

## TECHNOLOGY AND TIPS FOR PAPER-TO-CD/DVD-ROM CONVERSION OF SELECTED COLLECTION IN ACADEMIC LIBRARIES

---

**K. S. Rawat**  
P. K. Kelkar Library  
Indian Institute of Technology  
Kanpur  
Email: ksr@iitk.ac.in

*Academic libraries need to retain collection of age-old books, journals, etc., for longer period. Preservation of select highly useful and valuable materials on optical discs has higher benefits than merely reducing space problem. The chaff can either be consigned to remote warehouses or weeded out with other unwanted editions.*

### INTRODUCTION

Preservation of shared knowledge and information, ever since the dawn of human history, has made basically few shifts as far as storage media are concerned. The earliest records of knowledge have been found on rocks and clay tablets, followed by papyrus rolls, animal hides, strips of wood, etc., and finally, on paper, believed to have been first invented in China. Except for the paper, each previous medium must have been difficult to handle and write on with calligraphic effect, therefore, inconvenient and inadequate to record the variety of knowledge produced in the age it belonged. Paper, especially after the invention of printing press, has been the most convenient medium to store and spread all kinds of knowledge and information. However, reckoning the size of information produced in the later half of the 20<sup>th</sup> century, and the scale at which it is exploding now, paper also has perhaps outlived its archival use any longer. The greatest problem with paper media is that of storage and access.

### NEW STORAGE MEDIA

The end of the twentieth-century has seen the invention of advanced and new storage media

technologies such as: magnetic tape cartridges, floppy disks, computer hard discs, and finally, CD-ROM and DVD-ROM discs. Microforms apparently do not seem to figure any more as storage media. While magnetic tapes permit only sequential access, floppy disks have only about 1.44MB storage capacity. Large hard drives may be good for current online bibliographical databases. Therefore, removable CD/DVD-ROM discs, logical progressions from floppy disks, maybe the preferred choice for use of document preservation in academic libraries needing to retain, economically, some of their age-old, brittle and pest-eaten but still valuable collection of books, journals and other reading material.

### OPTICAL DISC DENSITIES

Each CD-ROM disc is capable of storing about 650 MB of information, equivalent to contents of 3,25,000 pages of text (at the rate of 500 pages per MB). Therefore, CD-ROM discs may be good to preserve contents of one or more individual books. However, to preserve the contents of large back volumes of journals, like "*Chemische Berichte*", which is brought out since the year 1880 till date, what could be the media? Either large, costly hard drives, or economical DVD-ROM discs. Each DVD-ROM disc for text use comes with storage density from 2.6GB to 4.7GB, and much more, [1] [2] [3]. One GB can store as many as 500,000 text pages [4]. Therefore, a single DVD disc is enough to preserve the contents of one or more

journals in libraries. Until lately, DVD-ROM (Digital Video) discs have been in use only for audio-visual data, such as motion pictures, etc. But the same, now described as Digital Versatile Discs can store large text and graphics as well.

### **WHERE TO BEGIN IN DATA CONVERSION?**

Libraries may begin data conversion first with PhD theses, internal publications and reports, etc. Later they may go for back volumes of serials, books, etc., which do not come under copyright. At the end, selected copyrighted material may be taken up for conversion depending on priority. Whenever necessary, it should be done only after taking proper permission from the concerned publishers.

Fortunately, copyright laws in most countries are liberal to libraries in respect of making "fair use" copies-digital or otherwise—which includes archiving lost, stolen, damaged or deteriorating works. In particular, Indian copyright laws [5], too, though dealing with only print works, are equally unambiguous in this respect. Therefore, generally, libraries may not face any legal hurdles in proceeding with the suggested data conversions.

### **DATABASES ON WEB OR ON OPTICAL DISCS**

Before starting out, it needs to be checked out, whether some of the books to be preserved on optical discs are already available free on the Internet for full-text access. Many of them are religious texts, classics in humanities, social sciences, science and engineering, etc. which are, mostly out-of- copyright or are "public domain" works (Generally, all works published before December 31, 1922 are in the public domain). For instance, the website "www.academicinfo.net" provides a good deal of them. New Zealand Digital Library is another non-commercial portal providing full-text reports and other literature for free. Users may find out many more sites for free digital documents. Each such title may be left out of the "Project Paper-to-CD/DVD Conversion".

Secondly, libraries should figure out whether it is more expensive to organize a book or journal on CD/DVD disc with efficient access mechanism than to purchase the same from the market, as many publishers have already brought out optical disc editions of their rare and archival publications for sale worldwide. They are also making CD/DVD editions of most of their current publications. In the author's opinion, the latter must be more economical.

Besides, many dictionaries, encyclopedias, directories, atlases and a variety of other reference tools are available, either on the Web for free access or on optical discs for instant acquisition into a library's digital collection. For free Web editions, one may carry out a directory search on "www.dictionary.com", "www.yahoo.com", "www.google.com", "www.altavista.com", etc., and publishers' catalogues for optical disc editions. Libraries need to be very careful and selective in purchasing optical disc editions, and prefer them to print editions for the attached tangible advantages.

### **ADVANTAGES OF CD/DVD DATABASE HOLDINGS**

A major advantage is that access to a CD or DVD database is fast, accurate and thorough. A proper search by a researcher may get the information needed in seconds. In contrast, traditional print services may take months for the same task, with much of information left out. Libraries cannot afford to continue with the rather slow manual information services in the fast paced environment of today. The present and the future belong to computer savvy professional users needing to stay on in the mainstream of their pursuits. Or, else they're bound to fall behind their peers elsewhere, particularly in developing countries like China, Korea, etc.

Another advantage of CD/DVD database holdings is frequency of access. With CD/DVD databases permanently mounted on jukeboxes and/or DVD-ROM towers, users could access

them with greater freedom. The same is not possible with Internet editions, whether annual subscription or per-connect-hour tariff for access is paid. For instance, if the library is using DIALOG or BRS resources, users are under tremendous pressure during connect time, since the library has to pay for every extra minute or second wasted.

On the other hand online databases have certain limitations. Firstly, not all journals or journal volumes are available on the Net. Secondly, the library has to subscribe to print or online editions to obtain access license. Thirdly, web journals permit only limited concurrent access. For instance, IEEE Xplore permits only four concurrent users at a time, of course, DIALOG charges extra tariff for each additional user password. Downloads and printouts cost extra. And, fourthly, if the library drops subscription of a journal, access permission to all its online editions is also closed, leaving users high and dry.

Some academic libraries currently lack building space. Paper-to-CD/DVD-ROM conversion project can reduce this immediate and compelling problem. The chaff can be consigned to remote warehouses or weeded out with other unwanted editions, to reclaim precious space for reading tables and fresh holdings. More than space saving, it is what users may gain in terms of information in real time that is important.

### **DISADVANTAGES OF OPTICAL DISCS**

There are some disadvantages with commercial CD/DVD-ROM databases. Firstly, they are not updated timely, while some web journals or indexing services are available online even before their print editions roll out of the press. For example, Engineering Index on disc (COMPENDEX\*PLUS) is updated quarterly, and Dissertation Abstracts on Disc (DAO), only once in an year.

Secondly, if not acquired from a single producer, CD/DVD-ROM databases require skills in using

different search engines. Making menu-driven, index searching, or single keyword searching at command prompt, may be easy to use in each system. But, users must make command or Boolean searches, to get at exact or less abstract hits. Sometimes users have been found keyboarding DIALOG syntax instead of SPIRS one to search into MathSciDisc or PsycLIT from SilverPlatter, and vice versa. Both the search engines are so dissimilar in constructs. Again, generally, users keyboard a collocation of keywords at the command prompt (including particles and other abstract terms, more often than not). They are flummoxed to discover that one access engine parses the search statement with "OR" operator between words, and the other with "AND" operator. The former results in too many hits, and the latter, in too few. Therefore, all new users require basic education in the techniques of bibliographic database searching [6]. Fourthly, if adequate jukeboxes are not provided for permanent mounting of optical databases, it is a great hassle for users to switch from one disc to another for either a different file or part of the same file. By comparison, on online services, like DIALOG's, the process is just a click away.

Yet another disadvantage with DVDs are, like CDs, they are vulnerable to damage by dust, fingerprints and scratches. Therefore, they need very careful handling. Libraries will have to make backup copy of each CD or DVD disc to be kept in safe custody, so fresh copies can be made.

### **COST FACTORS INVOLVED**

The cost of the "Project Paper-to-CD/DVD Conversion" will cover the development and maintenance of CD/DVD-ROM database collection. Major items will include cost of commercial CD/DVD databases, cost of computer hardware and software, cost of jukeboxes or CD/DVD towers and CD/DVD writers, cost of ergonomic furniture and appropriate physical environment to conveniently locate and use optical discs, writable blank CD and DVD discs, data conversion cost including cost of training staff

to design schema, handle metadata, write retrieval programs and make data entries, either by keyboarding or through optical scanners, etc., and cost of making backup copy of each CD/DVD disc, etc. Lastly, cost of user education programs on continuing basis too should be included.

The project, though initially expensive, may pay off immensely in the long run. Incidentally, the life-span of a recorded CD or DVD disc is believed to be about thirty years [7].

## REFERENCES

1. KURATA (Yokio). Trends of large-capacity optical disc. *Sharp Technical Journal*. 72; 1998; 9-12.
2. DUTTA-ROY (Amitava). Computers. *IEEE Spectrum*. 36; 1999; 6.
3. HUA (Zhong-Yi); CHEN (Guo-Rong); CHEN (Dian-Yong) and XU (Hua-Hua). New storage material for rewritable blue-light DVD. *IEEE trans. on consumer electronics*. 45, 1; 1999; 97-100.
4. HOLMSTROM (Dave). Examine your strategy, weigh your options in choosing removable storage. The Internet. Strategic Search Corp.
5. Ministry of Law (India). The Copyright Act, 1957. Manager of Publications, Delhi; p 20-24.
6. EFTHIMIADIS (Efthimis N). Study of end-user behaviour in searching CD-ROM bibliographic databases. Proceedings of the 15th National Online Meeting, New York. Learned Information: Medford, NJ. 1994; 113-120.
7. HOSMSTROM (Dave). Ibid