Historical ecology: LSU's electronic imaging laboratory

By Richard Condrey, Faye Phillips, and Tony Presti

Preserving historical records in electronic form

As society and scientists wrestle with uncertainty over global climate change, a wealth of potentially valuable information lies largely unused and deteriorating. These important links with our historical ecology are the written and illustrated record. They are contained in the rare and unindexed first-hand accounts of our ancestors. While the record is diffuse and sometimes misleading, it is also rich and insightful.

Because the current condition of these works is an impediment to their scientific study, a project is being completed to convert this record into a machine readable, searchable, and speakable form. Scientists and librarians at Louisiana State University (LSU), have successfully converted B. F. French's 1846, five volume, *Historical Collections of Louisiana* into an electronic text.

Funding

The electronic imaging of French's volumes was accomplished through special funding. In 1991 the Special Collections Division of the LSU Libraries, received a grant of $285,000 from the Louisiana Education Quality Support Fund (LEQSF) to establish an electronic imaging laboratory, a local area network, and a disabled users adaptive computer center. Because the LEQSF funds only equipment purchases and not staff, the project was staffed principally by graduate assistants funded from a variety of soft-money sources.

Importance of the project

French's *Historical Collections of Louisiana*, printed between 1846 and 1853, were chosen for the pilot project because of the information they contain, the relative rarity of the publication, and because the publication is in the public domain, no longer under copyright. French's volumes contain early travel logs of explorers to Louisiana and the Mississippi River in some of the first transcriptions from the French and Spanish. Co-principal investigator for the project, Richard Condrey, has used the publication extensively in his work on the historical ecology of Louisiana.

The process

The scanning and optical character recognition (OCR) was accomplished using Xerox's Kurzweil K5200 integrated system.

1. Each page was carefully turned and placed face down on the system's book-friendly scanner

2. A TIF image file of the volume was created at a rate of about 45 seconds per page

3. After a volume had been scanned, a verification file was developed, which "taught" the machine the correct identity of unclear characters (procedure took three to five hours)

4. With the OCR software a text file of the images was created

5. Text files were edited, using WordPerfect 5.1, in three steps:

a. the text file was run through the spell-check utility of WordPerfect, locating incorrect spellings. Because orthography in the original volume is not standard, not all of the "errors" detected by the spell-check program were actually scanning or OCR errors; each had to be cross-checked to the scanned image, a simple process of toggling from the text to the image files on the workstation.

Richard Condrey is associate professor, Coastal Fisheries Institute, Center for Coastal, Energy & Environmental Resources and Faye Phillips is head, Louisiana and Lower Mississippi Valley Collections, LSU Libraries, both at Louisiana State University. Tony Presti is with Key Systems Incorporated, Baton Rouge.

Pick a system. *Any* system.

At EBSCO we think one of the most important decisions you'll make for your library is that of choosing quality subscription service. Equally important is choosing the automated system that best fits your library's needs. When you choose EBSCO for subscription service, we respect your library automation decision and provide an array of electronic services that complements your system's capabilities. Our services are compatible with most major systems — so you can have superior subscription service *and* your system of choice, too. Call us today to learn more about our independence-oriented library automation services.

P.O. Box 1943 • Birmingham, AL 35201
(205) 991-6600 • Fax: (205) 995-1636

Where library automation is a liberating experience.

More Information, More Value— And A Brand New Look!

Wilson Library Bulletin—a must for every librarian.

With lively, engaging commentary on the developments, events, products, and trends that affect today's professional environment, *WLB* gives you more information and more value.

New in 1993!

- "Brazen Overtures," GraceAnne DeCandido's editorials on library and literary issues
- Lee Ratzan in "The Internet Cafe"
- "Coda," a closing essay on titles of topical interest.

Plus... *WLB* continues with Jim Rettig's reviews of reference titles, Patty Campbell's reviews of books from the independent press, Norman Steven's wry approach to "Our Profession," and the other *WLB* columnists whose words you rely on.

For more information call Toll-Free 800-367-6770.
Outside of the U.S. and Canada, call 718-588-8400. Fax 718-590-1617.

THE H.W. WILSON COMPANY *Sharing Knowledge Since 1898*
950 University Avenue • Bronx, NY 10452 • 800-367-6770

b. the text file was read, to search for correctly spelled but incorrect words, italics were added as needed and French and Spanish words were verified,

c. a draft was produced and cross-checked against the original text.

The product

The text is stored on erasable 1.0 gigabyte optical cartridges. These were used throughout the editing process. When the final edited version of the French volumes is completed, the master file will be stored on 940MB WORM (write-once-read-many) optical cartridges. These master files will be utilized as "archival" backups and as technology changes the data on the cartridges can be transferred to the next generation of storage medium whatever that might be.

The final product to be produced will be a CD-ROM containing parallel text and image files for all five volumes of the B. F. French. The CD-ROM will contain the navigational software necessary to operate the disk on standard CD-ROM drives. IBM BookManager BUILD utility will provide the necessary software capability for using the CD-ROM. Because the French volumes were never indexed, having a keyword or free-text searchable text file will make the publication a much more usable research tool. Pages may be printed out from the edited version or from the image version, or both. The CD-ROM edition will be available in December 1993 for approximately $40.

Electronic Imaging Laboratory

Equipment	*Notes*
1 *KSI ScanStation II*	80386-33 ISA buss computers, configured with 8MB RAM, 211MB SCSI-2 hard disk and 1.44MB and 1.2MB floppy drives in a super tower enclosure with a 400 watt power supply; bundled with Windows 3.1, DOS 5.0, mouse and enhanced keyboard; internal 1.0 GB multifunction optical drive, supporting the 650MB ISO 5.25" standard WORM format and 650MB and 1GB 5.25" erasable formats; storage subsystems interfaced via a SCSI-2 buss master controller with a 5.7MB per second transfer speed; a 19" 120 dpi monochrome monitor with glare reduction and 16 grey scale support provided by the matched high resolution controller.
1 *KSI ImageStation II*	identical to the above, except for addition of image coprocessor with 8MB RAM and videolink card.
1 *Hewlett-Packard LaserJet III* printer	
1 *Fujitsu VM2200* printer	
1 *Kurzweil K5000* Windows Document Reader:	with a book friendly edge on the scanner.
1 *Kurzweil K6200* Windows Document Reader	subsystem with a flatbed optical scanner.
1 *KSI ImageScan II scanner*	with 400 dpi resolution
1 *KSI ImagePrint printer*	
1 *Kurzweil K5200*	with Optical Character Recognition interface.
1 *Portable 80386-40 CPU*	with 4MB RAM, 105MB hard disk, 1.44MB floppy and a KSI internal 1.0GB multifunction optical drive, an 800 by 600 VGA backlit flat panel LCK with 32 grey scales and external VGA port. enhanced keyboard, Windows 3.1, DOS 5.0 and mouse, Logitec hand scanner and autostitch software.

Process costs

Combined time for scanning, editing, and correcting ran between 23 to 34 minutes per page. Personnel costs for scanning, editing, and correcting a 250-page volume were approximately $1,925.

Copyright

To date all projects planned for the EIL at LSU involve only those publications and materials in the public domain. Institutions must remain aware of copyright laws and comply with those applicable to their research products. As we investigate the electronic publication of copyrighted items, reports will be issued. *(con't page 448)*

PHONEDISC® TAKES THE LEGWORK OUT OF PHONE BOOKS.

The phone company's tip to "*Let your fingers do the walking*" works just fine when you're looking for someone in the same town. But this advice can get mighty frustrating when you're looking through hundreds of phone books for your Great Aunt Emma in Iowa. Or when your fingers are skipping across thousands of pages trying to find a business associate you lost track of after he relocated to some unknown state! Face it: Your fingertips were simply not meant for cross-country travel.

Fortunately, there is an alternative. **DDA's PhoneDisc® USA Residential** and **PhoneDisc® USA Business** CD-ROMs pack **80 million residential** names and addresses and **9.6 million business** listings in less room than your local directory! That's more listings than any other comparably priced telephone directory published on CD-ROM.

Imagine: The equivalent of several tons of printed residential and business phone books–weighing just ounces–at your beck and call. **PhoneDisc®** is a lot easier to use than paper directories, too: Simply type in the person's name, and the powerful built-in search program will find his/her address and phone number instantly!

USA Residential *USA Business*

PhoneDisc® USA Residential is a two-disc set (Eastern and Western United States) containing listed residential phone number in the country. If the person you are looking for is in a telephone directory—**PhoneDisc®** can help you. It is invaluable for locating people who have moved—like your former roommates, army buddies or (let's admit it) your old sweethearts. To save time, or eliminate duplicate names, you can limit your search to certain states, cities, zip codes or area codes.

PhoneDisc® USA Business lists large, medium, small and even tiny businesses. In addition to looking up a company by name, its reverse indexes allow you to search by business type, SIC code, street, address or phone number. Finding a supplier, starting a job search or checking references was never easier. **PhoneDisc® USA Residential** and **PhoneDisc® USA Business** are the most complete, accurate and economical telephone databases available today! And they'll stay that way, because they are updated quarterly.

Join the thousands of CD-ROM owners who have purchased **PhoneDisc.®** Call **Toll-Free** and order today! Only **$99** for the each or **$159** for both. You get full technical support, as well as an unconditional 30 day Money Back Guarantee.

If you need to search residential listings by more than last name, **PhoneDisc® Reverse** extends your search power! With **PhoneDisc® Reverse** you get the same high quality listings and powerful search software as **PhoneDisc® Residential and Business** plus reverse indexes for residences. Buy the entire U.S. on five discs for **$349**. Buy one region on one disc for **$129**.

See your nearest PhoneDisc® dealer or call to order.

TO ORDER: CALL TOLL-FREE 1-800-284-8353

DDA PHONEDISC CD-ROM

We Got Your Number.

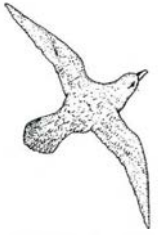
Digital Directory Assistance, Inc.
5161 River Road Bldg. 6
Bethesda, MD 20816

PhoneDisc is a registered trademark of Digital Directory Assistance, Inc.

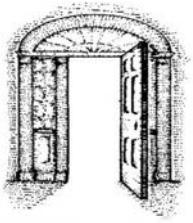
A Better Way To Search Databases

We started in 1985, database searchers committed to better search software design. We became the premier vendor of Medline, then expanded our catalog to other databases. Last year we won Information World Review's PRODUCT OF THE YEAR for faster, easier search software. But a better way means meeting the evolving needs – individual and campus wide – of today's library users.

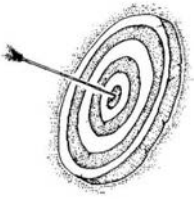
Announcing OVID: a database interface so flexible it molds itself to your search environment.



With OVID you're free to move from one operating system to another without retraining. OVID's Common User Interface assures identical functionality in DOS, Windows and UNIX.

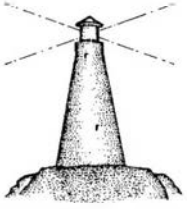


A haven for beginners, OVID's Easy Mode has on screen prompts. The more experienced can pull-down menus showing an array of search options. Experts will feel at home using online syntax.



Search with natural language if you like. OVID mapping cuts through the mystery of controlled vocabularies, homing in on precisely matching subject headings.

There's an unprecedented array of search tools – indexes, thesauri, limits and fields – many never before available in an interface. They're all standard OVID features.



HELP for every search function is context-sensitive and on screen, never more than a key-stroke or mouse click away.

OVID. A better way to search ERIC, Current Contents, PsycINFO®, Medline, Readers' Guide Abstracts, EMBASE and more.

CD PLUS Technologies

New York 800-950-2035/212-563-3006 London 44-(0)81-748-3777 Amsterdam 020-672-0242

Articles may include directions for reaching Russian and Eastern European sites by electronic mail, computer networks, databases, information technology, and the online industry within this area.

Newsletters

Subscription requests for newsletters are generally sent to the individual responsible for its distribution.

- CAROLINA@CSEARN, a weekly newsletter reporting on current events in the Czech Republic is published by the students of journalism at Charles University in Prague. Articles are in Czech or English.

A number of newsletters describe current events in the countries comprising the former Yugoslavia.

- Bosnet, published daily, covers news of Bosnia and Herzegovina. Articles are mostly in English and languages spoken in Bosnia. Contact Hozo Iztok, HOZO@MATH.LSA.UMICH.EDU.
- Croatian-News (English) / Hrvatski-Vjesnik (Croatian), offer news from Croatia, Bosnia and Herzegovina, and Slovenia. News sources may include transcripts from Croatian and other newspapers. Croatian-News is available from CROATIAN-NEWSREQUEST@ANDREW.CMU.EDU and Hrvatski-Vjesnik can be obtained from HRVATSKI-VJESNIK-ZAMOLBE@ANDREW.CMU.EDU.
- MILS-NEWS, published by the Macedonian Information and Liaison Service, is a digest of current events in Macedonia. It is issued daily in English and Macedonian via listserver MAK-NEWS@UTS.EDU.AU.
- RokPress presents news from Slovenia. Slovene is the principal language but news from Croatian and Serbian sources may be included. Subscriptions can be requested from Igor Benko, IBENKO@MAVERICO.UWATERLOO.CA.
- SII distributes news and discusses current events affecting Serbs. Contact OWNER@MOUMEE.CALSTATELA.EDU.
- VND (Vreme News Digest) is an electronic English-language version of the weekly newspaper, Vreme which serves all areas of former Yugoslavia. Selected articles from Vreme are translated into English. For subscriptions contact Dmitrije Stamenovic, DMITRIJE@BUENGA.BU.EDU.

Online Bibliographies

ABSEES Online, an electronic version of the American Bibliography of Slavic and East European Studies available at no charge via the Internet, is a source of citations for journal articles, government and research reports, book chapters, and books published in the United States and Canada. Contact Aaron Trehub, Editor, ABSEES, absees@uxl.cso.uiuc.edu. ■

(Historical ecology *con't from page 441)*

Future applications

Future applications of this technology include the following:

- a database of 90,000 scanned images of the historical photographs in the Louisiana and Lower Mississippi Valley Collections searchable by patrons in the reading room on a WORM optical cartridge; access will be through hierarchical or key word modes. (The authority file for the index will be the Library of Congress Thesaurus for Graphic Materials.)
- bibliographical and textual analysis projects, such as enhancing the comparison of literary text both within editions and across editions
- compilations of historical statistical data, which will be available both in simple tabular form and in datafiles suitable for manipulation by standard spreadsheet and database software
- electronic publication of a number of journals in the public domain, such as the *Fisheries Bulletin*
- access to other publications in the public domain, including a very large body of "grey" literature on the Gulf of Mexico (literature published in unrefereed publications)
- continue to serve as a Beta test site for new computer hardware and software developed by Key Systems Incorporated, a Baton Rouge company

Summary

The Electronic Imaging Laboratory has given the LSU Libraries an opportunity to make Special Collections materials available to a wider audience and to develop projects that will provide more and easier to use research materials in the future. Documents relating the early history of Louisiana are also being preserved with electronic imaging. ■