are usually loaded into plastic cartridges to protect them from damage during handling. To date only glass write once disks have been made available in any numbers and these recordable disks are relatively expensive at £400 each. They are intended for in-house use where not more than two or three copies of any disk will be made and not as a publishing medium.

If a publisher wishes to distribute a database or a software package on CD ROM he has to provide Philips or some similar mastering facility with copies of 0.5 inch magnetic tape carrying the database. Philips then process the data, splitting it into blocks of 2 Kbytes to which synchronisation patterns, headers plus EDAC (Error Detection And Correction) codes are added to form CD ROM sectors. The master recording system comprises a laser beam recorder (LBR), an encoder, a subcode processor, a digital tape recorder and a system controller which monitors all phases of the recording process. The master disc is placed into the LBR and the master tape is played on the digital tape recorder. The digital information is then fed into the encoder which carries out multiplexing, error correction, coding and modulation while the subcode information is converted into subcode signals and also fed into the encoder. The signals from the encoder are then passed to the acousto-optical modulator in the LBR and the laser beam is modulated and focused, under servo control, onto the sensitive coating of the rotating disc.

After recording the exposed areas of the disc's photoresist coating are etched away until the required pit geometry is achieved. Then, after development, the master disc is given a thin silver plating and undergoes final inspection before being processed in a nickel bath to form the first negative or father disc. From the father multiple positive mothers are made to produce a generation of sons which form the stampers used in the injection molding replication process described above.

The mastering and replication processes have to be done in tight clean room conditions and require the use of precision equipment which is extremely expensive to set up. CAD/CD ROM mastering and replication facilities cost many millions of pounds to establish. Currently, therefore, CD ROM mastering and replication will not be an in-house facility for most publishers and database providers, they will have to make use of external commercial facilities and there are very few of these within Europe.

Philips provide a CAD and CD ROM premastering service and use the Polygram facility in Germany to provide master and replicate CD ROMs but already capacity is almost fully booked and information providers could find it difficult to have small volumes of CD ROMs replicated within a tight turnaround period. However, there are many more companies in Europe mastering and replicating CADs at present and these all have a licence that enables them to switch to CD ROM production as well and outside of Europe there are a number of CD ROM mastering and replicating facilities in Japan and several starting up in the U.S.A. so the problem of where to have CD ROMs mastered and replicated quickly and cheaply should only prove to be a temporary one and by the time publishers are ready to launch full-scale information services based on CD ROM there should be no shortage of production facilities.

2.2 Ties With Compact Audio Discs

As noted above, the prime feature of CD ROM is that it is based on the worldwide CAD standard established by Philips and Sony. This has two main advantages. The first is that since CD ROM discs and players are virtually identical to Compact Audio Discs and players, CD ROM can benefit from the economies of scale which result from CAD's success in the consumer market. The second is that the physical attributes of CD ROM systems (disc diameter, data recording techniques, error correction and detection techniques, block size, track and sector layout) have been standardised - an essential requirement if CD ROM is to become a widely accepted publishing medium which can be used by libraries. We deal with the question of standardisation in detail below.

In 1986 some 4 million CAD players will be sold and the price of players has already come down from the introductory price of £600-£700 to £150-£200. The result is that most of the development costs associated with CAD players have already been covered and hence CD ROM drives can be sold at a substantially lower price than would otherwise have been possible.

We predict that CD ROM drives will eventually be available at prices ranging from £200 upwards. The exact price is difficult to predict, of course, as it depends very much on the final market size for CD ROM and what one includes in that price. £200 would be for the basic drive and there would be additional costs for hardware and software interfaces to link the CD ROM drive to a microcomputer. We itemise all the hardware and software elements involved in linking CD ROMs to microcomputers in detail below.

In addition, the disc production facilities required for CD ROM are already in place for CAD and the commercial success of CAD ensures that further production facilities will be introduced over the next few years leading to a widespread availability of CAD mastering and replicating facilities and hence a rapid drop in the cost of mastering and replicating both CADs and CD ROMs. At present the cost of mastering one CD ROM (assuming that the data is available on tape in the required format) is approximately £2,000 and the cost of replicating the discs is about £4 per disc in volume. In future, experts predict that the mastering costs will drop to roughly £1,000 and the cost of replicating a disc to approximately £1 per disc in high volume once Japanese CD ROM replication facilities come on stream.

Like CAD, a CD ROM is a single sided 4.72 inch (12cm) diameter disc with the data recorded on it in the form of pits in a spiral track. Without going into great detail, the essential features of CD ROM - its strengths and weaknesses - are as follows:

As the disc is relatively small and only single-sided the data is recorded on the disc in CLV (Constant Linear Velocity) mode in order to maximise the storage capacity (i.e. instead of recording the same amount of data on each track of the disc as in CAV mode, more data is recorded on the outside tracks which are longer, than is recorded on the inner tracks which are shorter). The total storage capacity obtained using CLV mode and the same data recording techniques employed for CAD is approximately 600 Mbytes (in excess of 240,000 pages of coded text or up to 10,000 raster scanned images). The trade off with CLV recording is that it restricts access speed as data cannot be addressed purely on a track-by-track

basis. Hence average access time per seek on a CD ROM is 1-2 seconds compared to approximately 100 milliseconds or less with Winchester disk systems or proposed CAV mode Optical Read Only Memories.

In general, the storage capacity will prove the key in electronic publishing applications but the relatively slow access times will place the onus on information providers and retrieval software suppliers for CD ROM systems to lay out their data and devise their retrieval procedures so as to minimise the number of seeks that have to be made (the number of movements of the laser read head) in order to carry out a search of a database and this will not be an insignificant task.

One significant advantage that CD ROMs have compared to magnetic diskettes is that, since they are read by a laser beam in the CD ROM player and they are a read only medium, replicate CD ROMs can be coated with a thick protective coating and there is no need for any contact with the read head - the laser beam is focused through the protective layer onto the recorded surface inside it. Hence CD ROMs, like CADs, are durable media and should tolerate rough handling and last for a minimum of 10 years. The exact longevity of the discs will not be known until they have been available for 20 or 30 years but it can be said with a fair degree of confidence that it will be more than sufficient for most database publishing applications.

2.3 CD ROM versus other Optical Read Only media

2.3.1 Hybrid videodiscs

The physical standardisation of CD ROM media and drives is its main advantage when compared with other read only optical disk systems that have been proposed. A number of companies such as Laserdata, TMS, Reference Technology and AT&T have proposed techniques for encoding digital data within the analogue signals of standard videodiscs - enabling the resultant hybrid videodisc to store a combination of digital data, audio and video. However, each of these companies uses a proprietary encoding technique and use of their system requires the use of either non-standard expensive videodisc drives or the use of proprietary control boxes or boards attached to standard videodisc players and the result is that there is no physical standardisation so these systems are unlikely to find any generic applications and secondly, the hardware element of the systems will be considerably more expensive than for CD ROM.

The main advantage of hybrid videodisc systems when compared to CD ROM systems are in applications where mixed still or moving video sequences need to be interspersed with coded text. With CD ROM there is no facility at present for storing still or moving video images and even if these are developed in future they will be limited due to the relatively small size of CD ROMs compared to videodiscs.

CD ROMs can be used to store coded data (text, programs, etc.), audio information and digitised (i.e. raster scanned) images not video information. Hence there will continue to be a specialised market for hybrid videodisc systems in the training, education and consumer market and in business applications where visual databases are of significant value. Recently

Philips have announced a new version of their Laservision videodisc player called Laservision ROM which permits some 324 Mbytes of digital data to be stored per side alongside the traditional 54,000 frames of video information and the two tracks of audio information. This could become a standard in future and has been adopted by the BBC for their Domesday publishing project but the hardware will be significantly more expensive than CD ROM drives and we believe that CD ROM drives will continue to be the main medium for largescale distribution of coded text and vectorised/raster scanned graphics due to its standardisation and the low cost of the drives.

2.3.2 OROM

The other possible challenger to CD ROM in the long term is the 5.25 inch CAV mode optical disk pioneered by 3M and Sony and called the OROM (Optical Read Only Memory) or Data ROM. This varies from CD ROM in many respects.

1. It is 5.25 inches in diameter based on the Winchester format.
2. It is double-sided
3. It is recorded in CAV mode and hence data can be addressed on a track by track basis and access times of 100 Msec are claimed compared to 1-2 seconds with CD ROM.
4. Because it is recorded in CAV mode the storage capacity is approximately 50% of CD ROM at 250-300 Mbytes per side.
5. The OROM is designed to play in so called 'universal drives' which could accept write once recordable 5.25 inch disks and also 5.25 inch erasable disks. One such drive, from Optotech, has already been shown.
6. Because it is being developed from scratch and does not use many of the parts common to CD Audio players and because the drives are being built with a record facility, the likely price of OROM drives will be well in excess of CD ROM drives - nearest estimates put the price of OROM drives at £1-2,000.

The last point indicates that OROM poses little threat to CD ROM as a largescale publishing and data distribution aid although it may well be adopted by some of the major computer suppliers, when it becomes available, as a magnetic disk or floppy disk substitute in applications where high volume storage is required.

2.4 How far can CD ROM standardisation go and how far has it got?

Standardisation is a key issue for publishers and information providers who are evaluating the potential uses of CD ROM. It is also an area where there are a lot of misconceptions, particularly among potential users of CD ROM systems who lack a sufficient understanding of the computer industry. The main misunderstanding appears to be that CD ROM systems are already totally standardised. This is not true. One level of CD ROM systems - namely the physical level which relates to the physical characteristics of the discs, the drives, the way data is recorded onto the discs, the error detection and correction codes

- has been standardised by Philips and Sony but other areas still need to be looked at to see whether standardisation would be beneficial or limiting. These include:-

- a. The logical file structures - how the information is organised on the discs;
- b. The hardware and software interfaces between the drives and the computers.

2.4.1 Logical file structures

The companies promoting CD ROM systems and the information providers both understand the significance of standardisation in this area and in the U.S. the National Information Standards Organisation (NISO) have set up a Compact Disc Format Committee to help develop an American National Standard for data publishing on CD ROM.

In a statement announcing the formation of the CD ROM committee they point out, as we have done above, that the lack of a standard logical structure for CD ROM has slowed the development of CD ROM based information products and services and their aim is to recommend a standard file structure which will enable:

- a. Publishers and other distributors of information to be able to master a single CD ROM disc and know that replicated copies will be readable throughout the world on most combinations of CD ROM drives, computer hardware and computer operating systems;
- b. Developers of CD ROM retrieval software to have a uniform file environment in which to design and implement applications;
- c. Manufacturers of CD ROM drives to be able to write a single version of file server software for creating device drivers for most computer operating systems;
- d. Mastering services to be able to validate replicated copies of compact discs using the directory and other standard information elements to verify accurate placement of data on the disc.

The scope of the Compact Disc Data Format Committee is to recommend, for publishing applications, a high level structure for directory and other information on compact discs. The following specific items are to be described:

Location and format of bootstrap blocks containing application-specific information including:

Title and description of contents

Character set(s) used for contents of disc

Designation if the disc is part of a multi-disc set

Logical block size

Length and format of the file name

Extent (if any) of data encryption

Location and minimum contents of the directory to the files on the disc

Location and minimum contents of the attributable fields associated with each file on the disc.

Shortly after the NISO statement, Digital Equipment Corporation announced that they have developed their own logical file format standard which they are calling Uni-File. DEC are making the standard available without licensing fees and simply make a charge of 250 dollars for the standards documentation. The DEC initiative has been welcomed by the NISO committee but they have no commitment to adopt it and other companies such as TMS and Reference Technology who have been working on alternative file formats have openly criticised the DEC proposal. One further problem with the DEC proposal is that it has been designed primarily to work with DEC's VMS operating system and not with the major competitive operating systems - MS DOS/PC DOS and UNIX.

Following the DEC announcement, representatives from some 12 major computer and software companies held a meeting at Lake Tahoe in November in order to discuss how data may be stored on CD ROM in a standard way. Companies included:-

Apple
DEC
Hitachi
Laserdata
Microsoft
Philips
Reference Technology
3M
TMS
Yelick

The group agreed that they will accept the physical structure of a compact disc as set by Philips/Sony and will address only data format issues. They also decided that the specifications should be hardware and software independent and that mixed modes and large files spanning multiple CD ROMs should be accommodated in the specifications. The members committed themselves to putting together a working version of a logical format standard by the end of the first quarter of 1986.

NISO are apparently now considering several logical file format proposals submitted to them including proposals from DEC, Reference Technology, TMS, Activenture and the Tahoe Conference Group.

In the wings, Microsoft is sponsoring a conference on CD ROM in Seattle from 3rd-6th March, 1986 at which the question of CD ROM standards and search software will be very much to the fore and in Europe a new body called the Optical Disk Forum invited representatives of European information providers and manufacturers to meet together with a view to developing a de facto standard. The first meeting will be in Luxembourg on 22nd January, 1986 and it is understood that at this meeting representatives from the U.S. industry standards initiative will be present and the aim will be to hammer out an interim standard that enough providers of information on optical disks will observe to enable information on CD ROM to be designed for a true broad market.

Given the recent birth of the European Group it would appear that what this actually means is that the European information providers and manufacturers will be asked to endorse whatever interim standard comes out of the various U.S. meetings described above if, indeed, one undisputed standard does emerge by then.

If no standard file format is agreed then the idea of CD ROM ever becoming a generic publishing medium and opening up a vast library market may be stillborn. Information providers would have to choose one retrieval software package and structure their database around that and accept that the choice would limit the size of their market or else they would have to form agreements with a range of retrieval software suppliers and create multiple versions of their database on different CD ROMs targeted at different computer users. A number of different file formats and retrieval software packages would tend to dominate and at the same time split the marketplace, each one supported by one of the major computer companies or software suppliers such as DEC or Microsoft.

2.4.2 Search software

Even if one standard logical file structure is established the next question is search software. Will it be adequate in the long term to simply modify existing search software or will radically new search software be introduced specifically for CD ROM applications. The U.S. Newsletter - CD Data Report - carried in its June 1985 issue a survey of U.S. firms with CD ROM search software products on the market or in development. It listed some 17 companies including a number of small start-up companies of little significance and some major software companies with a proven track record in the provision of search software for online database systems including:

- a. Battelle who have launched MicroBASIS for the CD ROM publication market. MicroBASIS is available on MS DOS and UNIX based systems as well as on the MicroVAX under VMS. The MicroBASIS software will be licensed to information providers and will be recorded onto the actual CD ROMs, accompanying the database information. MicroBASIS is used by DEC in their complete CD ROM system where it enables users of the DEC MicroVAX system to retrieve data on CD ROM that has been structured according to the Uni-File file format.
- b. BRS also now sells a micro version of their search software BRS SEARCH and has technology sharing arrangements with Laserdata and Reference Technology Inc. who have used the micro version to retrieve data from CD ROMs.
- c. TMS have developed their information research software for retrieving data on hybrid videodiscs or CD ROMs.
- d. Meanwhile in Europe, Harwell are known to be modifying their STATUS retrieval software for use with CD ROM and the French company, Telesystemes, are modifying their Questel software for use on microcomputers and for retrieving data from CD ROM.

Again it is early days yet to say whether any of the packages which have been shown are ideal for use with CD ROM. First impressions are that they are not, that a far more powerful database management system will be needed and that there will certainly be a wide range of software packages developed over the next few years which are more capable of dealing with the wide range of material that can potentially be stored on CD ROM. For example, the requirements made on retrieval software by a coded text database of abstracts will be different from the requirements made by a mixed mode database containing coded text and digitised images or vector graphics.

Many of the software packages described above are still being adapted for use on microcomputers and many compromises are having to be made. In the future there will be further trade-offs between requiring the user to have a powerful microcomputer with additional RAM or restricting the amount of data that can be taken from the CD ROM at any one time. We are still in the very early days with search software for CD ROM and no major breakthroughs can be expected, of course, until the file format standards are agreed and it is seen whether the first version of those standards deals with images as well as text or just text.

2.4.3 Hardware and software interfaces

If we now make the major assumption that a standard file format will be adopted and that a range of high quality software packages become available that are capable of running on all the most popular microcomputer operating systems, the last hurdle to acceptance by end users will be the question of physically interfacing the CD ROM drive to one of a range of microcomputers and, most importantly, how much this is likely to cost.

With production versions of CD ROM drives only just becoming available it is early days to make definite predictions about interfaces but we can look at some of the early products to have been launched by companies such as Philips, DEC and Reference Technology and point out all the elements involved in linking a CD ROM drive to a microcomputer.

Philips' CM 100 CD ROM player is now available at an end user price of £750 - quantity one. In addition, Philips can supply their CM 155 host resident CD ROM controller for the IBM and compatible PC-XT, PC-AT personal computers and lookalikes. Two of Philips' CM 100 - 600 Mbyte capacity drives can be attached and concurrently used with the CM 155 controller which is also priced at £750 including the serial interface cable. Included with the CM 155 is a software driver that works with both PC DOS and MS DOS and available with the CM 155 is file server software that allows the user to go beyond the implicit 32 MByte limitation of individual files under MS-DOS. The server software also allows the user to consider the combined storage capacity of the two drives as a multi-volume file. Data is transferred to the CM 155 at an average rate of 1.41 Mbits per second with a burst rate of 2.11 Mbits per second. Total cost of the Philips CD ROM subsystem is therefore £1,500.

DEC have produced a complete CD ROM package for users of their MicroVAX II super micro system. The system comprises a modified Philips CD ROM drive, a controller, a Q-bus interface, the 32-bit supermicrocomputer, the Battelle MicroBASIS retrieval software and a video display terminal. Price of the full system is \$18,840 in the U.S. Price for the CD ROM Drive, controller and cables is \$2,195. DEC have also recently made available the same CD ROM drive with a controller produced by Distributed Logic Corporation which can be linked to the IBM PC XT and AT.

The two different approaches adopted by Philips and DEC illustrate two possible approaches to the integration of CD ROM with microcomputers. Either one can link a basic CD ROM drive to a host resident controller such as Philips' CM 155 with CD Driver software via a proprietary device interface or one can fit a SCSI controller to the CD ROM Drive and then connect the Controller to a SCSI host adaptor within the PC via a standard I/O SCSI interface. The host adaptor then holds the SCSI driver software with a CD Module.

There are pros and cons with these two different approaches and today the performance (in terms of data transfer rate) of different drives and interfaces varies. This is not so critical when dealing with coded text but can be critical when dealing with raster images which require large amounts of data to be transferred from the CD ROM to the PC in streaming mode. The use of graphics on a CD ROM also introduces other technical problems as in order to display the images, graphics cards must be fitted to the IBM PC and, in some cases, high resolution terminals used to replace the standard PC monitor.

Finally, Reference Technology has recently launched their CLASIX Datadrive Series 500 CD ROM drive which is produced for them on an OEM basis by Hitachi. The drive can work with IBM PC, XT and AT or compatibles. The Series 500 system comprises the drive, the connecting cable, an interface adapter card, interface control software and documentation. In addition, Reference Technology have also adapted their STA/FILE software package in order to support the CD ROM discs. The price of the Series 500 including all the hardware needed to link it with the IBM PC range is \$1,535 in the U.S. and the STA/FILE software will sell at \$110 per user licence.

2.5 Marketing issues

Since the first announcements were made and the first prototype CD ROM drives shown in 1984, information providers and potential users such as libraries and major users of online and printed reference products have taken a considerable interest in the potential of CD ROM. A series of reports and market surveys have outlined the following major applications for CD ROM systems which we list in order of importance:

Database distribution
Consumer products (encyclopaedias, dictionaries accessible via home computers)
Software distribution
Interactive instruction
Periodical publishing
Legal information publications
Images
Medical information
Point of sale products.

However, as inevitably happens with any new information dissemination medium - microfilm, videotex, etc. - many of the predictions being made for CD ROM or optical publishing media in general are inevitably exaggerated and ill informed (tending to concentrate on the many advantages of CD ROM publishing over, for example, traditional paper, microfilm or even on-line information distribution systems and minimising the many current drawbacks to CD ROM publishing such as relatively high hardware costs, long turnaround times and the fact that there is, as yet, no installed base of CD ROM drives).

At present, for reasons stated above, information providers planning CD ROM services cannot assume that there will be one standard file format and one standard interface between CD ROM drives and PCs in the near future. In other words they cannot simply place their information on CD ROM together with the necessary search software and expect to sell the disc to a notional base of PC and CD ROM users in the way in which, for example, a publisher can today publish his information on microfiche at a standard reduction ratio and assume that all libraries will have the necessary equipment to read it.

For one thing, as there are such a wide range of choices open to the would be producer of a CD ROM database and so many potential pitfalls at each stage we would strongly recommend that the information provider enlist the services of one of the many companies who now offer a full premastering and mastering service to publishers. Many of these companies subcontract the actual mastering and replication but, importantly, are prepared to take responsibility for carrying a project through all the stages. They include DEC, Silver Platter Services, TMS, Laserdata, Reference Technology and in the U.K., Langton EPS.

The total costs of producing a CD ROM are almost impossible to estimate if the information product is being specially created for CD ROM. If it is a question of converting an existing textual database that is already available in machine readable form then Langton quote an average figure of £20,000 for data preparation, £2,000-£3,000 for mastering and then a run on cost for each replicate disc which depends on the volume of replicate discs required but is in the region of £4-£5 per disc in volume.

The early providers of CD ROM based information services, whether introducing a new service or offering CD ROM as a replacement/alternative to existing printed, microform or online distribution, must also realise that they will be introducing both a new service and a new technology. They will have to educate the potential users of the service in the advantages of CD ROM technology and offer, as part of a subscription service, to supply the user with the necessary hardware to access the data on CD ROM.

For the above reasons, the first applications for CD ROM will not be as a generic publishing medium like paper or microfilm but as part of closed user information delivery networks and in these applications the lack of standardisation will not be a significant limitation. For example, a publisher of legal or medical information - whether bibliographic or full text - could identify a percentage of their existing customers spending more than a certain amount in online search fees or in purchasing printed products and then offer to provide them, on a subscription basis, with the information on CD ROM, the necessary retrieval software package and the necessary hardware to playback the information. However they must be selling very valuable information and they must be sure that the CD ROM product is either cheaper in the long term than existing distribution media, or offers considerable value added features.

At present the majority of database providers who have shown CD ROM products or announced their intention to market all or selected portions of their database on CD ROM are adopting this approach and charging users an annual subscription to include one or more CD ROMs containing the database plus the retrieval software and the option of leasing a CD ROM drive with hardware and software interfaces to one or more microcomputers (usually an IBM PC, XT or AT). At present, however, the pricing of these services looks very conservative with database providers obviously worried that heavy users of their online services will migrate to CD ROM and not convinced that the CD ROM service will attract enough new users to cover this potential loss of income.

Unless database providers are prepared to take a more enterprising approach to pricing in future then the prediction that CD ROM databases will appeal to that sector of the market which currently cannot justify the cost of online searches and hence open up new markets for database providers will not be fulfilled. All that will happen is that the market for database information will be split three ways between printed, CD ROM and online services where once it was split two ways between printed and online products.

Once these pioneers have established the technical viability and the marketability of CD ROM systems, the next envisaged stage in the development of CD ROM as a publishing medium will be for a standard file format to be developed and retrieval software to be agreed upon which can interface with all the popular microcomputer operating systems. At which time, major publishers such as U.S. Government Agencies and current major micropublishers and publishers of printed reference works will begin to plan CD ROM services aimed at libraries and CD ROM may find a place as a generic publishing medium but this is far from certain.

One other assumption that is being made by those promoting CD ROM and information providers who are readying CD ROM products, which we would like to question, is that users already have the required IBM PC to access the data. If they are online users then they will probably have a microcomputer but it may well need a memory upgrade before being interfaced to the CD ROM and it may well be that it is already being heavily used for library automation functions, for searching other online databases or just for routine administrative functions. If users do not have existing PCs or are already making full use of the PC for other purposes then they will need to purchase both the PC and the CD ROM drive and interface and this could represent a considerable investment. Furthermore, while users

purchasing bibliographic databases on CD ROM may be satisfied with a single user station (i.e. one PC linked to one CD ROM drive), if the second stage is to become reality and CD ROMs are used as a generic publishing medium for libraries then libraries would need multiple user workstations and at one PC per CD ROM drive this would prove a very expensive hardware investment for material that is not traditionally heavily used.

3. CONCLUSION

We do not wish to appear unduly negative in our appraisal of the potential of CD ROM for information providers. Inevitably CD ROM will have a considerable impact for database providers and in many other fields such as the distribution of software and the distribution of technical manuals, parts lists and directory type information to closed user groups. We simply feel that as so much has been said in favour of CD ROMs, we need to point out with equal force some of the potential problems which could hinder the acceptance of the medium or at least lengthen the period of time that has to pass before CD ROM does become widely accepted. We also consider that it will take a long time and require greater cooperation than has ever before been demonstrated in the computer industry if CD ROM is to become a generic publishing medium for widespread use in libraries.

Inevitably, when trying to explain all aspects of a new technology it is hard to avoid lapsing into jargon or going into unnecessary technical details. We have tried to avoid this but realise that at 12 pages this account may provide more detail than is required by many with a non-technical interest in the potential of CD ROM technology. We have, therefore, included a second brief summary of the main points made in this introduction to CD ROM which follows.

APPENDIX I

EXECUTIVE SUMMARY

Introduction to CD ROM

CD ROM (Compact Disc Read Only Memory) is one of many laser optical storage devices which have been moving from the laboratory and into the marketplace in recent years.

CD ROM systems are based very closely on Compact Audio Discs which have proved very successful in the consumer market for high fidelity audio.

Like Compact Audio Discs, CD ROMs are single-sided, read only 4.72 inch (12cm) diameter discs. They hold up to 600 Mbytes of digital data (up to 240,000 pages of coded text or up to 10,000 raster scanned A4 images) and are essentially an optical publishing medium with potential for the distribution of digital databases, software packages or other forms of electronic publications required by users of micro/mini and mainframe computers.

To produce a CD ROM a publisher assembles his data on magnetic tape and sends it to a premastering facility. The data is then suitably reformatted and encoded and used to modulate a laser beam which records the data on a coated glass master disk. From the master disk, sub-masters are made and replicate polycarbonate CD ROMs are stamped out. The data preparation costs approximately £20,000, mastering of the disc costs £2,000-3,000 and replicate discs in quantity 1,000 plus cost about £4-5 each. The economics are therefore akin to those of conventional printing. The mastering and replication equipment is too expensive for in-house use so external facilities must be used.

CD ROM systems benefit from their ties with CADs in two ways. Firstly the physical characteristics of the media and most of the characteristics of the drives are standardised and secondly, CAD discs and players are in mass production now and the economies of scale benefit the producers of CD ROM drives which use identical components and CD ROM discs which are produced in the same way as CADs. The result is that CD ROM drives will eventually be available for approximately £200 instead of several thousand pounds.

The physical standardisation and the low cost of the drives will ensure that CD ROM becomes the dominant medium for optical publishing, shrugging off the challenges posed by hybrid videodiscs and Optical Read Only Memory (OROM) products which are described fully in the main account.

However, the standards referred to above relate only to physical attributes of CD ROM (diameter, data recording techniques, error detection and correction). There is another major area which needs to be standardised if CD ROM is to become an ideal generic publishing medium. This relates to how and where the information is laid out on the CD ROM - the logical file formats. If these can be agreed then different retrieval software packages working on different operating systems on different microcomputers can work with the same CD ROM. If not, publishers will have to produce different CD ROMs or use different software packages in order to market their CD ROM products to users of different computers and operating systems. DEC, TMS and others have proposed different logical file formats, each favouring one operating system and standards groups in the U.S. and Europe are attempting to reach a consensus.

Even without a standard file format, some 20 retrieval software companies have developed search software for CD ROM systems, the most notable of which are Battelle's MicroBASIS and BRS's Search. In Europe, Telesystemes are developing MicroQUESTEL and Harwell are developing MicroSTATUS for use with CD ROM.

The other area that has to be addressed is the hardware/software interface between the different CD ROM drives and the different microcomputers. There is no one standard interface and different companies are showing different approaches. Currently most CD ROM drives will interface to IBM PC XT or ATs or lookalikes and companies have produced interfaces to Apple computers and DEC's range of Rainbow, MicroVAX and VAX machines. Most microcomputers need a memory upgrade and the addition of controller cards and graphics cards and at present, in single evaluation quantities, the CD ROM drives plus the hardware and software needed to interface them to IBM microcomputers, cost in the order of £1,500-2,000. This price could drop to under £500 in volume.

In summary there are still some uncertainties and early producers of CD ROM databases will be acting as pioneers. In the absence of an agreed standard they will need to work to a de facto standard such as DEC's Uni-File and provide the end user with all the necessary hardware and software to bolt onto his existing PC. This will restrict the market to very high value products where the hardware costs can be lost in the subscription price and the benefits to the user of CD ROM are transparent. Heavy users of valuable databases such as medical and legal databases would be one category here. In the future, if file format standards are agreed and the hardware costs come down dramatically then a wider generic market for CD ROM publications in libraries and elsewhere may develop. CD ROM publishing at present offers potential and pitfalls in equal share.

AMH
Cimtech
January 1986

APPENDIX 2

COMMENTS ON EUROPEAN CD ROM STANDARDS FORUM

Introduction

An inaugural meeting of an unofficial group called the CD ROM Forum, which has been set up by by Learned Information, was held on 22nd January at the Commission. The main topic under discussion was CD ROM standardisation issues and the delegate list included senior representatives from the European information industry.

The work carried out by A. M. Hendley of Cimtech for DG XIII B in assessing proposals for CD ROM based information services had pointed to the crucial role that standardisation would play in the CD ROM area and to the need for European information providers to be better informed in this area. The Commission therefore invited A. M. Hendley to attend the meeting and report back to them on the main findings.

Learned Information and the Commission arranged for a full transcript of the proceedings to be made so there would be no value in detailing all the presentations here. We will confine ourselves to summarising the main points made by the speakers and in question sessions and to commenting on how the points made at the Forum and at the follow-up meeting held on 23rd January related to points raised by our assessment of the CD ROM proposals received by the Commission.

Content of the Forum

C. Vernimb - CEC DG XIII B

C. Vernimb of the Commission introduced the Forum by referring to the CD ROM proposals and the need to agree standards with the information providers providing the services. He therefore welcomed the initiative by Learned Information to provide a forum where standards issues could be discussed.

H. Collier - Learned Information

Harry Collier of Learned Information then introduced the forum by stating that there is a pressing need for some European input into the CD ROM standards making process and this was evidenced by the fact that over 70 delegates had come to this forum and a further 140 had asked for copies of the proceedings to be sent to them. For him the main purpose of the day was to explore all aspects of CD ROM standardisation and see what initiatives could usefully be made in Europe.

A. Proeme - Philips S & P Group

A. Proeme of Philips Subsystems & Peripherals Group then gave a presentation on established standards in the CD ROM area. This confirmed the points made in the attached report - namely that Philips and Sony have established de facto standards relating to the physical structure of CD ROM. There are two basic standards being readied. One is for business applications and the other is for the use of CD ROM in the home environment. The proposed physical format for data on CD ROMs for business use is contained in Philips' Yellow Book and they refer to this as Mode 1. The proposed physical format for data on CD ROMs for domestic use is contained in Philips' Red Book and they refer to this as Mode 2. In the home market Philips expect to announce a proposed physical and logical format standard and will launch a CD ROM player linked to a personal computer system - believed to be running MSX.

In the business field the next area for standardisation is the logical file structure which Philips, as a peripherals supplier, wishes to leave to the users of CD ROM systems to establish, i.e. computer hardware suppliers, operating system suppliers, software retrieval suppliers, information providers who wish to make use of CD ROM delivery systems and ultimately users of CD ROM products including libraries, computer departments, etc.

A number of proposed logical file formats have already been submitted to standards organisations in the U.S. by companies such as DEC, TMS, Laserdata, Reference Technology and others.

One other area which Proeme addressed was the need for standardisation of the hardware/software interface between the CD ROM drive and the microcomputer. At present these are proprietary but in the long run he felt that a de facto standard would emerge.

For the information providers in the audience, particularly those who have submitted proposals to the Commission for CD ROM products, the message from the Philips talk was definitely that while some aspects of CD ROM are standardised there is still much work to be done and some very major obstacles to be overcome before CD ROM products can easily be sold to a large generic market such as libraries or domestic users of PC's. At present the publisher would be restricting the size of his potential market if he opted for one particular file format, one retrieval software package, one CD ROM drive and interface and libraries would face having to buy different hardware and software to use CD ROMs from different suppliers.

J. Schwerin - InfoTech

Julie Schwerin then spoke on the need for CD ROM standards in publishing markets. She made the point that CD Audio is so successful precisely because it is standardised and any disc can be played on any player. She sees a need for a minimum standard in the CD ROM file format area on which publishers, software providers and others can then build value added features. At the base level are the physical standards governing the media which are established, then come the logical standards affecting the operating systems and drivers which are being discussed and then at the top level comes the application software which will not be standardised but tailored to specific applications to maximise the potential of CD ROM.

The main aim with CD ROM must be to build on the current installed base which most users have, i.e. an IBM PC or compatible running a standard operating system with a floppy disk drive. To that will need to be added a CD ROM drive, a controller, a device driver and a CD ROM with some standard file structure plus the necessary retrieval software.

At this point some discussion took place at which it was evident that for many the forum was serving an educational purpose with many delegates unclear of some of the issues that were being discussed. For information providers the standards question can be summarised as follows:

For specialised niche market publishing, i.e. distributing information to a set of users in-house, there is no real need for universal standards, they can simply follow one proprietary standard such as that proposed by DEC. For general publishing, however, it will be difficult for one publisher to sell CD ROM products and hardware to users unless they can demonstrate that other publishers in the same subject field are using the same file structure and hardware. For example, information providers in the chemical industry would have to agree to use one standard - unlikely since the IP's come from different countries and use different computer systems. Even this would not cater for the needs of general academic or public libraries or for large companies who would want to standardise on one CD ROM system.

Hence this session left us with the view that while CD ROMs could be used in a non-standard way, there would be much fragmentation of the marketplace and this would slow market growth and acceptance and might limit CD ROM usage to specialist information services.

B. Hatvani - Silver Platter Services

The final speaker before lunch was Bela Hatvani of Silver Platter Services who spoke on standards at the data location, format and labelling level. Much of what he had to say had already been covered in questions. He looked at the ideal requirements for CD ROM. The purchaser would want to be able to play different discs on one player and one computer system; the user would want one common set of command and search structures for a range of databases on CD ROM and the publisher wants one set of software which will run on all users' systems and access all his and other publishers' databases. He then went on in a rather unconvincing way to try and persuade publishers that all the hardware and software was available today to produce CD ROMs and his company could provide CD ROMs that can be read on any CD ROM drive and a range of PCs and operating systems.

The question that he did not address was what the trade-offs were when trying to develop universal CD ROMs at present with multiple directories so they can be used with a range of operating systems. More questions were raised than were answered by his talk and the discussion session ranged widely away from the standards issue and on to applications issues.

D. Martin - British Library

David Martin of the British Library was the next speaker, who gave a good practical talk, posing questions that he felt the forum should be addressing. The British Library have produced a demonstration CD ROM with bibliographic data on it and he made the point that as a publisher and user of such systems he would like to be able to ignore the lower level standardisation issues such as physical and file structures and concentrate on the applications level which will determine the value of CD ROM products in the future. It was generally concluded that this was the view of all publishers but unfortunately one of the lower level standardisation issues - namely logical file formats - was not yet resolved.

David's first point was that any lower level standards, however desirable for economic and other reasons, must not constrain those whose job it was to develop specialist applications software.

He, like Hatvani, posed an ideal scenario where the publisher would be able to produce a database in a form which could be used by users with a wide range of computers/drives/operating systems and retrieval software. The user could then buy data on CD ROM independently of the choice of software to access it.

However, unlike Hatvani, Martin granted that this was very unlikely to happen in practice and felt that for the foreseeable future publishers would have to assume a tight coupling between the choice of a CD ROM search system and the data layout or text content of the CD ROM.

At present the publisher has to choose an intermediary who will work to one file structure and offer one software package and then carry out all data formatting, premastering. For Martin it would be useful if, in the short term, the industry could agree upon a common format for publishers to provide their data in to these intermediaries.

A second requirement for Martin was a simple agreed process for what he called 'run and go' or the procedures involved in loading a CD ROM and setting the PC in CD ROM retrieval mode. This would be roughly akin to dial up and log on procedures for the on-line industry.

Thirdly, Martin pointed to the work that had been done in Europe on common command languages and recommended that existing sets should be used for CD ROM so users can control their own formats and build up their own command sets.

Finally, Martin looked at the data content and the presentation of files on the medium and their tagging and nomenclature. He felt that in future one attractive application for CD ROMs would be the publishing of many related small databases on one CD ROM and the provision of one standard access system for all those databases. He felt it might be possible to build on that and develop a simple set of guidelines for adoption by producers of such mixed database systems.

In conclusion he felt that in the short term there was no alternative but for publishers to provide both data and the retrieval software together on CD ROM as a complete package. This meant that the success of the package would depend as much on the software as on the data content and the size of the potential market for the product would depend on the degree of universality of the software or the number of different versions of the retrieval software that the publisher offered.

Martin's comments were very realistic and practical and hence of considerable value. Many of the points raised could usefully be handled in specific seminars or technical publications aimed at European information providers and it was felt that either the Commission or the ad hoc Optical Disc Forum should take note of these areas where further guidance was required. At the very least it would seem beneficial if information providers who have made proposals in the CD ROM area could be given guidance in these areas.

J. Schwerin, Infotech and P. Lord, Elsevier

The next presentation was on the issues in the U.S. for CD ROM standards and it was presented jointly by Philip Lord of Elsevier and Julie Schwerin of Info Tech. Julie described the involvement in CD ROM standards of four U.S. bodies:-

The National Information Standards Organisation (NISO) who are the official standards body and report to the American National Standards Institute

The unofficial but high powered High Sierra Group (HSG) representing intermediaries, software suppliers and computer suppliers

The U.S. Information Industry Association (IIA)

The American Library Association (ALA)

NISO first got together in February 1985 and reached consensus that there was a need for minimum file format standards for CD ROM that would allow interchangeability and optimal performance of CD ROM products and set a timescale of six months.

After six months the committee had only just been appointed and CD ROM products were already being launched. In addition, DEC and other companies had already proposed file format structures.

In October 1985 an ad hoc committee called the High Sierra Group met and decided to develop a working standard that they could then propose to NISO and use as an ad hoc standard until NISO finally recommended a standard.

The ALA have a technical standards committee and have set up an Optical Information Systems Committee to discuss CD ROM standards and draw up a set of functional requirements for library users.

The IIA Committee has also been set up to represent the U.S. information providers and has 3 main goals:

1. To work with other groups
2. To produce a list of functional requirements
3. To pass information onto their members

Overall, Julie highlighted three principles which were driving the push towards standardisation:

1. Interchangeability. One CD ROM should be usable on multiple operating systems so there is no need for multiple mastering of discs for different operating systems.

There should be no modifications necessary to the installed base of PC's - accommodation should be made for CD ROM at the device driver level.

One standard should allow for multiple boot blocks on one CD ROM otherwise publishers would need to provide a floppy disk with each CD ROM.

2. Performance. The standard should do nothing to add to the problem of slow access times for CD ROMs and any attempts to provide compatibility with write once and erasable disk systems should not restrict the uses to which CD ROM can be put.
3. Minimality. The standard will hopefully precede most applications and it will not be possible to envisage all the possible applications for CD ROM. Hence nothing in the standard should restrict possible future developments. The standard should be the minimum required to achieve the aims set out above.

R. Bilboul - Learned Information

Summing up the day, R. Bilboul pointed out that physical standards had been settled and at the logical level most activity in the U.S. is at the computer hardware/software level with publisher and user interest coming from the IIA and ALA. In Europe the only link at the hardware level was ECMA, there was no user forum which was why LI had called the Optical Disk Forum together.

The main point now was where did the forum and the European industry go from here? The options were to wait and see what standards emerged from the U.S. or to develop European standards for information providers. There was clearly a need for an on-going education process to point out the main important issues and to encourage co-operation between European information providers. Selected experts would be asked to stay over until the 23rd to discuss what initiatives should be made.

Looking at specific initiatives for the Forum to undertake, he felt there was a need to inform European information providers about standards activities in the U.S. and then to feed back any specifically European viewpoints to the U.S. standard making bodies while there was still time to influence decisions.

Bilboul wanted to see if the Optical Disk Forum could gain access to the IIA and ALA functional requirements and felt that a sub-committee of the OD Forum should add any European requirements to these and assess and comment upon any proposals that came from the High Sierra Group and then pass these findings on to the members of the Forum at a subsequent meeting.

These proposals were generally endorsed by the Forum on the understanding that the details of any plan of action would be worked out by the subcommittee on the following day and the forum dispersed.

In general the resolutions were valid. There is a need for a continuing educational programme for information providers - particularly those who have submitted proposals to the Commission. There is also a need for the European industry as a whole to be informed of all the standards issues and for the European viewpoint to be made known on the necessary standards bodies before it is too late and they are presented with a de facto standard that may discriminate against European interests.