The Current State of Primary Historical Sources Online

Richard Hacken

As a vital first step in substantiating and documenting historical details, there can be no substitute for a primary source derived from as close and contemporaneous an observation of a given event as possible. A historian unable to consult authoritative and honest voices from the past can verify little but is left to tinker with tradition and supposition. Until quite recently, the main mode of examining a primary source has been one on one—one scholar face-to-face with one original document in one physical space. Historiography has been slowed by travel expenses, time constraints, vagaries in obtaining permission, and other logistical difficulties standing between a historian and a source, wherever it may be housed. The steps of human progress in the arts and sciences of transcription, publication, photography, photocopying, and microfilming have been precursors to digitization, the latest boost that virtually places a document's image or essence before the critical eye of the scholar.

The vast majority of primary sources must, of course, still be sought *in situ*, be they locked away in a private stash or a public institution, at a monastery, in a special collection, a library, an archive, or another locale. We are at the very beginning of an imagined golden age of international online access. Even if some concerted and cooperative push à la Google Book—linking academia, museums, archives, and the corporate world—were applied to documents, letters, manuscripts, and other primary sources, it would still take decades before the numbers and varieties of sources available on your scholarly workstation were to reach a positive tipping point. To complicate matters, a largely invisible struggle now rages between fee-based documents on the one hand and sources with free and open access on the other. This review applies to open-access websites.

For ease of use, for preservation of materials, and not least for public relations, it is to the advantage of an institution to digitize historical sources unique to its own holdings. The most characteristic and rare documents will likely form a high priority on the digitization queue. As a local case in point, the Harold B. Lee Library has already added hundreds of nineteenth-century Mormon publications (books, hymnals, periodicals, etc.), along with pioneer and missionary diaries, to its digital collections and will be further augmenting these and other holdings unique to BYU.

As we look at this process worldwide, however, we see that funds and personnel vary greatly from one collection to the next. For a crass comparison, weigh the resources of Stanford University, which for years has been utilizing a Swiss-made robot scanner capable of digitizing over one thousand pages an hour,[1] against those of the St. Catherine Monastery on the Sinai Peninsula, which has hired a young Bedouin to operate a digital camera on a tripod inside an eight-by-eight-foot tent in the desert.[2]

One of the most practical questions for the historian or interested amateur is the following: How can I find which primary source documents are available to me online? All the digital labors are wasted if the results are hidden or extremely difficult to locate. One means of finding readily available, non-subscription-based historical documents has opened up through the so-called "Open Archives Initiative." This initiative, with the acronym OAI, offers a search engine called OAIster (pronounced "oyster"), whose slogan is to "find the pearls." This database[3] retrieves online documents from around the world: hundreds of institutions have made their digital sources OAI-compliant, that is, open and readable to the OAIster database. The Harold B. Lee Library online collections are among these searchable items. Thus, if you were to enter the search terms "Utah" and "history" and "sources" in the subject field of the OAIster search page, you would find links to a number of online documents offering primary sources for Utah history. One of them, cosponsored by the Lee Library and the Utah Academic Library Consortium, entitled "Trails to Utah and the Pacific: Diaries and Letters, 1846–1869," incorporates a large number of online diaries, maps, photographs, and illustrations. This important resource from the pioneer and immigrant era is now a constituent part of the Library of Congress's American Memory project.

In turn, American Memory, a major repository of online manuscripts, photographs, newspapers, sound files, and other means of documenting American history, has a customized search mechanism. At the time of this review, a search for "Nauvoo" in "American Memory" pulled up nearly a hundred online diary entries, transcriptions from issues of the U.S. Congressional journal, photographs, and other digital objects.

The search methodology used above—starting with OAIster and leading to "Trails to Utah and the Pacific" and then serendipitously backward to American Memory—is typical for current detective strategies of finding historical sources online. No single website can grant easy access to all the relevant sources that are online, not all the sources online are relevant, and most relevant sources are not yet online. "Too little, too early" sums up the situation.

General search engines help find a known web portal but will not, in most cases, satisfy those seeking individual historical source documents. Where OAIster gave us fourteen focused, digital documents for the terms "Utah," "history," and "sources," the same search in Google produces around 1,500,000 hits—the vast majority of which are irrelevant (pointing to analog rather than digital sources) and unfocused. As a result, resorting to "niche" web portals is currently a wise tactic. By "niche" web portals we refer to such websites as AMDOCS: Documents for the Study of American History; American Memory; Avalon Project: Documents in Law, History and Diplomacy; Digital Scriptorium (medieval and Renaissance documents); EuroDocs: Online Sources for European History; Gallica (French national digital library); Internet Ancient History Sourcebook; National Library of Canada Electronic Collection; Scripta Sinica (Chinese full text sources); World War I Document Archive; and similar sites that concentrate on given topics, geographical areas, or time periods while largely limiting themselves to primary source documents. The above list is just a sampling of web portals that have come online recently (since the mid-1990s) as guides to locating and linking digitized historical documents.[4]

Some criteria of quality and reliability worth considering when choosing and using an online version of a source document deal with digital provenance (Does the document come from a trustworthy provider, preferably the holder of the original?),[5] reliability (How helpful is the metadata? How correct, legible, and true to the original is the facsimile, transcription, or translation?), and stability of access (Can you count on the permanence and constant updating of the server where the document is posted?). The above-listed websites, sponsored by national and university libraries, or other critical and steady providers, were chosen for their careful adherence to these criteria.

An article in the *Wall Street Journal* last year remarked that the Internet provides "new troves of resources almost daily."[6] One exciting result of the advent of digital scanners, according to the reporter, is the "growing amount of primary sources such as journals, letters, photographs and other original documents." However, this excitement is mitigated

by the warning: "Now, if only it were all true." Erroneous transcriptions, typographical errors, or even consciously crafted counterfeits can be posted on websites. Due diligence is essential in locating and verifying the reliability of online sources. Five websites recommended in the *Journal* as efficient for searching through primary historical documents are ProQuest, a subscription database; the scanned items at the Library of Congress; the Avalon Project from Yale's Law School; a wax-cylinder sound recording site at the University of California at Santa Barbara; and the EuroDocs portal on the server of the Harold B. Lee Library.[7]

EuroDocs, compiled and tended by myself, offers links to digital facsimiles, transcriptions, and translations of mostly West European primary historical sources. The main page points to forty-six separate web indexes for countries and city-states of Europe, as well as to sites for "Prehistoric and Ancient Europe," "Medieval and Renaissance Europe," and "Europe as a Supranational Region." Most of the links are to external sources, but within the scope and linkage of EuroDocs are a number of digital treasures connected to BYU. Jesse D. Hurlbut of the French and Italian Department has assembled DScriptorium, an online image collection of medieval manuscripts. The Spain page features facsimiles and transcriptions of previously unpublished, late-sixteenth-century letters of King Philip II held in the L. Tom Perry Special Collections Library. In cooperation with an international research center housed in Turin, and initially as part of a sister-city relationship between the 2002 and 2006 Winter Olympic venues, we have begun to digitize accounts of American travelers to Italy—from Henry Adams to Washington Irving, from Harriet Beecher Stowe to Mark Twain. Monaco's history—from a barren rock fortress protecting the marine interests of the Holy Roman Empire to a principality sheltering high rollers and billionaires—is being documented online with BYU efforts in league with European archives. A number of EuroDocs pages point to the World War I Document Archive, a high-traffic attraction on the BYU Library server, among whose sources are found hundreds of locally re-imaged transcriptions of Austro-Hungarian documents on the 1914 outbreak of war. The library was also fortunate to get ownership of complete copies of the original Eisenhower communiqués, a collection of dispatches documenting the Allied advance across Europe after D Day in 1944, so that these typescripts and their searchable transcriptions could become a part of the HBLL's digital collections.

Obviously, we are only seeing the initial baby steps of a promising future for online historical documents. Likely areas of advancement will come in the documentary reproduction of local, regional, and family histories. As the volume of valuable historical materials increases online,

the finding aids and specialized search engines will undoubtedly become more sophisticated, intuitive, and precise. Plans are already underway for a "semantic web" that will utilize principles of artificial intelligence to dispatch "intelligent agents" into cyberspace to find, focus, classify, winnow, and display the sources we seek. Such a vision of the future will require historians not only to be consumers of electronic sources but also to be producers of computer ontology, to be describers and suppliers, to be toolsmith participants in the coming digital age of historiography.

———————

Richard Hacken (hacken@byu.edu) has worked for twenty-six years as European Studies Bibliographer at Brigham Young University. His recent online document projects include EuroDocs, World War I Document Archive, Letters of Philip II, and History of Monaco—Primary Documents. He received his PhD in 1975 from University of California, Davis.

1. John Markoff, "The Evelyn Wood of Digitized Book Scanners," *New York Times*, May 12, 2003, C8.

2. Suzanne Muchnic, "Saving a Treasured Trove, Ever So Slowly," *Los Angeles Times*, February 5, 2007, E1.

3. Given the pervasive presence of Google and other quick and effective search engines, I will not provide URLs (Web addresses) for OAIster and other websites, but will encourage you to discover them by name online.

4. Another history Web portal with search capabilities for primary documents on an international scale is the WWW-VL History Central Catalogue (WWW-VL stands for World Wide Web Virtual Library).

5. The websites of various national archives provide dependable sources of documents, as do "online collections" or "digital libraries" found on the websites of academic libraries and archives.

6. John Letzing, "Changing History," *Wall Street Journal*, February 13, 2006, R10.

7. Letzing, "Changing History," *Wall Street Journal*, February 13, 2006, R10.