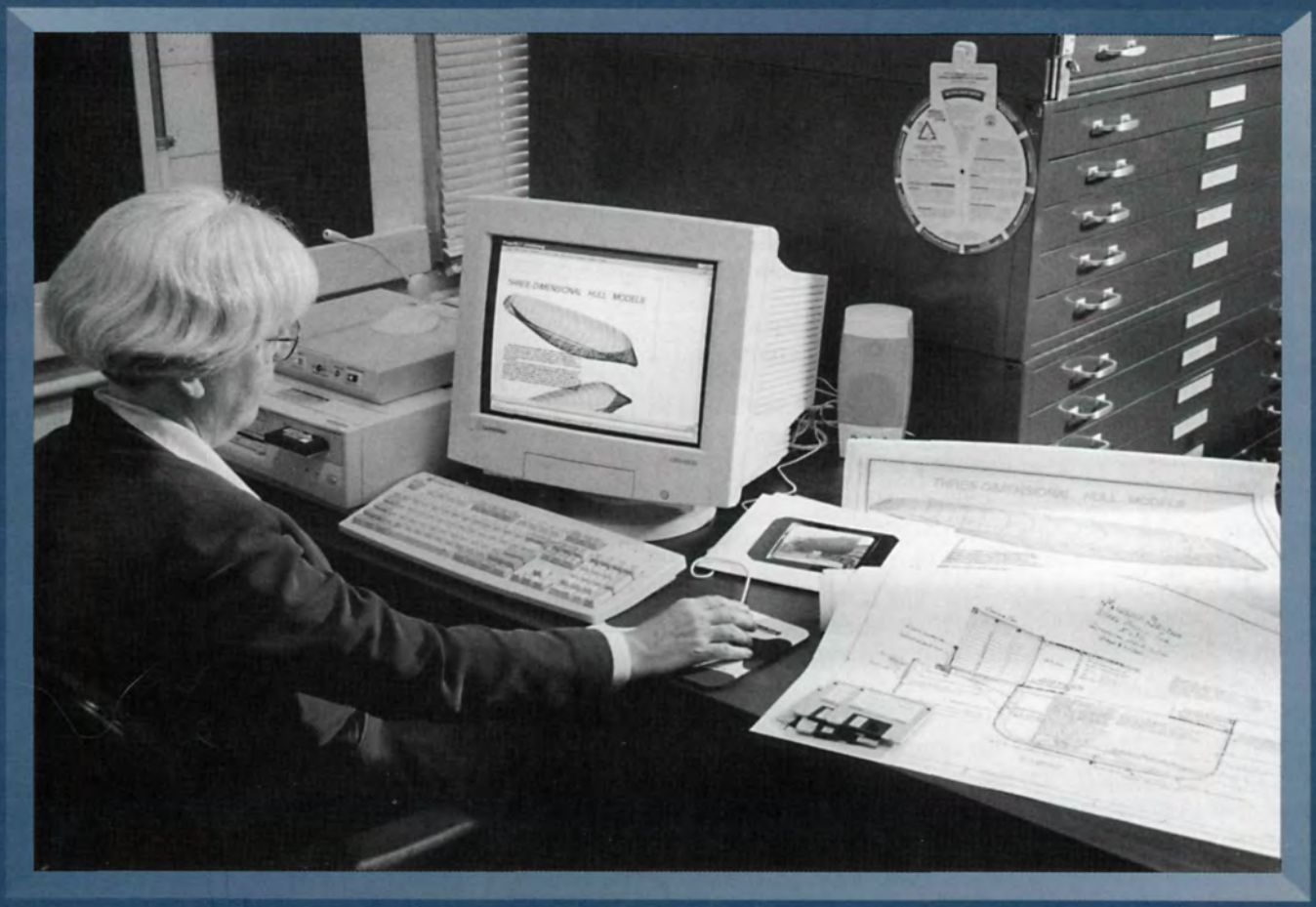
CRM

CULTURAL RESOURCE MANAGEMENT
Information for Parks, Federal Agencies, Indian Tribes, States, Local Governments, and the Private Sector

VOLUME 21 NO. 6 1998

The Information Ecosystem



U.S. DEPARTMENT OF THE INTERIOR
National Park Service
Cultural Resources

PUBLISHED BY THE NATIONAL PARK SERVICE

To promote and maintain high standards for preserving and managing cultural resources

DIRECTOR
Robert Stanton

ASSOCIATE DIRECTOR
CULTURAL RESOURCE STEWARDSHIP AND PARTNERSHIPS
Katherine H. Stevenson

EDITOR
Ronald M. Greenberg

GUEST EDITOR
Diane Vogt-O'Connor

ADVISORS
David Andrews
Editor, NPS
Joan Bacharach
Museum Registrar, NPS
Randall J. Biallas
Historical Architect, NPS
Susan Buggey
Director, Historical Services Branch
Parks Canada
John A. Burns
Architect, NPS
Harry A. Butowsky
Historian, NPS
Pratt Cassity
Executive Director,
National Alliance of Preservation Commissions
Muriel Crespi
Cultural Anthropologist, NPS
Mark R. Edwards
Director, Historic Preservation Division,
State Historic Preservation Officer, Georgia
Roger E. Kelly
Archeologist, NPS
Antoinette J. Lee
Historian, NPS

ASSISTANT
Denise M. Mayo

CONSULTANTS
Wm. H. Freeman
Design, Imaging, Production–Freeman Publishing Services
Janice C. McCoy
Editing–Editorial Notes

An electronic version of this issue of *CRM* can be accessed through the *CRM* homepage at <http://www.cr.nps.gov/crm>.

Contents

VOLUME 21 NO. 6 1998
ISSN 1068-4999

The Information Ecosystem

The Information Ecosystem 3
Diane Vogt-O'Connor

Using the Electronic Information Ecosystem for Research 7
Hugh O'Connor

Information-Seeking in Organizations and Archives 10
Mary Jo Pugh

Records Management in the National Park Service 15
Betsy Chittenden

Cultural Records in Question
Information and Its Moral Dilemmas 18
Michael F. Brown

Archiving Archeological Data in the Next Millennium 21
Harrison Eiteljorg II

The Stories You Could Tell
Using NARA Regional Resources for Site Interpretation 24
Kellee Blake

The Yellowstone Archives and its Affiliation with NARA 27
Susan Kraft

The Information Ecology of Archives 29
Richard Pearce-Moses

Information Skills and Primary Sources in Education 34
Susan Veccia

NPS Libraries Come of Age
Building the NPS "Virtual Library" 36
Amalin Ferguson

The National Register Information System
Preservation Strategies in the Data Center 39
John P. Byrne

Computerizing the National Register of Historic Places 45
Carol D. Shull

The Information Ecosystem Bibliography 49
Diane Vogt-O'Connor

Cover: Recording hull lines, the shape of the hull is critical to building or maintaining a vessel. On the computer are the three-dimensional hull models of Appleton Hall *produced by the Historic American Engineering Record in 1996 using a combination of hand measuring and computer-rectified photogrammetry. Photographs of the hull, visible next to the computer, drive the measurements. The photographs are digitized, solving for the three dimensional location of any point visible in at least two of the photographs. On the right is a copy of a hand-measured and hand-drawn Midship Section used to build the vessel c. 1914. Courtesy San Francisco Maritime National Historical Park. Photo by Danford/Campbell.*

Statements of fact and views are the responsibility of the authors and do not necessarily reflect an opinion or endorsement on the part of the editors, the *CRM* advisors and consultants, or the National Park Service. Send articles and correspondence to the Editor, *CRM*, U.S. Department of the Interior, National Park Service, Cultural Resources, 1849 C Street, NW, Suite 350NC, Washington, DC 20240; 202-343-8164, fax 202-343-5260; email: <ron_greenberg@nps.gov>.

Diane Vogt-O'Connor

The Information Ecosystem

Information Ecology, emphasizes an organization's entire information environment. It addresses all of a firm's values and beliefs about information (culture); how people actually use the information and what they do with it (behavior and work processes); the pitfalls that can interfere with information sharing (politics); and what information systems are already in place (yes, finally, technology).

Thomas H. Davenport, ***Information Ecology: Mastering the Information and Knowledge Environment,*** *(1997)*

The Information Ecosystem within our cultural resources community includes all those individuals and professions that create, manage, use, and adaptively re-use information in all forms. In an effective information ecosystem, data (facts and observations), information (data with purpose and context), and knowledge (valuable information plus human understanding) are all managed holistically as valuable professional and organizational resources of interest to an ever-growing international community. At the end of the 20th century, cultural resource managers have become knowledge workers.[1]

In Richard Lanham's *Electronic Word* (1993)[2] he states that in a knowledge-based economy, the scarcest commodity is human attention, not information. In this model, human attention is labor, which gives information structure, usefulness, and value—in effect making it knowledge. Lanham views information technology as a means to the end of capturing the interest of students and scholars and other information users. Technology democratizes access to information, leading to expanding markets for both knowledge workers and knowledge consumers.

At the turn of the millennium, our Information Ecosystem is both more complex and more vulnerable to neglect than ever before. No one organization working alone can preserve our knowledge and make it accessible to the huge audiences desiring it. If we are to learn how to create information efficiently, manage it effectively, and preserve it sufficiently, we must work together as a series of allied professions to meet the new challenges ahead.

This special issue of *CRM* supplements the course "Information Ecosystem: Managing the Life Cycle of Information for Preservation and Access." The Information Ecosystem course was offered at the National Archives facility in College Park, Maryland, March 10-13th, 1998, by the National Park Service, the Northeast Document Conservation Center, and the National Archives and Records Administration (NARA). Both the Information Ecosystem course and this issue of *CRM* focus on an integrated approach to the management of cultural resource information that builds upon the knowledge and expertise of archivists, curators, information resource managers, librarians, and records managers. A linked issue of *CRM, "Archives at the Millennium,"* will appear in early 1999.

Who are the Key CRM Players?

Ultimately it is the cumulative effect of many individual's small daily activities that determines whether or not we capture and preserve the staff knowledge, organizational information, and data that make up our cultural resources legacy. In Cultural Resources Management (CRM), there are many key players in the information ecosystem. This issue of *CRM* includes articles from most of them, including the following:

- ***archeologists,*** who excavate prehistoric and historic sites and produce documentation during excavations. An article by Harrison Eiteljorg II, Director, Center for the Study of Architecture and the Archaeological Data Archive Project on page 21 of this issue talks about archeology and archives.
- ***architectural historians,*** who work with archival primary sources, such as sketches, blueprints, and notes, and library publications to determine the history and original structure of a building and the history of changes to a structure which are recorded in further notes and drawings. An article by Keeper of the National Register Carol Shull discusses the history of National Register technological changes on page 45.
- ***archivists,*** who arrange, identify, appraise, describe, preserve, and provide access to the personal and family papers, corporate record, and organizational record of groups for scholars, students, publishers, and the general-

Continued on p. 5

The Challenges

- ***Tremendous growth in the size and nature of the national cultural resources record*** that we preserve and manage due to the increasing number of key organizations, people, and professions involved; growth in what cultural resources we protect; the increasing quantities of data collected by most organizations; the ease of information sharing among partners and cooperators in a digital world; and the increasing breadth of our cultural resource preservation, documentation, and outreach activities among partners and cooperators.
- ***Increasing demand for cultural resources information,*** including our own organization's appetite for 1) key summary information for management purposes such as the Government Performance Results Act (GPRA), 2) requirements to mount more data on our World Wide Web sites to meet the needs of the international community, and 3) the need to mount information on the Web to meet Freedom of Information Act requirements. Other insatiable audiences include scholars, publishers, vendors, online order fulfillment services, educators, school groups, colleagues, fellow professionals, and a growing community of "edutainment" producers of games, the History and Discovery Channels, and others.[1]
- ***Significant resource limitations*** which limit what cultural resources records we can preserve and make accessible and how (e.g., less conservation treatment, less detailed description). We are already masters of economies of scale; further improvements will shave very few cents off our budgetary dollar. Partnerships with outside for-profit organizations can help, but require time, cooperative agreements and memoranda of understanding, retraining, management policies, and sometimes special legislation.[2]
- ***Increasing costs of information management,*** including such factors as higher staff costs, preservation supplies costs, reformatting costs, and data migration and refreshment costs. Information may be more expensive to preserve in the 21st century than it has been in the 20th century. Figures from several major digital projects indicate that the cost of managing permanent digital files may be much greater (between 10 and 16 times greater—according to University of Maryland Professor Charles Lowry on the University of Pennsylvania Web site) than that of preserving and making accessible equivalent paper files. This is at least partially due to the need to maintain hardware and software and continuously migrate and refresh files.[3]
- ***Changes in professional standards, strategies, and techniques,*** which require major systems changes; data mapping or revision; and retraining of personnel including the need to learn metadata standards, the Encoded Archival Description standards, and similar professional expertise. In the past, the standards of one profession rarely impacted another; now, cross-fertilization is rife. For example: archivists need to learn how to preserve GIS data from data center staff; while most Cultural Resources staff need to learn how to produce more durable information formats from archivists and conservators.
- ***Challenges to standard archival and library access and use strategies and operating principles and key legislation,*** such as copyright, privacy and publicity legislation, and the concept of fair use, which determine what information we make accessible and how. The fair use of cultural resource materials on the Web is under attack by those who wish to support pay-for-view and similar services. Cultural, ethical, and moral challenges are being introduced by indigenous peoples who wish to preserve their privacy and maintain ownership of their cultural heritage information, including that information found in public archives, libraries, and museums.[4]
- ***The fragility of our electronic record,*** which requires that we migrate and refresh the data regularly; label it accurately and according to standards; prevent misuse; and manage, and upgrade the software and hardware as necessary in order to provide access over time.[5]
- ***Reorganizations of our cultural resource institutions,*** massive restructuring, downsizing, and retirements result in a loss of staff knowledge and institutional memory. The only effective ways to ward off a resulting institutional memory loss are long-term cross-training of staff, excellent records management, a functioning organizational archives, and effective oral and video history programs. Too often valuable files containing cultural resource management information are orphaned and inappropriately destroyed. This loss of the record results in a diminished institutional knowledge base, as well as a loss of sometimes-irreplaceable data. Staff depart, taking their knowledge with them.
- ***New, rapidly changing, and swiftly vanishing formats of information*** from new color photographic and laser printing processes and geographic information systems to the World Wide Web. The last two of which are revised so often that they frequently vanish before they can be permanently captured in a durable media for future use.[6]

Continued from p. 3

public. An article by supervisory Archivist of the San Francisco Maritime Museum Mary Jo Pugh on page 10 describes information-seeking behavior in organizations; while an article by Heard Museum Archivist Richard Pearce-Moses on page 29 describes how data, information, knowledge, and records are adaptively re-used in archives.

- ***curators***, who study the archival and library source materials and original objects and their documentation and conduct research for publications, exhibitions, and other outreach activities. An article by curator Susan Kraft of Yellowstone National Park describes their National Archives affiliated archives at Yellowstone National Park on page 27.
- ***historians***, who research individuals, groups, and themes via a combination of primary (archival manuscripts, photographs, electronic records, and motion picture footage, sound recordings, and videotapes), secondary (published monographs), and tertiary (textbooks, indices, reviews, and abstracts) sources in order to produce new documents and manuscripts. Archives are the key data used in the ever-changing narratives, debates, and discussions that are the products of historians. An article by NPS Bureau Historian Barry Mackintosh describing the value of archives to the NPS will appear in the upcoming issue of *CRM* on "*Archives at the Millennium*."
- ***conservators***, who preserve, treat, and reformat archival materials. Articles by conservators Jessica Johnson (NPS) and Steve Puglia (NARA) will provide an overview of NPS conservation publications and describe standards for the creation of permanent and durable information in the upcoming issue of *CRM* on "*Archives at the Millennium*."
- ***information resource managers, geographic information system staff, programmers, and systems analysts***, who capture and manage electronic data for current use. An article by National Register Computer Specialist John Byrne on page 39 discusses managing ever-changing information technology for cultural collections with a historical perspective.
- ***interpreters and educators***, who search the historical record for lively and telling stories to illuminate the past, which may be further captured in videotapes, articles, books, or notes. An article by Kellee Blake of NARA on page 24 explains how to use NARA Regional resources for interpretation; while an article by educator Susan Veccia of the National Digital Library (NDL) Program of the Library of Congress on page 34 explains how their archival resources are shared with millions online.
- ***librarians***, who provide reference services, and produce Web sites, library catalog records, and literature guides. NPS librarian Amalin Ferguson talks about the plans for the NPS Library Program on page 36; while Hugh O'Connor, Director of the American Association for Retired Persons Research Information Center, describes how to search the information ecosystem on the Web on page 7.
- ***records managers***, who locate, describe, appraise, and ultimately determine the final disposition of the miles of paper that exist within our organizations. A piece by NPS Records Manager Betsy Chittenden on page 15 provides her insights on the status and future of records management at the NPS.
- ***tribal cultural managers***, who research yesterday's activities in archives and libraries and record today's activities for placement in cultural centers for tomorrow's children. An article by Archivist Donna Longo DiMichele of the Nashantucket Pequot Tribe in the linked *CRM* issue on "*Archives at the Millennium*" describes an active tribal archival program; while a piece by Michael Brown, the James N. Lambert Professor of Anthropology and Latin American Studies at Williams College, describes the challenges to fair use of cultural materials taking place in archives internationally on page 18.

This issue of *CRM* pays tribute to the many professions that create and manage the information ecosystem that ultimately ends up in archives. For many of us, our discoveries, research, and records linger in our offices until we move on to another position, take on a different series of projects, or simply decide to clean up our offices and dispose of the files. We are often so closely tied to these records that it is hard to remember that this data forms an invaluable part of the informational legacy of our organization.

Without the data and information in the records that we have created, our organization's information base is impoverished and its ability to effectively manage our cultural resources over time is diminished. Each of us can either manage our information effectively, bequeathing to our professional heirs a rich legacy of data and information, or treat it as our personal disposable belonging. When we treat our informational legacy as a personal belonging we are ensuring that our contributions, knowledge, information, and data will be lost to those who come after us. Professional ethics, our interest in having our contributions remembered, and Federal Records Laws, all

demand that we responsibly manage our informational legacy for future scholars, educators, students, and the public.

In Summary

At the end of the 20th century the Cultural Resource Information Ecosystem is imperiled by increasing costs, decreased budgets, fewer staff, more users, burgeoning information, increasingly unstable information formats, changing professional information standards and practices, revised laws on fair use and copyright, and institutional restructuring and instability. Simple neglect alone is enough to ensure disaster.

No organization or profession working alone can preserve our knowledge, ensure the survival of our information and make it accessible to the insatiable audiences who demand it. We must work together as allied professions and organizations to share our expertise and resources if we are to ensure the survival of our data, information, and knowledge for future generations. This legacy, which safely stores our factual observations for future theorists and managers, our information for later adaptive re-use, and our professional knowledge for enhancement of our organizations and professions, is our greatest gift to the future.

With this knowledge intact our professions and organizations are empowered to move into the future with confidence and integrity. Without our informational legacy, our organizations lack vision, and a sense of confidence informed by history and experience. If our hard-won data and information is to survive for future re-use, we must individually and as professional allies care for our information legacy on a daily basis using the techniques and practices described in this *CRM* issue on the "*Information Ecosystem*" and in the upcoming "*Archives at the Millennium*" issue.

Tower Falls at Yellowstone. Photo by William Henry Jackson.

References

1 Davenport, Thomas H and Laurence Prusak. *Information Ecology: Mastering the Information and Knowledge Environment.* Oxford and New York: Oxford University Press, 1997.

2 Lanham, Richard. A. *The Electronic Word: Democracy, Technology, and the Arts.* Chicago: The University of Chicago Press, 1993.

References, page 4

1 Gilder, George. "Angst and Awe on the Internet" *Forbes ASAP* (December 4). Also Battin, Patricia. "From Preservation to Access: Paradigm for the Nineties." *IFLA Journal* 19 (1993): 367-373.

2 McCall, J.J. editor. *The Economics of Information and Uncertainty.* Chicago: University of Chicago Press, 1982.

3 Dollar, Charles M. *Archival Theory and Information Technologies.* Macerata: University of Macerata Press, 1992. Also Lesk, Michael. *Preservation of New Technology: A Report of the Technology Assessment Advisory Committee.* Washington, DC: Commission on Preservation and Access, March 1992.

4 Brown, Michael F. "Can Culture be Copyrighted?" in *Current Anthropology* 39(2), 193-222. Also Branscomb, Anne Wells. *Who Owns Information?* New York: Basic Books, 1995.

5 Conway, Paul. *Preservation in the Digital World.* Washington, DC: Commission on Preservation and Access, March 1992.

6 Lesk, Michael. *Preservation of New Technology: A Report of the Technology Assessment Advisory Committee.* Washington, DC: Commission on Preservation and Access, March 1992.

Diane Vogt-O'Connor is Senior Archivist, Museum Management Program, National Park Service.

This article was offered at the Information Ecosystem Conference, funded by the NPS Cultural Resources Training Initiative.

Hugh O'Connor

Using the Electronic Information Ecosystem for Research

Today there are many more sources for professional reference than there used to be. Many of these are entirely digital, without any print equivalent. In a sense, the Internet has introduced a "habitat" within which many new "species" of professional and popular reference tools have established a niche for themselves. This broadly increases the information available to knowledge workers, and also complicates the task of finding, evaluating and using all of it.

The major Web search services are dynamic in coverage and functions. Changes in site design and in the details of the services are made frequently, as each of the search engines, directories and subject guides compete to attract the traffic of the Net. Like the rest of the Internet, they evolve toward improvement.

What are some of these searching tools? Even though they are relatively rare compared with the totality of the Internet, there are still many of them from which to choose. One Web site (**World Wide Web Search Engines & Directories** at <http://www.lawresearch.com/cewwwd.htm>) lists nearly a thousand such searching services. The vast majority cost nothing to use.

In this discussion we are not talking about the classic fee-based database searching services that have been around, under one name or another, since the early 1970s. Although they continue to play a significant role in the information industry, their coverage of information on the Internet at large is still quite limited. They still index or contain the full-text of the premium, often print-equivalent information sources (newspapers, journals, and specialized data collections) that have been their specialty since their inception. They have, however, largely adopted the Web as the channel of access through which searchers now connect to those databases. A useful page of links to the Web incarnations of the major database search services can be found at On-line Inc.'s **SuperSites** page at <http://www.onlineinc.com/corporate/supersites.html>.

There are a number of Web-based search engines, using computer-based indexing methods, that attempt to offer very broad, if not comprehensive, coverage of the content of the World Wide Web. The most well-known and well-regarded of them (in alphabetical order) are:

Alta Vista <http://www.altavista.digital.com>
Excite <http://www.excite.com>
HotBot <http://www.hotbot.com>).
Infoseek <http://www.infoseek.com>).
Lycos <http://www.lycos.com>).
Northern Lights <http://www.nlsearch.com>
Open Text Index <http://index.opentext.net>
WebCrawler <http://www.webcrawler.com>

These services are extremely popular, but require some study to use them to best effect. Because they index words within some or all of the texts of the hundreds of thousands or millions of Web pages they index, they tend to retrieve hundreds and sometimes thousands of Web pages in response to the simplest search.

Each of them has some form of "relevance ranking" built into its search method, so that the pages identified as most likely to be relevant to a search are presented first. This is done to save a searcher's time. The expectation is that, as you go down a list of perhaps thousands of entries, their appropriateness and usefulness to you will drop off more or less rapidly, and you will soon reach a point below which the retrieved items are no longer likely to be of interest to you.

Although all the search engines listed above index the World Wide Web, they go about their work in different ways. Some index the words at a Web site's home page, and others go deeper than that. Some search engines cover a more selective set of Web sites than others. Also, the search rules are somewhat different from one engine to another. Since the search engines cover overlapping but not identical material; and since they operate differently, you will get different results from different search engines, at different times. Also, a lot of new material is always appearing on the Web, while the existing sites are always changing and being updated. The consequence is that you should be prepared to frequently re-run a search on any topic in which you are strongly interested—new material goes online every day.

Even the best indexes of Internet content are substantially behind the actual changes that go on constantly over the Net. Search engines work on the basis of indexes that are created by automated surveys of the state of the Web. The indexes have to be frequently rebuilt, since many of the Web's details will change sooner or later after they have been surveyed and recorded. Also, much new

material appears on the Web every day. One commentator has likened the task of indexing the World Wide Web to trying to paint an unfinished bridge that gets larger as you paint it.

This is not the only limitation on search engines' accuracy and range of coverage. Some Web sites, at their Webmasters' request and for whatever reason, are excluded from search engine indexes. Other Web sites contain part of their content in formats that are technically not indexable by the search engines.

Sometimes a Web page is not available even though it turns up in your search results. There are many possible explanations for this. The indexed site may be "off-line" for routine maintenance or because of a system problem, and will be available again soon. Popular sites get overwhelmed and may respond with a "busy signal." The indexed Web page may no longer be maintained at the address the search engine indicated. The page may have been moved, given a new file name, or removed altogether. This could happen at any time, because the maintainers of Web and other Internet sites have complete control over whether or not and how long they will keep their site available. In a sense, the price of the Net's attractive features—its creativity, spontaneity and decentralization—is that it changes so often as to be occasionally unreliable.

There are a number of Web sites dedicated to tracking and evaluating the major Web search engines. These sites are particularly useful because of their comparative approach. Some are addressed more to archivists and librarians and other professional researchers, and provide the results of in-depth performance tests of the search engines. Others are addressed more to Internet users in general, and are more succinct and less academic in their approach. All of them offer perspective on the search engines and advice on their use.

Search Engine Watch: News, Tips, and More About Search Engines <www.searchenginewatch.com>. This is probably the most comprehensive of the sites, and has information of interest to Web professionals as well as general readers. The section called "Search Engine Facts and Fun" is the part of the site most appropriate for novices and general Internet users, and contains information on how to use the search engines to best effect. A free monthly email newsletter, called the *Search Engine Report*, is offered to those who register with this site.

PC Magazine's Complete Guide to Searching the Net <www.zdnet.com/pcmag/features/websearch/_open.htm>. This is the online version of a feature article for the December 2, 1997 issue of the magazine. It consists of evaluative comments on each of the search engines, as well as some of the metasearch services. It is addressed to the general Internet user.

C/Net's Search Engine Shoot-Out: Top Engines Compared <www.cnet.com/Content/Reviews/Compare/Search2/?st.cn.fd.ccol.re>. Like **PC Magazine's Guide**, above, this is a recent (January 1998) comparative evaluation of search engines and metasearch services, aimed at the general Internet user. Note especially that this feature article includes a page from which you can test the major browsers for yourself.

The Search is Over: Search Engine Secrets of the Pros <http://www.zdnet.com/pccomp/features/fealO96/sub2.html>. This 1997 (otherwise undated) review of the search engines offers very specific tips and advice for using each of the major services. Since the search engine details may change as they are further developed, this article needs to be read circumspectly. If any of its tips do not work for you, this may be an indication of technical change in the search engines themselves.

Search Engine Showdown: Comparing Internet Finding Tools <http://www.imt.net/~notess/search/about.html). This excellent site is the product of Greg Notess, a librarian and author of many articles about online information searching. In his words, "This site summarizes, reviews, and compares the search features and database scope of the Internet search engines and finding aids." It is addressed primarily to other librarians and professional information searchers, and it contains a selected bibliography of print and online sources (with links) for those who want to become experts on the subject of search engines.

Because the retrieval results of the major search engines are at best statistical samplings of the Web's content, some other types of Web tools have an important role in Internet research. A few of these include:

ProFusion <http://www.designlab.ukans.edu/profusion/>. ProFusion is one of the so-called "metasearch" services, which deal with the variability in the search results of each of the major search engines by searching them all, and collating the results. This especially "intelligent" metasearch service will put your search terms into the various formats used by the search engines it employs, will eliminate duplicate citations retrieved by different engines, and will relevance-rank the search results from different engines into a single sequence. ProFusion offers a personalized search service, allowing users to register their regular queries. Those queries are periodically re-run, and the system tells you when there are new results.

The Argus Clearinghouse <http://www.clearinghouse.net>, **The World Wide Web Virtual Library** <http://www.w3.org/vl/>, **InfoMine**

<http://lib-www.ucr.edu/> and the **Virtual City of Alexandria** <http://www.alexandria-home.com/> are all attempts to organize the content of the World Wide Web into the subject arrangement of a library. Web sites like these are often referred to as "metasites." You choose your broad subject of interest from a list and these services connect you with a detailed page full of hyperlinks to the Internet resources offering information on your chosen topic. Coverage in these services is selective, but still very extensive. Because human judgment is involved, they offer greater precision—although perhaps not as much recall—than the search engines.

The Internet Sleuth <http://www.isleuth.com>. The Sleuth connects with "over 2,000 searchable databases" on the Web. It organizes those databases into broad categories like "News," "Sciences,," and "Travel" and permits you to search up to six databases in any given category at a time. Click the "About" button on the home page for a short, useful discussion on how to search the Internet Sleuth to best effect. Such gateways to Internet databases are particularly significant because search engines themselves do not usually index database contents.

Internets <http://www.internets.com>. This is another very large collection of links to databases, organized quite a bit like *The Internet Sleuth*. This service adds a current newsfeed and a planetary weather map to its home page. It claims to be "the biggest filtered collection of useful search engines anywhere on the World Wide Web."

Professional reference tools now include software. Along with the proliferation of Web sites offering guidance to the scattered resources of the Internet, there has been a parallel development of software products specifically geared to assisting in the process of Internet research. Some such programs include utilities that, for instance, help to manage and organize Web bookmark lists. Others are more elaborate, serving as "intelligent agents" that search, compare, analyze, compile and organize information from many Internet sources while their human client is doing something else. We are at a point in the flowering of the software industry in which many products are available—many inexpensive or free of charge—that help researchers to analyze Web sites, find new sites of interest, manipulate digital text, and do research generally in ways impossible only a few years ago.

Some Web sites have appeared that not only list programs in these categories, but serve as links to the sources of the software themselves. Although some of these programs are commercial products, many are freeware or at least shareware, and are downloadable. Some good lists to check include:

Web SearchUser Tools <http://www.zdnet.com/products/searchuser/tools.html>. This is a short listing of both major Web search sites and Web search software, with links to each.

Botspot <http://www.botspot.com>. This site is concerned specifically with intelligent agent and "bot" software.

Cool Tools(<http://www.cooltool.com>. This source is not exclusively concerned with software that affects Web and Internet searching. It is, rather, a good, general, critical review source on new Internet-related software in general.

TUCOWS <http://www.tucows.com>. TUCOWS (The Ultimate Compilation of Winsock Software) is one of the major compilations of Internet-related software of any kind. Search-related software is only a small part of what it offers.

Some software sites that offer interesting products for Web research, navigation and data compilation include the following. This is not at all a comprehensive list, and some or all of these items may be found listed at one or more of the sites indicated directly above. Most of these programs are available either free or in a free version from their sponsoring Web sites.

Alexa <http://www.alexa.com>
Autonomy <http://www.agentware.com>
Citizenl <http://www.citizenl.com>
WebFerret <http://www.ferretsoft.com>
WebTurbo <http://www.webturbo.com>

Hugh O'Connor is with the Research Information Center, American Association of Retired Persons. This article is adapted from his speech at the Information Ecosystem Conference.

Looking at the Mammoth bison corral at Yellowstone. Mounted ranger is visible inside the corral. Bunsen Peak is in the background. Photo courtesy Yellowstone NP Archives.

Mary Jo Pugh

Information-Seeking in Organizations and Archives

Searching for information in organizational archives is an extension of information-seeking behavior in everyday work in organizations. On a daily basis, staff members typically rely first on their own memory for needed information. Second, they rely on convenient reference tools or readily accessible records that document their knowledge and actions. Research suggests that people will use the most accessible information, regardless of whether it is the best information. One writer calls this the "Principle of Least Effort":

> ...most researchers (even "serious" scholars) will tend to choose easily available information sources, even when they are objectively of low quality, and, further, will tend to be satisfied with whatever can be found easily in preference to pursing higher quality sources whose use would require a greater expenditure of effort.[1]

This statement reflects the reality of information-seeking in most organizations, although it fails to recognize the limited resources and time pressures facing staff and administrators on a daily basis, and as we will see, it fails to recognize the structure of information flows in organizations.

Thus, when searching for information, a staff member is most likely to draw first on personal knowledge or on the records documenting his or her actions immediately at hand.[2] Records are utilitarian, created in the course of practical activities. As staff members order, direct, design, build, report, communicate, instruct, plan, evaluate, advertise, apply, announce, authorize, request, compensate, contract, or otherwise do their jobs, they create records. The title of a record often reflects the action that it creates it, such as application for leave, invitation, job order request, performance appraisal, change order, specifications, permit, and so forth. Recording technologies have proliferated, so that documents may be textual, graphic, photographic, audio, video, or electronic.

Documents are instruments for conveying information about actions in the organization and beyond from one place to another. Some documents, such as directives and instructions, flow from the top of the organization down; some, such as requests and reports, flow from the bottom up; still others, such as memoranda, flow laterally. Documents pool in filing systems in locations where the information is needed, so office files tend to have incoming documents, copies of outgoing documents, as well as notes, calendars, and other documents created and retained in the office. The filing structure is the primary mode of retrieval. Twentieth-century recording technologies, especially electrostatic copying, have increased the likelihood that copies will be found in many locations, but the aggregation of documents in any one location will be unique, reflecting the activities carried out at that location.[3]

For information beyond their own memory, files, and scope of activity, staff are likely to consult other people in the organization. One survey of university administrators found that 94% of all respondents cited other university staff members as their primary information resource.[4] Brown and Yakel found further that, "administrators rely most on human information networks resulting from years of experience and personal relationships built on trust and prior provision of reliable information."[5] It is natural that people trust information that has been selected and authenticated by a knowledgeable expert, and information that is given by the person responsible for the action. Thus, seeking information from other people in the organization is much more than simply following the principle of least effort.

Staff ask themselves, "Who would know or need to know about this problem?" To locate the right person, staff use their knowledge of organizational structure to identify the individual or office responsible for the sphere of activity, or they ask others more knowledgeable about the organization. They are likely to use the telephone in search of information. Staff directories and organization charts serve to guide people with questions to people with knowledge. Titles of both individuals and departments indicate responsibility for organizational functions.

The information seeker relies on the responsible official either to know the answer, to know their files in order to be able to find the information, or to refer the information seeker to another person, department, or organization. Staff members consult either their own memory or the memory of their actions embedded in their records, or

analyze the functions of the organization and then consult people or records resulting from that function. Often, however, this information-seeking behavior is so ingrained that staff do not think about these processes, and the search for information is so obvious that the process is transparent.

Information-seeking behavior in organizations is changing as information is increasingly recorded in electronic forms, especially in networked electronic environments. Archives at the millennium are faced with a paradigm shift comparable to the invention of the printing press 500 years ago; perhaps even comparable to the invention of writing itself five millennia ago. For the past 15 years, personal computers were primarily used to produce "fast paper," that is, people used software packages for word processing, database management, or spreadsheets to automate the production of paper documents. The flow of information continued largely through transmittal of traditional paper documents.

In the last five years, however, information is increasingly transmitted only in electronic form. Internal organizational information is distributed by electronic mail. Bulletin boards and discussion groups (list-servs) provide means to contact a wider pool of people than the telephone. Public information and reports that would have been disseminated via the printing press are now available instantly through the World Wide Web. Information once found in paper form in department files such as benefits information, customer records, library catalogs, archival finding aids, and other departmental databases are now accessible through Local Area Networks so that people can access them from their desktops, rather than calling the responsible official. Automated information services include shared cataloging through Online Public Access Catalogs, CD-ROM indexes, proprietary databases like Dialog, or full text databases like Lexis or Nexis.

In the last few years as computers have been linked in networks, either Local Area Networks within organizations or through the Internet to other organizations, users have come to rely upon powerful and convenient online tools and information resources. At the same time, however, many records are no longer captured in a tangible form. Records are created, communicated, filed, retrieved, or lost only in electronic form. **There is the tendency to think that if information is not in electronic form, it does not exist.** In some cases, the information resources on the Internet are so chaotic, information is as good as lost.

Information Seeking in Archives

With the passage of time, people move on, but the organization continues. With good records management, records documenting significant actions with continuing consequences are transferred to organizational archives so that later information seekers, whether later incumbents or others seeking evidence of past actions, can find them. **The mission of the archival profession is to identify records that have continuing usefulness, preserve them, and make the information in them accessible through time.** Records management for electronic records is still in its infancy, but for federal agencies the decision of District Court Judge Paul L. Freidman in *Public Citizen v. John Carlin*, October 1997, has given it greater urgency by making it mandatory. Although the National Archives and most state archives have begun programs to manage electronic records, most archival holdings consist of documents on paper.

The search for information in archives about past actions is similar to searching for current information in organizations but is more complex. Locating information about past actions depends on interaction among three archival functions: arrangement, description, and reference services.

Arrangement: Archivists use provenance and original order to capture the contextual information that made the records usable as they were created. Records are kept together as a group linked to the person, office, or organization that created them. Provenance links records to the functions that created them, reflects organizational functions, and preserves the lines of communication graphically outlined in the organizational chart. If the records are kept in the same order as they were filed, the location of each item in the filing structure can be predicted. Retaining provenance ensures that the evidence in the records is authentic. No later hand has added, subtracted, or moved the evidence from the actions that created them. In archives, unlike libraries, individual documents are not re-filed according to a predetermined subject scheme. If a later person, following a library model of information retrieval, rearranges documents or mixes them with documents from another organization, then the evidence cannot be trusted, nor can the content be predicted. **A chain of continuous custody from the creator to the user ensures the authenticity of evidence.**

Provenance is a powerful predictor of content and locator of evidence. Provenance and original order serve to retain the physical organization of records as found in offices when they are transferred to shelves in archives.

Description. As time passes, knowledge of functions and forms of records fades from personal memory. Or, records are transferred to outside repositories. Information seekers, whether archivists or researchers, need information about functions, forms, and content of records. Archival description focuses on groups of records and their

Finding aids describe and index historic documents to make the information in them accessible for use. The information in these paper finding aids will be migrated to the archives module of the NPS Museum Management Program's Automated National Catalog System (ANCS+) to provide interactive online access. Photo by Campbell/Danford, courtesy San Francisco Maritime National Historical Park.

relationships. Descriptive control over a collection is progressively refined from the top down, working from the broadest group to the appropriate level of description, whether series, file, or item.

Archival description is meta-data; that is, it is information about information, that leads a user to information and helps a user to understand it. It is also management information that allows archivists to acquire and preserve the holdings. A useful definition is:

> Archival description is the process of capturing, collating, analyzing, and organizing any information that serves to identify, manage, locate, and interpret the holdings of archival institutions and explain the contexts and records systems from which those holding were selected.[6]

Elements of information about records may be captured as they are created, acquired, arranged, and used. These data elements may be embodied in a number of products, such as donor records, accession lists, inventories, finding aids, catalogs, indexes, registers, card catalogs, indexes, databases, or guides, and displayed on paper or online. **An ideal descriptive system allows archivists to collect all data elements about a collection in one system and produce any number of products.** If such a system is based on national standards, information about collections can be shared with other repositories.

Over time, repositories have used many types of finding aids, but in recent years information about each record group has been standardized into finding aids that include an administrative history outlining organizational functions and series descriptions describing the forms and filing structures of the records. Most also include lists of the contents, for example, file titles for textual records, titles of videotapes, titles of oral history interviews, and identifications of photographs. The archivist also provides index terms for the record group. The index terms for all record groups are accumulated in a master index so that users who do not know the functions or forms of records can be pointed to the records likely to be of interest to them. Index terms can include subjects, personal names, corporate names, place names, as well as terms for functions, and terms for the forms of records, such as minutes, logbooks, architectural drawings, photographs, and videotapes.

Reference Services. Archives staff play a critical role in linking researchers, finding aids, and records. Reference assistance is often vital to the success of users in archives. School children are taught to use libraries, but most users are not familiar with archives, nor do they extend their insights from searching for information in their daily work to searching for information in archives. The reference process in archives has intellectual elements, administrative elements, and is complicated by the interpersonal dynamics of reference interaction.[7]

Intellectual elements. The most important function of reference services in archives is providing intellectual access. Providing intellectual access includes providing information about the repository; information about its holdings; information from its holdings; information about records creators, and referrals to other sources. To use records, users must know that they exist and how to find them. If users know the scope of the collection of a repository they can often predict whether the collections will have information or evidence for them. Researchers also need practical information about location, telephone numbers, public hours, services, and access policies. **Researchers find such information through national, regional, or thematic directories; publications such as brochures, signs, and guides; public programs such as lectures or workshops, and increasingly through the Internet.** A useful directory of archival Web sites, "Repositories of Primary Sources," is maintained by Terry Abraham at the University of Idaho.[8] Leon Miller at Tulane University maintains a site, "Ready,'Net,Go! Archival Internet Resources." This site includes lists of archival Web sites and well as links to tools for archivists, archival search engines, and professional sources.[9]

A number of sources locate information about holdings. Some repositories publish guides that summarize information about groups of records. Of most use to federal agencies is the *Guide to Federal Records in the National Archives of the United States* (1995). It includes extremely useful agency histories that identify the functions of government from its founding and their reporting relationships, as well as descriptions of the forms

and contents of records. It is extensively indexed. The National Archives maintains a very useful Web site. It includes the *NARA Archival Information Locator (NAIL)* a prototype searchable database of information about selected records and digital copies of some documents. Also maintained is the "Reference at Your Desk," by the National Archives Library and Information Center (ALIC), which includes links to laws, copyright regulations, legal resources, biographical, and geographical resources.

The first reference tool to describe and index manuscript collections from repositories throughout the United States was the *National Union Catalog of Manuscript Collections*, NUCMC, affectionately known as "nuck muck." From 1959 to 1993, the Library of Congress published descriptions of approximately 72,300 collections located in 1,406 different repositories in 29 annual printed volumes, which included approximately 1,085,000 index terms. The final printed volume was published in 1994. Cataloging for the volumes from 1986 to 1993 and all ongoing cataloging is available only online. To provide access to its online cataloging NUCMC provides free access through a Z39.50 Gateway to the Research Libraries Information Network-Archives Manuscripts Collections database (RLIN AMC). The NUCMC site also provides links to other Library of Congress resources, archival societies, archival education, electronic discussion groups and periodicals, bibliographical utilities, preservation, and the Encoded Archival Description (EAD) standard and its use by the archival profession.[10]

National bibliographical utilities, most notably the Online Computer Library Center (OCLC) and the Research Libraries Information Network (RLIN) also provide collection-level descriptions of archival holdings, structured according to the national standard for sharing information about holdings, in Library of Congress machine-readable cataloging format (MARC). Both began as utilities to provide the benefits of shared cataloging for published materials, OCLC to serve smaller college and public libraries, RLIN to serve large university research libraries. Both have now found that the databases are as useful for reference staff and patrons searching for information as for providing cataloging. RLIN in particular has developed as a cultural resources databases. Both charge for searches. Many library online public access catalogs (OPAC) include descriptions of archival holdings, and some are available though the Internet.

An increasing number of archival finding aids are now available online and archivists are creating a standard for storing this information so that it can be shared and migrated for preservation, the Encoded Archival Description (EAD) in Standard Generalized Markup Language (SGML). Collections of digitized documents, especially photographic images, are also available on the Web. A useful place to start is the Berkeley Digital Library SunSite, maintained by the University of California at Berkeley and Sun Microsystems.[11] It provides links to catalogs and indexes, including finding aids, as well as links to text and image collections both at Berkeley and elsewhere. It provides links to other services such as information for digital library developers, reports on research and development, software tools, and learning tools.

Chadwyck-Healey offers a subscription to *Archives USA* which includes the *Directory of Archives and Manuscript Repositories in the United States*, collection records from NUCMC, and indexing for its microfiche publication of finding aids, the *National Inventory of Documentary Sources*. It is available on CD-ROM or on the Web for subscribers.[12]

Interpersonal Dynamics of Reference Services

Providing intellectual access in the repository and in providing information from holdings has traditionally been predicated on personal interaction between archivist and information seeker. Archives are mysteries to most users, and reference services are often educational services in expanding the mental models of users to encompass the range of archival records, the variety of finding aids available, and the development of a search strategy for exploiting them. Information searches in archives are typically mediated through the archivist. In institutional archives, we have seen that administrators tend to use personal contacts for information, and Yakel and Bost confirm my experience that most do not use finding aids or records. They expect information to be extracted and packaged for them.[13]

Interpersonal dynamics are vital in the reference process. Although a library user may find information without ever interacting with a librarian, this is rarely true in archives. In most cases, an information seeker contacts a repository either in person or by phone, mail, fax, or email. Typically, a user begins with an initial interview with the archivist, whether in person or over the telephone, or through correspondence or email. **The initial interview begins with question abstraction, in which the seeker and the archivist identify the topic, delimited by time, place, and the seeker's intended use. Next, in question resolution, archivist and user analyze the topic and the intended use in terms of the archival resources available and form a search strategy, a plan for identifying the sources of information likely to answer the question.** This is an inferential process based on what is known

about the records and the functions that created them, an extension of information seeking in the creating organization. Archivists play a vital role in this process because of their understanding of the universe of documentation and how a user's questions fit that universe.

Question refinement is the third stage of reference interaction and is a continuing process as questions and topics are refined in light of information discovered during research. Ideally the reference interaction is closed with an exit interview in which the archivist has the opportunity to determine the success of the seeker and the patron can evaluate archival information and archival services.

Nonverbal communication both clarifies and complicates interpersonal communications. Information seekers often find it difficult to expose ignorance to a stranger when the response is unknown. This fact underscores the importance of building interpersonal relationships in institutional archives, so that administrators and staff will feel comfortable and confident in asking questions. Active listening is probably more important than talking. Taking time to draw out the full question and determining the level to which it needs to be answered is important for both seeker and archivist. The necessary administrative elements of providing physical access to archives, such as registration, procedures to ensure integrity and preservation of archival evidence, photocopying, and the like must be handled so that they do not hinder the building of trust and confidence necessary to intellectual access.[14]

Information seeking in organizations occurs in a complex environment of interpersonal networks, electronic networks, and record-keeping networks. If archivists understand the full range of information resources in organizations, regardless of their forms, implement a sound records management program for both tangible and electronic records, and respond to the complexity of information-seeking behaviors of administrators, staff, and the public, they can provide a vital service to their organization.

Notes

1 Thomas Mann, *Library Research Models* (New York, Oxford University Press, 1993), 91.

2 This fact explains why staff are often reluctant to send their records to the archives.

3 The role of documents in modern organization is discussed in JoAnne Yates, *Control through Communication: The Rise of System in American Management* (Baltimore: Johns Hopkins University Press, 1989).

4 Mary W. Sprague, "Information-Seeking Patterns of University Administrators and Nonfaculty Professional Staff Members," *Journal of Academic Librarianship* 19 (1994): 378. Cited in William Brown and Elizabeth Yakel, "Redefining the Role of College and University Archives, " *American Archivist*, (Summer 1996): 282.

5 Brown and Yakel, 283.

6 Victoria Irons Walch, *Standards for Archival Description: A Handbook* (Chicago: Society of American Archivists, 1994) 2.

7 Mary Jo Pugh, *Providing Reference Services for Archives and Manuscripts* (Chicago: Society of American Archivists, 1992, and "Illusion of Omniscience: Subject Access and the Reference Archivist," *American Archivist* 45 (Winter 1982):33-44, reprinted in *Modern Archives Reader* (Washington, D.C.: National Archives, 1984). See also Laura B. Cohen, ed., *Reference Services for Archives and Manuscripts* (New York: Haworth Press, 1997).

8 Repositories of Primary Sources <http://www.lib.uidaho.edu/special-collections/Other.Repositories.html>.

9 <http://www.tulane.edu/~lmiller/ArchivesResources.html>.

10 NUCMC <http://lcweb.loc.gov/coll/nucmc/nucmc.html>.

11 Berkeley Digital Library SunSite <http://sunsite.berkeley.edu>.

12 <http://archives.chadwyck.com>.

13 Elizabeth Yakel and Laura Bost, "Understanding Administrative Use and Users in College and University Archives," *American Archivist* 157 (Fall 1994) 596-615.

14 Linda J. Long, "Question Negotiation in the Archives Setting: The Use of Interpersonal Communications Techniques in the Reference Interview," *American Archivist* 52 (Winter 1989): 40-51, and Helen Tibbo, "Interviewing Techniques for Remote Reference: Electronic Versus Traditional Environments," *American Archivist* 58 (Summer 1995): 294-311.

Mary Jo Pugh is Supervisory Archivist, San Francisco Maritime National Historical Park.

This paper was written in part as a product of the author's participation in the 1997 Research Fellowship Program for the Study of Modern Archives administered by the Bentley Historical Library, University of Michigan and funded by the Andrew W. Mellon Foundation, the National Endowment for the Humanities, and the University of Michigan.

Betsy Chittenden

Records Management in the National Park Service

In March of 1997, a small reorganization in the NPS Washington administrative office brought the records management job to my office, a new entity called the Washington Administrative Program Center. When we began digging into our new area, what we found was not so much a functioning program, but fragments of a program, a program at its nadir. **The status of records management in the National Park Service at the end of the 20th century is that of a program that must be rebuilt nearly from the ground up, at the same time that it begins to tackle the immense challenge of electronic recordkeeping in a souped-up cyber world.**

Since the 1980s the National Park Service has struggled with flat budgets matched by expanding responsibilities, the downsizing of government in general, and fast fires (both real and figurative) that keep the staff busy. Not unexpectedly, the quiet activity of recordkeeping, whose customers are future managers and researchers, has not fared well. Recordkeeping practices in the National Park Service can be sublime (a professional archives at Yellowstone, officially affiliated with the National Archives, and with a master inventory of records available on the Web), but in most locations are marginal—subject files at desks, boxes in storerooms and attics, and no one around who knows what to do with it all. Particularly since 1994, as NPS has undergone a top-to-bottom reorganization under specific direction to downsize central offices, recordkeeping activities in headquarters and regions have been neglected out of necessity—no one to do them, too much other pressing business. In many places, records management has now been neglected so long that awareness of its basic elements—say, what a file code is, and why it is put on correspondence—has been lost. Staff motivated to tackle their records don't have much help, either—NPS-19, the internal guidance to NPS staff on records management, is bureaucratic and unhelpful, with a complex and outdated records disposition schedule. Records management has a clerical "central files" image, not intellectually connected with the information age or the NPS mission.

But as we have taken stock of records management over the last year, we have found good news as well. The Service has a long-term mission to preserve cultural and natural resources "unimpaired for future generations"—this creates a market and a direct mission-related need for records management for many types of park records. As an agency that has responsibility for cultural resources, the National Park Service has a small but vigorous community of historians, archivists, and curators. These professionals understand and strongly support the need for records management, and form a core group with expertise in many aspects of hands-on records management. The natural resources community of scientists also understands the need for long-term data retention and access. Around the Service, a handful of excellent records management and archival projects were ongoing in a few locations (see articles by Mary Jo Pugh and Susan Kraft elsewhere in this issue). The HABS/HAER program (Historic American Buildings Survey/Historic American Engineering Record), has its own records group number and functions as an affiliated archive (see article by Jerry Wallace in the upcoming archival issue of *CRM*). And the National Archives strongly supports our renewed efforts to revive the records management program.

Where to start? **We decided to start at the beginning, by rewriting entirely our guidance and records disposition schedule**, the old NPS-19, and work groups began work last summer. It was immediately clear that a new, "customer friendly" approach was needed to make records management simple, understandable, and worthwhile to people. Dry recitations of regulation and law are not enough; with many other jobs competing for precious staff time, records management, in essence must be "sold" to its customers as a responsibility worth their time and effort. Furthermore, with staff reductions, the day of the trained records manager, with few or no responsibilities other than records and files, has gone. Records management on the ground is going to be largely done as a collateral duty. The work groups began working on reducing and simplifying the immense and complex file codes and files disposition schedule, reformatting it into a new, user-friendly "plain English" (question and answer) format, and simplifying files disposition instructions.

It has also become clear we need to develop recordkeeping techniques and strategies that fit

NPS culture and the realities of park operations. One of the problems with "selling" proper records management in the NPS has always been that, in the end game, records were moved completely out of NPS control, and greatly reduced accessibility for NPS staff, while the long-term resource management mission of NPS requires that some records be kept and used almost indefinitely. The National Park Service has for many years been at loggerheads with the National Archives and Records Administration (NARA), the agency with the lead in managing records in the federal government, with the root of the dispute lying in a clash between the mission of these two agencies. In meeting its core mission of managing and preserving natural and cultural resources "unimpaired" for the next generation, NPS staff use many permanently valuable records on a daily basis. For example, the records of the National Register of Historic Places, dating back to 1966, are referenced daily by NPS staff. Collections of original photographs at the NPS Harpers Ferry Center—some from the WPA era, some by Ansel Adams—are used to prepare exhibits and park brochures. Yet federal records law requires that permanent records be physically and legally turned over to the National Archives after 30 years—and caring for those records is one of their core missions. **Two agencies, each with proud, long-term missions that dedicated staff were bent on fulfilling, found it difficult to come to a consensus on what to do.**

The fact that records law requires moving records out of the NPS and to another agency clashed with other parts of the NPS organizational ethic as well. The NPS is very proud of its own history and traditions, and this makes it difficult to go through a process which results in records being removed from direct NPS care and access. The geographic dispersion and isolation of parks also conflicts with standard NARA processes. For remote parks, moving your records to a records center or archives hundreds of miles distant may make them more accessible to the public, but makes doing your job as a ranger or superintendent more difficult. "Have the boxes sent back from the records center" is a degree easier in Washington DC, or Philadelphia, than in parks where a trip to the grocery store is an all-day excursion.

Clearly, for records management to ever be practiced on a wide scale in the National Park Service again, it is essential to develop alternatives to the traditional records-keeping paths that fit the NPS circumstance. Several NARA employees joined both our work groups and we began to explore options. **We found common ground by returning to the basics of NARA's mission in preserving permanent records: that they be well-cared for, and accessible to the public.** A number of options are now being explored that may allow NPS to keep more of its records close to home, if the National Park Service commits to caring for these records using archival standards, and to make them more accessible to outside researchers. For example, the NPS-19 work group is looking at creating a new NPS records category of "permanent active," to be applied to records that are permanent, but because they are in active use by the NPS, would remain in the custody of the NPS and not be transferred to NARA until they become inactive. For these permanent active records, the NPS would set standards for their care and public access that satisfy the intent of records management law and management accountability, and NPS managers would be required to make a commitment to meet those standards as a condition of maintaining records locally. Another alternative might be the development of in-house records expertise at NPS locations that could provide professional records services to small or isolated NPS parks without resources or facilities to care for their records on-site.

Records management is also on the difficult cusp between paper and electronic, a transition all enterprises are struggling with. What does it mean to the National Park Service—to any organization—to shift a large portion of its communications from letter and phone to email? Or to have a whole new medium of communication with the public open up on the Internet? The ubiquity of electronic documents, email, and the Web have thrown records management its greatest challenge since mankind stopped using clay tablets and had to learn how to preserve paper. **The electronic media evolve so quickly that the question is not what are archivally stable storage media, but what format can information be put into so that it can be read on available hardware and software 10 or 20 years from now.**

Email—which the courts ruled several years ago does constitute records—is exchanged in volumes that are exponentially greater than paper

The National Park Service Records Management Handbook Taskforce shown at their August 11-15 meeting at the Huff House in Roscoe, New York. Individuals included are: standing, left to right: Jerry Wallace of the National Archives, Carla Hahn, Phyllis Hahn, Elinor Aye of NARA, Debra Melton, and Mona Hutchinson; sitting left to right, Diane Vogt-O'Connor, NPS Records Manager Betsy Chittenden, Mary Beneterou, and Pinky Salley.

communications. With 10,000 mailboxes, the National Park Service is conservatively estimated to generate 10-20 million messages each year. All of them are not records—but legally, all of them need to be evaluated as to whether or not they are records. It's impractical to have anyone but the originators of those 20 million messages make the determination as to whether each message is a record—how will we teach all those employees to do that? And how are the thousands of records then indexed and stored? Guidance NARA issued in 1995 provided that email deemed to meet the definition of a record could be printed out and filed in paper recordkeeping systems. This was a clumsy solution, but at least marginally workable. There was no real alternative—satisfactory electronic recordkeeping software was not on the market and no federal agencies were in a position to handle filing email electronically. But in 1997, a Federal judge in another case ruled that this was not acceptable, that in fact, records created electronically must be stored electronically. "Simply put" the court held, "electronic communications are rarely identical to their paper counterparts; they are records unique and distinct from printed versions of the same record."[1] **We are now forced to face the reality of electronic recordkeeping for electronic records.**

At 75 million visits per year, the National Park Service's "ParkNet" Web site is now visited by more people than any single park in the system, and is rapidly approaching the 270 million visits annually to all parks combined. ParkNet <www.nps.gov> contains hundreds of individual pages about parks and National Park Service activities, from virtual tours to press releases to draft park planning documents out for public comment. Clearly ParkNet needs to be preserved as a record of what the National Park Service says and does. But ParkNet changes daily, as various programs, parks, and offices post and remove information. How can this be preserved? How often must a Web snapshot be taken? And again, how can something so technically complex, with videos, sound, and links to other sites, be archived in a way that it can be read and experienced in 20 years—or even 5?

To these questions there are no easy answers—no real answers at all as yet. **The National Archives and the Department of Defense are developing functional requirements for the first generation of true electronic recordkeeping systems, and the first commercial products are now coming on to market.** But these are add-on products, and the true solutions—recordkeeping and archiving built in to your email software, for example—seem far off. The U.S. Patent and Trade Office, which has serious recordkeeping responsibilities, wrote Microsoft directly to ask if they would work on electronic recordkeeping software—and were told no. The Web question bumps technical complexity up another degree of difficulty.

Yet the Web, and a new law, the Electronic Freedom of Information Act (E-FOIA) of 1996, also begin to suggest a path to solving two chronic recordkeeping problems—those of access, and resources. **The E-FOIA requires that certain types of agency documents, such as policy and guidance, all commonly used documents of interest to the public, be made available in "electronic reading rooms"—the Web.** The intent is for the federal government to become proactive, rather than reactive, in making available to the public the records that it is most interested in using. A permanent record posted on the Web is a record that is far more accessible than any paper document ever could be, and to some extent obviates the need for moving records to central archive locations. The new E-FOIA requirement is also focusing new attention—and possibly new funds—on records management, especially electronic records management. The administration's FY99 budget request now in Congress includes $1 million in permanent base funding for implementing E-FOIA in the NPS. If this request survives, it would go a long way toward assisting hundreds of NPS locations to make thousands of valuable records available electronically, and provide base money to start tackling the permanent electronic storage that the law now requires.

Records management in the National Park Service has a long way to go. The new user friendly NPS-19 and records disposition schedule aren't written yet, much remains to be worked out with the National Archives, the folk knowledge of recordkeeping requirements is still slipping away, and the E-FOIA money isn't here. But the extraordinary mission of the National Park Service, to preserve resources "unimpaired for future generations," means that records management only needs some creative thinking, some well-crafted solutions, and some culturally-sensitive marketing to bring it alive again. And the promise of the information age, which will let us bring extraordinary photographs and important park management plans to audiences that could never have seen them five years ago, makes it a wonderful time to be in the records business.

Note

1 *Public Citizen, Inc., et al., v. John Carlin, Archivist of the United States*, U.S. District Court, 1997 (Civil Action 96-2840(PLF))

Betsy Chittenden is the Records Manager of the National Park Service.

Michael F. Brown

Cultural Records in Question

Information and Its Moral Dilemmas

On a clear day in November 1997, as the late autumn sun warmed the mesas of northern Arizona, I sat in the crowded office shared by Lee Wayne Lomayestewa and Clyde Qotswisiuma of the Hopi Cultural Preservation Office, an agency of the Hopi Indians' tribal government. I had come to the Hopi reservation to discuss the future of cultural records held in the nation's repositories. In 1994, the chairman of the Hopi Tribe sent a letter to dozens of museums and archives requesting that they close Hopi collections to researchers who had not first obtained the tribe's written permission. Among other things, I wanted to know how public institutions responded to this request and, more important, why the Hopi harbored such strong feelings about documents that lie mostly unnoticed and unused in distant storage cabinets.

Our conversation began slowly. I was a stranger, possibly with an axe to grind, adding his questions to those of countless outsiders whose persistent curiosity baffles and sometimes exasperates the Hopi. But after an awkward silence the two men began to explain the tribe's policies regarding NAGPRA, The Native American Graves Protection and Repatriation Act of 1990, a law that has radically transformed relations between Indian tribes and America's museums. NAGPRA spells out procedures for handling Native American burials, grave goods, and items of religious significance, which can be repatriated to tribal claimants meeting certain conditions. As with most judicial processes in the United States, the law calls for claims to be substantiated with archival and testimonial evidence. This puts tribes in the awkward position of having to reveal secret religious knowledge in order to prove that contested items are in fact sacred. As Clyde Qotswisiuma observed, "Even something like a digging stick could have a ritual use, but we're not about to say what it is."

The dilemmas faced by the Hopi Cultural Preservation Office as it wrestles with the terms of NAGPRA have given rise to new conflicts likely to engulf archives in the United States, Canada, Australia, and elsewhere in the coming decades. **The central issue is the disposition of potentially sensitive cultural information**, including photographs, sketches, audio tapes, inventories of ritual objects, anthropological fieldnotes, and transcriptions of oral literature. The struggle pits native nations against the institutions entrusted with cultural records. **At the heart of this conflict are two irreconcilable views of information.**

The Moral Meaning of Information

American law, and the liberal democratic tradition in general, place a high value on the unfettered exchange of information. **It is an article of faith that openness fosters artistic creativity, encourages scientific innovation, and insures political accountability.** For reasons of personal privacy or national security, of course, access to information is sometimes restricted. But because history has shown time and again that governments readily hide improper behavior behind a screen of official secrecy laws, we work hard to restrict their scope and duration. Free access to information, in other words, is seen as a cornerstone of democracy and a key element of open societies.

Among many indigenous peoples, a different attitude prevails. The social fabric of native nations often consists of reciprocal spheres of knowledge, the boundaries of which are zealously protected. Elders preserve information that they share only with those who demonstrate required wisdom. Women and men have understandings unique to their gender, fostering complementarity that helps to keep spouses together in times of trouble. Specialized religious cults conserve practices that may extend back to a distant time when peoples with diverse traditions merged to form a single society. **The uneven distribution of information thus strengthens social bonds while insuring that powerful knowledge remains in the hands of those who know how to control it.** To outsiders, this patchwork approach to knowledge seems artificial, but to cultural insiders it is simply the way things were meant to be.

Indigenous attitudes toward knowledge were intensified by the colonial experience. In 17th-century Peru, for example, the Spanish priest Pablo José de Arriaga gathered information about the religious beliefs of local Indians not to preserve them for posterity but to "extirpate idolatry," his contribution to the Church's evangelization campaign. Closer to home, information about American Indian religions was used against practitioners

until quite recently, when Indian freedom of religion was guaranteed by federal law. **In the colonial setting, native peoples thus survived by protecting knowledge behind a wall of silence and sharing it only when necessary.**

Today many native groups perceive themselves as less threatened by overt persecution than by the rapid circulation of images of their cultures—sometimes accurate, sometimes wildly distorted—via the popular media. Particularly upsetting to American Indians are religious seekers, many involved in the New Age movement, who insist on performing ersatz versions of Native American rites, including sweat-lodge ceremonies and Medicine Wheel rituals. **Seeing their religions parasitized by outsiders, Indians feel a powerful urge to re-establish control over information about their cultures and, in particular, about traditional ritual practices, pilgrimage sites, and sacred stories.** Archives and other institutions that care for cultural records become lightning rods for this impulse because they, unlike the diffuse New Age movement and the culture from which it arises, are obliged to respond to criticism from members of the public.

The struggle over public knowledge about traditional religious life dovetails with broader concerns about the future of indigenous peoples' intellectual property, which is routinely appropriated by a majority culture hungry for novel ideas. Native art and music, local knowledge of medicinal plants and crop varieties, even the gene sequences of isolated populations—all remain largely unprotected by existing intellectual-property laws, making them easy targets for acquisitive outsiders (Greaves 1994).

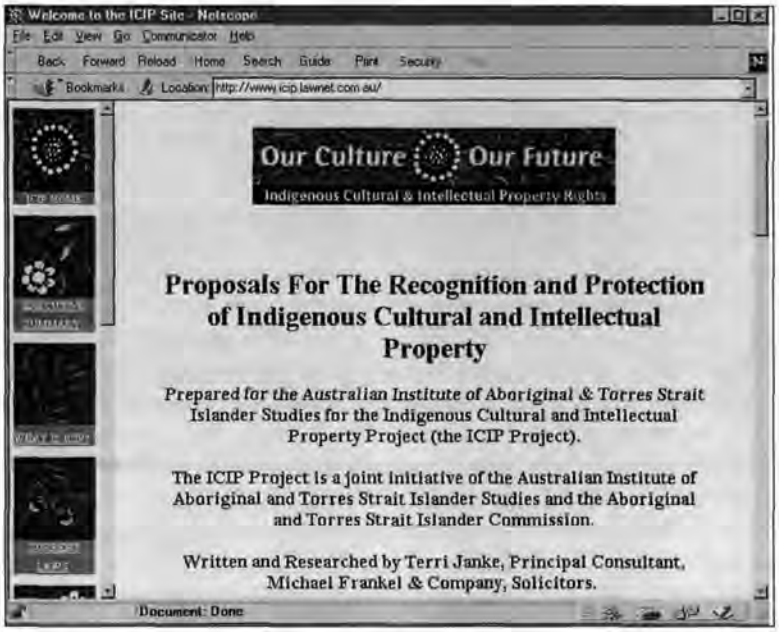
H. R. Voth preaching at Hopi, Winter 1901. Photo courtesy Mennonite Library and Archives, Bethel College, North Newton, Kansas.

Archival Ironies

For professionals who dedicate their lives to the conservation of irreplaceable cultural documents and who take understandable pride in their ability to make them available for public use, the growing firestorm of criticism comes as a shock. After all, archival materials have played a major role in countless legal decisions that have restored tribal lands, led to the protection of sacred sites, and helped native peoples assert their cultural sovereignty. Historical and ethnographic records continue to provide essential information for native groups trying to reconstruct the belief systems of their ancestors. **It is a particularly cruel irony, then, that repositories are now criticized for fulfilling their mission too well.**

Increased sensitivity to questions of historical injustice and colonial privilege forces us to confront the moral ambiguities of archival records. Consider, for example, the photographs and fieldnotes of the Reverend H. R. Voth (1855-1931), a Mennonite missionary and ethnologist who lived among the Hopi for more than 20 years. Voth's dual vocation presented a major conflict of interest, and at times it is hard to tell whether his goal was to destroy their religion or to preserve it through careful documentation. In any case, his photographs and first-hand observations of Hopi rituals are among the best ever recorded, and they figure importantly in most studies of Hopi culture published since the 1920s. Hopis are bitter about Voth's success in penetrating their ritual life. In the memorable Hopi autobiography *Sun Chief*, published in the 1940s, Don Talayesva remembered Voth as a "wicked man" who had "stolen so many of our ceremonial secrets." Fifty years later, sentiments remain unchanged. The Voth material, tribal leaders insist, continues to damage Hopi culture by making public a wealth of esoteric information that should be available only to authorized religious experts.

Yet would anything be accomplished by closing the Voth collection, however compromised its moral status? Voth's photographs and fieldnotes have been reproduced in countless books, so quarantining the originals would have little impact on the circulation of information about Hopi customs. **Mindful of precedent, archivists must also wonder whether the imposition of moral quarantines would set off a wave of similar requests by politically or religiously motivated groups offended by specific collections.** At risk is an honorable tradition of archival impartiality—one that led the Mennonite Archives of Bethel College, a major repository of Voth material, to grant permission to publish one of Voth's photographs despite my unflattering portrait of his work.



The Web site of the Indigenous Cultural and Intellectual Property Project <http://www.icip.lawnet.com.au> presents specific legal proposals designed to protect the cultural property of Australian Aborigines from inappropriate uses.

Prospects for the Future

In Australia, where the disposition of information about Aboriginal religion has long been a matter of debate, public repositories now informally limit access to sensitive documents and artifacts (Anderson 1995). Ironically, this often means that **Aboriginal** staff members are prevented from handling such collections unless they are members of the community from which the material comes. Aboriginal communities, it seems, worry more about transfer of ritual secrets to other Aboriginal groups than about their use by non-Aboriginals. In some cases, collections are also off-limits to female staff members, again in deference to community wishes. New intellectual-property laws currently under consideration would formally limit non-Aboriginal researchers' access to Aboriginal materials in Australia's archives (Janke 1997:54).

Given the anti-discrimination laws currently in force in the United States, it is hard to imagine that the Australian model could be followed here. **The personnel of most repositories would be reluctant to deny access solely on the basis of a patron's ethnic identity, gender, or religious affiliation. So how can they respond to native demands?**

Some archives are already following the common-sense practice of marking certain collections as sensitive and urging researchers to contact the appropriate Indian tribe before using them. Others are exercising greater care when preparing exhibits to insure that they do not contain religious information inappropriate for Indian children and uninitiated people in general. Most have opened dialogues with those communities that have the greatest claims on particular collections.

Nevertheless, pressure may be mounting for new legislation that would extend NAGPRA into the realm of cultural records (Nason 1997). Before this can happen, however, courts and legislators must answer difficult questions. Are some cultural records so morally contaminated that they should be closed to the general public? Does a culture "own" its traditions, or do they properly belong to the individuals who create and transmit them? In the interests of preserving indigenous societies, should free speech and freedom of information be curtailed by government edict? Finally, should we recognize an inherent right to "cultural privacy," a concept mentioned in a recent conference calling for fundamental changes in the relationship between the Hopi tribe and outside researchers *(Dialogue with the Hopi,* 1995)? Debate over these complex issues is likely to dominate conversations between archives and native peoples well into the 21st century.

References

Anderson, Christopher, ed. *Politics of the Secret.* Oceania Monograph 45. University of Sydney, 1995.

Dialogue with the Hopi: Cultural Copyright and Research Ethics, 1995. (Copies available from Paaqavi, Inc., P.O. Box 747, Hotevilla, AZ 86030).

Greaves, Tom, ed. *Intellectual Property Rights for Indigenous Peoples: A Source Book.* Oklahoma City, Okla.: Society for Applied Anthropology, 1994.

Janke, Terri. *Our Culture, Our Future: Proposals for Recognition and Protection of Indigenous Cultural and Intellectual Property.* Canberra: Australian Institute of Aboriginal and Torres Straits Islander Studies, 1997.

Nason, James D. "Native American Intellectual Property Rights: Issues in the Control of Esoteric Knowledge," in *Borrowed Power: Essays on Cultural Appropriation.* Edited by Bruce Ziff and Pratima V. Rao, pp. 237-254. New Brunswick, N.J.: Rutgers University Press, 1997.

Talayesva, Don. *Sun Chief: The Autobiography of a Hopi Indian.* Leo W. Simmons, editor. New Haven, Conn.: Yale University Press, 1942.

Michael F. Brown is the James N. Lambert Professor of Anthropology and Latin American Studies at Williams College in Williamstown, Massachusetts. He is currently writing a book on the future of cultural property.

Harrison Eiteljorg II

Archiving Archeological Data in the Next Millennium

A few years ago a colleague and friend of mine called from Turkey to the institution where he had once taught. He was looking at materials from an excavation he had conducted while on the staff at that institution some 30 years earlier, and he needed information from his own records—records that remained in the archives of his former employer. Within a few days he had his answer.

My own archival materials are different from my colleague's; they are electronic files rather than paper records. As a result, were I to call from Athens to my own office on a similar mission, in this case to get some information from my computer-aided design (CAD) model of the older propylon, I fear that the outcome would not be as successful. Although the computer file could surely be located, I would need to offer instructions at every step of the process so that the file could be opened and the information obtained. That is, I would have to explain not only what information I need and where it might be, but what program to use and how to retrieve the specific information required. **How, then will someone else find that information years from now?**

My concern about the difficulty of retrieving information from my own archival repository prompts a broader concern for archival storage as we approach the new millennium. Archeologists desperately need access to data from past excavations, not just the publications (which too often do not appear) but the raw data collected in the field. If electronic data complicates the archival process, then **our apparent progress in record keeping may be regression instead.**

I intentionally painted a sharp contrast in the preceding. There are doubtless things in paper archives that are all but impossible to find, and there are surely things in electronic archives that are easy to find. Nonetheless, there are special problems with computer archives that should concern all archeologists, issues that affect the ease with which data can be retrieved and, more important, the security of data in an archive.

At this point I can turn this discussion in either of two directions: technical or practical. I can talk about the technical issues that will bedevil those of us interested in electronic archives, or I can talk about the more practical problems that will crop up as we try to use the technology and preserve the records that are its fruit. As an admitted technological optimist, I will assert that the technical issues can be solved; furthermore, I think they are of little or no interest to the majority of scholars. The practical problems, on the other hand, can only be solved by those archeologists who are prepared to do things that may be unappealing. So I think the real-world difficulties encountered by archeologists who must create and preserve electronic data are more interesting to readers—and more significant for the profession.

The practical problems relate to two different parts of the archeological process—first, general dig planning and direction and second, treatment of the electronic files at the end of the line. Let me start with the issues that surround general dig planning and direction.

General Dig Planning and Direction

Excavation directors are a bit like orchestra conductors; both direct specialists, each of whom must be able to do much of his/her work without the direct intervention of the leader. Both worry about choosing the specialists, how they all fit together, and timing. As the conductor does not tell the violinist how to tune his instrument, so the excavation director does not tell the pottery specialist how to construct a database for the pottery. The director will watch over the utility of the end product, not the details. That is true for all the specialists.

Over time, directors have learned that there are some unexpected things that must be watched, though those matters may seem to be beyond their ken. For instance, how many worried about the kind of paper used for notebooks in the 1930s or, in later periods, the permanence of the film stock used by staff photographers? As those turned out to be critical items, so issues surrounding software used on sites—specific programs chosen and modifications made—are now important issues that I do not think all directors have recognized and confronted. When software is chosen, for instance, directors must ask their specialists such pertinent questions as "In what forms will/can the data be stored?" or "Can I use this data on a MAC/PC since

you are going to store it on a PC/MAC?" or, most important of all, "Can we integrate your data with that of the other specialists?"

Some excavation directors are asking those questions now, but issues surrounding archival storage of the data are still missed too often. Most important are choices of computer file format. If data are stored in uncommon file formats, then, at the least, extra work will be required when the files are ultimately archived, because the format will have to be changed. In extreme cases, the data may be all but useless in electronic form, and it may be necessary to print everything on paper to preserve the information. I should point out that, for very long-term excavations, these questions of file formats can have more immediate repercussions. As a dig progresses, computer power surely will change, but software chosen may not. Therefore, consideration of file formats may be very important for the day-to-day operation of a dig. Should the chosen software cease to be the best for the work, it should—no, must—be possible to bring along the data as the underlying software is changed.

As I see it, then, excavations directors, as they plan and direct the recordkeeping process, should concern themselves with issues of data storage. They do not currently ask their architects what pens or papers they use, but they do ask about permanence of the drawings. The same concerns—with a few wrinkles due to the technology—apply to electronic data storage; so directors must make certain that the data files created are in useful, modern formats and can be moved, if necessary, to standard formats for archival storage or data transfer.

Preparing Files for the Archive

Now I want to turn to the problems with the treatment of electronic files at the conclusion of a project. I will assume that the data have been stored in appropriate formats. When the project is complete, the director is responsible for making the archival arrangements. I think the biggest problem here is that the director must take an active role in this process, though that may not have been required for archival preservation in the past. It may once have been possible to assume that archivists would, as they must, accept what they received (or retrieved from vacated offices) and make the best of it. Electronic files simply cannot be treated that way.

I will use my CAD model of the older propylon as an example here—a 3D computer model of the remains of the entrance structure for the Athenian Acropolis before 437 B.C. Leaving aside the potential problem of finding the file on its hard drive and recognizing it as worthy of archival treatment, an archivist must recognize the electronic file format and know whether it is current or not, whether there are more appropriate formats, whether there are standard formats for the particular data type, whether it must be transferred to one

Upcoming National Park Service Courses

If you are interested in either of the following two courses contact Gay Tracy, NEDCC, 100 Brickstone Square, Andover, MA 01810; Tel: 978-470-1010 or ccmail <tracy@nedcc.org>

Afterimages: Reformatting Visual Materials in a Digital World ($100 for a limited number of NPS staff, $275 for all others) will be held September 16th-18th, 1998, at the National Archives and Records Administration Archives II Facility in College Park, MD, and is co-sponsored by NPS, NARA, and NEDCC. **Afterimages** will teach archivists, curators, historic preservation specialists, librarians, and other CRM staff how to:

- plan and manage projects to reformat visual materials, such as cellulose nitrate and acetate negatives, including contracting with an outside vendor
- select and prepare visual collections for reformatting including preservation and handling
- understand best practices, benchmarks, and quality control for photographic and digital imaging
- manage legal issues, contracts, cost containment, and cost benefit analyses

School for Scanning: Issues of Preservation and Access for Paper-Based Collections ($255 before October 15, 1998, $325 thereafter) will be held December 7-9, 1998, at the Le Petit Theatre du Vieux Carre in New Orleans, LA, and is co-sponsored by the NPS, the Getty Information Institute, the Historic New Orleans Collection, SOLINET, and NEDCC.

This course will teach participants:

- the basics of digital technology, including deciphering digital jargon
- content selection for digitization including legal issues
- text and image scanning, including costs, quality control, metadata, and multi-versioning
- electronic publications

of those formats, and which format (if any) can be expected to remain current for a reasonable length of time.

Once the question of format has been dealt with, the archivist must confront questions relating to the use of the model by others. For instance, my model has more than 200 different data segments. In-situ stones of cut marble, with specific date span, and lying in the stair of the entrance structure are in one data segment; the nearby tripod base, also of cut marble and with the same dates, is in another. The particulars of these data segments are not important for this discussion, but the model cannot be used effectively without an understanding of the segments and the way they have been named. That information is not implicit in the model; it must be supplied in a set of documentation that I must have prepared.

Included in the model are blocks that I measured with tapes and line levels as well as blocks that I surveyed with photogrammetric techniques. As a result, there are different levels of confidence to be placed in different parts of the model. Users of the model need to know that so they can assess the accuracy of specific data. However, the difference between the parts of the model measured with tapes and the parts surveyed with photogrammetry are not apparent to a user of the model, and the difference cannot be determined with the model alone. Again, I must have supplied information about survey methods if the model is to be used to maximum advantage. There are also data files attached to the CAD model. The formats of the files, the fields used, the limits on terms used for the data, and much more must be given to users so that they can use these files as well.

Once more, I must have supplied that information along with the files themselves. **In short, I, as supplier of the data, must have supplied considerable documentation along with the data files.** In an ideal world, there would have been similar documentation to accompany paper files, and that documentation has often been missing as well. However, there is a critical difference between the paper and electronic files. The paper files, by and large, can be used without the documentation. The terms can be teased out of internal relationships and usage; the organization can also be determined. It may take time, but it is possible. In the case of computer files, on the other hand, the relationships are often impossible to find, and the documentation is much more central to the utility of the files. Time simply may not be enough.

At the conclusion of a project, then, the director must produce the documentation required for archival storage—and he/she must do so quickly. The need for documentation has already been spelled out, but I believe that delay is an important problem as well. Not only is it easy to put off the work and, in the process, lose track of important information, the time lag also creates problems peculiar to electronic data. The longer the delay, the more likely it is that the files will be compromised by neglect. File formats may become obsolete, files may decay, or they may simply be lost. I do not believe they are safe when stored for long periods on institutional mainframes or servers; nor do I believe it is safe to leave them on an individual hard disk, not to mention a floppy disk. Time is of the essence.

Conclusion

So the new millennium is coming. Along with it come new forms of data storage. We can't afford the old, casual processes for archiving in the new millennium with its new electronic forms of data. Individual scholars and institutions must examine their priorities, assign a higher level to archival storage, and insist on meeting the ethical requirements of archival preservation. The technical problems, I believe, are small. However, particularly if our past record—in terms both of publication and past archival preservation processes—is a guide to the future, our success is far from assured. I fear that in this work, as the cartoon character, Pogo, once said, "We have met the enemy and he is us."

Harrison Eiteljorg II is Director, Center for the Study of Architecture and the Archaeological Data Archive Project.

Assistant Chief Ranger Trischman of Yellowstone at Upper Lamar, Yellowstone. Anonymous.

Kellee Blake

The Stories You Could Tell
Using NARA Regional Resources for Site Interpretation

The interpretive program of the Parkway has had only a good start; it is on paper for the most part. The delay occasioned by the war has been costly in deterioration of many of our exhibits. We must take up the work soon or all will be lost ....

Stanley W. Abbott
Resident Landscape Architect, Blue Ridge Parkway, 1946

A unique, colorful, and often overlooked perspective on National Park Service sites awaits NPS historians, staff, and other researchers who take up the work of interpreting these national treasures. Interpretive decisions, tourism efforts, CCC and WPA programs, local culture, site nomenclature, neighborhoods, industry, even issues of race, class, and gender are just some of the topics captured in archival NPS records held by the regional facilities of the National Archives and Records Administration (NARA) Office of Regional Records Services.

Though typically created to document site development, administration, and use, the NARA regional records also provide a remarkable treatment of many cultural, environmental, and historical issues. These files (some dating from the turn of the century to the early 1970s) regularly include maps, reports, building plans, news clippings, and photographs, but the files and their contents are as varied as the parks themselves. If the records were only used, the stories they could tell!

Some NPS site researchers may presume that most materials created by or about a site remain at the site, are at NPS headquarters in Washington, or have been forwarded to an NPS specialized office. Over the years, however, many individual NPS sites and NPS regional offices transferred their permanent materials to NARA regional facilities responsible for permanent federal records from specified areas. For example, the NARA-Northeast Region facility in New York serves New York and New Jersey. Records of the NPS are maintained in NARA as Record Group 79. **Note**: The inclusive dates, scope, content, quality, and quantity of NPS materials varies from region to region, and all the types of records described below may not be available at all NARA regional facilities.

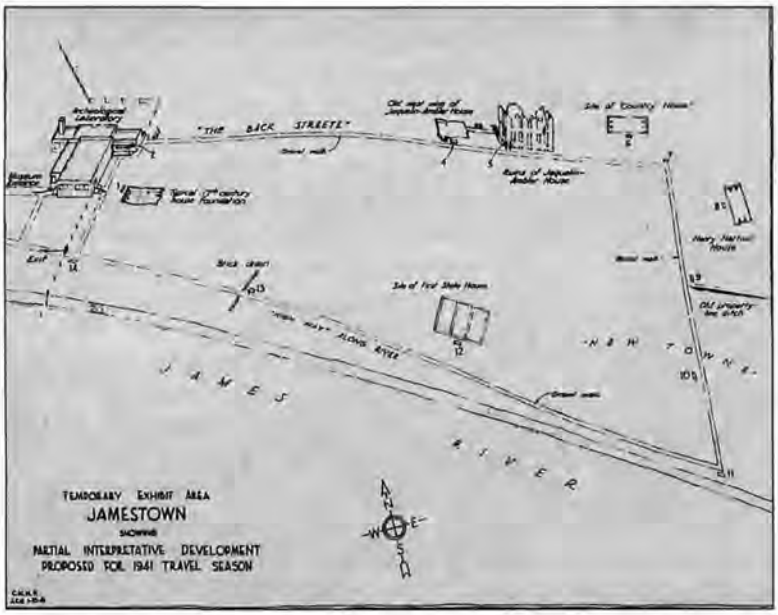
NPS Records

NPS site files often begin with projects to acquire land. Negotiations for the acquisition of sites usually reveal substantive information about prior use and ownership, related assets, and the people of the community at the time of the government's earliest interest. Early site appraisals often include photographs, maps and drawings, descriptions of current residents or neighborhoods, mineral and water surveys, agreements with municipalities, reports on proposed land use, debate over essential property, and offers to sell land to the United States or challenges to the acquisition. When landowners would not sell and the United States condemned the property, the NPS files often note the Federal court condemnation case number. Most NARA regional facilities maintain the Federal court condemnation cases for states in their region, and these can provide additional information regarding property provenance, genealogy, and social relationships as the court clarifies title or reconciles disputes over land values.

The NPS records capture the earliest ideas about how parks would be administered, from where, and by whom. They also provide a unique window into some of the first decisions made about site interpretation. Should, for example, the Great Smoky Mountains National Park be interpreted as a scenic natural area or as a focal point for the pioneer experience? How could the naturalistic aspects be reconciled with the lived-in look of occupied areas? Could the awe-inspiring beauty be shared with, yet protected from, thousands of visitors? How should the people of the area be represented? These are just a few of the interpretive issues documented in

UNITED STATES
DEPARTMENT OF THE INTERIOR
NATIONAL PARK SERVICE

NPS records at NARA regional facilities may include copies of newsclippings, event programs, and brochures such as this 1940 Blue Ridge Parkway brochure found in Central Files of the NPS Region I office. (Record Group 79, NARA-Mid Atlantic Region).



Plan for partial development of exhibit area at Jamestown Island proposed for the 1941 travel season, as found in records of the Colonial Natioal Parkway. (Record Group 79, NARA-Mid Atlantic Region).

records of the Great Smoky Mountains National Park held by NARA facilities in Atlanta and Philadelphia.

Early decisions about cultural resource interpretation often gave rise to debate over the appropriateness of existing structures and which structures, even areas, should be rehabilitated or set aside. Building appraisals in the files may be very general or they may provide detailed descriptions of structures, including inventories of site-specific furnishings, photographs, or drawings. **The files may also provide evidence of earlier preservation or restoration efforts, or earlier treatments applied to the site.** New construction is inconsistently documented. In some files construction is described in general terms for higher offices, while others contain surprisingly thorough construction plans, photographs, drawings, contracts, and maps.

No less compelling are records relating to natural resources and landscape interpretation. Complete or partial landscape development plans may be found in the files, as well as plant lists, garden plans, vegetation surveys, or procedures for improving fields and woods. Instructions for the pruning of trees at the Andrew Jackson National Historic Home site in Tennessee, for example, provide early evidence of what trees were on the site at the time of acquisition. Depending on the type of park, wildlife surveys and records relating to regional and site-specific wildlife programs may be also be present.

National Park Service employees may be especially amused as they read of the experiences of those who went before them. Files may include copious instructions regarding local procedures, uniforms, staffing, and salaries—even the replacement of ranger badges. Any park employee or affiliate that has wondered "where did that idea come from?" may find the answers here. The files can be inspirational—vividly capturing the legacy of folks who fought to establish sites or dedicated themselves to their service. The records speak of the triumphs and tragedies of real people whose lives and the parks were connected.

Who could not be moved by the unfolding drama at Appomattox when townspeople protested the use of black CCC camps for park construction in 1940? Imagine the social climates in which Pine Spur on the Blue Ridge Parkway was established for the exclusive use of Negro visitors or in which Japanese evacuees were housed at former CCC barracks at Death Valley National Monument during World War II.

Challenges of Using NPS Records in RG 79

Researchers inspired to use these records should be forewarned of their inherent challenges. While many records were sent directly from individual NPS sites to NARA, many more were first forwarded to NPS *regional* offices and then accessioned by NARA as records of the NPS regional office. Few agencies have had a more complicated set of regional boundary changes than the National Park Service.

NARA regions are not necessarily the same as NPS regions. NARA regional facilities may have records outside their prescribed state holdings because the NPS regional offices administered sites beyond those states. **In some instances researchers may need to contact more than one NARA facility for a complete record of a site, especially if the NPS regional office has changed over time.** For example, some records of Olympic National Park can be found at the facilities in Seattle and San Bruno; Grand Canyon National Park is documented at facilities in Laguna Niguel (CA) and Denver; and records of Isle Royale National Park in Michigan are found at NARA facilities in Chicago, Kansas City, and Philadelphia.

Other Records for NPS Site Research

In addition to the NPS records in RG 79, NARA regional facilities hold materials created by other federal agencies that may also document parks or the areas, people, and activities relating to them. Most NARA facilities have a complete set of U.S. Population Schedules 1790-1920 for all states in which the federal census was taken and is extant, not just the states they serve. A snapshot of the nation taken every 10 years, federal census records can capture building occupants and neighborhoods over time

The value of federal court condemnation cases has been noted above, but other categories of federal court cases (RG 21 or RG 276) may also be useful for site interpretation. During the Civil War,

Arlington House. Photo by Jack E. Boucher, NPS.

Inventory of Genl. Lee's personal property at Arlington House Aug 27th 1863	
11 Sofa bottom chairs	2 Marble top Stands
4 Arm red plush seat chairs	2 Hair Matrasses
1 Bussels Carpet	1 large Wardrobe
3 Red plush seat Sofas	1 Mops Matraps
7 large Paintings in Hall	2 Straw "
1 Extention Dining table	2 fancy Candlesticks
1 large Painting of Washington + his officers on the Battlefield	2 Bureaus
	1 fancy glass case Clock
	7 Paintings
4 Both cases	7 engravings
6 Cane bottom chairs	2 Piano Stools
9 Bedsteads	1 Music Stand
2 Wardrobes	1 Shovel, tongs & bellows
3 Hush Matraps	1 Secretary
3 Stands	18 guilt frames
1 Centre table	1 large Map of Va
1 Common Rock Chair	2 Guilt Window heads
12 Cane B Chairs	~~[illegible]~~ 1 Wash Stand
1 small Both case	2 drefs bureaus & glasses
2 Bureaus	1 Carpet, & 3 Boxes of Books
1 pair hand Irons	3 Stoves
4 Ward Robes	1 Marble top Wash Stand
1 Marble top drefs bureau	1 large secretary
2 Marble top W Stands	4 Stands
2 Side boards	3 Stands
1 Marble top Centre table	2 tables

Inventory of Robert E. Lee's personal property at Arlington House at the time of confiscation by the U.S. District Court in Alexandria. (Record Group 21, NARA-Mid Atlantic Region).

for example, the federal court at Alexandria, Virginia, confiscated Robert E. Lee's Arlington estate and the extant case file at Philadelphia includes inventories of furnishings in the house at the time of confiscation. A Ninth Circuit Court ruling found at the San Francisco facility documents concerns about logging and mining at the Grand Canyon forest reserve in 1918. Federal criminal cases at the Fort Worth facility colorfully capture the work of Fort Smith's Hanging Judge Parker, and federal bankruptcy cases invariably enumerate the property and debts of bankrupt individuals and businesses.

The maritime aspect of many NPS sites may be documented in federal admiralty cases (RG 21), customs or lighthouse records, or lifesaving station logs (RG 26, 36, and 41). For example, the Chicago facility holds life saving station logs from Sleeping Bear Point Life Saving Station, now part of the Sleeping Bear National Lakeshore in Michigan. The Philadelphia facility has lighthouse drawings from the Assateague Island Lighthouse and many of the lighthouses on the Cape Hatteras National Seashore.

The NARA facility at Boston, like several other facilities, holds Bureau of Public Roads (RG 30) materials relating to national park road projects. Inmate and administrative records among Bureau of Prisons records (RG 129) at the San Bruno (CA) facility reveal a different view of Alcatraz Penitentiary. The facility at Anchorage maintains Alaskan Aerial Survey Expedition records (1926-29) in RG 57, providing early aerial survey records of Glacier Bay National Park. Several regional facilities have park-related materials created by other agencies including the Army Corps of Engineers (RG 77), Naval Districts and Shore Establishments (RG 181), War Assets Administration (RG 270), and Coast and Artillery Defenses (RG 392). *There are many more examples.* Careful considerations of how NPS sites were used over time and how sites, once established, interacted with other federal agencies should reveal additional NARA research possibilities.

Access to Archival Records in NARA Facilities

Regional reference staffs are equipped to discuss prospective researcher projects and direct researchers to appropriate sources, but they cannot undertake major research for them. All facilities have finding aids to their RG 79 and related records and some have detailed box and folder listings. **Reference queries can be made by phone, Internet, or in writing, and researchers are always welcome to view material in regional facility research rooms.** Prospective researchers are encouraged to call ahead prior to a research visit. General information on each facility, its holdings, and services is also available on the NARA Web site at <http://www.nara.gov/regional/nrmenu.html>.

Conclusion

The stories that **could** be told **should** be told. The quick samples illustrated here represent a fraction of the compelling stories found in more than 3,500 cubic feet of archival RG 79 materials held by the NARA regional facilities. Whether high drama or routine administration, these records provide a unique window into NPS sites over time for interpretation. They certainly should not be silent bounties of history tucked in Hollinger boxes. Those who now take up the work of site interpretation are encouraged to use them.

Kellee Blake is an archivist at the NARA-Mid Atlantic Region. A version of this paper was presented at the 1997 Linear Parks Conference.

Susan Kraft

The Yellowstone Archives and its Affiliation with NARA

Through special arrangement with the National Archives and Records Administration (NARA), Yellowstone National Park is classified as an affiliated archives of NARA. This arrangement, reached through a Memorandum of Agreement (MOA) in 1978, means that the park is permitted to retain physical custody of all those records that would normally be transferred to a central NARA repository in accordance with normal federal records disposition procedures. Records covered by the MOA document the administration and operations of the park from 1882 to the present and include records from the era of Army administration and records created or received by the National Park Service, certain other government entities, and some park concessionaires. The photograph collection includes work by such noted photographers as William Henry Jackson, Henry Bird Calfee, W.I. Marshall, and F. Jay Haynes. Video and audio tapes, maps and drawings also fall under the agreement, and are subject to NARA standards for the care and management of records.

The MOA, which is the only such agreement that has been reached between NARA and the National Park Service (NPS), came about largely because Yellowstone found it impossible to meet park management goals and the numerous and extremely varied demands of outside researchers without having its records onsite. Each year, the park archives and museum staff receive nearly a thousand reference requests. Many researchers visit the park personally, as their projects involve both field work and research in the archives. Studying a park feature—whether a geyser, waterfall, historic hotel, or a long-abandoned hotel dump—then having immediate access to the photographs, reports and other records that document its history, is an opportunity cherished by outside researchers and park staff alike.

Approximately half of those using the archives each year are park employees. Landscape architects rely heavily on the archives, having used its holdings most recently to draft Design Standards for the park (to ensure that future design, construction and maintenance projects respect and harmonize with park resources); to document and teach about changes in cultural landscapes over time; to document the existence and extent of nurseries and other abandoned agricultural operations within the park; and to discover important details of historic cultural landscapes, such as lighting fixtures, fences, benches and landscaping. Maintenance and historic preservation workers study historic building materials and techniques, sometimes referring to actual samples from the museum collection in conjunction with photographs. Both park interpreters and concessions employees with education responsibilities make extensive use of the historic photograph collection in creating slide programs for the public. Exhibit planners continuously draw on the archives in developing and designing waysides and other interpretive media. Park rangers consult historic photographs before undertaking maintenance or alteration of historic backcountry cabins. Having park records at hand and readily accessible has also improved the park's ability to produce reports, plans and other documents on its most contentious issues, including bison and wolf management, grazing on the park's Northern range, and

Tourists in car and bus at Jupiter Terrace.

Yellowstone Rangers at Mud Volcano. Skis and poles on ground behind them.

the illegal introduction of non-native fish into Yellowstone Lake.

Scholarly researchers, including college and graduate students and professional historians, make extensive use of the archives year-round. Histories of fire in the Yellowstone ecosystem; park patrol cabins; nearby Jackson Hole, Wyoming; and the Canyon Hotel, an extraordinary arts-and-crafts-style building designed by Robert Reamer (architect of the Old Faithful Inn), completed in 1911, but sold to wreckers in 1959 for reasons that remain controversial, are just a few of the research projects currently underway. **In recent years, hundreds of film makers, documentary producers, journalists, and other members of the media have used thousands of images from the archives in their films, television specials and articles focusing on Yellowstone, the NPS and the history of national park movement.** The history of the visitor experience and the history of resource management in Yellowstone have been dramatically illustrated through photographs from the park's archives.

Photos courtesy Yellowstone National Park Archives.

In addition to its many routine, practical uses, **the archives plays a role in raising funds in support of park goals.** Information and images from the archives relating to the park's historic partnerships with various companies is in the hands of the recently-formed, non-profit Yellowstone Park Foundation, in order to assist its staff in its fundraising on behalf of the park.

The MOA with NARA requires the NPS to provide trained professional staff to care for and manage the park's archives. Currently, a full-time, permanent Archivist and Supervisory Museum Curator are assisted by part-time librarians funded by the park's cooperating association; term and seasonal museum, archives and library technicians on special projects funding; and an assortment of volunteers, interns, and Student Conservation Association (SCA) Resource Assistants. The archives and library staff manages approximately 2,000 linear feet of records, as well as an oral history collection of over 200 tapes, and other magnetic and electronic media which are as yet unmeasured. The photograph archive contains nearly 100,000 images—some of which fall under the MOA and some of which were donations to the NPS—and is managed as part of the park's museum collection.

In entering into the MOA, the park also had to demonstrate to NARA its ability to care for its records. The basement of the Albright Visitor Center in Mammoth Hot Springs (park headquarters) was refurbished to provide a secure home for the park archives, museum collection, and research library. In recent years, however, both collections and researcher demand have outgrown this space. On busy days in the summer and early fall, it is not unusual to see researchers working on every available surface in the library, archives and curatorial workroom, and including staff desks and even, on occasion, the floor. The facility also became the subject of an Office of the Inspector General audit which found it lacking adequate environmental controls. **To address these problems, the park has begun planning for a new museum, archives, and library facility in Mammoth Hot Springs.** In addition to providing additional storage, the building or buildings will include exhibition space and visible storage that will allow more of the collection to be displayed to the public.

In the meantime, the park staff is seeking other ways of making collections more accessible to the public. Last year, during Yellowstone's 125th anniversary, reproductions of various photographs and documents were provided to museums mounting exhibits with Yellowstone components, including the National Gallery of Art and the Autry Museum of Western Heritage. Yellowstone's archives inventory was added to the park's official Web site last year, and plans are underway to add images to the site as well. Park staff also collaborated on a 125th anniversary photo album book. Published in 1997, *A Yellowstone Album* contains hundreds of rare historic views of the park. Proceeds from sales of the book go the Yellowstone Park Foundation and benefit park projects.

Susan Kraft is staff curator at Yellowstone National Park.

Richard Pearce-Moses

The Information Ecology of Archives

The "information explosion," about which so much has been said and written, is to a great extent an explosion of misinformation and badly organized information . . . . The digital revolution has only made the problem more acute.

—Murray Gell-Mann
"Information versus knowledge and understanding,"
From *Information Ecology* by Thomas H. Davenport

Archives are often perceived as dim, dusty depositories of arcane knowledge and minutia haunted by historians fascinated by the past. Like all stereotypes, there's a grain of truth in this belief. But, what's often overlooked by many managers is that archivists (and their colleagues in the library) have significant skills in analyzing and synthesizing knowledge.

In many ways, information technology (IT) departments are reinventing archives as they seek means to preserve and provide access to electronic information. Unfortunately, the staff in IT departments often know programming and computers better than principles of information management.

I believe that in the evolving high-tech information ecosystem, a savvy manager will look at the strengths of these two disciplines and forge a new alliance between them. After all, nothing requires that electronic records be kept in a separate electronic archive.

If an organization is going to effectively manage its information resources, it's going to have to look at the individuals who create, use, and preserve that information. Thomas Davenport argues in *Information Ecology*[1] that few organizations have a well-developed information management plan. Well, that's putting it nicely. He says, "For the most part, . . . information environments are appalling. [Organizations] don't know what they know or what they need to know" (p. 7).

Davenport argues that many of the individuals responsible for managing information resources tend to see information as data—as simple facts—not as knowledge or understanding that makes the data meaningful. Because the information managers come from a computer background, they tend to look to technology for solutions; but until machines possess judgment, they will only be able to respond with data, not knowledge.

I hope to give you an appreciation for what really goes on in an archive. Archivists don't just put papers in boxes and consign them to oblivion. They ensure the quality of information through a process of appraisal. They arrange that information so that it remains useful. They house the records to preserve the information from deterioration. They describe the records, analyzing and abstracting the information so it's easy to locate. And finally, they work with researchers who want to reference the records to ensure that they find all relevant information. The traditional principles of archives as embodied in these five steps can serve as a model for managing information resources and remain useful in the age of electronic records.

Appraisal and Acquisition

Given that we simply don't have the resources to save all the data that comes into an organization, what shall be saved and what shall be lost? With electronic records, this question is even more compelling, as we must actively work against rapid deterioration and obsolescence.

When determining what records to preserve, archivists have traditionally looked for continuing administrative, fiscal, legal, or historical value. But what is administrative and fiscal value? When making a business decision, a manager wants to ensure that his or her information is accurate, timely, accessible, and relevant (see Davenport, p. 117-120). An archivist uses these same values when determining which information to save.

Accuracy. Archives have always placed a premium on acquiring authoritative information. Did these records come from an office charged with collecting or creating this information, or did they come from some other source? Is the source trustworthy? Are the records complete?

Good managers—like good historians and good journalists—look to more than one source of information to confirm facts. Are your managers consulting your organization's past experience to verify information from other sources and to check the credibility of those sources?

Timeliness. A good archivist reads the company newsletters, talks to people throughout the organization, and knows what's current. The archivist will tell people outside the archives about potentially useful information before it's too late. Is your archivist in the information loop? Do you encourage your decision makers to consult with the archivist?

Accessibility. Archivists dislike acquiring collections with restrictions. They want to see that the information in their care is used. And, they take great pains to develop a variety of tools to make that information easy to find so that it can be used.

Do your managers have their own "archives" in their offices? In the best case, where these records are well managed, you might ask why they're spending time doing the archivist's work and hoarding information from others in the organization that might need it. In the worst case, the manager can't find the information in their own file cabinets, nor is it accessible to anyone else. Consider implementing a strong records management program, rewarding individuals who deposit their corporate knowledge in the archives where it can be managed and made available to all people in the company who need it.

Relevancy. Archivists are more concerned about the present and the future than they are the history of your organization. They know it's easier to try to collect historical information when it's fresh and available. So they seek to acquire information relevant to your corporate mission and current activities, knowing it will become part of history.

Finally, archivists talk about the primary and secondary value of records (not to be confused with primary and secondary sources). These distinctions have little to do with the usefulness of the information; the secondary value of records may, in some instances, be more useful than their primary value. The primary value distinguishes the information the records creator intended to capture in the course of business; secondary value refers to other information captured as a by-product of the records creation process.

Archivists ask the same kinds of questions about the value of information in records, regardless of the records' format. While people tend to associate archives with textual records, archival collections include photographs, maps, sound recordings, and three-dimensional objects. After all, "text" no longer means just paper records; word processing files on disk are textual records.

Determining the value of records is an inexact art. Some of the most interesting uses of archival materials comes from the innovative use of records. Information captured for one purpose re-used in a novel way can help us understand something in a new light or with renewed appreciation. As much as possible, archivists try to add value to their collections by suggesting ways in which their collections can be used.

Arrangement

Libraries generally keep their books in a single large collection, which is organized according to a system established by the library—typically LC or Dewey call numbers. On the other hand, archives typically follow the principles of provenance and original order; collections from different sources are kept apart, and the records are kept in the order used by the creator. Practically, respect for provenance and original order saves the archivist time reorganizing the materials into some artificial order developed by the archive and allows the repository to exploit any existing access system developed by the records' creator (which one assumes to be useful, especially for the purpose for which the records were created).

Provenance is the organization or individual responsible for creating the records as a whole. That doesn't mean the provenance is responsible for authoring every letter, memo, or report in the collection. The collection is created through the aggregation of documents from a variety of individuals and organizations as the by-product of routine activities.

Respecting provenance—keeping the records of different sources separate in the archives—is a useful tool for accessing the information in those records as provenance provides a good clue to the type of materials likely to be in the collection. When looking for a certain type of information, one thinks of an agency that would have generated or tracked that information in the course of business, then consults the records of that agency.

Moreover, the nature of the office may suggest the character of information to be found in those records. For instance, if someone were researching AIDS and had access to the archives of the Center for Disease Control, Jerry Falwell, Dell Computers, and President Reagan, what's the most likely source of information? The CDC jumps out as the "right" answer because it is a leader in studying infectious disease. But the papers of Ronald Reagan might be more useful in studying responses to the disease in federal policy, and the papers of Jerry Falwell might be more useful to gain a perspective (albeit it slanted) on socio-religious attitudes toward the disease. But, the likelihood of the corporate archives of Dell Computers containing much relevant information is pretty slim.

Original order is the organizational system used by the records' creator, and generally reflects the routine activities in which the records were used. Original order is not necessarily the order of the records as received at the archive; some records may have been misfiled in the office, and order may have been disturbed during transfer. The archivist will determine how the records were organized in the office of creation, then arrange the records according to that order.

The relationship among the documents in the collection is itself a significant bearer of meaning. If the records are filed chronologically, an undated document may be dated by looking at adjacent records. Gaps in original order may be important clues to missing materials.

As noted above, maintaining original order is a time saving strategy that exploits any inherent retrieval system. As such, an archivist may describe the manner in which records are organized before describing the records themselves. The archivist will ensure that any indices, finding aids, or other access tools received with the collection are readily accessible.

Note that sometimes the records have no order, often because the records were not organized in the office of origin. And, respect for original order does not extend to respect for original chaos. In these situations, the archivist may impose a simple organizational scheme on the records.

Housing and Preservation

Because the records in archives are of enduring value, the archivist must preserve the physical carrier of the information. Acidic paper used for blueprints and field notes becomes brittle with age and must be reformatted. Photographs of important events fade. Electronic media holding accounting information, oral histories, and retirement records suffer bit loss and media deterioration. **Temperature, humidity, air borne pollutants, insects, and—most dangerous of all—humans, constantly put archival collections are risk.**

Archivists take several measures to minimize the damage. Documents are transferred from their original file folders and boxes that might damage the documents into archival quality containers (you don't want a collection of documents on the floor when the box fails). In the process, staples and paper clips may be removed, papers unfolded, and fragile materials may be photocopied onto acid-neutral paper or encapsulated in Mylar. Materials that are very dirty may be cleaned. Electronic records may be copied onto new media to refresh the data and to circumvent obsolescent formats. Records are stored in environmentally controlled vaults away from vermin of all types.

However, the archivist is not a conservator restoring documents to like-new condition. Rather they are trying to stabilize the records against deterioration and protect them from damage or theft.

Description

Archival description creates a verbal abstraction of the collections, enabling researchers to browse the materials on paper rather than having to rummage through hundreds of boxes. Description summarizes a collection's organization and essential details so that researchers don't have to read every document to know what's in the records.

Archival description does more than make it easier for the researcher. Because researchers look at just those portions of collections they need rather than entire collections, less staff time is spent pulling and retrieving boxes. And, because fewer materials are in the reading room, there's less chance that the records will suffer wear and tear or be stolen.

Archival description documents the archivists' experiential knowledge of the records. It summarizes all the facts the archivist has learned about the collection in a fashion that will be meaningful to those who may need that information. **Description, at its best, is the process of transforming the raw data of the records into knowledge.**

Archivists generally use three tools to describe their collections: a repository guide, summarizing all the holdings; finding aids for each collections, detailing their contents; and an index that complements provenance as an aid in identifying relevant collections. These tools work together progressing from a bird's-eye view of the forest, to maps of specific groves, to pointers to specific trees.[2]

A repository guide gives a researcher the big picture through a summary listing of all the collections in the archive. Because an archive organizes its collections by provenance, the repository guide is a list of the sources of collections with a brief description of the records from each source. The guide may include a short note about the provenance and a brief description of the materials.

A repository guide is a coarse sieve that helps researchers locate collections likely to contain records relevant to their research. In some

Archivist and records creator discuss transferring records from the Park Central File to the Historic Document Department archives. Maintaining the provenance and original order of the records will ensure their continuing usefulness as evidence of past actions for the administrative staff and for future users. Photo by Campbell/Danford. Courtesy San Francisco Maritime National Historical Park.

instances, the archive may have only a few records from a source, and the repository guide is the only place where those records are described.

Once researchers have identified promising collections, they ask to see the finding aid for that collection. The finding aid describes the collection's contents in greater detail than the repository guide. It facilitates access by bundling similar bits of information into manageable units and embodying the archivist's experiential in notes about the context in which the records were made, their creator, their significance, the reason for their creation, their contents, and their organization. These introductory notes—which may run several pages—give researchers a more complete sense of the collection, confirming the relevance of the materials or suggesting that time might be better spent on other collections.

Finally, a finding aid lists all the series and folder headings. Reading this list is equivalent to opening the drawer of a file cabinet. When researchers decide which folders they want to see, they request them from the archivist.

In some instances, collections that contain very valuable materials may list every item. An item-level inventory is useful evidence for security control, and is generally done only for items that would have a high market value. Ironically, highly detailed description is often a barrier, rather than a benefit, to increased access, as the researcher must read through much more text rather than a good summary. Although once the norm for archives, especially literary manuscript collections, item-level description is so time consuming, it is now relatively rare for entire collections to be described at this level, although description of a few selected items of great value is not uncommon.

Repository guides and finding aids are not perfect access tools. They emphasize a collection's primary value, but most collections contain interesting but tangential information. The hierarchical nature of repository guides and finding aids works well when the researcher is familiar with the names and organizations related to their subject. However, provenance is less useful when valuable information was captured by unexpected offices and, especially in the cases of personal papers, when even the best researchers are not going to be familiar with every source.

To complement traditional name-based top-down archival research methods, archivists index the repository guide and finding aids so that researchers can find relevant information hidden in unexpected places. The index provides more direct access to the contents of the collections than the hierarchical model of provenance and original order.

Description is enormously important for electronic records. While a researcher can call for documents and browse through them to find relevant materials, electronic records are not eye-readable; researchers cannot easily "browse" floppy disks and tapes. Because it's often hard to find the right software and hardware to read older electronic records, effective description is essential to help archivists and researchers know if the information contained in these electronic vaults is worth cracking.

To a large extent electronic records can be described using the same approach as their paper equivalents. Often electronic records contain raw, empirical data. In these instances, the archivist will try to describe the data fields; the software that originally created these records will probably not be serviceable in five to ten years, and the kind of information that the original software provided is likely not the same information needed for subsequent analysis. Because the data structure was documented, it's possible to write a new program to re-evaluate the data.

Reference

Once arranged, housed, and described, a collection is ready for researchers to use. How do those researchers find the collections that might contain useful information? How do researchers know which repository will hold the records?

The archivist is—in many ways—the most important means of locating materials. Although archivists try to translate their experiential knowledge into access tools, it's impossible to describe all possible uses of materials. The conversation between an archivist and researcher often has a collaborative nature, blending the researcher's novel view of a topic with the archivist's knowledge of the collections.

As noted earlier in sections on arrangement and description, provenance is the principal mechanism for locating relevant materials. Even in repositories with well-indexed collections that allow researchers to look up specific subjects directly, provenance remains useful as an access tool. Looking for a subject by identifying the names of those individuals or organizations potentially responsible for collecting information on that subject forces researchers to think about their topic more broadly. When researchers look up a subject in an index and find no entries, they logically assume that the archive has no records on that subject. In fact the archive may hold records with relevant information, but that information is was not immediately recognizable during description. **Researchers are often so motivated by a time deadline that they are more focused on getting an answer than asking the right questions;** approaching the subject indirectly through prove-

nance helps discipline researchers to formulate their queries with care.

Of course, researchers have to know which archive to visit. Provenance remains the best first step in knowing which repository will hold relevant records. Probate records for someone who died in Austin will be at the Travis County Recorder's office. Records of the Phoenix Indian School are probably at the Bureau of Indian Affairs or, for older records, at the National Archives branch in Laguna Nigel. But, while the individual may have died in Austin, his estate may have gone through probate elsewhere, and an important body of Indian School records may have been disposed in accordance with the records schedule but wound up at a museum.

Finally, it's anyone's guess where an individual's personal papers might wind up. Aunt Hattie's photographs may be at the city library, the county historical society, the state archives, or a university special collection library.

Archivists have developed a number of tools to help researchers identify which repositories hold relevant collections. The National Union Catalog of Manuscript Collections (NUCMC) is similar to a repository guide, but includes collections from many different repositories. The guide has now been migrated to the World Wide Web <http://lcweb.loc.gov/coll.nucmc/nucmc.html>.

A little more than 10 years ago, archivists began using online databases to share information about their collections. OCLC and RLIN, two bibliographic utilities with union catalogs of books and other materials, began assembling online archival union catalogs. Unfortunately, access to those utilities was problematic; the interface was difficult to use and access was limited to specialized terminals usually found in libraries. Possibly a more significant barrier was the fact that searching these databases was not free; even now that the utilities have better Web interfaces and are widely available on the Internet, access fees continue to reduce their use. Finally, the collection-level descriptions of these records were often so general that searches for information in the series and folders was not included, limiting the retrieval value of these databases.

In the last five years, archivists have begun to use the Internet to publicize their collections. **The Web is radically changing how archives provide access to their holdings.** Repositories that were not members of OCLC or RLIN could make their finding aids widely available at little or no expense. Because the entire finding aid was available, information in series and folder headings was accessible. And, the Web has made it easy to provide access to the documents themselves through digital images. With the exception of the Heard Museum's homepage, the Webpages for archival collections are downloaded more than any page on the entire site. The museum has started receiving many more inquiries for use of the archival collections. Fortunately these queries have not significantly added to the reference workload; because the entire finding aids is available the queries generally refer to specific materials, so that the reference archivist doesn't have to consult the finding aid for the researcher.

Ultimately, many researchers find out about archival collections through word of mouth. Often it's through conversations with colleagues or through footnotes in articles. However, the archivist is often an excellent source of information about collections in other repositories.

The Archives of Tomorrow

Possibly the aphorism of the 21st century will be "Death, taxes, and technological development." The last 15 years have seen enormous changes in archives, largely driven by ready access to automation. When PCs made computing inexpensive, archivists immediately adopted them to produce finding aids and indices. Now that desktop machines have the power of 1970s mainframes, we'll be seeing more and more innovation in access as archivists develop more sophisticated mechanisms to manipulate data.

As archivists acquire more and more word processing documents and databases, it will be more common to provide access to the entire contents of the archival collection. As search engines become more sophisticated, locating relevant documents will be easier. And archivists that are specialists in retrieving data will be able to "push" information about relevant collections to researchers who have registered their interests with the archive.

But in the long run the core functions of the archivist, if not unchanged, will remain essential: acquiring and appraising information, housing and preserving it, describing it, and helping researchers reference it. **Throughout this process, transforming data into knowledge will also remain the principal strength of archives.**

Notes

1. New York and Oxford: Oxford University Press, 1997.
2. Readers may want to look at complete samples of the access tools developed by the author on the Heard Museum's Web site <www.Heard.org/library/rcguides/ >.

Richard Pearce-Moses is the Archivist for the Heard Museum in Phoenix, Arizona. This article is adapted from his speech at the Information Ecosystem Conference.

Susan Veccia

Information Skills and Primary Sources in Education

"These resources have made history come alive for my students and have allowed them to become 'historians' working with primary documents."

Mike Federspiel
1997 American Memory Fellow

The Library of Congress is on the move in secondary education. The Library's online American Memory collections are being eagerly discovered by school teachers and media specialists who have found a wealth of Americana primary source material previously unavailable to them. For a glimpse at these photographs, documents, maps, films and recordings and how they are being used in education, visit the Library of Congress Learning Page <http://learning.loc.gov/learn/>. Here you will find descriptions of the collections with correlations to school curricula, search guides, sample lesson plans, and student activities. By harnessing technology to an ambitious outreach program, the Library serves a new constituency—one that we feel is important to the vitality of this nation.

How It Started

An evaluation project during 1991-1993 systematically introduced the early American Memory collections to various audiences: colleges and universities, public libraries, special libraries, state libraries, and schools. Then distributed on CD-ROM and videodiscs, these collections were of great interest to educators using primary sources to supplement history and social studies curricula. This growing interest coupled with the rapid growth of Internet connectivity in the nation's schools presented an opportunity for the Library to use technology to make these materials more accessible. Like many cultural institutions, the Library of Congress began preparing its American Memory special collections for presentation on the World Wide Web <http://memory.loc.gov/>.

The success of the pilot program enabled the Library to launch a major educational outreach initiative. Initially funded with a three-year grant from the W.K. Kellogg Foundation, we began in 1995 to define what role the Library should play in education via its Web site and the American Memory collections. Understanding that education is a national issue but a local affair, we identified 25 leading teachers and media specialists from across the country to help shape our outreach effort. **It became clear that while some teachers were using a variety of primary sources in their classrooms, because these resources had been difficult to find and acquire, most teachers had limited experience integrating them into the curriculum in a sustaining way.** Our educational advisors observed that it was not enough just to know about the Library of Congress American Memory collections on the Web; teachers needed hands-on training using the materials. Teachers also needed contextual material to help them understand the structure and provenance of the collections and major themes represented.

The American Memory Fellows Program

American Memory Fellows Program and the Learning Page became the focus of our work during 1996 and 1997. A week-long summer institute formed the centerpiece for this program, which brought 25 two-person teams to the Library of Congress for the purpose of developing sample American Memory lesson plans. Teams were selected based on experience using primary sources to motivate students, access to and comfort using technology in teaching, and professional leadership skills. Twenty states were represented among these teams. Facilitated by the EDC Center for Children and Technology, the institute provided the Fellows with hands-on experience searching the collections and developing sample activities that drew upon primary source materials from the Library's American Memory collections. Throughout the 1997-1998 school year, the Fellows tested their lesson plans in the classroom and participated in an online discussion group.

The institute was designed as a collaborative laboratory. We designed a four-week pre-institute online conferencing environment plus a six-month post-institute listserv. These components proved to be an invaluable part of the design. Participants got to know each other online, shared project ideas, and explored the Library of Congress Web site prior to arriving in Washington. Likewise, the listserv served as an easy way to facilitate peer review of the developing lesson plans after the institute.

In addition to hands-on training, participants spent a day at the Library of Congress meeting with curators and discussing technical challenges associated with building a digital archive. This experience sparked ideas about teaching children how to build an archive, how to recognize the part from the whole, the point of view of an item or collection, and the physical presentation of an archival collection.

Energy, enthusiasm, and imagination flowed as American Memory Fellows brainstormed teaching ideas, located resources, and

began to work on their lesson plans. Some chose to follow "the essential question" model, establishing a broad research question to focus student exploration of the online collections. Others emphasized visual literacy techniques to sharpen student observation skills and help students identify what they need to know to "read" a picture. Many engaged their students in oral history exercises—interviewing members of their own communities about a specific topic or time in history. Questions of "then and now" surfaced as educators wrestled with the best way to engage students in the primary source materials. As Fellows became more familiar with the collections, the wealth of material was almost overwhelming. "The only challenge connected with the collection is its sheer size and scope...that having [been] said...what a wonderful challenge it is!" noted one of the Fellows.

The Learning Page

Recognizing this challenge and the need to provide contextual information about the collections, the Library in 1996 launched the Learning Page—an online gateway to the American Memory collections designed for teachers and students. This is a good starting point for educators who wish to "Learn More About It." From these pages, the archival collections are described from an educational perspective with links to topics covered in the school curriculum. Items from the collection are highlighted along with suggested search words enabling teachers to dip into the collection with successful search techniques. Topical pathfinders enable educators to get a bird's-eye view of subjects represented in the collections. Sample lesson plans, curriculum ideas, feature presentations, and technical information round out the Learning Page into a guide to effective use of the American Memory site.

Ultimately the goal of any teacher is to engage students in "the hunt" as well as in the substance of the primary source materials. The "Historical Detective" and "The Big Picture" are two student activities designed for these purposes. In the first, an item from the collection is presented with "clues" about its identity. Students are encouraged to think like detectives in a historical framework and find that item within a collection on the Web site. The second activity is a jigsaw puzzle that changes weekly. Again, an item from one of the collections is presented in puzzle "pieces." The goal is to reassemble the puzzle within a specific time frame. Puzzles over the course of each month provide clues to a "Big Picture," which highlights a monthly theme using materials from the American Memory collections. As time goes on, more descriptive materials for teachers and activities for students will be added to the Learning Page.

What's Ahead

Our experience of the last three years points to the continuing need for teacher training in the use of primary source materials. Although this is a new area for many educators, primary sources are finding their way into established curriculum guidelines. In North Carolina, for example, for the first time the use of primary source materials is incorporated into the information skills curriculum. We see professional development as a collaborative effort between the Library of Congress and educational practitioners. This collaboration is paying off as our 1997 American Memory Fellows fan out into their own communities, speak at professional conferences, and train others in their schools. Their enthusiasm for the American Memory collections is contagious.

With help from the Ford Foundation and other generous donors, the Library of Congress is moving ahead with an ambitious plan to reach teachers in all 50 states by the year 2000 through a combination of online and in-person professional development workshops. We will be hosting an American Memory Fellows institute this summer and again in 1999. And, we will expand the content of the Learning Page to include teacher-developed sample lesson plans and professional development resources to facilitate informal "train the trainer" workshops. Our most ardent supporters are the teachers and media specialists who have helped us shape the program. We look forward to helping this new constituency blossom in the years ahead.

Two young women posed with a loaded car at Yellowstone. Photo courtesy , Yellowstone National Park Archives.

Susan Veccia is an Educational Specialist at the National Digital Library Program of the Library of Congress.

Amalin Ferguson

NPS Libraries Come of Age

Building the NPS "Virtual Library"

There are nearly 400 libraries in the national park system, most of them managed by park staff without library training on a collateral duty basis. There are currently no central funding sources for backlog cataloging or preservation treatment and local budgets for library operations are minimal. The number of book titles range from 150 to 60,000 (Morristown National Historical Park), with the average size being 3,000 volumes. When non-book items are taken into account, the National Park Service (NPS) library holdings are estimated to total 5.7 million pieces; however, the majority of these are not cataloged, at least not in a standardized, electronic format. These historically persistent conditions severely impact discovery, retrieval, and sharing of a significant component of the research and educational materials constituting the NPS information base, not only within individual parks, but across organizational lines.

The NPS library collections contain a wide variety of media and formats: books (including rare and out-of-print); reports and conference proceedings (many unpublished); theses and dissertations; whole journals and journal article reprints; newspaper clippings and other ephemera (many from the 19th and early-20th centuries); oral histories; maps, charts, and plans; microfilm and audio visual materials; CDs and computer files; photocopied reproductions of historic manuscript materials; note files; species lists, etc. The comprehensive and highly focused scopes of the NPS library collections may be considered their primary value. The following selected highlights are intended to give only a very partial view of the range and depth of materials to be found in NPS libraries.

Screen capture images used with permission of Endeavor Information Systems, Inc.

C&O Canal National Historical Park. Associate Supreme Court Justice William O. Douglas oral history tapes and related newspaper clippings.

Frederick Douglass National Historic Site. Valuable collection of books and other materials highlighting the life and career of the most notable African American of the 19th century.

Fredericksburg and Spotsylvania National Military Park. Regiment lists and xeroxes of correspondence and other historic documents relating to the Civil War.

Harpers Ferry Center. An enormous photo collection covering many parks in the System and specialized book collections in material culture, NPS history, technical conservation, design, etc.

Independence National Historical Park. Notecard file of historical data relating to people, events, and sites from the park's historic period.

Lowell National Historical Park. Large collection of textile industry trade catalogs, technical manuals, and other reference works from the 19th and early-20th centuries.

Saint-Gaudens National Historic Site. Focused collection of books on 19th century and early-20th century art.

San Francisco Maritime National Historical Park. Over 2,000 subject files containing Congressional reports, newspaper clippings, journal articles, obituaries, and staff research notes on maritime history.

Yellowstone National Park. Scrapbooks (1895-1940) containing periodical excerpts covering such topics as President Warren Harding's 1923 trip to Yellowstone and President Calvin Coolidge's trip in 1927.

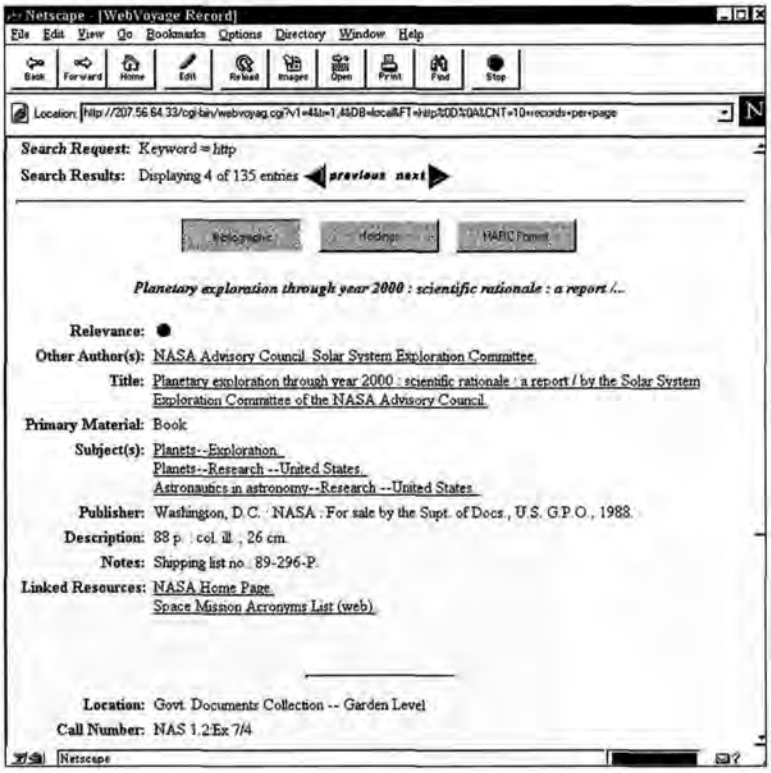
The National Park Service Union Catalog

To promote and enhance discovery, retrieval, and sharing of NPS intellectual resources (at least, the metadata describing it), the NPS Library Program will merge NPS bibliographic records that are in electronic format to a centralized database

which is fully compliant with professional library and information standards and is deliverable via the Web. Initially, the NPS Union Catalog will be accessible via the NPS Intranet only, and will contain approximately 130,000 records from individual and regional NPS library catalogs and subsets of the NPS natural resources and cultural resources bibliographies. To advance the automated cataloging efforts in the field and further develop the NPS Union Catalog, the NPS Information & Telecommunications Center (ITC) purchased 300 copies of ProCite bibliographic software for distribution to parks. ProCite users receive the program and software upgrades, along with user aids and technical support.

Using an adjunct program called BookWherePro, parks can build library catalogs cheaply, efficiently, and professionally by obtaining free records matching their holdings from library catalogs on the Internet ("copy cataloging"). The records may be downloaded into ProCite import format. BookWherePro also enables downloading of subject and name authority records from Library of Congress databases for "original cataloging" purposes (when cataloging copy cannot be obtained). Golden Spike National Historical Park recently cataloged 960 of their library of 1,000 titles in 50 staff hours using this model. Eventually, the NPS Union Catalog will be accessible to all Internet users and may include references or links to the following NPS information base components:

- all of the library holdings in NPS parks, support offices, central offices, and service centers, including selected format subsets, e.g., trade catalogs, and oral histories;
- selected holdings in NPS museum collections, e.g., archival collections, archives finding aids, rare books, and books from the personal libraries of historical figures;
- the holdings of the NPS Denver Service Center's Technical Information Center, containing in-house plans, drawings, and planning documents;
- NPS servicewide, program area bibliographies, e.g., Natural Resources Bibliography, Cultural Resources Bibliography; Trails Bibliography, and others (built using ProCite);
- "ThemeCats" merging the references to library holdings in parks sharing similar themes/resources, e.g., African-American history, Native American history, Civil War, women's history, etc.
- books sold by the NPS book stores (i.e., cooperating associations);
- all NPS and NPS-related publications that are available for distribution upon request, including those sold through partners and friends' groups (e.g., curriculum packets and unigrid park brochures).



Endeavor's Voyager OPAC

The software that will be used to manage and search the NPS Union Catalog is Voyager, by Endeavor Information Systems, Inc. Primary access to the OPAC (online public access catalog) will be through the Web. Scoped to research and special libraries, Voyager is a state-of-the-art library software program developed by seasoned library and information industry professionals. **The software design exploits to full advantage currently available technologies and standards,** including Web capabilities, the Z39.50 (v.3) information storage and retrieval protocol, multi-tiered client/server architecture, distributed networking, platform-independent design, and open database connectivity. Some of its features are listed here:

- default (customizeable) buttons on Voyager's main toolbar include links to the local catalog (i.e., the NPS Union Catalog), library information, other library catalogs, non-library databases and other electronic resources, and help;
- in addition to indexed field searching (author, title, subject, call number), supports searching for keywords located anywhere in the records with user-assigned weighting of search terms and returned items ranked for relevance;
- can perform command line searching or will build searches with pulldown menus, using boolean operators in either case;
- supports searches limited by date, type, medium, language, etc., as defined by system administrator;

- leads users to broader, narrower, and related terms when performing subject searches via cross-reference structures;
- configures author, title, and subject headings as hypertext links;
- can link to described or related electronic resources from within the catalog record;
- supports searching and management of course reserve materials (a function which could be commandeered to deliver NPS digital documents);
- can perform simultaneous searching of multiple library catalogs on the Internet, along with the NPS Union Catalog, and merge search results into one set;
- searches the NPS Union Catalog for users of other library catalogs with the parallel search capability via their native client interface;
- Voyager may be integrated with Endeavor's ImageServer, an electronic document management program with full multi-media capability and OCR optional module supporting automated indexing.

The NPS Virtual Library

The NPS Virtual Library Web site (working title) is envisioned as a first stop clearinghouse for discovering what information resources (excluding administrative data) are managed by the agency, regardless of format, scope, location, and subordinate body responsible for it. The Web site will provide direct links to NPS electronic resources served on the Internet or the NPS Intranet, but will not override any security measures implemented by the "owners" of the respective resources. The URL for the NPS Virtual Library Web site will be submitted to Internet search engine databases and other relevant Internet "hot lists" to enhance discovery of NPS information resources. A second goal of the NPS Virtual Library Web site is to maximize staff research time and achieve economies of scale by offering a centralized point of access to relevant electronic resources and information services (free and fee-based) external to the NPS. The implementation of the NPS Virtual Library will occur over a two-year phase-in period, beginning in summer of 1998, with limited access beginning in the fall of 1998. In addition to the Union Catalog, the NPS Virtual Library Web site may provide links to the following products and utilities, many of them through the Voyager front-end main toolbar:

- a GILS database (Government Information Locator Service), containing descriptions and selected hot links to NPS electronic and non-electronic resources: Web sites; publications; databases, e.g., GIS, natural and cultural resource data sets, directories of libraries, ProCite users, subject experts;
- a gateway to the Department of the Interior GILS and other GILS servers;
- thematic "pathfinders" linking a variety of NPS information formats managed by various program areas, e.g., African-American history, Civil War;
- centralized inter-library loan (for internal circulation only—professionally staffed NPS libraries may elect to lend their materials outside the organization);
- "Ask an NPS Librarian" reference service;
- "NPS Library Resource of the Month" feature;
- library management and ProCite guidelines, training materials, and other user aids for NPS staff;
- access to CARL Uncover, a public access database of current tables of contents for over 17,000 journals, searchable by subject, with some capability for fee-based document delivery;
- site-license access to DIALOG, a commercial service that provides abstracts and some full-text of thousands of journal articles, dissertations, conference papers, reports, and monographs, which are organized into hundreds of topical databases (e.g., GEOREF, BIOREF, American History & Life), now searchable via a Web interface;
- link to the IPL (Internet Public Library) which offers free reference service and access to hundreds of general and specialized reference sources on the Web (directories, atlases, encyclopedias).

The NPS Library Program

The NPS Library Program is funded by, and operates under, the auspices of the NPS Information & Telecommications Center (ITC) in Washington, D.C. The NPS Library Program Coordinator position is supported by ITC, the Pacific-Great Basin System Support Office, and the San Francisco Maritime National Historical Park, where the position is duty-stationed. Working with the 10-member steering committee (professional NPS librarians) of the NPS Library Advisory Council, the Program Coordinator plans, develops, and implements program products and services.

Amalin Ferguson is the Library Program Coordinator, NPS.

For more information about the NPS Library Program, or to request library assistance, contact Program Coordinator, Amalin Ferguson, via any of the following methods: voice at 415-556-0238, fax at 415-556-3540, cc:Mail under her name, Internet email at <amalin_ferguson @nps.gov>, or post to #BB Inbox: Library Link via cc:Mail.

John P. Byrne

The National Register Information System

Preservation Strategies in the Data Center

This tale is the saga of trying to preserve technological investments at the National Register of Historic Places in a sea of change while integrating some technological changes that have come in on the trade winds.

To set the stage, the National Park Service, in comparison to other bureaus within the Department of the Interior, or other federal agencies, is a modest user of computers. I work for the National Register of Historic Places . . . basically we keep a list of historic places. Our story goes back to the late 1960s, but we shall pick it up in the mid-'80s when a new effort was begun to implement an information system in what was then a state-of-the-art Hewlett Packard minicomputer (HP 3000) environment running a "network" database. The tale is told from the perspective of an ordinary user of technology, trying to use what is available to him, not someone with lots of exotic needs or access to lots of specialized expertise—perhaps not unlike yourselves.

In nautical terms we have been obliged to "tack" as we move forward. While not completely replacing our original database, we integrated relational databases, supplemented our third generation language code with fourth generation languages, replaced the hardware with a box two times faster at a fraction of the cost, added local area network (LAN) access, and got thousandfold increases in speed by using specialized indexes for keyword searches. We added a Windows interface for ad hoc searching and adopted Internet protocols for public access (Telnet, FTP, Web). **Throughout it all there is usually the need to run the old along with the new**—and remember, if this is not your requirement now, it will be later.

Preserving Your Investment

One of the major questions is how do you preserve your information system investment. Since the 1980s our needs at the National Register have grown, but they have not fundamentally changed. Technology has progressed, however, and finally our desires are within the bounds of what we can reasonably do. As compared to airplanes, which we expect to last for more than 30 years if well maintained, most people are lucky to get a couple of good years out of hardware or software before some kind of upgrade or replacement is necessary. The National Register has been fortunate in this respect because the vendor has been able to keep the technology up-to-date and we have been able to afford some upgrades. After 25 years of incremental improvement by HP our legacy platform is still going strong, and by comparison to some other platforms, takes little effort to manage. As a consequence, **we have been able to swallow change in sips instead of Super Big Gulps and we have even had some time left over to pursue some newer technologies on other platforms.**

One of our earliest changes was going from dumb terminals to personal computers (PCs) on the desktop. The Windows/Intel (Wintel) revolution both simplified and complicated computing. Given the tendency, relative to other computer architectures, for Wintel PCs to require more support and to be less reliable, early on we chose to integrate PCs rather than use them to replace our host database architecture. While we enthusiastically implemented Windows-based terminal emulation and reporting tools, we skipped a wholesale migration to a client/server architecture for the simple reason that it would slow down our data entry and unnecessarily complicate our processing. Lately, PCs have become more reliable and manageable than ever. Management solutions now include the Network Computer, the NetPC, the Windows terminal, and Managed PCs. In combination with the Web these technologies have a lot more appeal for broadening access than expensive client/server architectures.

Think creatively about how you can continue to get your money's worth out of your investments by evolution rather than revolution. You have more choice when you pick and choose. One disarmingly simple strategy is to move PC software, like terminal emulators, to the Web. Or wrap the output in the garb of the Web—recently the Montana Department of Public Health and Human Services put almost 400 screens worth

of mainframe forms data on the Web using a "screen scraping tool" for the job.[1] Word is that while there are many reasons to move to new technologies like Windows NT, cutting costs is not necessarily one of them.[2] Just as the historic preservation movement argued in the 1960s that there had to be a better way to improve the housing stock than to indiscriminately level neighborhoods, so too there has to be a better way to improve the information stock than to blindly eradicate software platforms.

Vendor Selection

Understand the offerings from the vendor you have before you buy into the devil you have not met. This includes reading the literature—your vendor may be changing in ways you do not know. In particular, I recommend unbiased reports like Datapro, compendiums of the trade press like Computer Select and even the white papers on the Web pages of vendors—after all, if nothing else the vendor at least knows how his product actually works. You may decide you want to extend what you have, at least for the time being. You may not. If you are a decision-maker be open to understanding the big picture beyond your desktop; if you are the one who evaluates technology simply keeping your boss up to speed will help.

Given that 80% of the data processing dollar is spent on maintenance rather than acquisition, give careful consideration to support issues—you may be entering into a decade long relationship with vendors, and you want ones you can trust. Is support a part of their culture, or do they make their money on volume? Software support is changing. Whereas once we typically bought unlimited support for mission-critical systems for a flat rate, now support is often priced on a per incident basis and sometimes not even available from the software vendor. Is this appropriate for your organization? In smaller shops it may not be. For example, we buy combined hardware/software support available from HP on a 24x7 basis for our mission-critical system because this is the most cost efficient way to get guaranteed access to immediate high level expertise. In another case, we spent a ton on connectivity software to allow us to connect just about any computer to any other computer, but the vendor, WRQ, throws in award-winning telephone support for the life of the product. Finally, is the vendor generally responsive—in other words has the user community had a meaningful impact on product development? When was the last time you got to vote on enhancement requests? Did it make a difference? A single vote may count for more here than in any other part of our democracy.

Be clued in to the important junctures in a product's history so you know when to hold them and when to fold them. Drawing from my own environment, I can say that within the last year the HP 3000 has undergone a small renaissance. While noted for reliability, a 25-year record of backward compatibility, and for embracing open standards, whether it would prosper into the next century was an open question. Recent decisions to port the Java virtual machine and a commitment to go to 64-bit computing breathed new life into the box. While it is in no danger of derailing the NT juggernaut, these were signs to our community that the existing investments were being preserved. In many older, more mature computer environments, the success of a platform independent, Internet-oriented language like Java will be the test of whether diversity will, in the long run, survive. When the vendor puts the appropriate technologies in place the organization can once again be in the driver's seat—pursuing change on its own terms.

The reality for most of us is that in our organizations we have mixed computer environments. You may have software that will not be rewritten any time soon or you may have merged with other parts of your organization which use different software. **Key to making it all work together is sufficient adherence to computer standards, both de facto and de jure, to allow pieces to inter-operate** rather than flocking to this or that package recommended in the latest computer magazine. As in the architecture of cities, with proper design the old and the new can co-exist gracefully as good neighbors. Few organizations have the manpower to implement continual technology changes, and even if they do, it is wasteful because there is never enough time to amortize the investment. Of course, you should still move forward with cutting edge, even bleeding edge, technology but you may want to try it out first in non-mission-critical environments and **migrate your bread and butter systems only after you have gotten your money's worth out of them.** Trust your intuition. If complicated technology like client/server never seemed like a good idea to you maybe it was because, for many purposes, it wasn't a good idea. On the question of complete reliance on a single software vendor I recommend you acquaint yourself with the history of antitrust legislation in this country before you come down for or against a company like Microsoft.

Web/Database Solutions

From a database point of view, the lure of the Web is that it provides, in the browser, a single interface to which all users have access. Right off the bat that solves the problem of the existence of client software. Yet for this great leap forward, a price has to be paid. Two steps forward and one step back. **The Web was not originally designed to do databases well**, and a host of issues present

themselves. Many of these issues are being addressed by standards, and being solved by vendors, but it is important to understand why sophisticated database processing has been slow to come to the Web.

As originally designed, the World Wide Web is stateless and connectionless which means that when a link on a Web page is activated, the browser makes a connection with a Web server, a document is sent and received, then the connection is closed. Permanent connections were never in the game plan. **This scheme poses problems for sophisticated database processing which is heavily dependent on "state" and a persistent connection.** It is axiomatic that when a user connects to a database for a transaction the host database knows, no matter how many screens that user goes through, who that user is and what that user is doing. Database transactions are meant to be atomic—all or none events—you should not be allowed to open a financial application, get halfway through some money matters and wander off to a baseball Web site leaving data in an inconsistent state. You have to carry the ones and zeros, so to speak. Somehow, a pseudo-state and a persistent connection have to be created so each user is bound to a specific running process. Being stateless is a great benefit in serving up static html pages because it is efficient—you can handle lots of hits because not much is asked of each connection. Go to a page, click, get a new connection. But it is an unacceptable scenario for conducting typical order entry or financial applications.

Early Web database designers tried to finesse the problem by focusing on database applications that did not require state—for example, by having the user enter everything on one Web page. **While state can be resolved in a variety of ways, every serious database application has to have a way to maintain it as well as a place to store information entered from preceding pages while a user moves forward, a means for maintaining the security profile throughout the application, and a scheme for guaranteeing persistent database access without having users logging in and out of the database all the time.**

In offering simple Web/database access to the public we chose to invest in a Windows NT host. Seemed like a clean way to get our feet wet in new technologies. While keeping our existing system, periodically we transfer our database to SQL server, a Microsoft database, and offer a means for users to access dynamically created Web pages from database tables. The strategy employed by the Web database software, Speedware Autobahn, is to have the host software impose state by assigning a unique data session identifier at the outset, to which it can refer later, and then incorporating the Web into an already highly developed fourth generation programming language. As powerful and ingenious as this is, as Web database processing evolves, ultimately this will not be the preferred strategy. The Web merits its own development environment.

In addition there is a basic problem in how the Web interacts with databases. The protocol of the Web includes a way to talk to databases called Common Gateway Interface (CGI). CGI is not an efficient protocol, because it wants to open and close a process for each database access request. Because of the need to spawn or "fork" an instance every time a user needs to issue a database call a database application can be overwhelmed by CGI calls. And when data is brought back from the database server to the Web server, and then fed into the browser, there can be a significant wait—orders of magnitude longer than legacy ways where all processing is host-based. To get around these problems, Web server vendors like Microsoft and Netscape have developed their own application

Sylvan Pass Lodge at Yellowstone. Photo courtesy Yellowstone National Park Archives.

programming interfaces (API) which speed things up, but cause other problems like making the developer write different versions for different Web servers. Kinda defeats the purpose. Some of these problems could be resolved by the Java language but Microsoft and Sun are squabbling over Java. Java is a promising computer language which is still only a couple of years old and thus still lacking in a lot of the features that languages which have been around for a couple of decades have.

A classic Web problem is that when data is entered into a Web database the information goes in all at once and then comes back—there is no edit checking on a field by field basis.

In the early days of mainframes and minicomputers when bandwidth and CPU were expensive this was known as block mode processing. The fields were processed as a block and thus when you pressed the enter key you got the attention of the host and went from there. With current Web technology sometimes you do not even have a good way to have the host database communicate intelligible error messages back to the client. For many of us who remember, this side to Web database development seems like a step back in time.

It is not a question of whether the major database vendors adapt their software to the Web; it is only a question of how they are doing it. The advantages are too great to ignore. Unlike mature relational databases, which are often surprisingly alike on the back-end these days, these are the pioneer days of Web database technology, and we expect lots of variation in implementation. Already, there are hundreds of Web database tools out there. Buyer beware, it is unreasonable to expect the variety of methods or the multiplicity of vendors to survive a shake out.

Query and Reporting Tools in the Data Mart

Often it is management's vision, especially in the public sector, that once information has been automated it should be relatively easy to serve it up to the public such that the inexperienced user can easily find the answers to questions of their own design. Rarely is this the case—almost all interfaces are "canned." **Rarer still (at least I have never seen it) is the case where a user can jump on the Internet, point to a database of interest, write a reasonably sophisticated query, and within an acceptable amount of time get an accurate answer.**

Why is this the case—wasn't this the whole point all along? Traditionally the impediments have been:

- the prospect of introducing this capability strikes terror into the hearts of information professionals because, at the very least, this capability compromises the consistent response time which they have worked hard at providing to users of their all important production databases. In the worst case scenario an innocent user can issue the "query from Hell" which completely locks up a system. I should know. I accidentally learned how to do this in my environment in the mid-eighties with just three words;
- all data structures and all software have a learning curve which you cannot reasonably expect the casual user to have mastered;
- obtaining answers to complex questions typically takes too long, "costs" too much in terms of the query optimizer, and often produces an inaccurate answer because the wrong question was asked in the first place.

To the extent this capability has been provided at all it has usually been done through dedicated client/server solutions which require expensive software for each client. I am happy to report this is changing due to the migration of sophisticated query and reporting tools to the Web and the growing acceptance of the value of data marts and data warehouses.

While most of us are familiar with query and reporting tools, you may not be familiar with data marts and data warehouses which store the data to be analyzed and without which, the tools do not work well. A trend begun in earnest in the 1990s, data warehouses are today the subject of more than thirty books. While there are many ways to define and implement a data warehouse, for our purpose we can view it as a separate, read only, integrated database optimized to answer questions. It differs from an operational database in that it is subject-oriented rather than geared to accommodate business processes, tends to have more historical data than is needed in an operational database, is organized to pump data out rather than get data in, and is reasonably current but not necessarily up-to-the-minute. A data warehouse can get large, in the terabyte range, take several years to implement, cost millions of dollars, and draw information from all over the enterprise. By comparison, data marts are smaller, in the gigabyte range, quicker to set up, cost tens of thousands of dollars, and typically aggregate information from a single department. In either case you can buy packaged solutions or roll your own.

I will use my own experience in building a poor man's data mart to illustrate the principles. **The first thing you have to do is to determine what kinds of information your users want.** In my case, I found that the summary, drill down data so often talked about in data warehousing is not as important to my users as easy, powerful, and flexible access to the full range of data. This influenced the design of my data mart—no need for a massive

redesign. Next was deciding on a database platform. While I did not have an extra server, I did have a host database package, HP ALLBASE, more suitable for data warehousing and for manipulation by Windows programs than the operational database—which is best at efficiently processing high volumes of incoming data. **The next step was to write software that grabs the necessary data from the various databases** in which it is stored, make all necessary transformations, and then to write batch load routines. Fortunately, I did not have data stored in too many different places and I already had, as a part of my development environment, a module specifically designed for extracting and reformatting data, and populating databases. Next, I had to provide access from a Windows client. Since I only had one license for Cognos Impromptu, a leading query and reporting tool commonly used in data warehouse applications, I elected to install the software on a dedicated PC for use within our Division. After working out the kinks in accessing a remote database from a Windows client I then performed a one-time download of the database description to the client. Then I modified the catalog slightly to make sure every file was joined exactly as needed to every other file since users would have trouble doing this. Next, I set the governor, and wrote some sample queries in a syntax-free report writer which others could modify in a point and click environment for their own purposes. To my amazement it worked and people actually use it.

Key to successful implementation is being able to give users access to a simplified data structure and for having a means to prevent the runaway query. While you should not delude yourself that everyone will then be able to stand on their own from here on, this was the first time in our environment a non-programmer could accurately and efficiently write their own queries. They do not get into trouble because they cannot. The user is not afforded a means to establish improper relationships or ask unreasonable questions.

The obvious flaw to this approach is that there is no Web component to this architecture—instead there is a costly piece of software which has to be licensed by each user. The vendor has recently incorporated Web database technology and is now selling a Web-enabled version of this software which we are evaluating. Preserving an investment can be as problematic for the vendor as it is for us because there was little carryover from the Windows-based code to the Web code. Cognos had to buy another company's technology to get it quickly.

Object and Object-Relational Databases

In the eighties sometimes you would see bumper stickers urging you to "get relational with your database," a hint that relational databases were the only friends worth making. As with most things, the reality was more complex as many decision-makers never took the time to really understand what a relational database is, the extent to which certain products were or were not "truly relational," nor even what relational databases did and did not do well.

As it turns out, relational databases are good at most of the things we need databases for. Modern relational databases can get quite large, service many users, and are good at keeping junk out of the database. They understand a common grammar for searching and can be accessed through standards compliant software. But they are also reductionist. Everything tends to get crammed into rows and columns—there was never much consideration paid to the need to store and manipulate what has been called the "complex" or "rich" or "unstructured" data that comprise so large a part of what we—especially in the humanities—might have. Unstructured data might be a photograph, a map, or a long piece of writing in a word processing format.

Being the inventive creatures that we are, and perhaps more importantly, databases being the multi-billion dollar market that it is, the relational database was modified to accommodate such non-standard data types. **Just as the network database was extended to become more relational so the relational database was extended to become more "object-oriented."** The first development was to add Binary Large ObjectS, or BLOBS as a remedy. A BLOB is an undifferentiated database type that can hold almost anything and can be used in a database in a couple of ways. The BLOB can contain a pointer to a file external to the database or it can contain the data itself. It was seen as a way to link, for example, maps or image processing systems to databases. Despite the advance, an immediate problem presented itself. The BLOB is a black box about which almost nothing is known by the database so it is difficult to search and even more problematic to search in combination with structured data. Moreover, the database itself does not know how to manipulate the contents—that has to be done with middleware or application programs on the client, by most reckonings an inelegant solution. Standard functions, like summing the contents, all of a sudden become irrelevant when dealing with photographs. And, truth be told, it takes lots of grinding to make it all work together—performance suffers because data has been flattened into 2D tables and reassembly eats CPU cycles. In the meantime, in a far, far cruder way, Web pages already appeared to be doing some of these things; users were getting restless and wanted to fill the Web with so much Web content

that the only sensible solution was a database approach.

Many vendors decided to extend the row and column format to embrace a lot more data types out of the box—such as extended text, video, audio, image, geo-spatial, time series, and even fingerprint data. They could price each add-on separately— and even add in the ability to let ambitious customers roll their own data types. **Add to the data type the ability for each type to know what it is and how to present itself to the user and you have solved a major problem—the recipient of the object does not need special software.** You ask the object to do something; you don't do something to an object. The object contains the code, you just pull the plug on the bottle and the genie appears.

The advantage of these products is that you can build on what you already have and you can plug in other data types and then use a single query in a familiar grammar to get an answer. That way you are preserving all the R&D that went into developing the relational database (the data type participates in backup and recovery, security, integrity, concurrency and query optimization schemes) and the data type belongs to the existing relational database, even though it may be indexed differently.

Using another major database vendor, Oracle, consider how spatial data can be integrated with relational databases. Like image processing systems that have often existed as separate systems, traditionally spatial data has been locked away in geographical information systems (GIS) and not integrated with business processes or the daily work of organizations. For actually doing the work of GIS that is fine, it is the specialized tool designed for the job, but in terms of the big picture it is limiting in that it assumes all questions are fundamentally geographic. **GIS also exacerbates the problem of storing data in more than one place and thus interferes with the ability to have a single way to ask a question.** An answer provided by Oracle is to offer ways to provide a subset of GIS capabilities in "geo-enabled" databases which allow common spatial queries such as what points fall within a rectangle drawn on a map or where a pipeline might cross a river. Geographic data can live as one layer along with many other kinds of information that can be queried in conjunction with traditional data elements.

For many years there have been "pure" object-oriented databases which are not extensions to the relational database but rather creatures unto themselves. There are some conceptual advantages to having a database designed from the ground up to support objects but only recently has there been a mainstream vendor offering for such databases. Now, with the introduction of Jasmine, by Computer Associates, a four billion dollar software company, such offerings are certainly worth considering, especially for those folks who are not heavily vested in the relational database model.

For anything extending beyond the traditional relational database, standards are an issue though not to the extent they are in pure object-oriented databases. Ironically, despite the power of these new technologies, sales of object-relational and object databases are still quite modest—perhaps we are all too busy digesting changes from the last time around.

Conclusion

In these heady times of technological progress when normal time has supposedly been compressed into Internet time, and parents feel they are being left in the dust by their kids, parallels with an earlier period in history come to mind. **We are building databases with the same energy with which the medieval cathedrals were erected.** While modern databases may not be monuments to God, they are intended to be the means toward our secular salvation: "If only all of this were automated." As happened to many of the 12th-century European cathedrals built before the ribbed vault, **most of our databases will fall of their own weight over time**, though the raw materials need not go to waste. Good information, like sturdy stone, can be re-used. And the faith will no doubt live on.

Like the philologist, we are convinced of the need to master special tongues in order to make progress. The debates over Greek and Latin translations are echoed in the debates between Sun's Java and Microsoft's Active X. Desiderius Erasmus went back to the Greek sources for, to him, "Greek is the stream truly running with gold." To modern practitioners, the "100% pure Java" movement is taken up with the same fervor. Now that the Web has been invented, we find ourselves in an infancy, not dissimilar to the one which followed the invention of the printing press—from the years 1450 to 1501—during which 20 million pieces of printing appeared in Europe. **We are still scratching our heads trying to figure out how best to deploy and use the new technology.** By most accounts the initial period of printing was a restless, highly competitive, free-for-all dominated by itinerant printers. One of the first uses of printing was propaganda, followed by circulating broadsides lampooning a person or institution. Sound like Web pages and Usenet groups? Just as printing matured to the point where it could offer us the dictionary and the encyclopedia, so too the Web is evolving into ways to offer convenient means to provide database information. **The struggle goes on—like the alchemists, we are trying to turn raw data**

(lead) into meaningful information (gold). Sift it, scrub it, purify it, and live off the nuggets of pure gold that your competitors would dearly love to have.

We now experience the need to adorn our databases with nontextual information just as the monks felt the need to "illuminate" books. Perhaps, what Pope Gregory the Great said of book illumination is true of the Web, "Painting can do for the illiterate what writing can do for those who can read." In the carrels of the scriptoria the monks worked, stopping only to complain "with two fingers I toil." An eerie parallel to the modern Internet worker, working in one of the cubicles pictured in Dilbert, mindful of carpal tunnel syndrome. You may wonder, however, whether the Web will ever host as enduring a work as the Book of Kells. If the Book of Kells is never produced, it will not be through want of effort for the young programmer with his mantra, "When I am not sleeping I am working" has a regimen more intense than the Benedictine monk with his prescribed life of one third prayer, one third sleep, and one third intellectual and manual labor.

Notes

1 Tim Ouellette, "Human Services Get Tech Boost," *Computerworld*, December 29/January 5, 1998.

2 Laura DiDio, "NT Server doesn't Come Cheap," Computerworld, October 20, 1997, p. 1.

3 IBM, "Object-Relational DB2", <www.ibm.com> September 1996, p. 10.

John P. Byrne is National Register Database Manager National Park Service. This speech was given at the Information Ecosystem Conference.

Carol D. Shull

Computerizing the National Register of Historic Places

Anyone anywhere in the world with Internet access can find out what is listed in the National Register of Historic Places, our nation's official inventory of buildings, sites, districts, structures and objects significant in American history, architecture, archeology, engineering, and culture. Since 1986, the National Park Service (NPS) has had a computerized index, the National Register Information System (NRIS), which contains information on the nearly 80,000 historic places that are either listed in or determined eligible for the National Register. Now available on the Internet, this automated index has made expanding and maintaining the National Register more efficient and opened to the public a wealth of information about heritage resources for research, planning, policy analysis, public education, and tourism. The way that the NPS has gone about creating the NRIS and related initiatives and the lessons learned along the way may be of value to others in planning and carrying out computerization projects.

The NPS considered automating the National Register as early as 1968, soon after the passage of the National Historic Preservation Act. National Register and Advisory Council on Historic Preservation staff worked with IBM to design the first nomination form. A report titled "An Information System for the National Register" was completed in 1969. Diane Miller, who formerly managed the NRIS, writes in her excellent summary of the history of efforts to computerize the National Register, published in *CRM*, that this prescient report stated that "only an automated file system can assure adequate storage, retrieval and presentation for the volume of entries (over 100,000) anticipated."[1]

The NPS actually began the development of a computerized index in 1974 and had an operational system by 1977. The bureaucratic disruptions caused by the transfer of the National Register program from the NPS to the newly created Heritage Conservation and Recreation Service (HCRS) in 1978 and its subsequent transfer back to the NPS when HCRS was abolished in 1981, and staff turnover resulted in the abandonment of that system and preparation of a revised functional requirements document in 1983. Data was reentered in the new system maintained on a Hewlett Packard minicomputer, and data entry in the NRIS of all listings up to that time was finally completed in 1986. **It is not uncommon for bureaucratic changes to negatively impact database planning**

and development. This puts a great responsibility on managers to carefully examine the effects of their decisions to assure that valuable work is not lost.

The NPS benefitted from preparing a detailed functional analysis to identify requirements and objectives of the proposed system prior to making final decisions. In planning the NRIS, we tried to select as few data elements as possible to provide an effective index to National Register properties and access to descriptions, statements of significance, bibliographical references, maps, and photographs on each property in the National Register files. Information management systems that collect a large number of data elements for numerous cultural resources can become too expensive to continue and maintain over time. The NRIS includes all 45 data elements for each of the over 68,000 listings in the National Register, any of which can be searched in combination, a manageable number that allows it to be used effectively for policy analysis, research, preservation planning, and public education.[2] **A cost benefit analysis is helpful to determine how important each proposed data element is, how often it is likely to be queried and for what purposes, and to determine whether the costs of collecting and entering data and maintaining the proposed database can be supported over time.**

To determine which data elements to collect, the National Park Service had an interdisciplinary team select draft data elements and then sought advice from a variety of users such as State Historic Preservation Offices, federal agencies, professional organizations, and individuals before making a final selection. Seeking input from users was very helpful in making decisions. Nominating authorities collect the information on National Register nomination forms from which data is entered into the NRIS, which also helps keep the NPS' costs down.

In the early 1980s when the data was being entered, the NPS hired graduate students and other trained professionals in preservation related disciplines to review the paper documentation on each property and key in the data. A system to monitor their work for accuracy was put in place, and subsequently nominating authorities have reviewed the NRIS for errors, which the National Register staff corrects as they are identified.

The National Register nomination and determination of eligibility processes were analyzed in preparing the functional analysis so that the NRIS could become integral to the processing of nominations and determinations of eligibility. Data is entered on new nominations and determination of eligibility requests as they are received. National Register staff all have access to the database at their desks, so they can answer information requests quickly and efficiently.

Although all data entry is currently done by the National Register staff, the NPS has just completed the development of a Microsoft Access software package for completing the National Register nomination form that includes instructions for filling out the form and is downloadable from the National Register's Internet site. The data elements are automated as information is entered by private citizens and government officials filling out the computerized form. The completed forms can be sent to nominating authorities electronically or on a disk, and nominating authorities, or federal officials requesting determinations of eligibility as part of the planning of a federal project, can transmit the completed forms in the same way to reviewing agencies and the NPS for final processing. This software builds on the NRIS and is the first step in the direction of transmitting and processing nominations and determinations of eligibility electronically.

The software also makes it possible for federal, state, local and tribal government agencies with computerized cultural resource inventories to dump the computerized data elements into their own databases if they are compatible and the NPS' National Register staff to do the same without re-keying. **The NPS has been encouraging government agencies to incorporate the NRIS data elements as core data elements in their own cultural resources inventory databases so that cultural resources data can be made more accessible in a consistent manner nationwide.**

The categories of information on cultural resources collected on National Register registration forms have remained fairly consistent throughout the history of the program, and the National Register has made only conservative revisions to the forms over time.[3] The same categories of information are recorded on all resource types. **Organizations that do not collect data consistently for all resource types will be hampered in their efforts to automate.** Historic preservation programs that have used a variety of survey and inventory forms collecting different data may find it more difficult to enter consistent data elements on cultural resources into an automated system.

Some organizations have developed entirely different databases for archeological sites and other resources types. One of the most effective aspects of the NRIS is that consistent data is entered into one database that includes all cultural resource types. This is particularly important for property types that may have a variety of values, such as districts that contain resources of historic, architectural, and archeological significance. Having consistent data in one database makes it

possible to do comprehensive nationwide searches of National Register properties.

The National Register staff generally has resisted the urge to add new data elements that increase the costs and would not be available for all of the entries. Although the NRIS can serve as a guide to many subjects there are many categories of information it does not contain. To aid the researcher who wants to explore such subjects, **the National Register has an Internet site that permits downloading of the entire database for in-depth analysis.** Researchers then can add their own cross references on listed properties.[4] Database managers, who have not already done so, might consider doing something similar with their own automated systems to make them as useful to the public as possible.

The National Register requires properties to be located on U.S. Geological Survey maps, and the National Register form records universal transfer mercator (UTM) references from USGS maps on each property. Because UTM references are computerized, National Register properties can be mapped using Geographic Information Systems (GIS). **The ability to map properties using GIS technology has greatly enhanced the usefulness of National Register information for research and planning and made it possible for those developing GIS systems to incorporate information from the NRIS and integrate it with other environmental and geographic data.** As global positioning technologies continue to develop and are used increasingly to collect location information in cultural resources surveys, this data will become even more precise.

The security of the NRIS has been a high priority both to ensure the accuracy of data and preclude unauthorized changes and to allow the NPS to withhold information that might lead to a significant invasion of privacy, risk harm to the historic resource, or impede the use of a traditional religious site by practitioners. The NPS is authorized to withhold information that might lead to any of these situations under a section of the National Historic Preservation Act.[5] Passwords are required for data entry and to allow different levels of access, and the database is programmed to flag properties for which data should be withheld and to conceal sensitive information except from those authorized to obtain it.

The master database continues to be maintained on the Hewlett Packard minicomputer, which also provides some Internet searching capabilities to the public. The NRIS is transferred to a Windows NT server at regular intervals for public access. An FTP version for downloading and a Web version for simple, preprogrammed searches are available. A goal now is to increase the public's ability to do sophisticated queries over the World Wide Web using their own combinations of data elements, as we are able to do at the National Register.

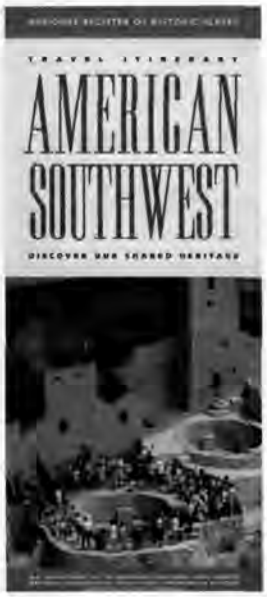
The NPS has an ongoing process of correcting the database and continues to upgrade the NRIS as money becomes available to finance improvements. One initiative is to add some searchable text in the form of summary paragraphs for each property, and this has been done for a small number of properties. In addition, this year we are beginning a project to scan the full text of registration documentation to facilitate responding to the thousands of requests the NPS receives for copies of National Register files each year, to make the full documentation more accessible electronically, and to help preserve the fragile paper records.

The NRIS has made the National Register of Historic Places more valuable and useful in many ways. Because the National Register has so many entries, searching the records by hand is simply not practical. The NRIS can be queried by such useful data elements as name of property, location, archi-

Researchers and writers use the NRIS to quickly obtain the names and locations of various types of resources listed in the National Register. The database was an invaluable tool for members of the National Register staff when identifying properties to include in *African American Historic Places*, recently reprinted by John Wiley and Sons Publishers.

tect/builder, architectural style, criteria for evaluation, dates, periods and areas of significance, significant persons, historic and current functions, nominating authority, as well as a name search of anything in the database and other pieces of information. **Anyone can use the NRIS on the Web to find which historic places in a community are listed.** Because the location, types and numbers of listings, and contributing resources within them can be quantified, meaningful policy analysis can be done. This has been useful in preparing reports to evaluate the impact of federal laws and legislative proposals, such as those dealing with the federal tax incentives for rehabilitating historic buildings that are listed in the National Register and for workload and other analyses. **The NRIS facilitates preservation planning** by allowing government agencies and others to find historic properties and the National Register documentation on them and to identify gaps and concentrations in cultural resource data. Government agencies and others can take all or portions of the NRIS to create or add to their own computerized databases. Individuals working to evaluate the significance of cultural resources are able to identify similar properties to which they can be compared. **The NRIS is the starting point for a variety of research projects and public education initiatives that identify related historic properties.** In one, the National Register staff queried the database and found more than 800 listings associated with Black Americans as the first step in preparing a book on African American historic places.[6] The National Register travel itinerary series includes GIS maps that link listed properties of interest to tourists,[7] and travellers anywhere in the world can use the NRIS to plan their trips. A first step in preparing a recent article on Young Women's Christian Association (YWCA) buildings began with a keyword search of YWCAs in the NRIS by name.[8]

The NRIS plays an important role when creating National Register travel itineraries for the "Discover Our Shared Heritage" series. National Register staff are able to search the database using such data elements as area of significance, period of significance, resource type, and location, ensuring that a diverse collection of historic places is included in the travel itineraries.



Much of America's rich cultural heritage is manifested in the historic places listed in the National Register and in the information about them that is in our national inventory. The National Register Information System has made it possible for the NPS to provide greater public access to these cultural resources. We plan to use the advances in information technology to continue to improve that access and to protect the invaluable records on these places that have so much to teach us about our collective heritage.

Notes

1. Diane Miller, "23 Years of Automating the National Register," in *CRM*, Vol. 14, No. 4, 1991.
2. Some properties determined eligible for the National Register, but not listed, have fewer data elements in the NRIS.
3. Explanations and definitions of the categories of information required on National Register forms, many of which are included in the National Register Information System, can be found in *How to Complete the National Register Registration Form* (Washington, D.C., National Park Service, National Register Bulletin 16A, 1991).
4. The address for the National Park Service, National Register of Historic Places Internet site is <www.nr.nps.gov.
5. See section 304 of the National Historic Preservation Act, as amended (16 U.S.C. 470w-3).
6. Beth L. Savage, ed., *African American Historic Places* (Washington, D.C.: The Preservation Press, National Trust for Historic Preservation,1994).
7. The published National Register of Historic Places "Discover our Shared Heritage" travel itinerary series is cosponsored by the National Park Service and the National Conference of State Historic Preservation Officers and partially funded by American Express. The series has been expanded to produce additional online itineraries on the National Register's Web site <www.cr.nps.gov/nr>.
8. Antoinette J. Lee, "Supporting Working Women—YWCA's in the National Register" in *CRM*, Vol. 20, No. 3, 1997.

Carol D. Shull is Keeper of the National Register of Historic Places, National Park Service. This speech was given at the Information Ecosystem Conference.

Diane Vogt-O'Connor

The Information Ecosystem Bibliography

Much excellent writing has recently been published on the challenges and perils of the information ecosystem from an unusually diverse crew of authors from academia, business, federal agencies, foundation and non-profit organizations, and similar groups. New and groundbreaking pieces are being issued daily. Keeping up can be a full-time job. Below are a few of the better writings on the information ecosystem, particularly pieces on electronic records preservation, intellectual property rights, information technology, and management of information including archives, electronic records, and paper.

Ackerman, Mar S. and Roy T. Fielding. "Collection Maintenance in the Digital Library" *Proceedings of Digital Libraries '95*, Austin Texas, June 1995, 39-48.

Association of British Columbia Archivists, Small Archives Committee. *A Manual for Small Archives*. British Columbia: Association of British Columbia Archivists, 1988.

Battin, Patricia. "From Preservation to Access: Paradigm for the Nineties." *IFLA Journal* 19 (1993): 367-373.

Bearman, David, ed. *Archives and Museum Information Technical Report No. 13: Archival Management of Electronic Records*. Pittsburgh, Pa.: Archives and Museum Informatics, 1991.

Bellardo, Lewis and Lynn Lady Bellardo. *A Glossary for Archivists, Manuscript Curators, and Records Managers*. Chicago: Society of American Archivists, 1992.

Branscomb, Anne Wells. "Public and Private Domains of Information: Defining the Legal Boundaries," in *Bulletin of the American Society for Information Science* 21 (2), pp. 14-18.

Branscomb, Anne Wells. *Who Owns Information: From Privacy to Public Access*. New York: Basic Books, 1995.

Burk, Cornelius and Forest W. Horton. *InfoMap: A Complete Guide to Discovering Corporate Information Resources*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

Conway, Paul. "Digital Preservation: Paper and Microfilm Go Electronic," *Library Journal* 119 (February 1, 1994): 42-45.

Conway, Paul. "The Implications of Digital Imaging for Preservation," In *Preservation of Library and Archival Materials*, 2nd ed. Edited by Sherelyn Ogden. Andover, MA: Northeast Document Conservation Center, 1994.

Conway, Paul. *Preservation in the Digital World*. Washington, DC: Commission on Preservation and Access. March, 1996.

Cook, Michael. *Archives Administration: A Manual for Intermediate and Smaller Organizations and for Local Government*. Folkestone, England: William Dawson & Sons, 1977.

Cox, Richard J. *Managing Institutional Archives*. New York: Greenwood Press, 1992.

Cronin, Blaise. "The Management of Intellectual Capital: from Texts to Markets," in Blaise Cronin, ed. *Information Management: From Strategies to Action*. London: Aslib, 1991.

Curtin, Bonnie Rose. *NAGARA (National Association of Government Archives and Records Administrators) GRASP Guide to Resources for Archival Strategic Preservation Planning*. Atlanta, GA: NAGARA, 1990.

Daniels, Maygene and Timothy Walch, eds., *A Modern Archives Reader: Basic Readings on Archival Theory and Practice*. Washington, DC: National Archives Trust Fund Board, 1984.

Davenport, Thomas H and Laurence Prusak. *Information Ecology: Mastering the Information and Knowledge Environment*. Oxford and New York: Oxford University Press, 1997.

Davenport, Thomas. *Process Innovation: Reengineering Work Through Information Technology*. Boston: Harvard Business School Press, 1993.

Dollar, Charles M. *Archival Theory and Information Technologies: The Impact of Information Technologies on Archival Principles and Methods*. Macerata: University of Macerata Press, 1992.

Dukas, Reuven, ed. *Cognitive Ecology: The Evolutionary Ecology of Information Processing and Decision-Making*. 1998.

Ellis, D. et. al. "Information Audits, Communication Audits and Information Mapping: A Review and Survey," *International Journal of Information Management* (1993), 134-152.

Ellis, Judith, ed., *Keeping Archives*. 2nd ed., Australia: Australian Society of Archivists and D.W. Thorpe, 1993.

Fox, Edward A. et al. "Digital Libraries: Introduction," *Communications of the ACM* 38 (April 1995): 23-28.

Getz, Janet, et al. *Oversize Color Images Project, 1994-1995: Final Report of Phase I*. Washington, DC: Commission on Preservation and Access, 1995.

Gilder, George. "Angst and awe on the Internet" *Forbes* ASAP (December 4). Also available: <http://homepage.seas.upe.nn. edu/~gaj1/ ggindex.html>.

Gill, T. Grandon. "High-Tech Hidebound: Case Studies of Information Technologies that Inhibited Organizational Learning," *Accounting, Management, and Information Technology*. 5:1 (1995), 41-60.

Graham, Peter S. "Requirements for the Digital Research Library," *College and Research Libraries* 56 (July 1995): 331-339.

Haas, Joan K., Helen Willa Samuels, and Barbara Trippel Simons. *Appraising the Records of Modern Science and Technology: A Guide*. Cambridge, Mass: MIT, 1985.

Ham, F. Gerald. *Archival Fundamentals Series: Selecting and Appraising Archives and Manuscripts*. Chicago: Society of American Archivists, 1992.

Hasenyager, Bruce W. *Managing the Information Ecology: A Collaborative Approach to Information Technology Management*. Greenwood, Conn: Quorum Books, 1996.

Henderson, John and N. Venkatraman, "Strategic Alignment: Leveraging Information Technology for Transforming Organizations," *IBM Systems Journal* 32:1 (1993) 4-17.

Henson, Steven L. *Archives, Personal Papers, and Manuscripts: A Cataloging Manual for Archival Repositories, Historical Societies, and Manuscript Libraries*. 2nd ed. Chicago: Society of American Archivists, 1989.

Kenney, Anne R. "Digital-to-Microfilm Conversion: An Interim Preservation Solution," *Library Resources and Technical Services* 37 (1993): 380-401.

Lanham, Richard. A. *The Electronic Word: Democracy, Technology, and the Arts*. Chicago: The University of Chicago Press, 1993.

Lesk, Michael. *Preservation of New Technology: A Report of the Technology Assessment Advisory Committee*. Washington, DC: Commission on Preservation and Access, 1992.

Levine, Melissa Smith. "Electronic Publishing: A Legal and Practical Primer," *CRM*, 18 (9), pp. 23-26.

Levy, David M. and Catherine C. Marchall. "Going Digital: A Look at Assumptions Underlying Digital Libraries," *Communications of the ACM* 38 (April 1995): 77-84.

Lynch, Clifford. "The Integrity of Digital Information: Mechanics and Definitional Issues," *Journal of the American Society for Information Science* 45 (December 1994): 737-744.

Lyytinen, Kalle and Rudy Hitschheim. *Information Failure—Surveys in Information Technology* (1987): 257-309.

McCall, J.J. editor. *The Economics of Information and Uncertainty*. Chicago: University of Chicago Press, 1982.

McKinnon, Sharon and William Burns. *The Information Mosaic*. Boston, MA: Harvard Business School Press, 1992.

Mohlhenrich, Janice. *Preservation of Electronic Formats: Electronic Formats for Preservation*. Fort Atkinson, WI: Highsmith, 1993.

National Archives and Records Administration. "Digital-Imaging and Optical Digital Data Disk Storage Systems: Long-Term Access Strategies for Federal Agencies." *Technical Information Paper No. 12*. Washington, DC: National Archives and Records Administration, 1994.

National Archives and Records Administration. *Disposition of Federal Records*. Washington, DC: NARA, Washington, DC: NARA, 1992.

National Research Council. *Preserving Scientific Data on Our Physical Universe: a New Strategy for Archiving the Nation's Scientific Resources*. Washington, DC: National Academy Press, 1995.

National Research Council. *Study on the Long-Term Retention of Selected Scientific and Technical Records of the Federal Government:*

Working Papers. Washington, DC: National Academy Press, 1995.

Ogden, Sherelyn, ed. for the Northeast Document Conservation Center. *Preservation of Library and Archival Materials*. Andover, MA: NEDCC, 1992.

Online Computer Library Center. *Tomorrow's Access-Today's Decisions: Ensuring Access to Today's Electronic Resources Tomorrow*. Proceedings of the Fourteenth Annual Conference of Research Library Directors. Dublin, Ohio: OCLC, 1996.

O'Toole, James M. *Archival Fundamentals Series: Understanding Archives and Manuscripts*. Chicago: Society of American Archivists, 1990.

Pelikan, Jaroslav. *The Idea of the University: A Reexamination*. New Haven: Yale University Press, 1992

Peterson, Gary M. and Trudy Huskamp Peterson. *Archives and Manuscripts: Law*. Chicago: Society of American Archivists, 1985.

Pugh, Mary Jo. *Archival Fundamentals Series: Providing Reference Services for Archives and Manuscripts*. Chicago: Society of American Archivists, 1992.

Ritzenthaler, Mary Lynn, Gerald J. Munoff, and Margery S. Long. *Archives and Manuscripts: Administration of Photographic Collections*. SAA Basic Manual Series. Chicago: Society of American Archivists, 1984.

Ritzenthaler, Mary Lynn. "Preservation of Archival Records: Holdings Maintenance at the National Archives," *National Archives Technical Information Paper No. 6*. Washington, DC: National Archives and Records Administration, 1990.

Ritzenthaler, Mary Lynn. *Archival Fundamentals Series: Preserving Archives and Manuscripts*. Chicago: Society of American Archivists, 1993.

Robinson, Peter. *The Digitization of Primary Textual Sources*. Office for Humanities Communication Publication, 4, Oxford, England: Oxford University Computing Services, 1993.

Rothenberg, Jeff. "Ensuring the Longevity of Digital Documents," *Scientific American* 272 (January 1995): 42-47.

Smithsonian Institution, National Archives and Records Administration, Library of Congress, and National Park Service. *A Primer on Disaster Preparedness, Management, and Response: Paper-Baed Materials*. Selected Reprints. Washington, DC: The National Archives, 1993. [Available free from Preservation Services, NARA, 7th & Pennsylvania Ave NW, Washington DC 20408.]

Task Force on the Archiving of Digital Information. *Preserving Digital Information*. Draft Report, Version 1.0, Washington DC: Commission on Preservation and Access and Mountain View CA: The Research Libraries Group, August 23, 1995. Also available: <http://www.rig.stanford.edu/ArchTF>.

Van Bogart, John. *Magnetic Tape Storage and Handling: A Guide for Libraries and Archives*. Washington, DC: Commission on Preservation and Access, 1995.

Vogt-O'Connor, Diane. "Archival and Manuscript Materials at the NPS" in *CRM* (1994) 17:05, pp. 33-35.

Vogt-O'Connor, Diane. "Finding Funds for the NPS Museum Collections" in *CRM* (1995), 18:04, pp. 9-13.

Vogt-O'Connor, Diane. "Why Keep Archives?" in *CRM* (1995)17:09, pp 9-11.

Walsch, James P. and Gerardo R. Ungson. "Organizational Memory," in *Academy of Management Review* 16:1: 57-91.

Waters, Donald J. *Realizing Benefits from Inter-Institutional Agreements: the Implications of the Draft Report of the Task Force on Archiving of Digital Information*. Washington, DC: The Commission on Preservation and Access, 1996. Also available: <http://arl.cni.org/arl/proceedings/127/waters.html>.

Tourists in an auto camp at Yellowstone. Two men ready supplies for preparing dinner. Photo courtesy Yellowstone National Park Archives.

Waters, Donald J. "Transforming Libraries Through Digital Preservation," In *Digital Imaging Technology for Preservation: Proceedings from an RLG Symposium held March 17 and 18th, 1994, 115-127*. Edited by Nancy E. Elkington. Mountain View, CA: Research Libraries Group, 1994.

Weibel, Stuart. "Metadata: The Foundations of Resource Description," *D-Lib Magazine*, July 1995, URL:<http://www.dlib.org>.

Wilson, Patrick. *Second Hand Knowledge: An Inquiry into Cognitive Authority*. Westport, CO.: Greenwood Press, Contributions in Librarianship and Information Science, Number 44, and 1983.

Wilsted, Thomas and William Nolte. *Archival Fundamental Series: Managing Archival and Manuscript Repositories*. Chicago: Society of American Archivists, 1991.

Zurkowski, P.G. "Integrating America's Infostructure," *Journal of the American Society for Information Science* 35:3 (1984): 170.

Periodicals—Archives

The Abbey Newsletter: Bookbinding and Conservation. (ISSN 0276-8291). Abbey Publications, 7105 Geneva Drive, Austin, TX 78723. Tel: 512-929-3992; Fax: 512-929-3995. Issued eight times a year.

The American Archivist (Official Journal of the Society of American Archivists, ISSN 0360-9081) The Society of American Archivists, 527 S. Wells St., 5th Floor, Chicago, IL 60607. Tel: 312-922-0140; Fax: 312-347-1452. Published quarterly.

Archival Outlook (Official Newsletter of the Society of American Archivists, ISSN 0091-5971) The Society of American Archivists, 527 S. Wells St., 5th Floor, Chicago, IL 60607. Tel: 312-922-0140; Fax: 312-347-1452. Published six times a year.

Archives and Museum Informatics: Cultural Heritage Information Quarterly (ISSN 1042-1467) published by Kluwer Academic Publishers, Order Dept., P.O. Box 322, Dordrecht, Netherlands 3300AH. Tel: (31)78 639 2392; Fax: (31)78 654 6474; Email: Services@wkap.nr. Published quarterly.

Diane Vogt-O'Connor is Senior Archivist for the Museum Management Program NPS.

Printed on recycled paper

CRM

U.S. Department of the Interior
National Park Service
Cultural Resources (Suite 350NC)
1849 C Street, NW
Washington, DC 20240

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE $300

FIRST CLASS MAIL
Postage & Fees Paid
U. S. Department of the Interior
G-83

VOLUME 21 • NO. 6
Cultural Resources
Washington, DC