

NEWS LIBRARY News

25th Anniversary
News Library News
1978-2003

Vol. 26, No.1/Fall 2003

NLN FEATURE

Digitizing microfilm: An overview

By SANDY LEVY, BALTIMORE SUN

With storage costs dropping and optical character recognition (OCR) getting more accurate, every industry is going digital these days. It is simply the fastest, most efficient way to distribute information. Entire runs of newspapers are available for full-text searching and viewing right from your computer. It's only a matter of time before it reaches you and your paper.

About 25 years ago, the federal government really went all out to preserve newspapers through the United States Newspaper Program. The program was a cooperative national effort among the states and the federal government to locate, catalog, and preserve on microfilm newspapers published in the United States from the eighteenth century to the present. Funding was provided by the National Endowment for the Humanities, and the Library of Congress furnished technical assistance. While most states have finished processing the collections, one state, in particular, has taken its project beyond microfilming to digital - Utah.

The Utah Digital Newspapers Project, orchestrated by the University of Utah and its technical partners, iArchives Inc., and Dimema, Inc., is gathering up all the newspapers that were published in the state and digitizing them. They're learning the hard way



Photo courtesy of Sarah Tews/Wisconsin State Journal

what techniques work best for this complex task. The big papers in Utah - *The Salt Lake Tribune* and *The Deseret News* - are not included in the massive project as yet.

According to the Online Computer Library Center (OCLC) Web site, newspapers on microfilm are preserved for 500 years, but using microfilm to find an article is tough.

Using microfilm is such a horrendous job,

and newspapers are so poorly organized that it discourages many researchers from even consulting it. Imagine the change that digitized full text (with some indexing) would make. Besides leaving news librarians breathlessly grateful, there are many people who want easy access to the social history and legal records that newspapers provide.

Electronic text frees you from the unhealthy practice of "rolling the microfilm,"

Continued on page 8

THE INSIDE STORY

THE CHAIR SAYS/ 4

Linda Henderson urges us to be activists for news librarianship.

BLACKOUT/ 7

Patti Graziano reports on the blackout magic performed by news librarians.

VENDOR LIST/ 9

Linda Chapman has compiled a handy list of vendors involved in digitizing microfilm.

LIBRARY PROFILE / 12

Rebekah Azen takes you on a tour of the *Santa Fe New Mexican*.

NEWS DIVISION EXECUTIVE BOARD

CHAIR, **Linda Henderson**, *Providence Journal*, Providence, RI, 401/277-7887, e-mail: lhenders@projo.com

CHAIR-ELECT, **Jennifer Small Evert**, *LexisNexis*, Miamisburg, OH, 800/227-9597, x58037, e-mail: jennifer.everet@lexisnexis.com

PAST CHAIR/DIRECTOR-AWARDS, **Michael Jesse**, *Indianapolis Star*, Indianapolis, IN, 317/444-6293, e-mail: jesse@indystar.com

SECRETARY AND CONTRIBUTING EDITOR, **Leigh Poitinger**, *San Jose Mercury News*, San Jose, CA, 408/920-5972 e-mail: lpoitinger@mercurynews.com

TREASURER, **Justin Scroggs**, *Atlanta Journal-Constitution*, Atlanta, GA, 404/526-7550, e-mail: jscroggs@ajc.com

DIRECTOR-EDUCATION/PROFESSIONAL DEVELOPMENT CHAIR, **Jim Hunter**, *Columbus Dispatch*, Columbus, OH, 614/461-5039, e-mail: jhunter@dispatch.com

DIRECTOR-PUBLICATIONS, **Elizabeth Donovan**, *Miami Herald*, Miami, FLA, 305/376-3404, e-mail: mdonovan@herald.com

COMMITTEE CHAIRS

NEWS DIVISION COMMITTEE CHAIRS

ARCHIVIST CHAIR, **Teresa Leonard**, *News and Observer*, Raleigh, NC, 919/829-4866 e-mail: tleonard@nando.com

BROADCAST CHAIR, **Vacant**

DIVERSITY CHAIR, **Kee Malesky**, *National Public Radio*, Washington, DC, 202/513-2356, e-mail: kmalesky@npr.org

EMPLOYMENT CHAIR, **Rosemary Nelms**, *Commercial-Appeal*, Memphis, TN, 901/529-2782, e-mail: Nelms@gomemphis.com

GOVERNMENT RELATIONS CHAIR, **Heidi Yacker**, *Congressional Research Service*, Washington, DC, 202/707-7490, e-mail: hyacker@crs.loc.gov

INTERNATIONAL RELATIONS CHAIR, **Wil Roestenburg**, *PCM Landelijke Dagbladen*, Rotterdam, (NL), 31-(0)10-406-7741, e-mail: wroestenburg@home.nl

MEMBERSHIP CHAIR, **Alice Pepper**, *Detroit Free Press*, Detroit, MI, 313/222-5135, e-mail: pepper@freepress.com

NOMINATIONS CHAIR, **Peter Johnson**, *Los Angeles Times*, Los Angeles, CA, 213/237-3349, e-mail: peter.johnson@latimes.com

PUBLICITY-PUBLIC RELATIONS CHAIR, **Dana Gordon**, *Newsweek Inc.*, New York, NY, 212/445-4012, e-mail: danag@newsweek.com

SMALL LIBRARIES CHAIR, **Marcia MacVane**, *Portland Newspapers*, Portland, ME, 207/791-6318, e-mail: mmacvane@pressherald.com

TELLER/ELECTIONS CHAIR, **Marilyn Bromley**, *Bureau of National Affairs*, Washington, DC, 202/452-6380, e-mail: mbromley@bna.com

WEBMEISTRESS, **Jessica Baumgart**, Harvard University, Cambridge, MA, 617/495-4739, e-mail: jessica_baumgart@harvard.edu

NLN MANAGING EDITOR

Ron Larson

Wisconsin State Journal/The Capital Times
1901 Fish Hatchery Rd.
Madison, WI 53713
608/252-6113
e-mail: rl Larson@madison.com

NLN STAFF

CONTRIBUTING EDITORS

PEOPLE

Kathryn Pease

Newsday
Melville, NY
631/843-2333
e-mail: kpease@newsday.com

INTRANET PROFILES

Leigh Poitinger

San Jose Mercury News
San Jose, CA
408/920-5972
e-mail: lpoitinger@mercurynews.com

LIBRARY PROFILES

Jennifer O'Neill

Florida Times Union
Jacksonville, FL
904/359-4184
e-mail: joneill@jacksonville.com

Total membership: 725

News Library News
Fall 2003
Vol. 26, No. 1

News Library News (ISSN 1047-417X) is the bulletin of the News Division of the Special Libraries Association. SLA Headquarters address is: SLA, 1700 Eighteenth St. N.W., Washington, DC 20009. The phone number is: 202-234-4700. *News Library News* is published four times a year by the division. Reproduction in whole or part without permission is prohibited.

Special Libraries Association assumes no responsibility for the statements and opinions advanced by the contributors of the association's publications. Editorial views do not necessarily represent the official position of Special Libraries Association. Acceptance of an advertisement does not imply endorsement of the product by Special Libraries Association.

Subscription price is \$30 (U.S.) annually. Address all subscription correspondence to:

Ron Larson
Wisconsin State Journal/The Capital Times
1901 Fish Hatchery Rd.
Madison, WI 53713
Library
Phone: 608/252-6113
e-mail: rl Larson@madison.com

To place advertisements or to obtain advertising information, you can contact **Ron Larson**, the managing editor, at his e-mail address: rl Larson@madison.com

Copy for bylined columns should be submitted to the column editor. All other copy may be submitted to **Ron Larson** at his e-mail address.



RON LARSON

Welcome to the new online version of *News Library News*. The format may have changed but the efforts at providing our members interesting and informative content have not.

Sandy Levy, **Linda Chapman** and **Rhonda Hollimon** have joined forces in this issue to offer a glimpse at a very hot issue among news librarians - digitizing microfilm. They provide a very good overview on this topic, laying out a good roadmap to follow for those of you who are ready to investigate this technology further.

Also in this issue, **Rebekah Azen** welcomes us into her library at the *Santa Fe New Mexican* with a delightful profile, while **Patti Graziano** reports on how news librarians managed during the massive blackout this past August. And once again, **Kathryn Pease** has compiled another interesting report of news librarian happenings.

It never ceases to amaze me what talented and energetic people we have in the News Division. We are a dedicated group, not only to our news organizations and libraries, but to each other. We are constantly offering our expertise to one another through NewsLib, *News Library News* and one-on-one conversations.

But, as News Division chair **Linda Henderson** mentions in her column, we need to go one step further by becoming activists for news librarianship. Compared to just a few years ago, membership in the News Division is down by over 200. That drop is quite alarming and is probably, in part, a reflection of the bad economic times. But with the collective creative talent we possess, we can turn that trend around.



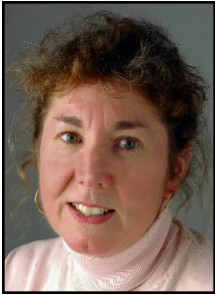
This past summer, my youngest son became very involved in following Major League Baseball, especially his two favorite teams, the Milwaukee Brewers and Seattle Mariners.

One morning, he told me he was going to check the previous night's games on the Internet. I told him all of the scores could be found in the morning newspaper. His response, much to my horror, was, "Why do I need that? I have the Internet."

I quickly opened the sports section and showed him that all of the box scores and standings were neatly arranged on just one page, no clicking required. I am relieved to say that each morning thereafter, he bypassed the Internet and checked the scores in the newspaper. There were even a few mornings where he was up first and walked out to the driveway to get the paper!

Newspaper 1, Internet 0

NOTES FROM THE CHAIR



BY LINDA HENDERSON

Are you an activist for news librarianship?

After considerable mental downtime this summer recovering from the rigors of conference planning and executing, I have had a harder than usual time refocusing on the business (any business) at hand. Because the air was already autumnally crisp by early September this year, I can't even use the continuation of summer weather as an excuse for the inability to snap to attention. So column ideas were elusive until I started hearing a buzz about needing more participation in the Division's activities. Shazam!

Let me start on the assumption that there are many more of you who are partaking of *News Library News* now that we are publishing an electronic edition. And let me assume that many of you participate in discussions (or just lurk) on the NewsLib listserv. You may even use the resources at the News Division Web site. But I know that lots of you readers and listers and lurkers have not become active in the News Division.

So here, in newspeak, is the nut graf - The News Division needs you!

You can start by becoming a member of SLA and the News Division. If your boss is already a member, ask if your organization will cover the cost of additional memberships. If you're the boss, talk to your editors about the value of an organization that focuses specifically on our tiny profession.

The benefits of membership in SLA are both tangible and intangible. Along with the membership directory, which gives you direct access to your colleagues around the world, the most obvious perk is the annual SLA conference. Yes, attendance is expensive, but even if you can only get to conference once or twice in your career, it's worth it. It recharges the batteries and reminds you why you are in this profession in the first place.

And although discussions on hot topics in news librarianship may begin on NewsLib, the issues are tackled in depth at conferences. You also get to meet a lot of the colleagues that you know from NewsLib discussions. NewsLib is a great resource, but it's a lot easier to call on a colleague for help or just to bounce an idea around if you can put a face and a voice with a name.

Once you are a member of the News Division, you can also help to mold the direction of the profession by becoming active in the leadership of the Division. Obviously, I'm a little biased on this subject, but I have thoroughly enjoyed serving the Division in almost every

capacity that I could. And it has raised my professional profile with my editors, especially when we discovered that I am the fourth Division chair to come from the *Providence Journal*. Even the publisher thought that was pretty neat! (See Barbara Semonche's compilation at: <http://parklibrary.jomc.unc.edu/nwsdivchairs.html>.)

We in the Division leadership are having discussions now about how to enhance the value of Division membership. If you have any suggestions or comments, we would very much like to hear from you. If you want to see more of what SLA and the division are all about, go to our Web site at: <http://www.ibiblio.org/slanews> or the main SLA page at <http://www.sla.org>.

And if you would like more information about how to become a member, please contact **Alice Pepper**, our membership chair, at pepper@freepress.com.

Reach out

When I started in newspapers a hundred years ago, the first thing I did was to contact **Homer Martin** at *The* (Bergen County, NJ) *Record*, who was just up the road from my little paper in central Jersey. After strongly recommending that I join the News Division (which I already had), he was gracious and supportive and offered all the help I could ever have asked for. (I'd like to think that we're ALL this welcoming to newcomers).

This brings me to another way that you can contribute to news librarianship. Reach out to others in the profession in your area. Find out if there is a regional news librarians' group and if not, try to get one started. There are a lot of our colleagues at small news organizations who might not even be aware that there are so many of our kind out there, much less have the wherewithal to travel to annual conferences.

For example, here in New England we formed an informal group, grandly called it the New England News Libraries Association (NENLA) and vowed to try to meet at least once a year. Although we are spread out from Maine to Connecticut, we have a pretty good turnout for our meetings and, more importantly, have a great time sharing ideas and problems.

So why not celebrate the 80th year of the News Division by becoming a news library activist. Get connected by joining the Division. Get out and organize in your area. And join us in the leadership of the News Division and news librarianship as a whole.

Who's your portal partner?



VIGNETTE®



They're ours too.

Dialog Portals

Now it's easy. Integrate world-class Dialog content into your portal. Right at your corporate desktop. Right along side your proprietary content. Right from the very first click.

INFORMATION TO CHANGE THE WORLD

Free Trial
Find out more at
www.dialog.com/promotions

THOMSON
DIALOG

©2003 Dialog

PEOPLE



BY KATHRYN PEASE

On the Move

The *U.S. News & World Report* Library has a new addition. **Jill Konieczko** has joined their staff as the Assistant Director of the Library. Additionally, Jill is the Chair of SLA's Leadership & Management Division, and she's the new DC Chapter's Membership Chair.

Anne E. Feiler has joined the staff at *The Denver Post* Research Library. Anne is fluent in German (a native speaker), French and American Sign Language. At the front desk of the Chicago Public Library, her language skills were put to good use in aiding an ethnically and culturally diverse clientele. While there, she also developed one of the country's first service programs for deaf patrons. Its success made the library a gathering point for thousands of deaf and hard of hearing members of the greater Chicago community. More recently, Anne was a training specialist with netLibrary, Inc. in Boulder, a distributor of electronic books via the Internet.

Welcome to the News Division, Jill and Anne!

Awards & Acknowledgments

In the Summer 2003 *NLN* People column, I forgot to mention that **Carol Taylor** of the *Daily Camera* in Boulder, Colorado won the first annual C.B. Hayden fellowship during June's SLA Conference. Named after the former Director of Research at ABC News, the fellowship provides a grant of up to \$500 for the purpose of job-related continuing education. Carol used her grant to study the

history of Colorado at the University of Colorado, Boulder.

Newsmakers and Events

We have a couple of librarians who made the news over the summer:

News Division Webmaster **Jessica Baumgart** of Harvard's Office of News and Public Affairs was interviewed by Helen Jung of the

Associated Press in September for a wire story that ran about the infamous librarian action figure. For those of you who missed the hubbub, you can read the article at: <http://www.tribnet.com/business/story/3854886p-3878999c.html> or you can check out Jessica's blog for more details:

<http://blogs.law.harvard.edu/jkbaumga/2003/09/05#a269> and [http://blogs.law.harvard.edu/jkbaumga/newsItems/viewDepartment\\$Action%20Figure](http://blogs.law.harvard.edu/jkbaumga/newsItems/viewDepartment$Action%20Figure).

John Martin of the *St. Petersburg Times* shared a byline with reporter Chris Goffard for the story, "Double Life,"

about a Tampa road contractor's two families. It's a fascinating read that can be found at: http://www.sptimes.com/2003/06/04/Tampabay/A_double_life.shtml. For their efforts, John and Chris won the monthly newsroom contest: "How I wrote the story" (non-deadline division). Additionally, John discussed the story on a couple of talk radio programs - WIBR in Baton Rouge and radio station 2GB out of Sydney, Australia - and the story was also featured on CNN and Good Morning America. It should also be noted that *Times* researcher Cathy Wos contributed to the report.

University of Minnesota New Media Center scholars and researchers **Kathleen A. Hansen** and **Nora Paul** have a new book out titled, "*Behind the Message: Information Strategies for Communicators.*" Published by Pearson Education, Inc., Barbara Semonche had this to say: "This book is not exclusively about/news librarians and media researchers. Rather, it is an updated conceptual framework for journalism educators and students. It's just that those of us who happen to be in news research will probably find this book very useful. If you are searching for an excellent foundation in types of media messages and types of delivery formats that affect information search strategies not only in journalism but in the broad range of communication industries, then this is a book for your personal and professional collection."

Speaking of Barbara, it's hard to believe, but on September 15, NewsLib celebrated its 10th anniversary. A big thanks to Barbara, our tireless list administrator, and to everyone who makes NewsLib the invaluable resource it is. It's hard to imagine our jobs without it.

While we're on the topic of milestones, on September 2, the News Research Department at the *San Antonio Express-News* answered their 10,000th research request since they started tracking requests in March 1999.

Captive Audiences

Shira Kavon, Senior Librarian at CNN's New York Bureau, participated in a panel for SLA student members at Queens College Graduate School of Library and Information Studies on October 1. She also hosted a group of Library and Information Science students and faculty from Drexel University on a tour of the CNN New York Library and newsrooms on October 10.

Further east, on October 14, I participated in a panel for future special librarians at C.W. Post, which is part of Long Island University.

Continued on page 7

Blackout Magic

BY PATTI GRAZIANO, *THE PLAIN DEALER*

On August 14, a massive blackout occurred at 4:15 p.m., knocking out power to nearly 55 million people in an area ranging from parts of New England to Michigan and extending into Canada.

How did newspapers and other news media and, more specifically, news libraries cope with this crisis? It seems that most of us fared pretty well.

In New York City, where major power outages have occurred in the past 30 years, **Laura Harris** of the *New York Post* library thought at first the blackout was a Midtown occurrence due to inefficient power companies. Soon she realized that it had affected more than just New York City. Fortunately for the *Post*, their back-up generator kicked in. Harris said they were able to use some computers to do research and, by using flashlights, they searched the clipping files for information on blackouts. The *Post's* presses had no back-up generator so the paper printed at a plant in New Jersey. Only a 40- or 50-page edition came out. According to Harris, it was quite a wild experience.

Other New York City news librarians seemed to handle the situation with aplomb. **Kerry Pendergast**, NBC Information Center, said that they had auxiliary power because of a back-up generator and were able to do research. Their main concern was sending material to the users in New Jersey where they were without power.

At the *New York Times*, **Linda Amster** felt that the researchers did a "good job in using the clips and electronic resources in the dim light." The back-up generator gave them some electricity and lights, making one PC operational but not the printers. Fortunately, the clips on the previous blackouts were available. Amster said that full power was not restored until noon the next day. Times Square had power before the newspaper did. The *New York Times* still managed to publish a regular full paper.

At the *Staten Island Advance*, **Melinda**

Gottlieb experienced a damaged archiving system from the power outage. The UPS was severely damaged and subsequently crashed the archiving system. The system was also the output system for photos to the pagination system. Several days' stories were lost. As for research, Gottlieb said they had access to the Internet. The paper plans to install a back-up server for the archiving system sometime next year.

In the Midwest, the *Detroit Free Press* lost all power when the blackout occurred. They had no back-up generators. **Patrice Williams**, one of the librarians, saw people scurrying to get flashlights and candles. Reporters could use laptops and phone lines. Williams said that no research could be done since the library had no access to working computers. She said that power was restored the next day at about noon. The paper managed to publish an eight-page edition with ads.

Here in Cleveland, Ohio, *The Plain Dealer* lost power but managed to produce an eight-page edition with the aid of our competitor, the *Akron Beacon Journal*. We have a reciprocal agreement with them for help in such situations. There was no backup generator at either the downtown office or the production plant. As for the research aspect, we survived by using clipping files, books and telephones. One news researcher took her laptop to one of the bureaus that had power and did some research there. The news research center was able to provide a flashlight and a modem line so that editing could be done on a laptop.

Needless to say, this crisis showed us how vulnerable we are. The newsroom is now being wired for back-up generators and getting other equipment and supplies in case a similar situation happens again.

As for our Canadian colleagues, most of them seemed to experience longer periods of power outage than in the United States.

The *Ottawa Citizen* was forced to use a back-up generator for a week until the nuclear power plants went back online. The paper did

publish a full edition the day after the blackout. According to **Liisa Tuominen** at the *Citizen* library, everyone at the paper went into an "energizing mode" and remained at that level during the crisis. She said it was frustrating, however, for them to look across the Ottawa River and see the glitter of the lights in the Quebec region.

Meanwhile at the *Toronto Star* library, **Andrea Hall** said they had some emergency power. The newsroom had set up a laptop for the librarians to do research. Fortunately, the laptop stayed on long enough for them to find names of experts and to get information on the power grid. They also utilized the print sources and photos so they could do a timeline on the 1965 power outage that severely affected the Toronto area. The paper was able to publish an abbreviated edition.

My impression is that news librarians in general performed admirably well under the circumstances. Those who worked at papers that had back-up generators fared better in terms of being able to do more extensive research.

People

Continued from page 6

Wish You Were Here

On a vacation note, the *San Francisco Chronicle's* **Richard Geiger** and his wife, Susan, sailed an 18' dory (that's a small, narrow, flat-bottomed boat for all you landlubbers) to Catalina Island and back - 26 miles across the sea. Three other open boats made the passage with them. For some great photos, see: <http://thegeigers.org/>

Not to be outdone, the *Chronicle's* **Judy Canter** had a busy year taking four European excursions. In January, she and family members ventured to Chamonix in the French Alps with side trips to Geneva and Vienna. The end of June saw Judy in Tuscany a la Diane Lane followed by another family trip to southern France before heading off to the Netherlands in September.

Digitizing microfilm: An overview

Continued from page 1

transporting you to a world of lightning searching and information retrieval of details and facts that would have stayed hidden forever in its current state. Digitized microfilm opens up the possibility of multimedia, addition of metadata and field searching, concept searching, special additions to your Web site, and last, but certainly not least, revenue. Reaching the genealogy market alone, reported to be the second largest hobby in the world, would seem to be worth the effort of scanning the microfilm.

Large newspapers with a national scope are lucky enough to be scanned for free by Proquest Information and Learning, the company in Michigan that also keeps the newspapers' microfilm via the Bell & Howell division. So far, Proquest has worked its magic on *The Christian Science Monitor*, *The New York Times*, *The Wall Street Journal*, and *The Washington Post*. *The Los Angeles Times* and *The Chicago Tribune* are in the queue and are due to be finished in 2004. Proquest revenues for the historic newspapers have been less than expected this year, due to tight budgets for academic and public libraries. The historic newspaper digitization project is the most ambitious Proquest has ever undertaken, and officials predict the business will ultimately earn the company \$40 to \$50 million in revenue annually. Another major competitor for large newspapers is Gale Group, which is digitizing *The Times* of London.

Other newspapers can turn to Cold North Wind's Paper of Record, the company that began their move in Canada, for digitization of microfilm. Royalties for the papers are 10%.

Heritage Microfilm in Cedar Rapids, Iowa, also is in the business of digitizing historic newspaper microfilm. The company partners with the papers by hosting the finished product on the Heritage Web site as well as the newspaper's Web site. Pricing and royalty splits vary based on circulation. Currently,

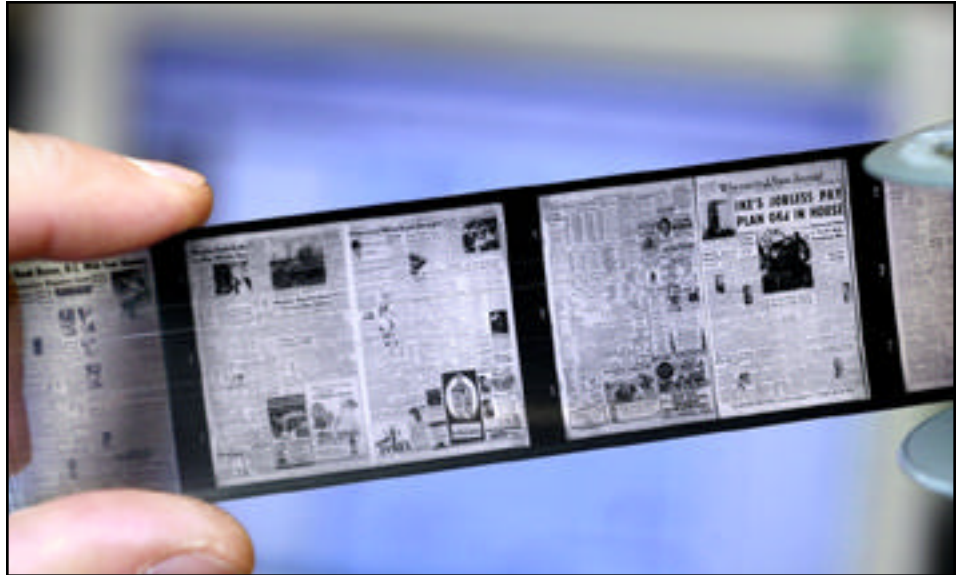


Photo courtesy of Sarah Tews/*Wisconsin State Journal*

Heritage is working with the International Press Association to bring advertisers to the site, which should lower the costs for newspapers.

The Daily Oklahoman, among others, has gone with Olive Software, Inc.'s Active Paper Archive product to digitize its historic archive.

International efforts to corral and digitize microfilm newspapers include those in Britain, Norway, and Australia.

One of the stumbling blocks to making money with digital historic newspapers is copyright. The copyright term for newspapers is 95 years and is retroactive to anything published from 1923 forward. Papers published prior to 1923 are in the public domain. To allow for this, Proquest offers a one-time purchase of its product for the out-of-copyright years to libraries.

For more on copyright, check out the Copyright Office site <http://www.loc.gov/copyright/> and Professor Laura Gasaway's chart at the University of North Carolina <http://www.unc.edu/~unclng/public-d.htm>

The other stumbling block is the Tasini decision, which protects freelance writers and photographers from 1964 onward and makes the digitization messier. To comply with Tasini, the digitizing vendor must prevent the reader from seeing freelance copy apart from

the complete page. The vendors are all figuring out how to do this.

The job of digitizing microfilm itself is tremendous. Broadsheet newspapers are cumbersome, and the print is tiny. The quality of the scanning is vital, or there's no point in doing the work. Some old microfilm is nearly impossible to read, so it's a struggle for the scanners to work. This is why newspaper clippings were preferred in the early days. Newspapers that have kept print copies have a big advantage, and newspapers that are preparing to do this in the future would do well to save electronic versions of the full pages.

If microfilm is really bad, it becomes necessary to rely on manual entry, which is often done overseas. It is also important to scan each column separately - called zoning - which takes into account the jump of an article. The ubiquitous PDF document seems to be the preferred method for storage and viewing. More technical advice can be found at the Northeast Document Conservation Center www.nedcc.org or the Library of Congress <http://memory.loc.gov/ammem/techdocs/index.html>.

Keeping up with digital versions so that the digitized format will always be available will

Continued on page 9

Issues in digitization of microfilm

BY LINDA S. CHAPMAN, *PATRIOT LEDGER*

Besides investigating these companies as you would any other, and testing their demo sites (for some you will need a vendor-provided password), there are many things to consider when selecting a firm for your microfilm digitization project. Here are some of the variables:

Types of companies

Some microfilm production and storage firms do digitization. If your current microfilm company is one, you may benefit by it already having access to your master copies (for best quality scanning.) If not, your microfilm business might be required as part of the digitization contract.

Some firms offer a wide range of digital archiving solutions and other software products besides microfilm digitization. Some have an aggregate product which will include a gateway to your digitized content (another access point would be your paper's Web site). An aggregate product offers you the benefits of serendipitous hits. Consider if this collection would be the best fit for your newspaper. Watch contract language for who gets the rights to your digital content, how subscriptions are sold, and how the revenue is shared. A firm may offer standalone hosting and e-commerce without an aggregate product. With others you will do your own hosting, subscription sales, e-commerce, etc. You may pay a fee for licensing their software annually.

Companies will need access to your masters or a clean print; this could be an additional expense. One uses film from the Library of Congress.

Scanning & other production processes

Although they are all using some form of optical character recognition (OCR) scanning, many vendors have patented software and techniques. This is the feature you'll hear the most about when you talk to them. Most offer different grades of service at different

prices. Some firms tout antique font recognition, important for older papers; some easily handle differing page sizes; some offer headline and byline searching; meta-tagging is usually available.

Pay attention to level of human interaction and quality control. The human factor can increase the quality of the end product but also add heavily to the cost. Ask if they check for missing pages or key garbled content from poorer-quality papers (or sections thereof), or other special services. Some companies outsource internationally. Determine if the people working on your newspaper speak and understand the language.

Format

There are two basic types of presentation: page or article.

The page view is usually a PDF of the original print page, with the searched text highlighted. Vendors with this format cite the value (and desire by patrons) of seeing the article in its original context; the avoidance of copyright violations; the lower cost. Disadvantage is the slow loading speed of the PDFs.

The article format will bring up specific articles with hit-highlighting (as well as the ability to see full pages); articles may be HTML or XML, sometimes linking jumps. These services offer copyright management, more flexibility in delivery, faster search results. Disadvantage is generally higher cost.

Time frame

Check their claims on how fast their turnaround time is. Some firms offer to do your content in batches, with some desirable content done first (i.e., Civil War, WWI or WWII; a local event of general historical interest that your paper recorded, such as the civil rights era, etc.) for faster return on investment. A few companies pro-rate royalties to help you start up (a "subsidy," as it were.)

Search engine and interface

This aspect is something librarians excel at critiquing.

Ask about and test the interface and search engine capabilities of companies who host. How intuitive are the search screens? How robust is the search engine? If you want to sign and retain subscribers, you want the easiest-to-use design for average consumers, but one that allows more complex searches for adept searchers - with the fastest retrieval rate that you can afford.

Continued on page 11

Digitizing microfilm: An overview

Continued from page 8

be just one more technical detail to watch in the digitizing microfilm movement. Since microfilm is required for copyright and other legal reasons, is a stable long-term medium, and is the most reliable image of the printed newspaper, it should be saved even after digitization.

Even with all these electronic full pages zipping around, there will still be room for the text archive and the image archive, unless the digitized microfilm indexing gets very good. Reporters and others will just want the article. They won't want to sift through advertising to get a story, and photographers and graphics artists will just want the image.

Digitizing microfilm currently is technically complex, takes a lot of time, and costs a lot of money. Still the promise of being able to find anything ever printed in a newspaper is getting closer than ever.

Additional Resources:

OCLC

<http://www.oclc.org/digitalpreservation/>

The United States Newspaper Program

<http://www.neh.gov/projects/usnp.html>

Utah - Utah Digital Newspapers

<http://www.lib.utah.edu/digital/unews/>

Digitization is a success for the AugustaArchives

By RHONDA HOLLIMON, *AUGUSTA CHRONICLE*

In late 1992, I decided it was time to trace my family's roots. Never could I have imagined what I was getting myself into. It takes about an hour to look at a month's worth of newspapers, and back in the "old days" only the year of death was listed on the tombstone. So, it could feasibly take 10-12 hours to find one obituary. Since my grandfather and grandmother each had seven siblings, and several of those siblings had ten children, one can easily see my dilemma.

After several months of searching the microfilm in the traditional way, I decided there had to be a better way to do this. Faced with over 200 years of data, I had to find a way to get every page into a searchable format.

I considered re-keying everything, but that would be extremely costly and would never be completed within my lifetime. I am not a patient person and wanted something immediately.

By late 1993, I thought scanning was the way to do this. I brought in a company to inventory some of our clipping files, and give a price quote for digitizing the clips. After about a week, the quote arrived. The cost for digitizing ten years' worth of clip files was going to be \$300,000. There was no way I could justify this cost, so it was back to the drawing board.

In late 1997 or early 1998, I stumbled across Convera's (then Excalibur) Web site. The company boldly proclaimed that it was "Turning information into Knowledge." I thought about that statement for a moment, and knew that I, too, wanted to turn information into knowledge.

AugustaArchives, or at least the concept, was born that day. I knew that if I could get the archives into a digitized format, run an OCR application, and apply the RetrievalWare search engine, I would be in business. After several months of planning, and about of year of digitization, we were in business.

There were so many things to consider during the evolution of this system. Here is a review of the major decisions that must be made:

1. Digitization - Do you do the scanning in-house or is this something you outsource? We looked at both options. To do this in-house, we would have to buy the equipment at a cost of over \$100,000. Then, we would have to add one or more staff persons. They would have to be trained, and would only be able to work so many hours per day. The cost of adding staff and equipment made this option less attractive. And, as I said, I am impatient. We would have to order the equipment, wait for it to arrive, train the staff, and have one or two people working to get over 1 million pages into a digitized format.

We decided to partner with a vendor already set up to do provide this service. With equipment in place, and a trained staff, we could start immediately.

2. Digitize the entire newspaper at once, or digitize in increments - We decided to digitize the entire run of the newspaper at once. The cost of start-up is really the bulk of the expense, so why not get the entire newspaper out there at once. Some newspapers have chosen to digitize ten years, and later digitize ten more. During these uncertain economic times, one can understand this decision. But you have to realize that to put ten years online, you have to spend the money for all of the basics (server, search engine, interface, e-commerce piece, and digitization), and this will make your per image cost a great deal higher.

In addition, revenues will be low because only customers interested in that specific time-frame will be interested in using the system.

3. Host the site or have it hosted? - We knew we would host the site ourselves. Our corporation owns several computer technology and Internet divisions, so this was not an

issue for us.

For others, it will be a huge issue. The size of our database, 1.2 million pages, is almost half a terrabyte.

Without the hosting infrastructure already in place, our cost would have been extremely high. Servers and extra staff would be needed.

After the decision to host the site is made, equipment and software (servers, database, search engine) must be purchased and installed. While this is in progress, the scanning can begin. By the time everything is installed on-site, CDs or DVDs should begin to arrive for upload into the system. In the case of The Augusta Chronicle, we used RetrievalWare's Fileroom, and the library staff uploaded the data as it arrived.

4. Format of the data - TIFF, PDF, or other formats. We looked at JPEG (too large), GIF (way too large), PDF, and TIFF. Our new media director was a proponent of TIFF, so we decided to go with TIFF.

We tested 25 years' of data in-house, decided the project was going to be successful, and started on the Web interface design and e-commerce portion of the project.

During weekly meetings, we were updated on the various aspects of the project.

On September 4, 2001 at 9 a.m., AugustaArchives was opened to the public. Two years, and almost 3,000 registered subscribers later, we have learned a great deal. From a technical standpoint, we learned that we needed to spend the money to have the search terms highlighted. We also learned that as OCR technology improves, we will need to update our data to provide the best available searches.

Customer complaints have been relatively few. Of course, there will always be someone who thinks all information should be available at no charge, even if it did cost almost \$500,000 to make it available.

The only recurring complaints are that the search engine doesn't work, and that the

Continued on next page

Digitization is a success for the AugustaArchives

Continued from previous page

images won't load.

The search engine does work, it just doesn't work like a Yahoo or a Google engine. Those engines are searching clean text and are searching left to right on a page. With a newspaper database such as this, if you are searching on John Smith, you may have John on one line and Smith on the next. A proximity search is not effective. Once we talk customers through this, and explain to them how to be creative, they are usually able to find what they are looking for.

When images won't load, the problem is usually with the Web browser. Customers may not have cookies or JAVA enabled, or maybe they have an old version of a browser. All are easily rectified, and again, the customer is happy.

AugustaArchives has subscribers from 48 states, and 8 foreign countries. A reporter from Australia used the site to find information about a person on the lam in the late 1800s, who ended up in Augusta. He would never have found this without the digitized archives.

We have an annual subscriber from Nevada who called me one day to say that she was in Atlanta for a cousin's wedding and wondered if she could drive down to talk to me. When she walked in, she had a list of things she liked, and a list of things she wanted improved. She had lived two streets from a cousin she had not seen in over 30 years, and both were doing genealogical research. She saw the cousin's name and town in a relative's obituary, and they reconnected.

We get a lot of heart-warming stories like this. There are the not-so-nice stories as well. One day a lady came into the library to use the archives because she did not have a computer at home. She knew only that her mother had died when she was really young. Her family would not tell her when her mother died, or where she was buried. She searched the archives, and found that her father had murdered her mother. The discovery was a horrendous moment for her, and for the library staff as well.

People are using the site and finding things they never dreamed of finding. My grandfather had a brother who was killed in a wreck when in his early thirties. I knew the year, but nothing more. Using the archives, I found that it was a front-page story, and that he had been hit by a drunken driver. It was a horrific accident with two fatalities.

Reporters use the archives on a daily basis, and the site has spawned a couple of feature sections. We did a series of "Remember when" stories, on places people would congregate in the "old days." Every day, we do a "Back in Time" feature in the metro section, a "this day in Augusta history" type brief.

Schools subscribe so students may use the archives for history-related projects, and public libraries subscribe for genealogical and research purposes.

In 2002, we decided it was time to upgrade the OCR and add a highlighting feature. The new version was launched on September 15. The complaints we receive are the same as before, and once we train them on use of the viewer and the search engine, most everyone ends up happy.

Issues in digitization of microfilm

Continued from page 9

Price

Prices are usually per page. Ranging from free up to \$1.80 per page, prices are dropping as firms get more competitive and technology improves. Estimates of pages per reel are from 700-1000. Ask how they handle zoned editions.

Here are some of the companies digitizing newspaper microfilm

Cold North Wind

<http://www.coldnorthwind.com>, <http://www.paperofrecord.com>

Kerry Slagle, President and CEO, Atlantic Syndication

1970 Barber Rd., Sarasota, FL 34240

Phone: 941-371-2252

Fax: 941-342-0999

kslagle@atlanticsyndication.com;

[web: www.atlanticsyndication.com](http://www.atlanticsyndication.com)

In Canada:

205 Catherine St., Ottawa, Ontario,

Canada, K2P 1C3

Phone: 613-722-9886

Fax: 613-722-923

Heritage Microfilm

<http://www.newspaperarchive.com>

John Kuiken, Heritage Microfilm

4049 21st Ave. SW, Cedar Rapids, IA 52404

Phone: 319-390-9442

Fax: 319-390-9448

john@heritagemicrofilm.com

Morris Digital Works

<http://www.morris.com>, <http://www.morrisdigitalworks.com>

Bill Ferguson, Business Development, Morris Digital Works

753 Broad St., Eighth Floor Augusta, GA 30901

Phone: 706-823-3610

bill.ferguson@morris.com

Newsbank Inc.

<http://www.newsbank.com>

Tina M. Owens, Business Development Director, NewsBank Media Services

397 Main St., Chester, VT 05143

Phone: 603-887-0800

Fax: 603-887-0801

Cell: 603-490-4928

towens@newsbank.com

Continued on page 13

Accomplishments add up; Santa Fe library enjoys a steady ride

By REBEKAH AZEN, *SANTA FE NEW MEXICAN*

It's not the most ideal closet space for a library, with a coin counting room situated dead center, a paper shredding machine at one end, and a door that opens out to the shipping and receiving dock at the other, but I've learned to appreciate it.

Space is at a premium at the *Santa Fe New Mexican*; we've outgrown our quarters downtown on Marcy Street, a location we've been at since before 1949. There is an office or operations of some sort going on in every nook and cranny. Actually, the library is often surveyed for another patch of open space, but that availability is rapidly shrinking. Witness the 10 drawer microfilm cabinets now stacked on top of each other, almost reaching the ceiling; the boxes of papers stored for temp archival purposes that do indeed reach the ceiling; and the friendly Web guys who moved in about 6 months ago to share our workstation, complete now with weather monitoring equipment!

A news librarian was recently congratulating us on how far we've come since my "basic training" days at Poynter in '98 and said, "You didn't even have your own desk and phone!" I thought it over for a minute and then decided to tell her, "I still don't!" The News Library desk and computer space have by tradition been the domain of the archivist and I, being hired as a News Library Consultant, have worked mainly from home. It's beginning to change however, and I do work at the News Library more, in between the Web guys and the archivist.

I know this situation may sound daunting to some, but it's been a great challenge. Humor and flexibility, along with the ability to adopt unorthodox library practices are the norm here, and in the long run we've managed quite well and prospered.

The *Santa Fe New Mexican* is somewhat of a marvel, like the "little engine that could." It is still an independent paper, having been in



Space is at a premium at the *Santa Fe New Mexican* News Library, but Rebekah Azen can look forward to moving her library into more spacious quarters in the near future.

business since 1849 and is not owned or operated by the media moguls. It is fondly known as "the oldest paper in the west." We serve the capital city of Santa Fe and the surrounding communities, with a circulation figure of about 27,000.

I entered the scene in 1997 quite by accident, having just returned to New Mexico from Minnesota as the library director at a tribal college. I just needed a job doing anything and it didn't even occur to me that there was a news library at the paper. Well, someone heard I was a librarian, and the rest is history.

It was fairly clear when I first gazed upon the library, that it was a morgue, and I instantly understood the essence of the terminology. The library had been sadly neglected, but there were signs that noble attempts at order had been made, on and off over the years; there was a microfilm collection, a clip file collection going back to the 80's, scattered books and resources, and an archiving system

that had been installed and in use since 1994.

Management requested slow and steady improvement in the news library, not a quick fix with a huge investment, so I was hired as a news library consultant. I remain to this day a consultant, working part-time, along with an archivist who works part-time as well. Our two jobs together equal one full-time position. At first glance, one may think this arrangement unworkable, but actually it's been quite successful, and fiscally prudent. The combination of a professional librarian working part-time to manage and direct, along with a competent (let me emphasize competent here, as it is a critical factor) archivist is a pretty good arrangement for a paper of our size on a limited budget. I would prefer that the archivist worked full-time as a library technician to take on more responsibilities but we haven't reached that optimum place yet. In the meantime, we're handling

Continued on next page

Accomplishments add up; Santa Fe library enjoys a steady ride

Continued from previous page

what needs doing and making slow but steady progress.

I do a multiplicity of things, which makes the job interesting. The majority of my time is spent with data issues and data quality with regard to the online archives, mostly because we are still in a transition phase from an old archiving system to a new one, along with a new editorial system that was installed about three months ago. Along with this go a lot of technological challenges that have to be ironed out, which takes a ton of patience. At the top of the list as well is archival management, both with the physical newspapers and microfilm. We are one of the few newspapers left in the country that has an original archival run going back to the mid 1800's and we are ensuring the survival of the record by taking a multi-pronged approach to preservation that may well include digitization.

At any moment I am available for reference or research questions. I do bibliographic training for the newsroom with regard to the online archives, DIALOG, NM Technet, Accurint, Google, localized databases, Proquest, etc. We don't get inundated with public queries, but we do get them and I attend to that as well.

There are many other areas I'm engaged in,

such as collection development, management of collections, acquisitions, policy development, formulating procedures, Web site issues with respect to the archives, overseeing contracts and attending to vendor relations, upgrading services, creating a newsletter, etc.

Besides having a competent archivist, I've found that things work best when there is collaboration among departments and people. However, that objective is not always met, leading sometimes to serious difficulties. It may have something to do with the nature of the business, news, which is always accosted by a deadline, creating a mentality of urgency that permeates everything, leading to rushed decision-making.

There is one other essential piece to successful operations I have found, and this is superior technology and technological assistance. Poor technology and absent or incompetent technological help is disastrous. We now have a devoted IT person and a responsive department, as well as excellent archiving technology. The decision to purchase Software Construction Company software was a primo decision on the part of management and probably the best decision we ever made. The decision was arrived at after thorough research.

I can't say that any of our achievements are earth shaking, but they are significant, however humble. Little accomplishments of greater quality and improvement add up. The archival record is being safeguarded now.

Reporters can get professional research help and bibliographic instruction. The public is being well served. Revenues from royalties are significant (they top my salary plus!), the online archives have improved 100%, the new databases we subscribe to have the information reporters need, the vertical newspaper shelving system we purchased makes life so much more manageable now in the news library for all, the newsletter got out with information that the newsroom can use, etc.

Plans for the library include:

- Archiving photos and pdf's
- Making physical collections available through an online catalog
- Creating an intranet for the newsroom

And now, for the best news, the Santa Fe New Mexican is building a new facility and we have already broken ground! There is a REAL news library in the plans and I may even get my own computer and workstation. No more living in the closet. Watch out because I'm coming out! On my bike that is. The new facility is just a few miles from my front door. Such a fit analogy to end with for the SFNM News Library, moving along at the speed of a bicycle. It takes a long time and a lot of work, but the ride is enjoyable and we do get there.

Rebekah Azen is News Library Consultant for the *Santa Fe New Mexican*. Feel free to contact her with questions or comments at 505-424-9475 or by email at library@sfnewmexican.com.

Issues in digitization of microfilm

Continued from page 9

Olive Software

<http://www.olivesoftware.com>

Kim Dail, Vice President of Sales & Marketing
Olive Software
1642 South Parker Rd., 2nd floor
Denver, CO 80231
Phone: 720-747-1220
Fax: 720-747-1217
kim@olivesoftware.com

Progressive Technology Federal Systems, Inc. (PTFS)

<http://www.ptfs.com>

Jim Mullican, Senior Account Manager
PTFS, Inc.
7315 Wisconsin Ave., Suite 1200W
Bethesda, MD 20814
Phone: 301-951-4300
Fax: 301-951-5934
Cell: 301-928-0145
jmullican@ptfs.com

ProQuest Information and Learning

<http://www.il.proquest.com/proquest/features/feature-04/default.shtml>

Cliff Pierce, Director, Publisher Relations
ProQuest Information and Learning
300 N. Zeeb Rd.
Ann Arbor, MI 48103
Phone: 405-285-2341
cliff.pierce@il.proquest.com

Southwest Micropublishing, Inc. (SMI)

<http://www.swmicropublishing.com>

Victor R. Rivera, President
Southwest Micropublishing, Inc.
2627 E. Yandell Dr.
El Paso, TX 79903
Phone: 800-367-1273
Fax: 915-562-4527
sales@swmicropublishing.com