Can the Web Really Do It All? Perceptions of Historical Research on the Internet
Author(s): Julie A. Corley
Published by: University of California Press on behalf of the National Council on Public History
Stable URL: http://www.jstor.org/stable/3378867
Accessed: 27/03/2013 13:36

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.
Contractor-Client Relations

Can the Web Really Do It All?  
Perceptions of Historical Research on the Internet

JULIE A. CORLEY

TELEVISION, MAGAZINES, AND NEWSPAPERS have run advertising campaigns creating the illusion that a vast amount of detailed and specific historical information is available via the Internet. If one believes the media, technology contains the answers to all questions, and a computer and a modem can quench the thirst for knowledge. Microsoft, for instance, would have us believe that we can study the mysteries of the pyramids or the holdings of the greatest museums in the world in depth via the Internet. Pacific Bell attempts to convince us that resources similar to those formerly found in the ancient library at Alexandria are at our fingertips, and IBM regularly suggests that companies can save billions of dollars by using the Internet. Many people have taken these claims at face value, in part because the allure is great. One historian states on his web page, “Imagine a library so large that it cannot be stored in one place. . . . Imagine . . . everything from stock quotes to the latest research on dinosaurs. All of this and more is available here on the World Wide Web.”

JULIE A. CORLEY received an M.A. from Arizona State University in 1992. Her work in the field of public history has included study of the community at St. Mary’s Church in Phoenix, Arizona, and extensive research in issues surrounding land-use history. She has been employed by PHR Environmental Consultants since 1992, and is currently a project manager with that firm.


© 1998 by the Regents of the University of California and the National Council on Public History
There can be no argument that a vast amount of historical information is available on the Internet. The University of Kansas, for example, has created an on-line index of resources for historians which provides connections to 2,400 Web pages containing historical information on various subjects. Other Web sites, such as History On-Line and The History Net, provide historical information in a popular-magazine format aimed at enthusiasts rather than professional researchers. Sites such as these present the text of documents like the Gettysburg Address and the Declaration of Independence. The vast amounts of data available have made publications such as *History Highway: A Guide to Internet Resources* a standard resource for people with an interest in history.

The ability to communicate easily with a large number of peers, conduct library searches without leaving the office, and identify relevant manuscript collections on almost any given research topic has made the Internet a valuable tool for professional historians. Technological advances, however, have taken this research tool a step farther and have provided a way for documents and photographs to be "published" on the Internet in digital libraries and archives. A growing number of photographs, documents from manuscript collections, and oral interviews have already been so replicated and are available on-line. Projects such as the California Heritage Digital Image Access Project, *American Memory* from the National Digital Library Program, and the Center for Electronic Text & Image from the University of Pennsylvania contain examples of the newly created digital archives. In the California Heritage Collection, on-line researchers use finding aids that contain embedded digital representations of primary sources. The Bancroft Library has selected materials such as the diary of Patrick Breen, scanned them, and made them available on the Internet. The University of Pennsylvania has made select company annual reports available in its digital archive, and photographs of the women's suffrage movement are presented at the *American Memory* Web site. Researchers equipped with a computer and modem can see images of the brown-stained pages and read about the hardships endured by the Donner party in 1846 or learn about the operations of Scovill Manufacturing Company between 1939 and 1942.

The growing numbers of these types of Internet sites suggest that the technology and use of the Internet continues to expand. The cost to scan

8. http://www.library.upenn.edu/etext
documents, coupled with the volume of materials held in archival collections, however, makes it impractical for the digital archive to be much more than a preservation tool or a method of expanding public relations and outreach programs. Few institutions have established formal criteria for the inclusion of materials in digital libraries and archives. Those who have consider issues like local interest, copyright ownership, preservation of fragile originals, and the "do-ability" of a project. Major programs, such as the Archives of Industrial Society at the University of Pittsburgh, have established planning work groups to develop guidelines for the selection of materials for digitization. Other institutions, however, are charting a cautious path and are taking on smaller projects. The John Oxley Library, Queensland, Australia, for example, determined that its users were interested in immigration history. The library started its imaging project by scanning photographs of ships that brought immigrants to Queensland. The library intends to incorporate manuscript and archival materials into the digital collection, and will show, for instance, a ship diary linked to the appropriate photograph. An archivist from another facility, Georgetown University, stated that "we [scan] what we want, but when we really get into production mode, we will be thinking of a function combining interest [and] rights, . . . and protecting fragile originals from excess handling." As of November 1997, the institution has digitized approximately 1,500 images/pages.

**Perspectives of a Public History Consultant**

The scope and types of information and resources available on the Internet vary greatly. The low cost of placing a Web page on the Internet has made it affordable for anyone to construct and maintain one. By late 1997, according to Carol Leita, director of Infopeople, a site run by the University of California, Berkeley, to help librarians teach people how to use the Internet, over 100 million Web sites have been posted on the Internet. Some of the information contained on some of the Web pages is accurate, and a lot of it is not. The sheer amount of information has spawned search engines, indices, and numerous publications that do nothing but review Web pages. Other resources available on the Internet include subject-related mailing lists, which enable group members to keep up to date on new or particularly useful Web pages and which serve as a forum to foster


communication on any given topic. Archives, public history, private investigation, history of technology, and museum management are only a few subjects with active discussion groups. Thousands of these groups are available, and if, by chance, one needed to search farther for others interested in a subject, thousands of newsgroups are accessible.\textsuperscript{13}

Although the availability of information and expansion of education should be encouraged, the digital library, as well as the media hype about the Internet, makes our job as consultants much harder. The Internet is a resource tool which provides a little information about complex subjects. Individuals with a casual interest in history find its features educational, whereas those with a need for obscure primary source material can use it to help locate relevant materials. Potential clients with complex research questions have asked, “Isn’t this just a matter of doing a couple of Internet searches?” or “Can’t you just do some database search?” In the private sector, consultants have been explaining the value of using historians to conduct research for a number of years. When presented with the opportunity, clients usually acknowledge the advantage of using researchers who have been trained in historical methodologies. The value gained by hiring someone who understands the source material, who can analyze the value of information, and who can interpret its meaning by placing it in its historic context is appreciable. Nonetheless, the perception that all relevant information can be found on the Internet is alluring and is often seen as a shortcut to solve the pending problem; but like another advertising campaign suggests, perception is not always reality. Let’s consider the Internet search results for the following three public history research questions.

1. An aqueduct and series of canals were constructed at Mission San Antonio de Padua, located in Monterey County, California, sometime prior to 1900. Our client wanted to reconstruct the aqueducts, but was unsure of their location. Desiring to be historically accurate, the client’s question was, “What is the history of the construction and where were the canals located?”

The Infoseek search engine produces 1,726 matches for the term “Mission San Antonio.” By adding “de Padua” as a qualifying term, the search results yield seventy-four matches.\textsuperscript{14} A quick scan of the first twenty of these

\textsuperscript{13} Mailing lists are accessed by subscription and have established guidelines or rules for participation. The subject matter in newsgroup discussions tends to be more general in nature and the postings to it can be accessed any time.

\textsuperscript{14} Each search engine has its own strengths and weaknesses and its own set of search tools. Several years ago, when the Web was in its early stages of development, programmers began to create computer programs that would search the Internet automatically and produce results to the queries presented. The trend at that time was to build a program that could handle the vast amounts of information available through the Internet and present this information in a usable form. Each entity that sponsored a search engine developed its own database. Then, when a query was posed, the search engine reviewed information in its database and “looked”
citations reveals several general-interest articles, such as road trips to California’s missions, but does not suggest that any useful information will be located. Attempts in other search engines including Lycos, Excite, Alta Vista, and Yahoo produce similar results. The best information acquired was an overview summary of the history of the mission, which noted that the padres relocated it in 1773 to a site that offered a better water supply. Sometime after the relocation, a dam was built on the San Antonio River and long aqueducts were constructed to bring water to the mission. The article about the mission was not cited, and no authoritative information was available about the source—except that it was a part of a series of Web pages about the California missions and that a company called BGe Internet, Intranet & Multimedia Services created the Web pages. The information is useful, but hardly something a public historian would submit to a paying client.

The search engine query did not provide citations for digitized collections; however, additional searches for information on the Mission San Antonio de Padua, including one of the libraries of the University of California and the RLIN manuscript database, produced twelve promising citations for archival materials held at the Bancroft Library and several general publications about the Mission San Antonio de Padua. A search of California Heritage Digital Image Access Project produced two undated photographs; a query posed to the SunSITE Finding Aids Web site yielded the Bancroft Library’s inventory of the San Antonio de Padua Mission documents. The usefulness of the Internet search, in this case, is that it pointed to some of the primary source collections—including the papers of the padres who settled the mission in the 1770s. These materials, along with maps prepared by Monterey County surveyors in the mid-1800s, were ultimately used to compile the history of the aqueduct and its canals.

2. The St. Louis and San Francisco Railway was a prominent rail line that operated in the Midwest and West in the mid-nineteenth and twentieth centuries. The railway company no longer exists under that name. The client’s question was, “What company is the corporate successor in interest to the St. Louis and San Francisco Railway?”

at the first paragraph of text on each Web page and categorized its usefulness to the query. As the Internet grows, the usefulness of these types of search engines is coming into question. A number of newer search engines are being developed that are based on summaries compiled by individuals rather than by a computer. The drawback to such search engines is that the databases are very small. Most of the search engines rank or sort the responses so that those most likely to be on target appear at the top of the list. Most also provide for boolean searches, and a few can handle adjacency and truncation in the keyword searches. See, for instance, Paula Mackinnon, “If They Only Had a Brain: Reviewing the Web Reviewers.” James Rettig, “Beyond ‘Cool’: Analog Models for Reviewing Digital Resources.”

The Infoseek search engine for the name of the company nets over twenty-one million citations.\(^{16}\) The first item on the list, however, is a notation that a collection of the company’s records (1859 to 1980) is located at the University of Missouri, Rolla. Although that material looks helpful, it will probably be worthwhile to see if a quick answer can be found for the specific research question at hand. Item twenty-eight on the search results leads to a cross-reference to Burlington Northern Santa Fe Corporation which states that the St. Louis and San Francisco Railway Company merged with Burlington Northern, Inc. The information suggests that the search is headed in the right direction, but that further details are needed. A search for the Burlington Northern home page leads to a summary history of the company, which states that the St. Louis and San Francisco Railway Company “was acquired and merged into Burlington Northern.”\(^{17}\)

Despite Burlington Northern’s Web page assertion that it acquired the St. Louis and San Francisco Railway Company in 1980, documentation that is more reliable will most likely be required. Our searches have shown the name of the company as the “St. Louis and San Francisco Railway” and the “St. Louis and San Francisco Railway Company,” and given the number of years involved (mid-1800s to 1980), it would be prudent to document the connection further. A search of the Moody’s *Manual of Railroads and Corporations* is recommended, but, unfortunately, the materials are not available on-line.

According to Moody’s, the St. Louis and San Francisco Railway was incorporated in Missouri in 1876. On June 30, 1896, the company was reorganized and chartered under the laws of Missouri as the St. Louis and San Francisco Railroad Company.\(^{18}\) In 1916, the St. Louis and San Francisco Railroad Company was sold at foreclosure proceedings according to a plan of reorganization.\(^{19}\) The St. Louis and San Francisco Railway Company was incorporated on August 24, 1916, and purchased the properties of the St. Louis and San Francisco Railroad Company.\(^{20}\) In 1977, Burlington Northern, Inc. and St. Louis and San Francisco Railway Company proposed a merger of the two companies, subject to approval of stockholders and regulatory agencies.\(^{21}\) On November 21, 1980, the companies merged by exchanging 1.9 shares of the Burlington common stock plus one-half share

---

16. The search terms can be manipulated in a number of ways, but most produce too few matches for a preliminary information search. For instance, a simple search using “St. Louis and San Francisco Railway” produces over 25 million hits. By changing the query to “+St. +Louis and +San +Francisco +Railway” (Infoseek’s method to ensure that each word will be found in the entry) the search result is one record—the manuscript collection at Rolla.


of no par preferred stock. According to Moody’s, the acquisition was accounted for as a purchase.22

In 1996, Burlington Northern, Inc. and Santa Fe Pacific Corporation merged, with the Santa Fe shareholders receiving 0.27 of a share of Burlington’s common stock for each Santa Fe share. The new company became Burlington Northern Santa Fe Corporation.23 If desired or needed, additional documentation could be gathered. These materials might include documentation of the reorganization of St. Louis and San Francisco Railway, foreclosure proceedings for St. Louis and San Francisco Railroad Company, Articles of Incorporation and associated filings, and records from the Securities Exchange Commission, none of which is available on-line. The usefulness of the Internet for this example is that it allowed us quickly to identify a likely successor for the St. Louis and San Francisco Railway.

3. During World War II, many heater and furnace manufacturers converted their operations to accommodate the federal government’s need for munitions. Did Carrier Corporation, a large heating, ventilation, and air conditioning manufacturer, produce furnaces or heaters prior to the war? If so, did it retool its operations to produce munitions? If so, which of the company’s plants were affected and has a history of these operations been written?24

The Lycos search engine produces over 800 matches for the terms “Carrier Corporation” and “history,” most of which appear to be useless for this task. A similar search in the Excite search engine provides several summaries on the history of the company.24 A quick review of several of these Web pages suggests that Carrier Corporation may have produced heaters and furnaces prior to 1941, but the listings are not definitive. The Carrier Corporation home Web page requires software not currently on my computer, so that source cannot be used without upgrading the system.

On-line searches of various library catalogs do not indicate that a comprehensive history for Carrier Corporation’s activities during World War II has been prepared. Likewise, further Web page searches do not provide a definitive answer to the questions at hand. A search of the RLIN database shows that Cornell University holds company records dating from 1875 to 1964. At this point, it appears that the Internet is not going to provide the answers to the research questions. An e-mail message to the archivist at Cornell University yields a response that a finding aid for the collection has

24. The Excite search engine provides links to different categories in its response. For instance, the Carrier Corporation query results include a match with press releases about the company, another with information about a distributor, another about pending litigation, and another with a link to the company’s annual report. The researcher can view the cited Web page, or can choose to look at similar documents.
been developed and can be obtained by sending a check for $28.80. Before ordering the finding aid, I wanted further verification, so working on the premise that the company was large enough to have had contracts with the government during the war, I conducted an Internet search of library collections for compilations of government contracts and industrial plants, locating two promising sources, *Alphabetic Listing of Major War Supply Contracts* and *War Industrial Facilities Authorized, July 1940 to August 1945*. According to the *Alphabetic Listing of Major War Supply Contracts*, Carrier Corporation had government contracts to produce a wide range of refrigerating equipment and air conditioners, as well as contracts for ordnance equipment and gun parts. The listing of contracts also indicates that the company had two plants, one located in Bridgeport, Pennsylvania, and another at Syracuse, New York. Review of *War Industrial Facilities Authorized, July 1940 to August 1945*, reveals that the company also produced aircraft engine mounts and Navy Mark X Projectors. The finding aid for the Cornell University collection of company records indicates that there are a number of files on "rifle shot" dating to the World War II era and several items that may provide context on the company’s operations during the war.

The Internet findings suggested that the company may have manufactured heaters and/or furnaces. Review of the *Alphabetic Listing of Major War Supply Contracts* and the Cornell finding aid, however, suggest that the company predominantly manufactured refrigeration products. Library research showed that the company had war contracts and produced some munitions, but it was unlikely that the company’s facilities were extensively retooled for this production. In this case, the Internet information was somewhat misleading and was not particularly helpful.

**Historians or Information Managers?**

As public historians conducting business in the private sector, we usually compete for work with others, including paralegals, lawyers, engineers, graduate students, investigative firms, and other historians. In many cases, clients select the person who appears to demonstrate the highest level of competence. Unlike historians operating in academia, we rarely have the opportunity to specialize in only one area of history; we must be versatile. Thus, we need to be able to identify the availability of relevant research materials quickly, be prepared to address a wide range of subject matter, and convince the potential client that historical research has merit. The

---


Internet is one tool that can help identify research materials and educate—to a certain degree—on a wide range of subject matter. Convincing potential clients of the merit of historical research, however, still requires human interaction. Clients want to be sure that any consultants they hire are well versed in the tools of the information age.

In 1993, E. J. Valauskas enthusiastically described the Internet as the “ultimate reference tool.” More tempered descriptions consider it part of a “mediated reference service” useful in certain situations. The information contained on the Internet can be useful, and historical consultants need to maintain a level of proficiency with on-line sources in order to ensure that clients are receiving the best possible service. Historical research requires a variety of skills and a willingness to improvise. Technology cannot replace basic training in historical research methodologies. Researchers should augment their skills with new tools, but not expect to replace them with the latest gadget or piece of software. The Internet contains much information that still requires verification for use in professional situations. Anyone can set up a Web page and post information on it. Such information must be evaluated carefully before it can be considered useful to a professional consultant.

Awareness of the Internet has made the historical consultant similar to what science fiction writer Robert A. Heinlein called an intuitive analyst. In his book Friday, the protagonist is an analyst whose duties include answering the esoteric research questions posed by her co-workers. Using only a computer terminal, she is able to investigate the complex corporate structure of multi-national companies, define the signs of a sick culture, and provide analysis on the historical relationships between men’s beards, women’s skirts, and the price of gold.

Current technology falls far short of that depicted in Heinlein’s novel, where books and documents were stored in a nitrogen environment and could be “read” from a computer terminal by using keystrokes to turn the pages. Archives and libraries selectively choose the materials for the digital collections, and these materials do not necessarily represent an accurate record of the facility’s holdings. The high cost of digitizing documents and photographs automatically means that some will never be made available in this format. Clients need to be reminded that even if some relevant materials have been digitized and are available on the Internet, it is likely that additional pertinent items may be located in that or other repositories.