The dank cellar of electronic texts[1]

Peter Shillingsburg
Department of English, Loyola University Chicago, USA

Correspondence:
Peter Shillingsburg, Department of English, Crown Center for the Humanities, Loyola University Chicago, 6525 N. Sheridan Road, Chicago, IL 60626, USA.
E-mail:
pshillingsburg@luc.edu

Abstract

The dank cellar surveys rather critically the litter of casualty electronic editions and the false bases and limited goals that informed so many early—that is, current—efforts; and it points hopefully to the best early, though still inadequate, efforts to provide electronic texts responsibly and with added scholarly value. It looks at some problems of representing Victorian fiction.

This essay was written in 1997. It has not been significantly updated but reflects that particular moment in the debate on electronic text.

The invitation to make this presentation reached me at the moment that I was reading 'RootCellar' by Theodore Roethke about a place, 'dank as a ditch',[2] where the remains of stored vegetables served as the foundation for a rich complex of new developments, not all of which seemed very attractive; the olfactory sensations, the 'congress of stinks', did, however, represent on-going life. It happened, also, that the last paper on electronic texts I was asked to present had focused on the remarkable expansion of available electronic texts on the internet about which, after a two or three week survey, I worked very rough estimates and guessed that about one-tenth of 1% of the available texts on the internet were reliable for scholarly work—99.9% of the texts were who knows what. The word 'cellar', I believe, means storehouse. And when it comes to root cellars, the word 'dank' is not necessarily pejorative. But there is something antiseptic in popular images of electronic texts, archived, as they seem to be, in a luminous box or cellar above a keyboard: they are dry, they resist handling except through some remote medium, one does not press the flesh of electronic texts, and therefore one does not leave on them an ever accumulating deposit of body oils and odours as one does on books in the library. So, 'dank' in reference to electronic texts brings an unfamiliar, organic, biological ambience into this antiseptic world, suggesting that electronic texts might breed, or grow, or develop molds or viruses of various sorts.

I do not know if this notion of a dank cellar is at all illuminating to anyone but myself; so perhaps I should drop it and get on to the ideas that it generated for me.

One thought is that we are being overwhelmed by texts of unknown provenance, with unknown corruptions, representing unidentified versions. These texts are the results of enthusiasms of various sorts concerning computers and the internet in particular. It is now easy to scan or type a text (preferably one that is out of copyright) and post it on the World Wide Web and become thereby a publisher. The reasons to do this may be egotistical, but for many, I am sure, the reason is altruistic: to make available for free the fruits of some labour and some technology where the alternatives might be to purchase a book or visit a library. The first great, obvious advantage of having computerized texts is that they are searchable, excerptable, and indexable without the use of 3 by 5 cards (is there anyone left who remembers 3 by 5 cards?); but these excellent qualities of electronic texts have tended to generate an enthusiasm that neglects concerns for textual accuracy and provenance. The rank glut of such texts on the internet produces a perfume of fecundity with a distinct under-scent of decay and disease. But I suggest we do not dwell on that phenomenon of electronic texts. In any case, this view merely replicates objections voiced in every generation since the fifteenth century—objections to the

Literary and Linguistic Computing, Vol. 24, No. 1, 2009.
doi:10.1093/llc/fqn031 Advance Access Published on 3 December 2008

unbridled proliferation of corrupt and corrupting texts.

A companion to that thought is that our students are therefore exposed to texts that are untethered from their origins, from their original dates, original publishers, original typefaces, and original page arrangements and weights. These are the residual marks of what I ask leave to call the 'eventness' of texts, the clues to the cultural contexts that informed the writing and reading of the works when they were fresh. And even if such matters are known by our students because they have access to other sources of information, the texts of most electronic editions have often been poorly prepared or improperly vetted, or not vetted at all; so that inadvertent readings lurk unmarked like verbal land mines to sabotage the pursuit of learning. Students eager for free texts and happy with the advantages of the computer medium are not naturally inclined to ask the necessary questions about provenance and accuracy. The unsophisticated replication of texts on the internet as well as the proliferation of relatively cheaply printed paper texts reflects the widespread assumption employed in, for example, the construction of classroom anthologies: that a literary text consists of words and punctuation that will mean the same thing wherever and, however, it appears. But communication theory and critical reflection suggest that each bibliographical event, like each verbal utterance, is significantly affected by its constituting context and medium. As students of texts, we care about provenance, contexts, histories, bibliography, and the accuracy of texts because all these affect how we read and how we understand the text. We must not neglect to teach our students these things anew every year. Of course this criticism of student texts, like the first criticism, above, is very familiar; for we have voiced them of cheap reprints for years.

A third thought is that contemplating the growing disaster of electronic pulp texts (if you will pardon the mixed metaphor)—contemplating these matters might generate a hysterical attempt on our part to police the internet or to set up a national or international text-vetting bureau or to publish in some official organ a list of approved electronic editions. After all, there is precedence for doing such things in the Modern Language Association's Center for Editions of American Authors and Committee on Scholarly Editions, about which I could say a few things. But such organizations in the past have not stemmed the tide of unreliable reprints. Nor will they, in my opinion, make a dent in the flow of bad or unsafe or merely undocumented electronic texts.

Instead of that desperate remedy, I suggest that the activity of scholarly editing in electronic media be conducted with firm attention to its own house, to its own root cellar. There is, after all, not much difference between the electronic world and the print world concerning the relative availability of corrupt or at least unvetted texts and reliable ones. Unreliable, unvetted reprints outnumber scholarly editions in the print world by proportions similar to the ones I am guessing at for electronic texts.

I turn, then, from the at-large world of electronic texts to the smaller one of electronic scholarly editions, where a good deal of thought and a developing body of practice presents both a more hopeful and yet problematic picture. I do not mean to be the historian of this movement, but I can use a thumbnail sketch of it as a background for my concerns today. Electronic texts date back twenty-five years and more, but the electronic book as an end in itself—that is, not as a temporary form to use, for example, in creating a concordance or for typesetting a book—is a phenomenon of the very late 1980s and 1990s. Its short history is remarkably varied. Its early manifestations now make us smile: ftp sites from which one could download an ASCII text file without italics or other formatting seemed adequate to many people who thought their personal libraries would be greatly increased at little or no cost. Then virtual worlds arrived and 'Saturday Night Live' satirized the virtual book with images of a book shown on a screen where keyboard commands turned codex pages, scarcely as good as a movie of a book with someone turning pages. The first of these electronic schemes reduced the rich complexity of the codex into a flat stream of ASCII characters; the second demonstrated a paucity of imagination trapped in the world of physical objects and photography.

In 1993 with the help of a number of people in modern literature, I drew up a description of what I

then thought would be an ideal electronic book, one that took advantage of developing software to provide textual experiences not available in codex forms and stepping gingerly beyond developing software to imagine other possibilities not yet available in prototype form. It was my first excursion into fantasy fiction. That description was widely circulated on the internet, presented at an MLA convention, and subsequently expanded and published in *The Literary Text in the Digital Age*.[3] I am happy to say that some of what we imagined then has since become possible and other desirable attributes of electronic texts that I did not imagine then have been developed. But other parts have not yet been developed.

The most ambitious development spawned by the demands of sophisticated electronic editions has been, I think the SGML/TEI markup system, which has attempted to provide a language and logical structures with which to build electronic texts that accommodate the rich complexity of language and the physical structures of books. Among the important practical projects implementing that markup language and doing in the short term what can be done to make the fruits of scholarly editing available in electronic form is the Model Editions Partnership conducted by David Chesnutt, Susan Hockey, and Michael Sperberg-McQueen. Other very important projects are the Rossetti archive being developed by Jerome McGann, the Chaucer project by Peter Robinson, the Emily Dickinson project by Martha Nell Smith, Ellen Louise Hart, and Marta Werner, two Ezra Pound projects by Patti Cockran and by Richard Taylor, the Women Writers Project at Brown University directed by Julia Flanders, the William Blake project by Joseph Viscomi and Morris Eaves, and the Piers Plowman archives by Hoyt Duggan.[4] There are others, but I mention these to call attention to serious, imaginative projects that make available research tools of genuine value, not only for the foundational scholarly work on the authors and texts involved but because they are pushing the envelope of electronic possibilities, asking more and more of the medium in an attempt to conceive and define what is meant or can be meant by the electronic edition, as opposed to a print edition.

But it was both thrilling and sobering to me recently to hear Jerome McGann and Peter Robinson focus their attention, not on the glitzy achievements of their projects, though both could have done so, and not on difficulties and long hours that have dogged their work, though they could have.[5] But what they noticed was the way in which their projects had forced them to think about texts in new ways, to ask new questions about textuality, and perhaps most interestingly, to ask new questions about what the archive was to do.

I found this aspect of their work especially interesting for two reasons. The first was nostalgic. In 1977–78 when I spent a year with a programmer developing the nine programmes known as CASE (Computer Assisted Scholarly Editing)—programmes developed to aid the textual research of transcription, collation, and textual apparatuses and to prepare input for typesetting machines—the main thing I discovered was that my rationales for scholarly editing were not carefully thought out. I discovered that, whereas in science if one is careless one's lab explodes, in the humanities if one makes a mistake it generally goes unnoticed until one has to develop computer programmes that will replicate or organize the work. Then one discovers the meaning of the word 'discipline'.

I heard in McGann's presentation that the difficulties of accommodating the Rossetti material to the SGML markup language was stretching his understanding of both the structural complexity of the works he was handling (that is, he was learning to look at the paintings in new ways) and the structural gaps in SGML. (That is a point, incidentally, that is debated by the developers and other users of SGML/TEI.) One great value of modern work on electronic archives lies in its process, in what the work teaches the worker, rather than in the accomplishment of the product. I will come back to that thought later.

In Peter Robinson's presentation, as he focused on what did not please him about his very impressive *Wife of Bath* prototype for the electronic Chaucer, and as he outlined what he will do differently at the next stage, there was a similar emphasis on how doing the work and struggling with the inadequacies of the software led to new

understandings of the significance of Chaucer's work and of the responsibilities of the editor. Let me make that idea less abstract. What I understood Robinson to say was that the neutral, objective, CD presentation and description of eighty-eight Chaucer manuscripts with approximately ten million hyperlinks (most of which were generated automatically by his software) ran the high risk of burying the user. His proposed solution is to provide the reader with explanations of the significance of the parts, to provide frankly interpretive signposts, to express editorial opinions about the value and significance of the materials presented.

That is an important development for the debate on editorial theory and practice. It brings me to the second reason I was pleased to hear of developments in McGann's and Robinson's projects. It used to be, thirty years ago, that an editor imposed his or her understanding of the textual situation onto the edition, edited it boldly, and expected appreciation from readers for having provided, as a finished product, an edition on which readers could rely without having to investigate. The justifications for such editorial actions stressed the objectivity of the work and the honesty of the apparatus. But self-deception in scholarship never lasts for long, and idealist, intentionalist editing was recognized for the subjective, dominating, reductive activity that it was. In its place more recently arose the apparently 'truly' objective pursuit of the electronic archive with objective, high-resolution digitized reproductions of original texts, objective and antiseptic parallel texts, and hyperlinks to related materials. And now Peter Robinson—who to do him justice never said that what he was doing was objective, but who tried to minimize the presence of the editor in the *Wife of Bath* CD-ROM edition—now he says that such a neutral approach in a complex edition might bury the reader who needs the editor to exert a guiding hand. Amen. We have once again discovered that scholarly editing is not best served by editors whose main goal is to efface their presence from the project but rather by those who confidently and boldly assert their presence, demanding that readers recognize the scholarly edition as a contribution to criticism. Editing is by nature and by definition interference; it cannot be done objectively. It should be undertaken boldly, and should be reported straightforwardly. There is room for alternative editions.

In the last few years scholarly editors at conferences on electronic editions have been focusing their attention, naturally enough, on the ways to produce electronic editions more than on the implications of those ways to the study of texts or to the pursuit of academic interests that such a shift in medium entails. I believe this conference to be a good corrective to the general trend. It is only natural that the practical considerations of electronic text-production should absorb the attention of electronic text producers, for there is so much to deal with: choices of hardware (where every new year's model is capable of so much more than the model you bought) and choices of software (where the cleverest and niftiest capabilities entail, as likely as not, proprietary coding that locks the edition into one form of access). To these one can add the many problems entailed in developing or learning a text-encoding system that will overcome the problems of ever developing and evolving hardware and software. And it has been necessary for scholarly editors also to devote a great deal of attention to design: what shall the screen page look like? How shall hypertext links look and work? How should windows pop up and disappear or fade into the background? How shall we incorporate sound and motion in our texts? Where are the boundaries between scholarly editing, archiving, and pedagogy?

These important concerns are likely to continue to absorb a great deal of attention, and we must not be distracted from them by temptations to scream out against the rapid proliferation of unreliable texts prepared (I speak loosely, of course), prepared by people who have not bothered to think as long or as hard as scholarly editors do about textual matters.

But I have already said all that I know about hardware and software and book design for electronic texts in *Scholarly Editing in the Computer Age* and in my chapter of Richard Finneran's *The Literary Text in the Digital Age*. And I will be contented in reminding myself of a fundamental fact about scholarly editing that I fear has become a hobby-horse for me.

It is that no editorial task, whether in print or electronic medium is merely the reproduction of a text. Stated positively, all editorial tasks entail more than the mere transfer of words and punctuation or images from one form to another. Every such transfer creates something new and so radically different from its original that scholarly editors for centuries have felt the need to introduce their editions to explain what has happened. I leave aside here introductions that claim the new form is better or more accurate or more real or more useful; I call attention only to the need for an introduction that explains the differences between original and newly edited texts.

We need to remember or reinvent the reasoning which generates the need to introduce new editions. It has to do with our understanding of how texts operate—or to be more accurate, of how humans operate with texts. Texts are, after all, merely arrangements of symbols presented in certain formats. Humans undergo considerable training in interpreting the symbols, and frequent use teaches them much about the implications of the formats in which they encounter the symbols. Meanings are generated by readers who have learned to deal with symbols and formats. Change the symbol and the meaning changes; change the format and the implications are changed; change the contexts of interactions with texts and the importance and significance of the text changes.

Two summers ago I had these truths confront me in a physical way. I visited the extensive Victorian fiction collections known as the Metzdorf, the Parrish, and Taylor collections at the Firestone Library in Princeton and, two weeks later the even more extensive Wolff collection at the Harry Ransom Humanities Center at the University of Texas. I was working on a paper concerning the appearance, the formats, of Victorian fiction and was fortunate to be introduced to the stacks of the Parrish and Wolff collections and to have multiple copies of single titles brought to my work table. The sensations, the look, the feel, the odours, the textures, the colours, shapes, and sizes of the books all have a bearing on what follows, though the most overwhelming sensation I felt as I looked at 7000 Victorian novels in Texas was that of ignorance. I know next to nothing about Victorian fiction, for that large room of books is merely representative of the 40,000 or 50,000 novels published between 1830 and 1900. And that vast array of books stood in stark isolation from the books that they jostled in the days when they were fresh off the press. Where were the books of agriculture, history, geography, exploration, navigation, mining, economics, art, music, biography, science, religion, cookery, and domestic economy? How could I come to know the meaning or importance or significance of any one of these books without knowing what other books were published by its publishers or what other books were reviewed together with it? Must I go to Austin, Texas, and get permission between the hours of ten and four to see no more than five books at a time to get a sense of what Victorian fiction was? I tasted and smelled and felt enough of original books in the context of their peers to know that it makes a huge difference in our experience of texts—a difference that does not and will not manifest itself in paperback reprints, in scholarly print editions, or in electronic books as currently conceived.

Nor can the interests I developed in the subject be easily satisfied by existing bibliographies or descriptions of collections, though the Wolff bibliography, Michael Sadleir's *Nineteenth Century Fiction*, and descriptions of the Rauri MacLean collection in Toronto are illuminating, impressive, and mind bending once one has developed a hunger to know. Nor can existing histories of writing, publishing, and reading, or even a few visits to major collections of Victorian fiction provide more than appetizers, experienced at too long intervals. Additional descriptions of the 'faces of Victorian fiction'[6] of the whimsical sort I wrote will not do much to provide students or colleagues with the materials needed to develop a sensitivity to what Victorian fiction looked like or what its appearances meant to contemporary readers.

Further reflection on my visits to the Metzdorf, Parrish, Taylor, and Wolff collections and perusal of descriptions of Sadleir's books and of MacLean's collection led me to the conclusion that these rich resources have two very serious flaws: the first is that they are by necessity sequestered from most

students and especially from me, an interested party who lives near to none of the honey; the second is that each collection is, by virtue of its selection principles, a distortion of the history of Victorian fiction. What bookstore in Victorian England shelved only first editions of novels? How can such collections represent the 'eventness' of Victorian fiction?

Robert Lee Wolff's collection of Mary Elizabeth Braddon's works, for example, included many reprints of her best-selling *Lady Audley's Secret* and it managed to fail to include a first printing of the first edition. But this is an unusual condition for special collections. For most of his authors, Wolff's collection holds only the first edition. His collection fails to include all the reprints, even of *Lady Audley's Secret* and therefore his bibliography of Victorian Fiction, which is really a catalogue of the books he owned, leaves out much of the printing history of Braddon's book. This is not a criticism of Wolff's magnificent work. It is a description of the state of affairs that developed at a time when collectors' interests were driven by concerns different to those that drive my conception of this project.

Gordon S. Haight's description of the editions of George Eliot's *The Mill on the Floss*, for a different kind of example, provides a comprehensive record of 'authorized' productions of that work from 1860 to 1888[?]. It provides a better visualization and understanding of the production objects and texts of *The Mill* than can be found in any one library in the world. But even such a listing and description and analysis are bounded by the compiler's interest—in this case, in the 'composition and revision life' of the work. Haight's account does not include the evidence of the very active continued 'reading life' of the work represented by the multiple unauthorized editions in America and Germany nor the multiple unauthorized editions everywhere that proliferated after the author's death.

Of course for most Victorian fiction there was only one edition ever. But there is no bibliography of Victorian fiction that is comprehensive as a record of Victorian consumption of native fiction. Obviously, the boundaries of this project are artificial and could be extended.

My long-standing commitment to editing the works of W. M. Thackeray and my newly discovered interest in the 'Faces of Victorian Fiction' led me to imagine a resource that is the kernel of an impossible dream proposal. It is to pursue systematically the establishment of an electronic archive of Victorian Fiction, beginning with images of original editions in contemporary bindings, and then including digitized texts, and then expanding to include textual, production, and other historical annotations, and texts newly edited by bold editors. I imagined an opening page, not unlike a storage root cellar or library door into which the user entered a virtual bibliography: a shelving area that appears to hold lifelike books arranged—arranged at the touch of a button chronologically (so I could see all the novels published in 1859), or alphabetically by author, or grouped according to publisher, each publisher's output arranged in the order published or by author, or perhaps grouped according to original publisher's price, or by format (all two volume books sold at twelve shillings a volume separated from all the three volume novels sold for 36 shillings, for example). In my dream I saw a merging of Parrish, Taylor, Wolff, Metzdorf, Sadleir, and MacLean, with additions from the Bodleian, and the British Library. And from my keyboard I could pull any book from the shelf and read, and search, and see parallel texts, and read historical and textual annotations, and have textual cruxes in verbose or background mode, and hyperlink to—to you name it.

And yet that is not enough to satisfy my craving. Special collections in libraries distort our sense of the marketing and textual history of Victorian Fiction. Every book ever published had a first edition, but the most widely read books were produced in multiple editions. Special collections tend to emphasize two qualities that belie the historical record: first editions and fine condition. The result is that the evidence of reading is not preserved. Books that are much reprinted are seldom preserved in those reprinted forms. Copies that are much read, are seldom preserved because of the used condition. So, my dank cellar of electronic Victorian fiction must include cheap reprints and shaken and mauled editions as well. And before waking from this nightmarish vision of delight I should note that this proposal does not pretend to incorporate

the contexts of Victorian fiction or literature on a broad scale as does George Landow's Victorian Web. I imagine there being a variety of links between the Victorian Web and the 'Victorian Fiction' archive as well as other relevant web sites.

But I come back from that vision of an end product to a point I promised to return to: to the research involved in constructing such a library, to the bibliographical investigation, to the careful observation of the text in its various physical forms each representing textual acts at points in time. There is no bibliography we can merely scan into a data base; there is no collection of texts we merely need to photograph. What this grandiose scheme requires is what every electronic scholarly edition requires in a smaller way: basic textual scholarship, bibliography, textual criticism, critical acumen in the writing of introductions and explanatory notes, sensitivity to voices in the text, and the courage to express the best thinking the scholar has about the significance of the evidence at hand.

To conclude, root cellars are very necessary units and they house the sustenance of life. But they need constant attention, annual cleaning, restocking, and shelf repair. Scholarly editing is not ever finished; and it is done well or ill and *re*done well or ill, and it is constantly being confused by lay persons, students, and even members of our profession with reprintings of every sort. Our task, I think, is to keep our cellar restocked, properly labelled, properly preserved, and as free of mold, viruses, and tainted fruit as careful scholarship can provide. I do not mean we should keep out the cheap reprints; I mean only that no text should pretend to be other than what it is. If it is a cheap, shoddily prepared text, let it present itself as such—at least in our root cellar—and not pretend to represent 'the work itself'. We cannot rid the world of mold and debris, but we can refrain from creating it, and we can keep it out of our cellar by circulating our stock, by providing accurate descriptions of text and provenance, by adding to old scholarly editions new ones, boldly and straightforwardly edited, not intended to last for ever but intended to feed the minds of the next generation for a few years until they see how best to feed themselves and their progeny.

Notes

1 This text is published in a slightly revised version as Chapter 6 of *From Gutenberg to Google: Electronic Representations of Literary Texts* (Cambridge: Cambridge University Press, 2006), pp. 138–50.

2 **Roethke, T.** (1966). *Collected Poems of Theodore Roethke*. New York: Doubleday, p. 38.

3 **Richard, J. F. (ed.).** (1996). Principles for electronic archives, scholarly editions, and tutorials in *The Literary Text in the Digital Age*. Ann Arbor: U Michigan P, pp. 23–35.

4 Model Editions Partnership, http://adh.sc.edu/, (last accessed 6 November 2008); Rossetti Archive, http://www.rossettiarchive.org/, (last accessed 6 November 2008); Chaucer, http://www.canterburytalesproject.org/index.html, last visited 6 November; Dickinson, http://www.emilydickinson.org/index.html, (last accessed 6 November 2008); Women Writers Project, http://www.wwp.brown.edu/, (last accessed 6 November 2008); Blake Archive, http://www.blakearchive.org/blake/main.html, (last accessed 6 November 2008); Piers Plowman, http://www.iath.virginia.edu/seenet/piers/, (last accessed 6 November 2008)

5 McGann's and Robinson's remarks were made in the 'Voice, Text and Hypertext at the Millennium Conference at the University of Washington', 29 October 1997 to 1 November 1997.

6 **Bornstein G. and Tinkle T.** (1998). The Faces of Victorian Fiction. *The Iconic Page*. Ann Arbor: U Michigan P, pp. 141–56.