Electronic Scholarly Editions

Kenneth M. Price

Many people have commented on the ongoing transformation of scholarly publication, with some fearing and some exulting over the move to digital production and dissemination. It seems inevitable that an ever-increasing amount of scholarly work will take digital form. Yet if we assess the current moment and consider key genres of traditional print scholarship — articles, monographs, collections of essays, and scholarly editions — we find that digital work has achieved primacy only for editions. The turn toward electronic editions is remarkable because they are often large in scope, requiring significant investments in time and money. Moreover, although editions are monuments to memory, no one is expressing great confidence in our ability to preserve them. Marilyn Deegan notes the challenge librarians face in preserving digital scholarly editions — this “most important of scholarly tools” — and suggests steps scholars may take “in order to ensure that what is produced is preserved to the extent possible” (Deegan 2006: 358). Morris Eaves candidly acknowledges the uncertain future of the project he co-edits: “We plow forward with no answer to the haunting question of where and how a project like [The William Blake Archive] will live out its useful life” (Eaves 2006: 218). Given the apparent folly of investing huge amounts of time and money in what cannot be preserved with certainty, there must be sound reasons that make this medium attractive for scholarly editors. This chapter explores some of those reasons.

Jerome McGann has argued that the entirety of our cultural heritage will need to be re-edited in accord with the new possibilities of the digital medium (McGann 2002: B7). He laments that the academy has done little to prepare literary and historical scholars for this work, thus leaving the task, he fears, to people less knowledgeable about the content itself: librarians and systems engineers. In general, humanities scholars have neglected editorial work because the reward structures in the academy have not favored editing but instead literary and cultural theory. Many academics fail to recognize the theoretical sophistication, historical knowledge, and analytical strengths necessary to produce a sound text or texts and the appropriate scholarly apparatus for a first-rate edition. In this fraught context, it seems useful to clarify the key terms and assumptions at work in this chapter. By scholarly edition, I mean the establishment of a text on explicitly stated principles and by someone with specialized knowledge about textual scholarship and the writer or writers involved. An edition is scholarly both because of the rigor with which the text is reproduced or altered and because of the expertise brought to bear on the task and in the offering of suitable introductions, notes, and textual apparatus. Mere digitizing produces information; in contrast, scholarly editing produces knowledge.

Many prominent electronic editions are referred to as digital archives, and such undertakings may strike some people as loose usage. In fact, electronic editorial undertakings are only imperfectly described by any of the terms currently in use: edition, project, archive, thematic research collection. Traditionally, an archive has referred to a repository holding material artifacts rather than digital surrogates. An archive in this traditional sense may well be described in finding aids but its materials are rarely, if ever, meticulously edited and annotated as a whole. In an electronic environment, archive has gradually come to mean a purposeful collection of digital surrogates. Words take on new meanings over time, of course, and archive in a digital context has come to suggest something that blends features of editing and archiving. To meld features of both — to have the care of treatment and annotation of an edition and the inclusiveness of an archive — is one of the tendencies of recent work in electronic editing. One such project, the William Blake Archive, was awarded an MLA prize recently as a distinguished scholarly edition. The Walt Whitman Archive also contains a fair amount of matter that, in the past, ordinarily would not be included as part of an edition focused on an
"author's" writings: we have finding guides to manuscripts, a biography, all reviews of Whitman's work, photographs of the poet, selected criticism from his contemporaries and from our own time, an audio file of what is thought to be the poet's voice reading "America," encoding guidelines, and other information about the history and technical underpinnings of the site. In other words, in a digital context, the "edition" is only a piece of the "archive," and, in contrast to print, "editions," "resources," and "tools" can be interdependent rather than independent.

Why Are People Making Electronic Editions?

One distinguishing feature of electronic editions is their capaciousness: scholars are no longer limited by what they can fit on a page or afford to produce within the economics of print publishing. It is not as if economic problems go away with electronic editing, but they are of a different sort. For example, because color images are prohibitively expensive for most book publications, scholars can usually hope to have only a few black and white illustrations in a book. In an electronic edition, however, we can include as many high-resolution color images as can be procured, assuming adequate server space for storage and delivery, and assuming sufficient staff to carry out the laborious process of scanning or photographing materials and making them available to users. A group of scholars who have sufficient resources (and who work with cooperative repositories) can create an edition of extraordinary depth and richness, an edition that provides both the evidence and the final product — a transcribed text and the images of the material they worked from in producing the edition. They can include audio and video clips, high-quality color reproductions of art works, and interactive maps. For multimedia artists such as William Blake and Dante Gabriel Rossetti the benefits are clear: much more adequate representation of their complex achievements. Likewise it is now possible to track and present in one convenient place the iconographic tradition of Don Quixote, a novel that has been repeatedly illustrated in ways rich with interpretive possibilities (<http://www.csdltamu.edu/cervantes/english/images_temp.html>). The non-authorial illustrations of this novel are part of the social text and provide an index to the culturally inscribed meanings of Cervantes' novel. Electronic editions have also deepened interest in the nature of textuality itself, thus giving the field of scholarly editing a new cachet.

The possibility of including so much and so many kinds of material makes the question of where and what is the text for an electronic edition every bit as vexed (if not more so) than it has been for print scholarship. For example, we might ask: what constitutes the text of an electronic version of a periodical publication? The Making of America project, collaboratively undertaken by Cornell University (<http://cdl.library.cornell.edu/moa/>) and the University of Michigan (<http://www.hti.umich.edu/m/mpagrp/>), is a pioneering contribution of great value, but it excludes advertising, thereby implying that the authentic or real text of a periodical excludes this material of great importance for the study of history, popular culture, commerce, and so on. A similar disregard of the material object is seen also in editions that emphasize the intention of the writer over the actual documents produced. This type of edition has been praised, on the one hand, as producing a purer text, a version that achieves what a writer ideally wanted to create without the intervening complications of overzealous copy-editors, officious censors, sloppy compositors, or slips of the writer's own pen. For these intentionalist editors, any particular material manifestation of a text may well differ from what a writer meant to produce. On the other hand, the so-called "critical" editions they produce have been denigrated in some quarters as being ahistorical, as producing a text that never existed in the world. The long-standing debate between, say, critical and documentary editors won't be resolved because each represents a legitimate approach to editing.

A great deal of twentieth-century editing — and, in fact, editing of centuries before — as G. Thomas Tanselle notes, was based on finding an authoritative text based on "final intentions" (Tanselle 1995: 15–16). Ordinarily editors emphasized the intentions of the "author" (a contested term in recent decades) and neglected a range of other possible collaborators including friends, proofreaders, editors, and compositors, among others. A concern with final intentions makes sense at one level: the final version of a work is often stronger — more fully developed, more carefully considered, more precisely phrased — than an early or intermediate draft of it. But for poets, novelists, and
dramatists whose work may span decades, there is real question about the wisdom of relying on last choices. Are people at their sharpest, most daring, and most experimental at the end of life when energies (and sometimes clarity) fade and other signs of age begin to show? Further, the final version of a text is often even more mediated by the concerns of editors and censors than are earlier versions, and the ability of anyone to discern what a writer might have hoped for, absent these social pressures, is open to question.

The long-standing concern with an author’s final intentions has faced a significant challenge in recent years. Richard J. Finneran notes that the advent of new technologies "coincided with a fundamental shift in textual theory, away from the notion of a single-text 'definitive edition' " (Finneran 1996: x). Increasingly, editors have wanted to provide multiple texts, sometimes all versions of a text. Texts tend to exist in more than one version, and a heightened concern with versioning or fluidity is a distinctive trait of contemporary editing (Bryant 2002; Schreibman 2002: 287). Electronic editing allows us to avoid choosing, say, the early William Wordsworth or Henry James over the late. Even if there were unanimous agreement over the superiority of one period over another for these and other writers, that superiority would probably rest on aesthetic grounds open to question. And of course we often want to ask questions of texts that have nothing to do with the issue of the "best version."

New developments in electronic editing — including the possibility of presenting all versions of some especially valuable texts — are exciting, but anyone contemplating an electronic edition should also consider that the range of responsibilities for an editorial team has dramatically increased, too. The division of labor that was seen in print editing is no longer so well defined. The electronic scholarly edition is an enterprise that relies fundamentally on collaboration, and work in this medium is likely to stretch a traditional humanist, the solitary scholar who churns out articles and books behind the closed door of his office. Significant digital projects tend to be of a scope that they cannot be effectively undertaken alone, and hence collaborations with librarians, archivists, graduate students, undergraduate students, academic administrators, funding agencies, and private donors are likely to be necessary. These collaborations can require a fair amount of social capital for a large project, since the good will of many parties is required. Editors now deal with many issues they rarely concerned themselves with before. In the world of print, the appearance of a monograph was largely someone else's concern, and the proper functioning of the codex as an artifact could be assumed. With regard to book design, a wide range of options were available, but these choices customarily were made by the publisher. With electronic work, on the other hand, proper functionality cannot be assumed, and design choices are wide open. Collaboration with technical experts is necessary, and to make collaboration successful some knowledge of technical issues — mark-up of texts and database design, for example — is required. These matters turn out to be not merely technical but fundamental to editorial decision-making.

The electronic edition can provide exact facsimiles along with transcriptions of all manuscripts and books of a given writer. At first glance, it might be tempting to think of this as unproblematic, an edition without bias. Of course, there are problems with this view. For one thing, all editions have a perspective and make some views more available than others. An author-centered project may imply a biographical framework for reading, while a gender-specific project like the Brown University Women Writers Project (<http://www.brown.edu/>) implies that the texts should be read with gender as a key interpretive lens. Neutrality is finally not a possibility. Peter Robinson and Hans Walter Gabler have observed that "experiments with the design of electronic textual editions suggest that mere assembly does not avoid the problem of closing off some views in order to open up a particular view" (Robinson and Gabler 2000: 3). Even an inclusive digital archive is both an amassing of material and a shaping of it. That shaping takes place in a variety of places, including in the annotations, introductions, interface, and navigation. Editors are shirking their duties if they do not offer guidance: they are and should be, after all, something more than blind, values-neutral deliverers of goods. The act of not overtly highlighting specific works is problematic in that it seems to assert that all works in a corpus are on the same footing, when they never are. Still, there are
degrees of editorial intervention; some editions are thesis-ridden and others are more inviting of a multitude of interpretations.

A master database for a project may provide a unique identifier (id) for every document created by a writer. In this sense all documents are equal. Lev Manovich has commented interestingly on the role of the database in digital culture:

After the novel and subsequently cinema privileged narrative as the key form of cultural expression in the modern age, the computer age introduces its correlate — the database. Many new media objects do not tell stories; they don't have beginning or end; in fact, they don't have any development, thematically, formally or otherwise which would organize their elements into a sequence. Instead, they are collections of individual items, where every item has the same significance as any other. As a cultural form, the database represents the world as a list of items and it refuses to order this list. In contrast, a narrative creates a cause-and-effect trajectory of seemingly unordered items. Therefore, database and narrative are natural enemies. Competing for the same territory of human culture, each claims an exclusive right to make meaning out of the world.

(Manovich 2005)

Manovich's remarks notwithstanding, the best electronic editions thrive on the combined strengths of database and narrative. William Horton has written that creators of digital resources may feel tempted to forego the difficult analysis that linear writing requires and throw the decision of what is important and what to know first onto the user. But users expect the writer to lead them through the jungle of information. They may not like being controlled or manipulated, but they do expect the writer to blaze trails for them … Users don't want to have to hack their way through hundreds of choices at every juncture.

(Horton 1994: 160)

The idea of including everything in an edition is suitable for writers of great significance. Yet such an approach has some negative consequences: it can be seen as damaging to a writer (and might be seen as counterproductive for readers and editors themselves). In fact, the decision to include everything is not as clear a goal or as "objective" an approach as one might think. Just what is "everything," after all? For example, are signatures given to autograph hunters part of what is everything in a truly inclusive edition of collected writings? Should marginalia be included? If so, does it comprise only words inscribed in margins or also underlinings and symbolic notations? Should address books, shopping and laundry lists be included? When dealing with a prolific writer such as W. B. Yeats, editors are likely to ask themselves if there are limits to what all should be construed to be. What separates wisdom and madness in a project that sets out to represent everything?

This section started by asking why people are making electronic editions, and to some extent the discussion has focused on the challenges of electronic editing. I would argue that these very challenges contribute to the attraction of working in this medium. At times there is a palpable excitement surrounding digital work stemming in part from the belief this is a rare moment in the history of scholarship. Fundamental aspects of literary editing are up for reconsideration: the scope of what can be undertaken, the extent and diversity of the audience, and the query potential that —through encoding — can be embedded in texts and enrich future interpretations. Electronic editing can be daunting — financially, technically, institutionally, and theoretically — but it is also a field of expansiveness and tremendous possibility.

Digital Libraries and Scholarly Editions

It is worthwhile to distinguish between the useful contributions made by large-scale digitization projects, sometimes referred to as digital library projects, and the more fine-grained, specialized work of an electronic scholarly edition. Each is valuable, though they have different procedures and purposes. The Wright American Fiction project can be taken as a representative digital library project (<http://www.letrs.indiana.edu/web/w/wright2/>). Such collections are typically vast. Wright American Fiction is a tremendous resource that makes freely available nearly
3,000 works of American fiction published from 1851 to 1875. This ambitious undertaking has been supported by a consortium of Big Ten schools plus the University of Chicago. As of September 2005 (the most recent update available) nearly half the texts were fully edited and encoded; the rest were unedited. "Fully edited" in this context means proofread and corrected (rather than remaining as texts created via optical character recognition (OCR) with a high number of errors). The "fully edited" texts also have SGML tagging that enables better navigation to chapter or other divisions within the fiction and links from the table of contents to these parts. Not surprisingly, Wright American Fiction lacks scholarly introductions, annotations, and collation of texts. Instead, the collection is made up of full-text presentations of the titles listed in Lyle Wright's American Fiction 1851–1875: A Contribution Toward a Bibliography (1957; rev. 1965). The texts are presented as page images and transcriptions based on microfilm originally produced by a commercial firm, Primary Source Media. It is telling that the selection of texts for Wright American Fiction was determined by a pre-existing commercial venture and was not based on finding the best texts available or by creating a fresh "critical" edition. The latter option was no doubt a practical impossibility for a project of this scale. In contrast to Wright American Fiction's acceptance of texts selected and reproduced by a third party, the editor of a scholarly edition would take the establishment of a suitable text to be a foundational undertaking. Moreover, as the MLA "Guidelines for Editors of Scholarly Editions" indicate, any edition that purports to be scholarly will provide annotations and other glosses. Wright American Fiction, in a nutshell, has taken a useful bibliography and made it more useful by adding the content of the titles Wright originally listed. One might think of the total record of texts as the great collective American novel for the period. Wright American Fiction has so extended and so enriched the original print bibliography that it has become a fundamentally new thing: the difference between a title and a full text is enormous. As a searchable collection of what approaches the entirety of American fiction for a period, Wright American Fiction has a new identity quite distinct from, even if it has its basis in, the original printed volume.

Wright American Fiction's handling of Harriet Jacobs (listed by her pseudonym Linda Brent) demonstrates some of the differences between the aims of a digital library project and a scholarly edition. The key objective of Wright American Fiction has been to build a searchable body of American fiction. Given that between the creation of the bibliography and the creation of the full-text electronic resource, an error in the original bibliography has been detected (a work listed as fiction has been determined not be fictional), it seems a mistaken policy to magnify the error by providing the full text of a non-fictional document. Incidents in the Life of a Slave Girl was listed as fiction by the original bibliographer, Lyle Wright — at a time when Jacobs's work was thought to be fictional — and it remains so designated in the online project even though the book is no longer catalogued or understood that way. This would seem to be a mechanical rather than a scholarly response to this particular problem. The collection of texts in the Wright project is valuable on its own terms, but it is different from an edition where scholarly judgment is paramount. Wright American Fiction consistently follows its own principles, though these principles are such that the project remains neutral on matters where a scholarly edition would be at pains to take a stand. Wright American Fiction is a major contribution to scholarship without being a scholarly edition per se.

In contrast to the digital library collection, we also see a new type of scholarly edition that is often a traditional edition and more, sometimes called a thematic research collection. Many thematic research collections — also often called archives — aim toward the ideal of being all-inclusive resources for the study of given topics. In an electronic environment, it is possible to provide the virtues of both a facsimile and a critical or documentary edition simultaneously. G. Thomas Tanselle calls "a newly keyboarded rendition (searchable for any word)" and "a facsimile that shows the original typography or handwriting, lineation, and layout ... the first requirement of an electronic edition" (1995: 58). A facsimile is especially important for those who believe that texts are not separable from artifacts, that texts are fundamentally linked to whatever conveys them in physical form. Those interested in bibliographic codes — how typeface, margins, ornamentation, and other aspects of the material object convey meaning — are well
served by electronic editions that present high-quality color facsimile images of documents. Of course, digital facsimiles cannot convey every physical aspect of the text — the smell, the texture, and the weight, for example.

An additional feature of electronic editions deserves emphasis — the possibility of incremental development and delivery. If adequate texts do not exist in print then it is often advisable to release work-in-progress. For example, if someone were to undertake an electronic edition of the complete letters of Harriet Beecher Stowe, it would be sensible to release that material as it is transcribed and annotated since no adequate edition of Stowe's letters has been published. It has become conventional to release electronic work before it reaches fully realized form, and for good reason. Even when a print edition exists, it can be useful to release electronic work-in-progress because of its searchability and other functionality. Of course, delivering work that is still in development raises interesting new questions. For example: when is an electronic edition stable? And when is it ready to be ingested into a library system?

Electronic editing projects often bring into heightened relief a difficulty confronted by generations of scholarly editors. As Ian Donaldson argues in "Collecting Ben Jonson," "The collected edition [a gathering of the totality of a writer's oeuvre, however conceived] is a phrase that borders upon oxymoron, hinting at a creative tension that lies at the very heart of editorial practice. For collecting and editing — gathering in and sorting out — are very different pursuits, that often lead in quite contrary directions" (Donaldson 2003: 19). An "electronic edition" is arguably a different thing than an archive of primary documents, even a "complete" collection of documents, and the activities of winnowing and collecting are quite different in their approaches to representation. Is a writer best represented by reproducing what may be most popular or regarded as most important? Should an editor try to capture all variants of a particular work, or even all variants of all works? As Amanda Gailey has argued, "Each of these editorial objectives carries risks. Editors whose project is selection threaten to oversimplify an author's corpus, [and to] neglect certain works while propagating overexposed ones... Conversely, editors who seek to collect an author's work, as Donaldson put it, risk 'unshap[ing] the familiar canon in disconcerting ways'" (2006: 3).

Unresolved Issues and Unrealized Potentials

Much of what is most exciting about digital scholarship is not yet realized but can be glimpsed in suggestive indicators of what the future may hold. We are in an experimental time, with software and hardware changing at dizzying speeds and the expectations for and the possibilities of our work not yet fully articulated. Despite uncertainty on many fronts, one thing is clear: it is of the utmost importance that electronic scholarly editions adhere to international standards. Projects that are idiosyncratic are almost certain to remain stand-alone efforts: they have effectively abandoned the possibility of interoperability. They cannot be meshed with other projects to become part of larger collections and so a significant amount of the research potential of electronic work is lost. Their creators face huge barriers if they find they want to integrate their work with intellectually similar materials. As Marilyn Deegan remarks, "Interoperability is difficult to achieve, and the digital library world has been grappling with it for some time. Editors should not strive to ensure the interoperability of their editions but make editorial and technical decisions that do not preclude the possibility of libraries creating the connections at a later date" (Deegan 2006: 362). Deegan perceives that tag sets such as the Text Encoding Initiative (TEI) and Encoded Archival Description (EAD) make possible interoperability, though they do not guarantee it. Figuring out how to pull projects together effectively will be a challenging but not impossible task. We face interesting questions: for example, can we aggregate texts in ways that establish the necessary degree of uniformity across projects without stripping out what is regarded as vital by individual projects? An additional difficulty is that the injunction to follow standards is not simple because the standards themselves are not always firmly established.

Many archivists refer to eXtensible Markup Language (XML) as the "acid-free paper of the digital age" because it is platform-independent and non-proprietary. Nearly all
current development in descriptive text markup is XML-based, and far more tools (many of them open-source) are available for handling XML data than were ever available for SGML. Future versions of TEI will be XML-only, and XML search engines are maturing quickly. XML and more particularly, the TEI implementation of XML, has become the de facto standard for serious humanities computing projects. XML allows editors to determine which aspects of a text are of interest to their project and to "tag" them, or label them with markup. For example, at the Whitman Archive, we tag structural aspects of the manuscripts, such as line breaks and stanza breaks, and also the revisions that Whitman made to the manuscripts, as when he deleted a word and replaced it with another. Our markup includes more information than would be evident to a casual reader. A stylesheet, written in Extensible Stylesheet Language Transformations (XSLT), processes our XML files into reader-friendly HTML that users see when they visit our site. A crucial benefit of XML is that it allows for flexibility, and the application of XSLT allows data entered once to be transformed in various ways for differing outputs. So while some of the information that we have embedded in our tagging is not evident in the HTML display, later, if we decide that the information is valuable to readers, we can revise our stylesheet to include it in the HTML display.

The combination of XML-encoded texts and XSLT stylesheets also enables projects to offer multiple views of a single digital file. This is of considerable importance in scholarly editing because editors (or users) may want to have the same content displayed in various ways at different times. For example, the interface of the expanded, digital edition of A Calendar of the Letters of Willa Cather, now being prepared by a team led by Andrew Jewell, will allow users to dynamically reorder and index over 2,000 letter descriptions according to different factors, such as chronology, addressee, or repository. Additionally, the Cather editorial team is considering giving users the option of "turning on" some of the editorial information that is, by default, suppressed. In digitizing new descriptions and those in Janis Stout's original print volume, such data as regularized names of people and titles often takes the form of a mini-annotation. In the digital edition, the default summary might read something like, "Going to see mother in California"; with the regularization visible, it might read, "Going to see mother [Cather, Mary Virginia (Jennie) Boak] in California."

These multiple views of the Calendar are enabled by the rich markup adopted by this project and point to some of the issues that editors must consider. Even while adhering to the same TEI standard used on many other digital editing projects, these editors' choices of which textual characteristics to mark up may nonetheless differ significantly from choices made by other project editors. When an effort is made to aggregate material from The Willa Cather Archive (<http://cather.unl.edu>) with that of other comparable sites, how will differences in markup be handled? Of course, unless projects make their source code available to interested scholars and expose their metadata for harvesting in accordance with Open Archives Initiative (OAI) protocols, it won't be easy now or in the future even to know what digital objects make up a given edition or how they were treated.

Clearly, then, issues of preservation and aggregation are now key for editors. In looking toward the future, the Whitman Archive is attempting to develop a model use of the Metadata Encoding and Transmission Standard (METS) for integrating metadata in digital thematic research collections. A METS document enables the Whitman Archive —and all entities that subsequently ingest the Whitman materials into larger collections—to describe administrative metadata, structural metadata, and descriptive metadata. For example, the Whitman Archive uses thousands of individual files in our project to display transcriptions and digital images of Whitman's manuscripts and published works. These files — TEI-encoded transcriptions, archival TIFF images, derived JPEG images, EAD finding guides — could be more formally united through the use of a METS document to record their specific relationships. The use of METS will help preserve the structural integrity of the Whitman Archive by recording essential relationships among the files. Additionally, we think that using METS files which adhere to a proper profile will promote accessibility and sustainability of the Archive and other projects like it, making them prime candidates for ingestion into library collections.

The proper integration of metadata is important because large sites are likely to employ
Redundant data is at least a workflow problem when it involves unnecessary labor. There is also the matter of figuring out what the canonical form of the metadata is. Currently, no METS Profile for digital thematic research collections has been developed, and there has not been a demonstration of the effectiveness of METS as an integration tool for such collections. In a report to the UK Joint Information Systems Committee, Richard Gartner argues that what prohibits "the higher education community [from] deriving its full benefit from METS is the lack of standardization of metadata content which is needed to complement the standardization of format provided by METS" (Gartner 2002). Specifically, he calls for new work to address this gap so that the full benefits of METS can be attained. In short, practices need to be normalized by user communities and expressed in detailed and precise profiles if the promise of METS as a method for building manageable, sustainable digital collections is to be realized.

Researchers and librarians are only just beginning to use the METS standard, so the time is right for an established humanities computing project, like the Whitman Archive, to develop a profile that properly addresses the complicated demands of scholar-enriched digital documents. In fact, a grant from the Institute for Museum and Library Services is allowing the Whitman Archive to pursue this goal, with vital assistance from some high-level consultants.²

Cost

A significant amount of scholarly material is now freely available on the web, and there is a strong movement for open access. As I have observed elsewhere, scholarly work may be free to the end user but it is not free to produce (Price 2001: 29). That disjunction is at the heart of some core difficulties in digital scholarship. If one undertakes a truly ambitious project, how can it be paid for? Will granting agencies provide editors the resources to make costly but freely delivered web-based resources?

The comments of Paul Israel, Director and General Editor of the Thomas A. Edison Papers, highlight the problem; his comments are all the more sobering when one recalls that the Edison edition recently was honored with the Eugene S. Ferguson Prize as an outstanding history of technology reference work:

> It is clear that for all but the most well funded projects online editions are prohibitively expensive to self publish. The Edison Papers has been unable to fund additional work on our existing online image edition. We have therefore focused on collaborating with others. We are working with the Rutgers Library to digitize a small microfilm edition we did of all extant early motion picture catalogs through 1908, which we are hoping will be up later this year. The library is doing this as part of a pilot project related to work on digital infrastructure. Such infrastructure rather than content has been the focus [of] funding agencies that initially provided funding for early digital editions like ours. We are also now working with the publisher of our book edition, Johns Hopkins University Press, to do an online edition of the book documents with links to our image edition. In both cases the other institution is both paying for and undertaking the bulk of the work on these electronic editions. (Israel 2006)

Over recent decades, grant support for the humanities in the US has declined in real dollars. By one estimate, "adjusted for inflation, total funding for NEH is still only about 60% of its level 10 years ago and 35% of its peak in 1979 of $386.5 million" (Craig 2006). If development hinges on grant support, what risks are there that the priorities of various agencies could skew emphases in intellectual work? Some have worried about a reinstitution of an old canon, and others have expressed concern about the possible dangers of political bias in federally awarded grants. Some worry that canonical writers may be more likely to be funded by federal money and that writers perceived to be controversial might be denied funding.

One possible strategy for responding to the various funding pressures is to build an endowment to support a scholarly edition. The National Endowment for the Humanities provides challenge grants that hold promise for some editorial undertakings. One dilemma, however, is that the NEH directs challenge grants toward support of institutions rather than projects — that is, toward ongoing enterprises rather than those
of limited duration. Thus far, some funding for editorial work has been allocated via "We the People" challenge grants from NEH, but this resource has limited applicability because only certain kinds of undertakings can maintain that they will have ongoing activities beyond the completion of the basic editorial work, and only a few can be plausibly described as treating material of foundational importance to US culture.3

Presses and Digital Centers

University presses and digital centers are other obvious places one might look for resources to support digital publication, and yet neither has shown itself to be fully equipped to meet current needs. Oxford University Press, Cambridge University Press, and the University of Michigan Press all expressed interest in publishing electronic editions in the 1990s, though that enthusiasm has since waned. On the whole, university presses have been slow to react effectively to the possibilities of electronic scholarly editions. University presses and digital centers have overlapping roles and interests, and there may well be opportunities for useful collaborations in the future, and some collaborations have already been effective. Notably, it was the Institute for Advanced Technology in the Humanities (IATH), under John Unsworth's leadership, that secured grant funding and internal support for an electronic imprint, now known as Rotunda, at the University of Virginia Press. It is too early to tell whether or not the Electronic Imprint at the University of Virginia will flourish — either way it is already an important experiment.

Digital Centers such as IATH, the Maryland Institute for Technology in the Humanities (MITH), and the Center for Digital Research in the Humanities at the University of Nebraska—Lincoln (CDRH) in a sense have been acting as publishers for some of their projects, though they are really research and development units. They ordinarily handle some publisher functions, but other well-established parts of the publishing system (advertising, peer review, cost-recovery) are not at the moment within their ordinary work. Nor are these units especially suited to contend with issues of long-term preservation. They can nurture projects with sustainability in mind, but the library community is more likely to have the people, expertise, and the institutional frameworks necessary for the vital task of long-term preservation. As indicated, many scholars promote open access, but not all of the scholarship we want to see developed can get adequate support from universities, foundations, and granting agencies. Presses might one day provide a revenue stream to support projects that cannot be developed without it. What we need are additional creative partnerships that build bridges between the scholars who produce content, the publishers who (sometimes) vet and distribute it, and the libraries who, we hope, will increasingly ingest and preserve it. A further challenge for editors of digital editions is that this description of roles suggests a more clearly marked division of responsibilities than actually exists. Traditional boundaries are blurring before our eyes as these groups — publishers, scholars, and librarians — increasingly take on overlapping functions. While this situation leaves much uncertainty, it also affords ample room for creativity, too, as we move across newly porous dividing lines.

Audience

Having the ability, potentially, of reaching a vast audience is one of the great appeals of online work. Most of us, when writing scholarly articles and books, know we are writing for a limited audience: scholars and advanced students who share our interests. Print runs for a book of literary criticism are now rarely more than 1,000 copies, if that. A good percentage of these books will end up in libraries, where fortunately a single volume can be read over time by numerous people. In contrast, an online resource for a prominent writer can enjoy a worldwide audience of significant size. For example, during the last two months of the past academic year (March — April 2006), The Willa Cather Archive averaged about 7,700 hits a day, or about 230,000 hits a month. In that period, there was an average of 7,639 unique visitors a month. The Journals of the Lewis and Clark Expedition (<http://lewisandclarkjournals.unl.edu>), which, interestingly, has numbers that correspond less to the academic year, had its peak period of the last year during January—March 2006. For those three months, Lewis and Clark averaged 188,000
hits a month (about 6,300 hits a day), and an average of 10,413 unique visitors a month.

In the past the fate of the monumental scholarly edition was clear: it would land on library shelves and, with rare exceptions, be purchased only by the most serious and devoted specialists. Now a free scholarly edition can be accessed by people all over the world with vastly different backgrounds and training. What assumptions about audience should the editors of such an edition make? This is a difficult question since the audience may range widely in age and sophistication and training, and the answer of how best to handle the complexity is unclear. A savvy interface designer on a particular project might figure out a way to provide levels of access and gradations of difficulty, but unless carefully handled, such an approach might seem condescending. Whatever the challenges are of meeting a dramatically expanded readership — and those challenges are considerable — we should also celebrate this opportunity to democratize learning. Anyone with a web browser has access, free of charge, to a great deal of material that was once hidden away in locked-up rare-book rooms. The social benefit of freely available electronic resources is enormous.

Possible Future Developments

Future scholarship will be less likely than now to draw on electronic archives to produce paper-based scholarly articles. It seems likely that scholars will increasingly make the rich forms of their data (rather than the HTML output) open to other scholars so that they can work out of or back into the digital edition, archive, or project. An example may clarify how this could work in practice. The Whitman Archive, in its tagging of printed texts, has relied primarily on non-controversial markup of structural features. Over the years, there have been many fine studies of Leaves of Grass as a "language experiment." When the next scholar with such an interest comes along, we could provide her with a copy of, say, "Song of Myself" or the 1855 Leaves to mark up for linguistic features. This could be a constituent part of her own free-standing scholarly work. Alternatively, she could offer the material back to the Archive as a specialized contribution. To take another example, in our own work we have avoided thematic tagging for many reasons, but we wouldn't stand in the way of scholars who wished to build upon our work to tag Whitman for, say, racial or sexual tropes.

Translation

In an intriguing article, "Scales of Aggregation: Prenational, Subnational, Transnational," Wai Chee Dimock points out that "humanistic fields are divided by nations: the contours of our knowledge are never the contours of humanity." She further notes that

nowhere is the adjective American more secure than when it is offered as American literature; nowhere is it more naturalized, more reflexively affirmed as inviolate. American literature is a self-evident field, as American physics and American biology are not. The disciplinary form of the humanities is "homeland defense" at its deepest and most unconscious.  

(Dimock 2006: 223)

We might be able to adjust our fields of inquiry if we, through editorial work on translations, highlighted the limitations of "humanistic fields ... divided by nations." For example, we could approach Whitman via a social text method — where versions produced by non-US societies are key objects of interest. The remaking of a writer to suit altered cultural contexts is a rich and largely untapped field of analysis. A truly expansive electronic edition — one more expansive than any yet realized — could make this possibility a reality.

Whitman said to one of his early German translators: "It has not been for my country alone — ambitious as the saying so may seem — that I have composed that work. It has been to practically start an internationality of poems. The final aim of the United States of America is the solidarity of the world One purpose of my chants is to cordially salute all foreign lands in America's name" (quoted in Folsom 2005: 110). As Ed Folsom has remarked, Whitman "enters most countries as both invader and immigrant, as the
confident, pushy, overwhelming representative of his nation ... and as the intimate, inviting, submissive, endlessly malleable immigrant, whose work gets absorbed and rewritten in some surprising ways" (Folsom 2005: 110–11). Scholarly editions, especially with the new possibilities in their electronic form, can trouble the nationally bounded vision so common to most of us.

In the near future, the Walt Whitman Archive will publish the first widely distributed Spanish language edition of Leaves of Grass. One of our goals is to make a crucial document for understanding Whitman's circulation in the Hispanophone world available, but an equally important goal is to make available an edition that will broaden the audience for the Archive, both in the US and abroad, to include a huge Spanish-speaking population. Foreign-language editions of even major literary texts are hard to come by online, and in many cases the physical originals are decaying and can barely be studied at all. Translations and other "responses" (artistic, literary) can be part of a digital scholarly edition if we take a sociological view of the text. In this social text view a translation is seen as a version. The translation by Álvaro Amendo Armando Vasseur that we are beginning with is fascinating as the work of a Uruguayan poet who translated Whitman not directly from English but via an Italian translation (apparently Luigi Gamberale's 1907 edition), and our text is based on the 1912 F. Semper issue. These details are suggestive of the development of a particular version of modernism in South America, and of the complex circulation of culture.

The translation project serves as an example of the expansibility of electronic editions. Electronic work allows an editor to consider adding more perspectives and materials to illuminate texts. These exciting prospects raise anew familiar issues: who pays for expansive and experimental work? Given that not all of the goals imaginable by a project can be achieved, what is the appropriate scope and what are the best goals? Scholars who create electronic editions are engaged in the practice of historical criticism. Editing texts is a way to preserve and study the past, to bring it forward into the present so that it remains a living part of our heritage. How we answer the difficult questions that face us will to a large extent determine the past we can possess and the future we can shape.

Notes

1 It could be argued that the Making of America creators didn't so much make as inherit this decision. That is, librarians and publishers made decisions about what was important to include when issues were bound together into volumes for posterity. One could imagine a project less given to mass digitization and more devoted to the state of the original material objects that would have searched more thoroughly to see how much of the now "lost" material is actually recoverable.

2 Daniel Pitti, Julia Flanders, and Terry Catapano.

3 The Whitman Archive is addressing the question of cost by building an endowment to support ongoing editorial work. In 2005 the University of Nebraska—Lincoln received a $500,000 "We the People" NEH challenge grant to support the building of a permanent endowment for the Walt Whitman Archive. The grant carries a 3 to 1 matching requirement, and thus we need to raise $1.5 million dollars in order to receive the NEH funds. The Whitman Archive is the first literary project to receive a "We the People" challenge grant. What this may mean for other projects is not yet clear. In a best-case scenario, the Whitman Archive may use its resources wisely, develop a rich and valuable site, and help create a demand for similar funding to support the work of comparable projects.

4 A Brazilian scholar, Maria Clara Bonetti Paro, recently found that she needed to travel to the Library of Congress in order to work with Portuguese translations of Leaves of Grass because the copies in the national library in Brazil were falling apart and even scholars had only limited access to them.

References and Further Reading


