CREATING LOCAL JOURNAL DATABASE FOR DOCUMENT DELIVERY

Mounir A. Khalil

Science Library

City College of CUNY

New York, New York 10031

E-Mail: Moucc@CUNYVM

Abstract: This paper describes the application of the optical technology for developing a local journal article database for electronic paperless document delivery service. The system provides instant access to requested articles. Includes price lists of hardware and software needed for the database. Copyright issues related to document delivery will be touched upon.

1. INTRODUCTION:

In the last two decades, we have seen important changes in the way we use information. Microcomputers have become indispensable tools in our daily life. Of most importance is the dramatic decrease in computer hardware costs and increase in computer hardware power. Nowdays,

287

microcomputers have speeds and memory capacities comparable to the most powerful minicomputers a few years ago. There is increasing use of small computer systems, such as personal computers and Unix-based multiuser workstations. Individual users of computer systems do not work in isolation and want to access data from several sources in preparation of a document. Use of telecommunications using networking is the compelling factor for interconnection among all these small systems. In addition, digital optical technology has made it possible to use optical disks for storing huge amount of data, for example, 12-inch disks can store up to 9 gigabytes of data (9), enough to store the contents of 150,000 pages.Stored data on optical disks can be easily replicated back to paper. The technological problems will soon disappear. What remains is how to plan and design an efficient system and educate the users about the power and potential of this new technology.

Information technology has enhanced many aspects of daily library operations by automating the organization, storage , and retrieval of information. Technologies such as public access cataloges and online databases have expanded access to bibliographic information, providing faster and more efficient access to wider materials.

However, despite the ease with which bibliographic refer-ences
are found, users often experience long delays in receiving the
original document for the following reasons:

1) Financial pressures which have forced libraries to
   shift from the acquisitions for local collections to
   providing access to those locate at remote institut-
   ion, such that larger percentages of documents must
   be requested and delivered from other sources, and

2) the application of information technology is prima-
   rily to the retrieval of bibliographic information,
   rather than the physical delivery of actual doc-
   uments to end-users. The discrepancy between bib-
   liographic access and document supply threatens to
   become even more acute as a result of the increase
   in bibliographic access engendered by recent develop-
   ments of end-user searching of locally mounted cit-
   ation database, CD-ROMs, and network-accessible
   library catalogues. (4)

The gap between access and supply can be poten-
tiallly overcome through electronic document delivery and
creating a local journal database of needed articles. This
will allow the automated process of retrieving full-text of
the article and its delivery to the user to be easy. In other
words this process will allow library staff and researchers

to select and obtain journal articles from their PC-work-
stations.

Integrated local journal article database system
will form a crucial part of the electronic libraries that have
been envisioned for development on the reserach computer
networks that link academic and research libraries throughout
the world. In America, President Clinton has the NREN
initiative to create information networks capable of providing
easy access to vast information resources including full-
text.

2- WHY IS IT NECESSARY FOR LIBRARIES TO HAVE SUCH A LOCAL
   JOURNAL DATABASE?:

Over the last few years , substantial increases in
the cost of serial publications have forced libraries to
cancel subscriptions to journal in many disciplines. At the
same time, scholars have been demanding greater access to more
publications. As mentioned previously, the developing of a
local journal database using, preferably an optical disk over
other disks, it will make it easy for the faculty and students
as well as any users to search the citation in the database
and retrieve a copy of full-text from the system. Copyright
fee could be paid either by the library or user. The library

could fax the copy of the article or send it electronically to their home/office computers.

While economic pressures may cause librarires to cancel journal subscriptions, the existance of such a system will make it partially compansate for those cancellations. Such a developed integrated system will fill the gap. Libraries have been trying to supplement their collections with compact disks containing indexes and abstracts from selected journals. With a CD-Index, users find more and more items they want, and it is more and more of what they do not have in their libraries. The development of such a local journal database system including journal articles of the highest demand for document delivery will benefit the users.

Traditionally, an ILL Division receives an ILL request and tries to locate the journal if the library subscribes to it , where it shelved and whether the item is immediately available. When the journal issue is found it has to be physically , retrieved from the shelf and taken to the photocopier in order to make a copy and send it by fax or mail. This is time-consuming.

3- ADVANTAGES:
Using the optical storage technology to store

large numbers of digitized journal articles received from different resources amd stored in a local database which could be integrated in other library database would make it easy for faculty and students to use without knowing the various steps of interfaces, search conventions or database formats. In addition, it will improve the library services by using the new technologies. The advantages are as follows:

1- Time and cost savings as compared with ILL and manual document deliery services.

2- Instant availability of the full-text of the requested article 24 hours a day, 7 days a week.

3- Quick full-text search which eliminates hours of library research through library indexes and shelves.

4- Retrieving documents by a number of users with access to the data communications network.

5- No storage or maintenance. For example, a jukebox-like system typically can hold papers that would occupy more than one thousand four-drawer file cabinets.

6- Gaining better understanding of how faculty and students use full-text database and how much use is made of online as to printing copies of journal articles locally.

7- If the library uses electronic journals, it can be integrated into the system.

8- Reducing the duplication of extensive library material among other libraries in the same subject areas.

9- Enhancements of resource sharings among libraries which will strengthen the academic program as better library services are developed.

10- Free viewing of the image of the journal article.

4-IMAGE DATABASES:

Nowadays image databases use optical disk storage capabilities preferably over other storage media. Optical disk technology provides a sensible solution not only to modernize document delivery but to transform it into an electronic paperless document delivery services. Users can retrieve their articles from their own computers in any place of the world over telephone line using a modem or network (Internet). Some document delivery vendors such as (ADONIS, UMI,...etc.) and Carnegie Mellon University, The University of Michigan and Cornell University are a few examples of libraries which use this technology in document delivery.

How is an image database different from full-text?

An image is an electronic photograph of each page composed of black and white dots. It contains all text, graphics, halftones and font data of the original page exactly as published. Each and every page of the complete article can be stored on CD-ROM or WORM disk. (9)

When the image is scanned by the scanner it produces the black and white dot composition of the page (300 dpi). Those dots are compressed (CCITT Group 4) and stored on CD-ROM. The process then is reversed when an image is retrieved from the disk for display or printing.

Don Willis (8),in his article " New Optical Technologies & Media for Information Access", indicates that it is actually less expensive to store high quality page images on CD-ROM than storing text only pages in ASCII format on a hard disk of a computer. For example, a typical image page requires 100,000 bytes of storage space, while that same page in ASCII text is 3,000 bytes. This means that ASCII has a 20 to 1 advantage in space over image. The cost differential favors CD-ROM because a hard disk costs 400 times more than CD-ROM storage. A hard disk comes in at about $ 6 or $ 10 per megabytes, while CD-ROM is about $0.02 megabyte. So storing the page image on CD-ROM is consider-ably less costly than storing the the same page in a character code

format on magnetic disk.

Libraries should use to store journal articles which they believe are cost-effective. Improved communications, automated ordering and finding procedures will allow other libraries to be delivered quickly and cost effective in the near future.

5-FUNCTIONS OF DIGITAL IMAGING SYSTEM : (1)

1-Journal article are captured on a scanner which converts image information into digital format. Index information is also generated.

2-The image is stored on an optical disk.

3-Application software and the stored index information reside in the computer database in order to manage the stored data.

4-Document sharing takes place over a computer network such as LAN, WAN, using a image-capable workstation, reducing the need to generate paper copies.

6- TECHNICAL REQUIREMENTS (HARDWARE & SOFTWARE)

As stated earlier, with today's advanced personal computers, a typical configuration will consist of a power-

ful PC with extended/expanded memory, A CD-Drive and laser.

Now there are many vendors in the market who can provide systems tailored to any library's budget or special need. Librarians can consult: <u>(6) The Handbook of Optical Memory</u> <u>Systems)</u> by C. Peter Waegemann which has a chapter on vendors, consultants for business, government, banks, insurance,....etc.).

Librarians can make their own CD-ROM databse if they need a few copies before large-scale replication. There are low-cost CD-ROM available in the market which costs less than $ 6.000 from Yamaha, JVC, Philips,SONY. In addition, a powerful system configuration (A 486 CPU, 8 MB RAM, 600 MB hard disk) is recommended. Besides, indexing and retrieval software, scanner, OCR ,CD-ROM Players, CD-recording (CD-ROM). It takes 60-72 minutes to record a files on CD-ROM. (Please check guidelines for ordering hardware and software for CD-ROMS).

Dr. Jeffrey W. Schram, President of Facstore, Inc., New Jersey, United States has provided me with the needed hardware and software and their prices. He has indicated that librarians can use their library micro-computers, laser printers and Novell networks and can buy scanner, imaging software, OCR software and fax card.

HARDWARE & SOFTWARE		PRICE
PC 486	Less than	$ 2,000
Scanner with automatic feed		$ 2,000
Laser Printer		$ 1,800
10 User Novell network		$ 2,000
Novell server with 2 gigabytes		$ 8,000
Imaging software and OCR		$17,000
Faxing capabilities with hardware and software		$ 2,500
Installation fee for 10 user Novell Network		$ 400
	Total	$35,300

Plus training fees   ?

As the demand increases and need more for storage space and speed, the hard disk may be replaced with optical jukebox. This replacement for a 15 Gigabyte Jukebox, will cost approximately $ 29,000. This includes two optical drives and 30 optical platters.

For larger libraries, A UNIX based SUN is recommended . The system will have the following:

SUN Sparc 10 Server $ 24,000

25 page per minute scanner with automatic feed $ 12,000

If the system uses PC's , Additional Pc's should be purchased
to act as X Windows servers. $ 9,000@

UNIX software for OCR of images amd full text search$ 39,000

This system might cost about $ 150,000.

If this system is bought it requires time and money to
maintain.

If the decision is to create CD-ROM by the institution
rather than using the service of a service bureau. The
following chekclist is designed to help librarians in
selecting the needed hardware and software for CD-ROM. The
follwoing items should be taken into consideration before
deciding the purchase:

1- Try to look for the software packages that will be
   necessary before selecting hardware for the system
   applications.

2- Take into consideration software, peripherals,
   operational and supply, and training costs.

3- Do not wait long for the perfect CD-ROM system
   because there are always rapid changes and constant
   changes in the computer industry. Choose the hard-
   ware and software that meet the library's present

needs and allows for expanding in the future.

As mentioned before, select software first before selecting hardware.

7- HOW TO SELECT A VENDOR ?:

1- Would one like to buy from a dealer, manuafacturer, distributor, or other vendor?

2- How long has the vendor been in the computer business? Less than one year, one to five years, five to ten years or more than ten years?

3- Does the vendor supply the whole CD-ROM system?

- hardware only

- software only

- hardware and software as a package.

4- Is the vendor finacially stable?

5- Is the vendor local or nation-wide?

6- How many CD-ROM systems has the vendor installed? Less than one hundred, several hundred, more than a thousand, or several thousand?

7- Does the vendor provide service?

8- How long does it take to get service? Less than twenty-four to forty-eight hours, more than forty eight hours?

9- Does the vendor provide technical support and training to library staff on continuous basis?

10- Does vendor provide free installation and intial
    training?

11- Ask a vendor to see a sample of license agreement.
    Is it satisfactory? Does it meet your expections?

8- Conclusion:

Retrieving journal articles from local image database
can be done within minutes. Actually the cost for retieving
article from the local database will be trivial in order to
pay for the copyright fee. Viewing an article is free unless
the user has to pay for the printing. In addition,a
distributed model of client and server will provide the
greatest leverage for wide use of image databases. Therefore,
faxing or sending an article over Internet will be much
easier. The technology is here today to researchers, other
users and librarians have the opportunity to have articles
faxed to them without leaving their location. It is widely
accepted that it would be a tremendous substitute to the
currently extensive interlibrary loaning of material.

This local journal database should be concentrated
on journal articles in the high demand subject areas to meet
the needs of the students and faculty. In the near future
researchers will have virtual access to local databases

throughout the world.

9-COPYRIGHT ISSUES:

Copyright is one of the main issue which should be resolved before building a local journal database. A word of advice, permissions from journal publishers should be obtained in advance before creating journal article database.

10-REFERENCES:

1- Greco, E.J.; "Hybrid strategy guarantees future document access" International Journal of Micrographics and Optical Technology, 10 (2): 51-54, 1992.

2- Lynch, Clifford "Electronic imaging in the university environment: issues and applications International Electronic Imaging Exposition & Conference, Mass, 1988: 596-601.

3- Stern, Barrie & Campier, H. "ADONIS- Document delivery in the CD-ROM age ", Interlending and Document Supply, 18 (3): 79-85 (1990).

4- Swain, Leigh & Cleveland, Gary "Electronic document delivery and libraries " INSPEL, 26 (2): 169-184, 1992.

5- Thomas, George & Walker, Frank L. "A Prototype electronic document delivery system" Proceedings , 53rd

ASIS Annual Meeting , 1990, 29-38.

6- Waegemann, C. Peter, The Handbook of optical Memory Systems, Mass, Optical Disk Institute 1991.

7- Walker, Frank & Thoma,George " A Hybrid system for retrieval of online biomedical citations and optical disk-based documents" Proceedings of Optical Publsihing & Storage, 1987, 179-191.

8- Willis, Don "Imaging: the information access tool of the nineties" Proceedings of the Thirteenth National Online Meeting, May 1992, 435-444.

9- Wust, Ruth "Optical disk based document delivery at the Congressional Research Service of the Library of Congress." International Essen Symposium, Essen 1991, 169-177.