

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Documentary Editing: Journal of the Association  
for Documentary Editing (1979-2011)

Documentary Editing, Association for

---

3-2002

## Access, Analysis, and Quality

Susan Belasco

University of Nebraska-Lincoln, sbelasco@unl.edu

Follow this and additional works at: <http://digitalcommons.unl.edu/docedit>

 Part of the [Digital Humanities Commons](#), [Other Arts and Humanities Commons](#), [Reading and Language Commons](#), and the [Technical and Professional Writing Commons](#)

---

Belasco, Susan, "Access, Analysis, and Quality" (2002). *Documentary Editing: Journal of the Association for Documentary Editing (1979-2011)*. 381.

<http://digitalcommons.unl.edu/docedit/381>

This Article is brought to you for free and open access by the Documentary Editing, Association for at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in *Documentary Editing: Journal of the Association for Documentary Editing (1979-2011)* by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

# Access, Analysis, and Quality

SUSAN BELASCO

In discussing the state of Web-based resources in my field, nineteenth-century American literature, my purpose is to provide some assessment of notable achievements and good works in progress, as well as to offer some comments about directions for the future.

Last April, I attended a memorial service in honor of an editor whom probably many of you knew: Joseph Slater. A long-time colleague of my husband's at Colgate University, Joe was the editor of *The Correspondence of Emerson and Carlyle*, and for many years the editor-in-chief of *The Collected Works of Ralph Waldo Emerson*.<sup>1</sup> This massive editorial project, which published its first volume thirty years ago, now stands at five of a projected ten volumes. The project has outlived most of the original editors and continues with a second and possibly a third generation of Emerson scholars. I mention the Emerson edition to make a point about the kind of time involved in many large editorial projects of the late twentieth century; it is little wonder that many contemporary editors and publishers are now drawn to the idea of distributing and publishing editions electronically. But few are fully aware that electronic editions may prove to be just as time-consuming as print editions. As Susan Hockey has said in her recent and excellent book, *Electronic Texts in the Humanities*:

It has been thought that electronic editions might be cheaper to produce, but . . . the amount of labour needed to produce a printed edition may not decrease for an electronic one. It still entails all the scholarly processes which are required for a print edition with the added complications of keeping up with new technology. . . . The editor is freed from the constraints of the shape and format of the printed book. This makes it pos-

sible to include much more material and also to avoid privileging one view or version of the text. It in fact it makes it possible to include every version of manuscript both in transcription and as a digital image. But the apparently infinite capacities of computers can lead to very ambitious projects which may not in the end have as much scholarly value. It is very tempting to provide as much source material as possible, but this can be at the expense of adding the critical analysis which makes the edition useful. The possibility of making changes over and over again to electronic information can also slow a project down, particularly if the changes have ramifications throughout the material.<sup>2</sup>

The complications and challenges of electronic editions, as outlined here by Hockey, are considerable. Several professional groups such as the consortium that forms the Model Editions Partnership, the MLA Committee on Scholarly Editions (now in revision), The Center for Electronic Texts in the Humanities at Rutgers University, and, of course, the ADE Committee on Electronic Standards have worked hard to establish principles and provide guidelines. These entities have provided us with clear and useful standards as well as helpful definitions.<sup>3</sup>

For example, the Prospectus for Electronic Historical Editions, developed by the Model Editions Partnership, advises us that:

Although historical editions share many common characteristics, each edition is unique. Letters, diaries, speeches, newspaper accounts, pamphlets, public records, and many other types of documents challenge the editor's skill and imagination in presenting representations of the texts which today's readers will find understandable and useful. Just as in today's book and microform editions, it is the editor who establishes a reliable text, provides the carefully chosen commentary necessary to understand that text, and creates the indices and other editorial devices which provide intellectual access to that text and commentary. Electronic editions will be no less challenging than creating today's book and microform editions, perhaps even moreso.<sup>4</sup>

---

Susan Belasco is Professor of English at the University of Nebraska, Lincoln, where she co-chairs the Text Studies Steering Committee. The author of articles and reviews on nineteenth-century American literature and culture, she is the editor of Margaret Fuller's *Summer on the Lakes* and Fanny Fern's *Ruth Hall*. She is the co-editor of *"These Sad but Glorious Days": Dispatches from Europe, 1846-1850* by Margaret Fuller; *Periodical Literature in Nineteenth-Century America*; and *Approaches to Teaching Uncle Tom's Cabin*. She is currently editing the poems Walt Whitman published in periodicals for inclusion in the Walt Whitman Electronic Archive.

The Committee on Electronic Standards of the ADE provides a straightforward definition of a quality text:

Simply rendering a text in electronic form does not constitute an electronic edition. The ADE-CES defines an electronic edition as primary source material prepared with 1) rigorous attention to the text, 2) explanatory annotation and 3) a description of the editorial practices used on the texts. There can be an enormous range of practice within the field of documentary editing, but all embody these characteristics. Electronic editions require the same burden of scholarship as print and microform publications do, and because of their format, additional issues must be considered.<sup>5</sup>

Despite the existence of these and other guidelines, many “editions” are being triumphantly posted on the Web, often incomplete, superficially proofread, and infrequently annotated. Sometimes these texts are little more than the would-be editor’s transcription of source material. Terms like *text* and *edition* are used loosely—even *archive* and *library* seem to have shifting definitions on the Web.

It is often very difficult to know where the texts or editions have come from and how they have been transcribed. My favorite example is *Leaves of Grass*, available in an unspecified edition, downloadable from the vast Project Gutenberg—no textual information of any kind is provided. A close second is Thoreau’s “Civil Disobedience,” available from the E-Server at the University of Washington. The note on the text says that it “originated with the Gutenberg version of the text.” “To improve online readability,” the editor goes on to say, “we have added margins and blank lines between paragraphs. We have also added Thoreau’s italics and other special characters from the original text, and divided the essay into two parts.” The *Gutenberg version* of the text? Two *parts*? What the editor means by the “original text”—the 1849 *Aesthetic Papers* where the essay version of this lecture first appeared, or the 1866 reprint in *A Yankee in Canada, with Anti-Slavery and Reform Papers*—is not disclosed. As I’m sure Emerson himself once said, “Editing is a fool’s paradise.”

Next to serving as an increasingly crucial communication device, the Web’s most stunning success is surely access to information—we can easily access headline news and the latest sports scores—as well as the images and texts that form the materials of our scholarly work. Gaining instant access to these materials—databases, indices, electronic libraries, library catalogues, and even electronic archives and texts—has become our daily expectation. But mere access is only a first step.

In the wake of the events of September 11, former secretary of state Madeleine Albright was among the many people interviewed on the “Today Show.” She made some trenchant observations about our inability to evaluate intelligence information. She observed that, because of computer technology, we have overwhelming and unprecedented access to information. But Albright was quick to point out that we have not yet developed sophisticated and effective ways to analyze what we can access.

The same point might well be made about literary texts available on the Web. The act of simply transcribing a text and mounting it on a website and the act of then reading a text on a screen hardly exploits computer technology. Creating access does not take significant advantage of the ways in which computers enable us to manipulate and analyze texts—what Susan Hockey and others have referred to as the real power of electronic texts. And besides, we all know that reading texts on screens—even as e-books on Palm Pilots—is not the same experience as reading texts in printed form. Putting aside the issue of convenience of accessing “editions” of texts, what are some of the projects that are based on sound editorial principles and practice and are providing something beyond mere access to materials?

In the appendix, I have listed a number of projects of various sizes and kinds that may be of interest to editors, and I want to examine a few of them. I have included some archives of electronic texts accessible online that include collections of texts, often the collective work of members of a related group, usually produced by libraries and institutions and occasionally by very hard-working faculty members who probably have taken on more than they realize. The major electronic text centers at Virginia, Princeton, Michigan, and California have led the way in this effort, and I suspect that most of you are familiar with their collections. In general, these collections provide images of original documents as well as transcriptions of the documents that include retrieval and search tools. Detailed documentation of their editorial practices is provided as well as evidence of conformance to relevant standards (such as using SGML as a format and adhering to TEI guidelines). In some cases supplementary materials are also available, such as biographical information about the authors, images of people and texts, bibliographic information, and, very importantly, textual information.

I want to comment on some of the resources listed in the appendix, demonstrating the kinds of documentation that some editors are using in their efforts to provide good information about the texts they are presenting

on their websites. My first example, however, is something of a negative one, and one that demonstrates both the potential of electronic text projects as well as the pitfalls. The Nineteenth Century Women Writers Web, Etext Library is maintained by a group of volunteers who submit text to the Society for the Study of American Women Writers website. There are no guidelines for the texts, and very little information is provided about transcriptions and editorial decisions. In the entry for Lydia Maria Child, for example, we can read the poorly proofread text of a short story, "Stand from Under," which first appeared in *The Liberator*, 28 January 1832. A "web-based version" introduces original art produced for this text, which was certainly not a part of the original publication and may lead readers to inaccurate conceptions of Child's story as well as *The Liberator* (which was not an illustrated newspaper).

Much more valuable is The Online Archive of Nineteenth-Century U.S. Women's Writings. This project, organized by Glynis Carr at Bucknell University, is similar in intention to the Nineteenth Century Women E-text Library, but it makes much more documentation available to users. A preface and an essay, "Editorial Practices and Textual Notes," provide clear information to readers about how the texts have been transcribed and edited for this project. In addition, each section of the library has separate textual information available. In this project, the section on Lydia Maria Child provides a collection of texts published by Child in the *Liberty Bell*, an antislavery annual published by Maria Weston Chapman in Boston from 1843 to 1858. Documents are in HTML (and not SGML); some illustrations from the *Liberty Bell* are included but not images of the pages. The editor also includes a chronological index, very helpful in using the microfilm version of *The Liberty Bell* available in the American Periodical Series. Professor Carr's site is limited by her use of HTML, which means, among other things, that the site is not searchable (except by single pages using the Find on Page feature of Netscape Navigator). Nonetheless, Professor Carr is careful to present her texts for exactly what they are. Her documentation is exemplary; the transcriptions have been scrupulously proofread, and as a project headed by a single professor with a busy teaching schedule, it is an excellent resource.

When we move to libraries with institutional sponsors, we enter a different order of sophistication, and I have listed sites in the appendix that are excellent examples of how effective electronic archives can be. Edited by William L. Andrews, the North American Slave Narratives is an ambitious project to include all the narratives

of fugitive and former slaves to 1920. Detailed information is provided about the digitization of the collection with short essays on standards, methodology, scanning of images, and encoding guidelines. A status report provides current information about the state of the project, with a production chart that shows various stages in the process of mounting the documents on the site. Since the texts have all been encoded in SGML according to TEI guidelines, the collection is searchable. For example, using the Google search feature, I can easily determine that "Haiti" appears twenty-five times in the collection, in texts ranging from *The Life and Times of Frederick Douglass* (1892) to Theophilus Gould Steward's *Fifty Years in the Gospel Ministry from 1864 to 1914* (1921?).

Similarly, the African American Women Writers of the 19<sup>th</sup> Century is a digital collection of some fifty-two published works by nineteenth-century black women writers. A part of the Digital Schomburg, this collection provides access to books and pamphlets written by women and published prior to 1920. Like the North American Slave Narratives library, the Schomburg editorial staff offers detailed descriptions of their editorial methods and encoding standards. Although the access to texts that are difficult to find is important, the Schomburg has also made use of the electronic medium to make the texts flexible; as the editors explain, "The Digital Schomburg African American Women Writers of the Nineteenth Century series is a contextualized collection of fully searchable electronic editions. In addition to the value of these works as individual electronic texts, their search capabilities and their interlinked arrangement with other collections allow them to be read within a historical context of multiple media. The reader may choose to focus on either a single work or to cross-read or search the whole collection."<sup>6</sup> The collection provides images of title pages and illustrations and has an especially good navigation system so that readers can go quickly to chapters and pages within an individual text.

Among the single-text projects that have been developed in nineteenth-century American literature, one is clearly a standout: *Uncle Tom's Cabin* and American Culture. This project, as described by the editor Stephen Railton, is a "multi-media archive." And it is an archive in the fullest sense; the editor has included not only electronic texts of other works, such as Stowe's *A Key to Uncle Tom's Cabin*, but also playable versions of songs and hymns mentioned in the novel; clips of movie versions of *Uncle Tom's Cabin*, including the 1903 Edison-Porter film; articles and reviews of the novel and its aftermath; other historical documents related to slavery, such as portions of the

text of Lydia Maria Child's *An Appeal in Favor of That Class of Americans Called Africans* (1833); and a collection of images of Uncle Tom that were used in advertisements in the nineteenth and twentieth centuries. The archive is a vast collection of material—obviously far more than any print text could offer. The text of *Uncle Tom's Cabin* is based on the first edition published in 1852; no notes are provided. However, users of the site can compare images of the extant manuscript pages alongside pages from *The National Era* (the newspaper in which the novel was serialized from 1851 to 1852) and the novel as published by John Jewett in 1852. In recognition that some users (especially students) might find the array of resources overwhelming, Railton has included some guides for teachers in the archive as well.

Several promising single-author projects involving electronic texts and editions are in various stages of completion on the Web. The Melville Electronic Library has some sample pages from *Moby-Dick* and *Typee* available for viewing, but the project is in the very early stages of planning and is currently seeking funding for development. The Charles Chesnutt Archive is currently a small project, funded by Berea College and edited by Stephanie Browner. A number of Chesnutt's works have been scanned or transcribed for the site; other texts, such as "The Goophered Grapevine," are provided via the University of Virginia E-Text Center. Access to Chesnutt materials is the major goal of the site; the texts are not searchable.

The two most developed single-author sites on the Web so far are The Dickinson Electronic Archives and The Walt Whitman Archive. Both archives have had a variety of kinds of institutional support (especially from the Institute of Advanced Technology at the University of Virginia) and have been awarded funding from the National Endowment for the Humanities and/or the Fund for the Improvement of Post-Secondary Education, as well as private foundations. And both archives have extensive documentation about the encoding and textual practices for the project and include descriptions of the intent and purposes of the archive. Both archives also include images of texts and manuscripts, as well as SGML-marked versions of manuscripts and editions. The Dickinson Archive, edited by Martha Nell Smith, Ellen Louise Hart, Marta Werner, and Lara Vetter, includes a critical edition of *The Writings of Susan Dickinson* (Dickinson's sister-in-law and most favored correspondent). Parts of the Dickinson archive have restricted access because of ongoing negotiations about permissions with the Houghton Library at Harvard University (the

vexed question of copyright and permissions about e-texts is another issue that all electronic editors must consider carefully). These valuable sections, especially "Emily Dickinson Writing a Poem," include images of manuscripts, transcriptions of poems, and notes that provide a rare opportunity to study in the manuscripts and the various ways in which editors have made decisions about what constitutes Dickinson's poems.

The Walt Whitman Archive, which is currently being revised with a new mainpage and an improved navigational system, includes all the editions of *Leaves of Grass* published in Whitman's lifetime, some of the prose, all the contemporary reviews, all the known photographs of Whitman, a large number of manuscripts, biographical essays, bibliographies, and a section of resources for classroom use. Edited by Kenneth M. Price and Ed Folsom, the Whitman Archive includes both facsimile and e-text versions of Whitman's work; the archive is fully searchable and thoroughly documented. The editors are now preparing facsimiles and e-text of Whitman's poetry manuscripts, which have never before been systematically collected and edited. In addition, plans are under way to add to the site the approximately 110 of Whitman's poems printed in periodicals during his lifetime; this addition will include images from the periodicals as well as transcripts of the poems.

What lessons might we learn from this brief tour of Web-based resources in nineteenth-century American literature? Clearly we are seeing a proliferation of electronic texts and archives on the Web. Even if we have not found them, our students have, and they are using them. If we want students to have access to first-rate, reliable resources, we need to participate in educating them about what they are using.

First, we need to promote, circulate, and advertise the guidelines and standards for the creation of electronic texts that the ADE and other organizations have formulated and encourage our colleagues and institutions to follow them as they prepare texts and resources. In addition, we need better mechanisms for evaluating these projects and providing them electronically, such as the critical survey of twenty archives and editions undertaken by the ADE Committee on Electronic Editions and Charles Green's excellent online review of Whitman sites.<sup>7</sup> In my view, associations, author societies, and professional organizations that are positioned to mount reviews of electronic projects and provide guidance to scholars and students should do so. And finally, we need more ways of educating ourselves about the theory and practice of electronic text technology. We need summer institutes, short courses,

and the creation of undergraduate and graduate programs that integrate computing in the humanities firmly into the curriculum. Without doubt, the Web has already changed the way we conduct research and teach our students. Our challenge is to make sure that we position ourselves, our students, and our institutions to take full advantage of this transformation.

## Appendix: Web Sources in Nineteenth-Century American Literature

### *Guidelines for Electronic Texts and Other Useful Resources*

- ADE Committee on Electronic Standards <http://www.nyu.edu/projects/sanger/ceshome.htm>
- Center for Electronic Texts in the Humanities at Rutgers University <http://www.ceth.rutgers.edu/>
- Hockey, Susan. *Electronic Texts in the Humanities: Principles and Practice*. Oxford: Oxford University Press, 2000.
- MLA Committee on Scholarly Editions <http://sunsite.berkeley.edu/MLA/guidelines.html> and <http://www.mla.org/> (go to Reports and Documents for revision plans of these guidelines)
- Model Editions Partnership <http://adh.sc.edu/>

### *Libraries and Archives*

- African-American Women: On-line Archival Collections <http://scriptorium.lib.duke.edu/collections/african-american-women.html>
- African American Women Writers of the 19<sup>th</sup> Century [http://digital.nypl.org/schomburg/writers\\_aa19/](http://digital.nypl.org/schomburg/writers_aa19/)
- Harlem: Mecca of the New Negro <http://etext.virginia.edu/harlem/index.html>
- Nineteenth Century Women Writers Web, Etext Library <http://www.unl.edu/legacy/Elibrary.html>
- North American Slave Narratives <http://docsouth.unc.edu/neh/neh.html>
- Online Archive of Nineteenth-Century U.S. Women's Writings <http://www.facstaff.bucknell.edu/gcarr/19cUSWW/index.html>

### *Single-Text Projects*

- How the Chimney-Sweep Got the Ear of the Emperor: A Critical Edition with Facsimile <http://etext.virginia.edu/users/tousignant/chimney/>
- Mark Twain's *Huckleberry Finn*: Text, Illustrations, and Early Reviews <http://etext.lib.virginia.edu/twain/huckfinn.html>
- *Uncle Tom's Cabin* and American Culture <http://jefferson.village.virginia.edu/utc/>

### *Print Edition Projects with Electronic Components*

- Mark Twain Papers and Project <http://library.berkeley.edu/BANC/MTP/index.html>
- Writings of Henry D. Thoreau
- <http://libws66.lib.niu.edu/thoreau/>

### *Single-Author Projects*

- William Ellery Channing Center <http://www.athens.net/~wells/indexwec.htm>
- Charles Chesnutt Archive <http://www.berea.edu/ENG/chesnutt/index.html>
- Dickinson Electronic Archives <http://jefferson.village.virginia.edu/dickinson/>
- Papers of Frederick Douglass (MEP Project) <http://adh.sc.edu/fd/fd-table.html>
- Melville Electronic Library <http://www.iath.virginia.edu/melville/>
- Elizabeth Oakes Smith Page <http://www.neiu.edu/~thscherm/eos/eospace.htm>
- Mark Twain in His Times <http://etext.virginia.edu/railton/>
- Walt Whitman Hypertext Archive <http://jefferson.village.virginia.edu/whitman>

### *Author Society Projects that Include Texts or Editions*

- James Fenimore Cooper Society
- <http://webserver1.oneonta.edu/external/cooper/>
- Edgar Allan Poe Society of Baltimore
- <http://www.eapoe.org>

## Notes

1. See Joseph Slater, *The Correspondence of Emerson and Carlyle* (New York: Columbia University Press, 1964); and Robert E. Spiller et al., *The Collected Works of Ralph Waldo Emerson* (Cambridge: Belknap Press of Harvard University Press, 1971–).

2. Susan Hockey, *Electronic Texts in the Humanities: Principles and Practice* (New York: Oxford University Press, 2000), 132.

3. The URLs for all sites mentioned in the text are provided in the Appendix.

4. See <http://adh.sc.edu/mepinfo/MEP-Docs/proptoc.htm#Section.2>.

5. See <http://www.nyu.edu/projects/sanger/ceshome.htm>, "Defining Electronic Editions."

6. See [http://digital.nypl.org/schomburg/writers\\_aa19/editorial.html#](http://digital.nypl.org/schomburg/writers_aa19/editorial.html#) The Digital Schomburg Editions: An Interlinked Collection.

7. See Charles Green's review at <http://english.ttu.edu/kairos/2.2/reviews/green/whitfirst.html>; the ADE Reviews are available at Survey of Sites, <http://www.nyu.edu/projects/sanger/ceshome.htm>. See also my "Surfing the American Renaissance: Internet Resources for Literary Scholars," *ESQ* 45 (Winter 1999): 67–95.