Commissioned Commentary

Current issues in making digital editions of medieval texts—or, do electronic scholarly editions have a future?

Author:  

Abstract

It has been more than ten years since the first digital editions began to see the light of day. This article examines the current state of and future possibilities for the digital critical edition. Despite great promise, the article argues, digital editions have not been as successful with the general scholarly community as was expected by early digital theorists. The author attributes this failure to two main problems: a lack of easy-to-use tools and a lack of support from major publishing houses. The result is that it currently remains far easier to make a print than electronic edition. This situation will not improve until the tools and distribution of electronic projects is such that any scholar with the disciplinary skills to make an edition in print can be assured he or she will have access to the tools and distribution necessary to make it in the electronic medium.

Keywords: critical editions, electronic publication, XML, tools, digital scholarly publishing, digital editions

How to Cite: Robinson, P., 2005. Current issues in making digital editions of medieval texts—or, do electronic scholarly editions have a future?. Digital Medievalist, 1. DOI: http://doi.org/10.16995/dm.8

The digital edition ten years on

§ 1 Scholars working in our area—broadly, texts from medieval western Europe—now have around a decade of experience of making digital editions. In 1994, Hoyt N. Duggan posted the prospectus for the Piers Plowman Electronic Archive to the IATH webserver (Duggan 1994/2003). SEENET, the Society for Early English and Norse Electronic Texts of which the archive is a major part, first appeared on the web about the same time and has published four CD-ROMs since (http://jefferson.village.virginia.edu/seenet/home.html). At the same time, Kevin Kiernan's work with the Beowulf manuscript was bearing its first electronic fruit in Kiernan 1994 and, ultimately, Kiernan 1999/2003. My own Canterbury Tales Project dates to the same time: we published the first volume of our Occasional Papers (effectively a prospectus for the whole project) in 1993, our first CD-ROM in 1996 (Robinson 1996), and five more discs in subsequent years (Solopova 2000; Stubs 2001; Bordalejo 2003; Lloyd Morgan 2003; Robinson 2004a).

§ 2 Ten years and the experience of so many is a reasonable platform for discussions about what has been (and might have been) achieved. I have recently written three articles that include various retrospective and prospective elements. The first, "The history, discoveries and aims of the Canterbury Tales project" (Robinson 2003), deals specifically with the work of that project over the last decade. The second, "The Canterbury Tales and other medieval texts" (Robinson
§ 3 These articles deal with specific questions which I do not need, therefore, to rehearse in this article. Robinson Forthcoming, for example, recommends strongly that electronic editing projects should make full electronic transcripts of manuscript sources, that these transcripts should be made according to explicit principles, and that the transcripts should be made available under an "open transcription" policy. Robinson 2004b addresses a wider issue, describing a model of collaborative scholarship and publication leading toward what I describe as "fluid, co-operative and distributed" editions. One could say that Robinson Forthcoming focusses on narrower issues of project organization, while Robinson 2004b argues toward a considerable, and very challenging, desideratum: a means by which scholarly work may be dynamically corrected, revised and augmented (a kind of scholarly Wikipedia, if you will).

§ 4 There are two issues which these articles do not face. Firstly, and particularly in the case of Robinson 2004b and Robinson Forthcoming, they assume as a given that the case for electronic editions does not need to be made: that such editions are of such self-evident superiority that one need not argue the point. Secondly, they have very little to say about the tools available to scholars for making electronic editions. Rather, there is an unspoken assumption that the benefits of the electronic medium are so great that one might bear unreasonable inconvenience in order to make editions for it. In light of the first assumption, indeed, one might even argue—wrongly, however—that the issue of tools hardly matters: one uses what one has to use to gain the great advantages of the digital medium, and that is that.

§ 5 In fact, I think these assumptions involve the two most important issues that those of us concerned with making digital editions of medieval texts must face. After more than a decade of experience, we are at a crossroads. Rather few of us have made rather few digital editions. Are we the pioneers, then, who are making a trail that now many more will follow? Or are we brave experiementers who have produced some very interesting results but at a cost so high that our work remains something that few will repeat? To put it another way: in the future, will the great majority of scholarly editions in our area be conceived, executed, and published in digital form, with print being restricted to a few specialized instances? Or will things remain instead much as they are now, with most editions being prepared for print publication, and digital editions restricted to a few specialized instances? As a long-time maker of digital editions, obviously I see myself as a pioneer, not an experiemnter. By definition, I think that most of the Digital Medievalist community also will prefer to see the first rather than the second vision of the future come to pass. But we will not achieve this unless we resolve these issues of perception and tools. The whole scholarly community needs to be persuaded that digital editions are indeed superior to print; and it needs to have access to tools so that any scholar who can make a print edition can make a digital edition instead.

The continuing dominance of print editions

§ 6 We should begin by being realistic about the degree to which print editions still predominate. Print editions, some of them quite major, are still being conceived and made. Existing major print editions commonly have been slow to convert to electronic form. We have become rather too familiar with what one might call the token electronic edition: the edition where the real meat is on display in the print version, and the electronic version consisting of flat, plain-text files, distributed on the internet or in CD–ROM. This is not because scholars and publishers do not know about digital editions. Consider the example of the Early English Text Society (EETS), perhaps the single most important body of editions for medievalists working with English materials. Over the last decade, it is known that the EETS board has discussed several times whether and how the EETS corpus should "go digital". The society has collaborated with several initiatives to put substantial segments of the EETS into electronic form—as part, for example, of the Middle English Compendium (MEC) (http://ets.umdl.umich.edu/m/mec/). There has also been an interesting experiment, led by Bella Millet of Southampton University, in preparing a specimen digital edition of part of the Ancrene Wisse with a view to creating a model of practice that might be used more widely by EETS editors (Millet
However, the EETS board has so far (to my knowledge) refrained from an absolute declaration that all future EETS editions will be prepared and published in digital form. And frankly, it would be irresponsible for the board, or any similar agency, to issue any such edict at present. Before any such declaration could be made by a major editorial group, we would need to satisfactorily address the two issues on which this article focusses: we would need to establish an overwhelming agreement within the community that digital editions are indeed the way to go; and we would need to have tools available so that any editor who had the skills to make a print edition could make a digital one instead.

§ 7 One can see the EETS experience repeated elsewhere. The Model Editions Partnership (MEP) was initiated some ten years ago with the aim of achieving for American documentary editors much what Millet’s EETS-sponsored project had in mind for early English: the creation of a model which could then be generalized to all American documentary editions (such as the large National Endowment for the Humanities (NEH)-funded editions of the papers of Franklin, Adams, and Twain). Despite considerable effort from an exceptionally well-qualified project team and some impressive prototype results (see http://adh.sc.edu/[1]; also Chesnutt 1995), the MEP does not appear to have shifted the paradigm of American documentary editions from print to digital. Crossing the Atlantic to the UK, the situation is if anything, worse. The two major academic publishers, Cambridge University Press (CUP) and Oxford University Press (OUP), both in the last decade actually have stopped publishing scholarly editions in digital form. This is the more remarkable given that, in the first half of the 1990s, both publishers made considerable investments in electronic publication of scholarly editions. OUP was first, with the massive project that eventually published some 20,000 pages of Wittgenstein's Nachlaß in digital facsimile and transcripts. But even before this was published, the press had decided to pull back from digital publication of scholarly editions. In 1993, I was looking for a publisher who might be able to publish my edition of the Wife of Bath's Prologue in electronic form. Informal soundings at OUP led to a polite but firm "no": the press was already concerned about the viability of such digital editions and was not willing to commit itself to any more. At around the same time, a group from Birmingham University approached the press with a proposal for a digital edition of Johnson’s Dictionary and received the same response. It happened that exactly at this time, CUP became interested in the possibility of scholarly editions in digital form. Discussions between myself and key staff at the press followed and bore fruit in the first two electronic publications of the Canterbury Tales Project (Robinson 1996 and Solopova 2000) and the CD-ROM of the Birmingham project’s edition of Johnson’s Dictionary (McDermott 1996). At one point, around late 1994, we all became rather excited about the prospects for digital editions (then called electronic editions): proposals were flowing in for electronic editions from so many scholars, that CUP even issued a prospectus for a Cambridge Electronic Editions series. The excitement soon faded, however, as CUP discovered what OUP had already learnt: that electronic editions cost no less than print editions to produce and require staff to be educated in the new possibilities. One can see this new scepticism in an article I wrote with Kevin Taylor of CUP on our experiences in publishing the The Wife of Bath's Prologue on CD-ROM (Robinson and Taylor 1998). By 1999, CUP too had had enough and informed me that our second planned CD-ROM in the series, Elizabeth Solopova’s edition of the The General Prologue, was the last they could commit to publishing. In need of an alternative publication venue, and having discovered that the University of Michigan Press had also deserted the electronic publication field, I set up a new electronic publishing house, Scholarly Digital Editions (SDE). We have been moderately successful in attracting other scholars interested in making digital editions, but hardly so much as to raise visions of a mass changeover from print to digital. Nor is it that editors have stopped editing or publishers publishing their work. While neither OUP nor CUP were willing to publish editions exclusively on CD-ROM, both presses have committed to new, large, and primarily print-based editorial projects such as the new Cambridge Ben Johnson and the new Oxford Jonathan Swift.

Have we gone wrong?

§ 8 In academic life (perhaps unlike in other spheres of human activity) to lead is to follow slightly in advance of
where others are already going. Scholars are not now heading in the direction of electronic editions and pronouncements from on high will not force them to go there. We have to conclude that many scholars are not persuaded of the advantages of digital editions—or at least that they are still sufficiently satisfied with print editions as to be happy to continue to make and use them. About a decade ago, there was a small spate of articles hymning the glories of electronic editions as we then imagined them (e.g. McGann 1995/1997 and articles by myself and others in Sutherland 1997, Landow and Delaney 1993, and Finneran 1996). Were we wrong then in our belief that the electronic scholarly editions to come would indeed be massively superior to anything that a print edition could ever be? We now have, as we did not then, several instances of full electronic editions. We no longer need to speak speculatively of what digital editions might be. It is now possible to compare what actually has been done with what we then thought might be possible. Is it that the actuality has fallen so far short of the hope?

What has gone right
§ 9 I do not think this is the case. We have seen several electronic editions which indeed do all that was imagined of them—by some of us anyway—ten years ago. We presumed then that we would find editions which presented the full text of important original sources alongside high-quality digital images: the Rosetti Archive (http://www.rossettiarchive.org/), SEUNET, the Canterbury Tales Project, and the Electronic Beowulf (among others) have achieved that. We presumed then that text could be presented in many forms, representing different levels of editorial intervention and of readerly interest: again, the projects published over the last decade have offered just this. We presumed then that we would be able to offer a multiplicity of new ways to present variation among texts, freeing readers from arid collations: the Canterbury Tales Project publications offer line by line comparisons between manuscripts, highlighting their differences, and much else. We presumed then that we would be able to offer readers a wealth of new tools to permit dramatic visualizations of relations among our texts and assist in their exploration: you can indeed find these now in, for example, the Blake archive (http://www.blakearchive.org/), Estelle Stubbs' edition of the Hengwrt digital facsimile (Stubbs 2001), Martin Foys' Bayeux Tapestry Digital Edition (Foys 2003), and my own edition of the Miller's Tale on CD-ROM (Robinson 2004a)—not to mention perhaps the most elaborate and ambitious of all current electronic edition projects, the 28th Nestle-Aland edition of the Greek New Testament (http://nestlealand.uni-muenster.de/). We presumed then that these texts would be searchable, in many different ways—and indeed they are. In one important respect, moreover, digital editions actually have greatly exceeded our original hopes: the extraordinary progress in computing hardware and software over the last ten years has made it possible to give editions interfaces of remarkable beauty and flexibility.

§ 10 From these examples (and there are several others) I think it is fair to conclude that digital editions can and do fulfill the hopes we had of them. True, print editions have not stood still over the period: computer-based typesetting and new production methods make it possible to create beautiful new print editions, such as Jim Mays' Coleridge (e.g. Mays 2001), or the Cornell Wordsworth (e.g. Parrish 1977). But I believe that the digital editions we have made over the last ten years now far exceed their print equivalents in what they include, in their ability to shape themselves to the reader's needs, in the routes they offer to understanding highly complex webs of knowledge. There is good evidence, too, that others share this view: in recent years, major academic awards have been won by digital editions. In 1998, the Beatrice White prize for an outstanding publication in Medieval and Renaissance studies was awarded to my edition of the Wife of Bath's Prologue (Robinson 1996), the first time the award was given to a non-print publication. In 2001, Kevin Kiernan's Beowulf (Kiernan 1999/2003) repeated the feat.

What has gone wrong
§ 11 However, there is one crucial respect in which digital editions have not achieved our expectations. Some ten years ago, there were at least three major publishers offering to publish them: OUP, CUP, and Michigan. As I have explained above, all three have now stepped out of the field—a loss that is not made up for by the creation of SDE, operating out of my spare bedroom. We thought then that we had a sound publication model for digital editions: major publishers would publish them, just as they have always done for print editions. But this has not happened. Further, we now
know anecdotally that many scholars remain sceptical of electronic publication. Combined with the movement by leading academic publishers away from this field, this scepticism leads rather easily to the opinion that electronic publication is not "real" publication at all. It is depressing to find cases where scholars do not use the digital editions one has gone to such trouble to make, even when they know of and have access to them. To give just one example: my edition of the Wife of Bath's Prologue and several later Canterbury Tales Project publications include Dan Mosser's descriptions of the Canterbury Tales manuscripts. These descriptions are the result of several decades of work by Professor Mosser, in the course of which he has inspected every manuscript and every complete incunable copy (and very many fragments too); consulted with every leading scholar; read every article of note; and built up a formidable expertise in palaeography, codicology, and watermarks. By all odds, these are not just the most recent, but also the most careful and comprehensive accounts ever made of the Tales' earliest texts. Despite this, I have come across several examples of work, even by senior scholars with access to Professor Mosser's research, where these essential resources have not been cited.

**The current state of affairs**

§ 12 Our discussion of the first of the two issues—the superiority of digital editions over print editions—thus leads us to two contradictory positions. Firstly, it is rather clear that well-made digital editions are better than print editions from the perspective of their users. I think the digital editions named above show this comprehensively and I am not aware of any study that attempts to argue the opposite. Indeed, of the many kinds of print objects produced over the last centuries, it is difficult to think of any genre that is so well adapted to the computer as the scholarly edition. The layers of footnotes, the multiplicity of textual views, the opportunities for dramatic visualization interweaving the many with each other and offering different modes of viewing the one within the many—all this proclaims "I am a hypertext: invent a dynamic device to show me". The computer is exactly this dynamic device. There may be reasons to doubt the superiority of digital forms of other types of print objects such as books and newspapers. But scholarly editions are something else again: if mass copying of the certain texts in the late days of the manuscript age prefigured the print age to come, then the elaborate print editions of the last decades have prefigured the digital age.

**Tools**

§ 13 Which leads us to the paradox: if digital editions are so manifestly superior, then why indeed are we in the state of affairs described above? Why are so many scholars, and so many scholarly projects, still making print editions? I suggest that the answer to this lies almost wholly in the second aspect to this problem: the availability of usable tools. Over the past two decades I have made two of the leading tools for making scholarly electronic editions. The first is the collation software Collate, which I first wrote as a set of VAX routines in the 1980s, and re-wrote into a Macintosh program in the 1990s. The second is the XML publication software Anastasia, which I initiated in the mid-1990s. Several of the electronic editions named above depend heavily on these two tools. One can assert that it is indeed possible to use them to make digital editions which offer all we could hope for. But as their creator I think I am uniquely qualified to note that they are not easy to use: if everyone who wanted to make digital editions was required to use these two tools, very few digital editions would ever be made. Both tools require very high levels of dedication (though not particularly advanced computer skills). They can be frustrating in the extreme: perhaps refusing to work at all, or (just as bad) working in quite unexpected ways for reasons which are far from obvious. One can spend hours—days even—pursuing some slight problem with the programs' output. Typically, an editor using these tools will need large amounts of time, dogged determination, disciplined organization, and considerable help. It is not quite true that I am the only person in the world who can use Collate and Anastasia (it would be an interesting exercise to trace all the humanities software that has been used only by its creators)—but it is uncomfortably close to the truth. The wonder is not so much that rather few projects have managed to go the whole distance with these tools, as it is that any have been able to do so at all.

§ 14 There are other tools about. The success of XML in establishing itself throughout the business and academic
communities as the universal encoding system has brought with it many new tools for preparing and publishing texts. Furthermore, for more than a decade now, the Text Encoding Initiative (TEI) has been preparing text-encoding guidelines that have achieved wide acceptance in the humanities computing community. Indeed, the TEI encodings are fundamental to several of the editions named above, editions which simply could not have been made without the TEI. Once more, I believe I am in a unique position to pronounce on these guidelines, and hence on tools based on the TEI scheme in terms of their appropriateness to the wide range of actual scholars who want to make actual editions. I was chair of the working group that was responsible for the chapter on textual apparatus in the first published TEI guidelines, and so wrote the draft which formed the substance of the chapter. Somehow, although I was not formally a member of the working group on manuscript transcription, I attended a meeting of this group and ended up writing a draft that became the substance of the transcription chapter too. I am therefore responsible for some 50 of the 1250 plus pages of the Guidelines as they were published in 1994. Further, these are the 50 pages of most concern to scholars making editions of texts based on primary sources. Several digital scholarly editions have indeed used these guidelines profitably, so it must be said that in terms of their immediate aim—to provide encodings which would support such editions—the guidelines were and are successful. But in terms of another aim, to provide a system which any reasonably competent humanities scholar can use (which, eventually, is the only aim that matters), the guidelines are a failure. One has only to see the look of dismay on a scholar's face when encountering their full horror for the first time to know this. One may contemplate, with equanimity, every complexity of Byzantine medieval military history but be quite defeated by the unfamiliar vocabulary of the mysteriously interconnected universe which is the TEI. All scholars bring the same two questions: where do I start? and where do I find what I need? But the answers each time are different, and even those expert in the TEI may struggle to find them—while engaging in intense theological disputes over the correct interpretation. Little wonder then that scholars choose to find other things to do—or to make print editions. I do not mean by this that the TEI is fundamentally wrong in its design and aims. In fact, the first priority had to be to make a system that worked, with usability a later concern. But this is the situation it has placed us in. Putting it simply: if a scholar has to understand how the TEI encodings work—and indeed, requires any more than a basic knowledge of how XML works—in order to make an electronic edition, very few will ever do so.

Of course, the TEI and XML are encoding systems, not tools. One needs tools to put the encoding in and other tools to publish the encoded texts. Here, however, we are even worse off. There are certainly exceptionally powerful tools about and, with the widespread success of XML, many to choose from. But once more, they are not for everyone. One may take as an example the Oxygen system. This is widely used by IT professionals, reasonably priced, and well supported and documented. But a glance at the vendor's website shows that it is not for the average academic working in the average department, with little or no specialist support. The first features listed at the site on January 15, 2005 were "Different perspectives: source editor, XSLT debugger, tree viewer/editor" and "Support for XML 1.0, XML Schema, Relax NG, DTD and NRL schemas". A scholar struggling to find time to work on an edition between teaching, administration, and what used to be called leisure is unlikely to go much further. Yet Oxygen is simplicity itself compared to some other XML editing tools: just look over the correspondence on the Emacs editor on any XML mailing list (e.g. on TEI-L for 14-17 January, 2005 [http://listserv.brown.edu/archives/tei-l.html]).

Even if our average scholar successfully does master Oxygen or some other XML tool, he or she then needs to seek out a publication system that is able to handle reasonably large and complex documents with appropriate searching tools and interface customization. In the past, a publisher might be expected to do this (as, indeed, publishers have done for centuries for print editors). But, as we noted above, the supply of publishers willing to support complex electronic scholarly publications has dried up. So this means that our editor is likely to engage in some form of self-publishing, with whatever help he or she can muster. And that is not easy. Our editor will need, first of all, access to a web server with full permissions to configure as needed (something many universities, rather reasonably, routinely deny to individual academics). Presuming he or she is actually given permission, he or she will then need to configure
the system. Here is a rather typical configuration put together by one humanities computing specialist:[3]

- Install the Apache webserver, set up so that it can be seen by the outside world
- Install the MySQL database
- Install the PHP package within Apache, permitting queries of the MySQL database for search purposes
- Install the Sablotron XSLT processor to transform the XML files to HTML when found
- Prepare scripts to translate the XML into the database, and further scripts for managing searching and display
- Set this up, make sure all the pieces are working together, keep it running.

§ 17 To be sure, many people have been down this path before. Many computer systems arrive with many of these pieces already in place; XSLT scripts to manage the crucial XML transforms are also widely and freely available. But even with all this—and with much better documentation and training than is now available—one cannot imagine the gulf being bridged between the technical skills possessed by most editors and those required by publication systems they are required to use. It could be argued that the examples I have been giving are of open source systems, which typically do require higher levels of computer skill. But even if he or she can afford a proprietary system (and that is a large if), our scholar is still going to need a high level of skill to set up and configure the software, and then customize the interface for his or her data. And even then, he or she may still find that the software cannot do everything required of it.

Publishers

§ 18 This gulf between the actual technical skills of scholars and those demanded of the publication systems should not come as a surprise. Translate the terms of this discussion to the print world. No-one would expect a scholar, having written a book, to set the type, make the paper, choose, configure, set up, run, and maintain the printing press, operate the binding machines, pack the books into boxes, store them, and finally take care of their marketing and distribution. These are exactly the tasks that, for five hundred years now, publishers have done for scholars. But the disappearance of publishers from the field of electronic scholarly publication has left us with a problem. Who, in their absence, is to do the equivalent for a digital publication?

§ 19 The answer that has emerged over the past decade is that these critical technical tasks—where they are done at all—are performed by institutional IT or humanities computing specialists. This means that the only scholars who are able to make digital editions are those few who are lucky enough to be at institutions which have capable experts.[4] And there are rather few institutions that do—probably less than a dozen in the English speaking world. It might be argued that this should be enough: after all, fewer than a dozen academic publishers are sufficient to publish the great majority of the world’s academic print output. But partly for reasons to do with academic politics, partly for reasons to do with simple logistics, the few institutions with such capability generally are unable to collaborate with many outside academics—and in some cases, with very few inside scholars either. It is not quite true that one needs to be tenured at the University of Virginia to make a scholarly electronic edition. Once again, however, it is too close to the truth for the health of us all.

§ 20 Suppose, then, you are a scholar—as most of us are—at one of the many institutions which does not boast a team of experts in TEI encoding and software tools. What are your choices? You have three:

- Become an expert in these things yourself. This will take time (very large amounts of time) that could have been spent on the scholarship itself
- Form a partnership with a group in an institution that has the necessary expertise. This requires attracting the attention of that group, which you will only do if your project is of overwhelming interest or (amounting to the
Comparing the digital and print worlds

§ 21 Should you choose this third option, a pleasant surprise awaits you. While tools for making and publishing electronic scholarly editions seem rather too difficult for the average scholar, tools for making an impressive print edition are far easier to find and use. There are several based around Donald Knuth’s venerable but beautiful typesetting system TeX, e.g. John Lavagnino and Dominic Wujastyk’s EDMAC program with Bernt Karasch’s Critical Edition Typesetter (see http://www.homepages.ucl.ac.uk/~ucgadkw/edmac/index.html); Typographica Academica Traiectina (http://www.typographica.nl). Others are based on extensions of the Microsoft Word family of software: e.g. Imprimatur (http://www.geocities.com/imprimaturweb/) and the Classical Text Editor (CTE) (http://www.oeaw.ac.at/kvk/cte/index.htm), which seems the most fully developed and supported, and has been used by more than 16 projects and editions.

§ 22 Other pleasant thoughts console the editor who chooses to work with print. We have had centuries of experience in the making of print editions. We know what they look like, how they might be published and used. Particularly, there is a wealth of practical advice to draw on: how to prepare data about variation between versions; how to prepare the edition, apparatus, introductions and notes for the publisher. Of course too for print we have many possible choices for publication. There are all the traditional academic publishing houses—several of which, including Elsevier, Oxford, and Cambridge, have centuries-long traditions of publishing critical editions, leading at times back to the beginnings of print. Now these are joined by many smaller specialist publishers, including some that use "on-demand" systems for books with very small print runs.

The way forward

§ 23 Given these circumstances, it is hardly surprising that electronic editions have not made the impact that we expected in the last years. But all the advantages we foresaw for them remain valid, and the few electronic editions that actually have been made have shown that these advantages are real and practical. I believe that it must still be our goal that most scholarly editions should be made and distributed in electronic form (often alongside a print counterpart) with comparatively few editions being made in print form alone. This is precisely the reverse of the situation at the moment. So how do we get there from where we are now?

§ 24 In the last decade, we have tried to promote the making of electronic scholarly editions by giving individuals the knowledge of the encoding schemes and of the software to exploit them. Thus the many courses in TEI encoding, usually with a software component, attached to humanities computing conferences or mounted by specialist centres. While these certainly have served some areas well (notably the digital library domain), they have had little impact on the scholarly editing community. I do not believe that putting on more such courses focussed on scholarly editing, attempting to produce better documentation, or further refining the TEI guidelines, will help.

Goals

§ 25 What will help, then? Our goal must be to ensure that any scholar able to make an edition in one medium should be able to make an edition in the other. Further, that an edition in either medium should be equally assured of appropriate distribution: just as once a library has bought a print edition it can be used by any member of the library for years to come, so too should it be for electronic editions.

Tools

§ 26 Let us deal with each of these goals separately. The first is a matter of tools. Fundamental to the model of electronic scholarly edition as it has developed over the past decade is the inclusion of full transcripts of all witnesses same thing) you have lots of funding. Most scholarly editions do not qualify on either count.

✱
to the text. These transcripts need complex structural and other encoding from which computer programs may generate full collations of all the witnesses. Further computer programs may offer different representations and analytic tools for studying the differences and similarities among the witnesses, and then present all these, typically alongside images of the manuscripts, within a single electronic interface. At present, to do all these tasks requires a range of different computer programs. A tool that brought all these into a single interface—that permitted submission of the transcripts, their collation, adjustment of the collation, presentation of collation analyses, presentation of the transcripts, and integration of editorial matter, all on-line—would go far to ease the way for scholars into electronic editions. One immediate benefit of this is that it would permit scholars to see, almost instantly, the effects of an editorial decision. At present, the long chain of software needed to make an electronic edition means that work is done in batch mode with each link of the chain being a discrete production stage: you do all the transcripts, then all the collation, then adjust the collation, then you put it all in a publication interface, and only then do you discover you made a mistake right back in the first transcription stage. So you go back, redo that, and go right through the chain once more to see if the fix worked. This wide gap between editorial decision and edition realization is perhaps the single largest problem with current systems (including my own) for making electronic editions. It is also the single area where systems such as TeX or the CTE are stronger: the editor makes a change and within moments, you can see the result.

§ 27 I am not the only scholar to recognize this weakness in current systems for preparing electronic scholarly editions, or even the first. At least two major efforts are under way which, to varying degrees, address the problem. The first is the Archway project, led by Kiernan with other scholars from the University of Kentucky (http://beowulf.engl.uky.edu/~kiernan/ARCHway/entrance.htm). Its declared aims are to "produce a system for building digital libraries of image-based scholarly editions for the humanities", and to "lay the groundwork for sophisticated technical tools to interpret, assemble, disseminate, and maintain image-based scholarly editions on a continuing basis". The second is the NINES initiative, led by McGann and others at the University of Virginia. Among the many aims of this project is the creation of JUXTA, "a text comparison and collation tool" (http://www.nines.org/tools/tools.html; see also http://www.nines.org/). Both initiatives, however, have several other aims and it remains to be seen how far the tools they make will serve the needs of those of us who want to make electronic scholarly editions. It will not surprise the reader, then, to learn that I am planning to make such a tool. Tentatively named EDITION (for Editing Digital Interactive Texts through Online Networks) this will amalgamate the collation algorithms, built up over the last two decades and currently contained within Collate, into the Anastasia electronic publishing system. To these we will add a CVS system for handling different versions of submitted transcripts, a relational database for storing editorial information relating to regularization and variant setting operations and phylogenetic and variant database tools for analysis of patterns of variation. The entire suite of tools will be bundled together so that scholars will be able to submit, validate, and collate transcripts, view and explore patterns in textual variation, and then see the collation, analyses, transcripts, and manuscript images all using a single web interface. Effectively, this will collapse all the many links of the chain by which we currently make editions such as The Miller's Tale on CD-ROM into a single operation. We are currently (January 2005) seeking funding for this, with the aim of commencing work on this in July 2005.

Publication

§ 28 Tools such as those described above might go some way to helping achieve our first goal: to bring it about so that scholarly electronic editions require no greater level of computer expertise than is required say for the CTE. But they will not satisfy the second goal: to provide a route to electronic publication which is as secure (and, therefore, likely to be as esteemed) as is now the case for print publication. Quite simply, making electronic editions and then just parking them on a website somewhere, hoping (firstly) that scholars will flock to your site and (secondly) the site will miraculously sustain and develop itself over the years, is not publication. There is no help for it. For people to learn about your site, and—especially—for the site to be kept alive as computer systems change, someone has to pay
money. No granting agency will ever commit the funds to do this, in perpetuity—and nor should they. It could be argued that the move towards "free" publication in the humanities—"free" meaning of course that someone else pays—might provide a platform for these publications. But scholarly editions are highly complex publications, far removed from the pdf files which are the staple of the Open Archives Initiative (OAI) and its congenitors http://www.soros.org/openaccess/index.shtml. I am very sceptical that systems designed for mass publication of scholarly articles and the like could cope with the editions we want to make (I am actually sceptical of the whole OAI concept, but that is another matter).

§ 29 On the other side, it can be said that scholarly editions of the kind we are making are among the "must have" items of any reputable humanities school. If we make them as they can be made, with the interface they should have, they will be used at every level: for reading, for exploration, for teaching, for simple fun. A successful model of institutional subscription has been established for electronic materials of this kind: witness the success of projects such as Early English Books Online (http://eebo.chadwyck.com/home), the MEC, and online versions of the Oxford English Dictionary (http://dictionary.oed.com/) and Dictionary of National Biography (http://www.oxforddnb.com/). These show that academic institutions (and, we can suppose, individuals) will pay reasonable prices for high-quality academic electronic materials. It seems a fair hope, then, that scholarly electronic editions can be published, maintained, and extended by a subscription model. Once more, I will be exploring this later in 2005 with SDE as we begin to set up online subscription to our publications. Perhaps publishers will return to the field of electronic scholarly editions, after all.

Conclusions

§ 30 Throughout this article, I have expressed what I think should be our aim: that some time quite soon scholars wishing to make scholarly editions will naturally choose the electronic form. It follows then that all major series of scholarly editions, including those now published by the major academic presses, also will become digital. There will be exceptions: there always will be a place for a printed "reader's edition" or similar. But we should expect that for most of the purposes for which we now use editions, the editions we use will be electronic. We should do this not just to keep up with the rest of the world, but because indeed electronic editions make possible kinds of reading and research never before available and offer valuable insights into and approaches to the texts they cover.

§ 31 But this will not happen simply because we will it, or because this conclusion is obvious. We need some things we do not yet have: software that does not exist and established online publication systems that have yet to be created. Let us not wait too long for these.

Notes

[1]. The MEP website has been inaccessible due to renovations since December 7, 2004.

[2]. Indeed, e-books and digital newspapers have in many areas, despite the expenditure of vast sums (far larger than has been spent on digital scholarly editions!), failed to impact on print publication. For doubts about the viability of e-books see Peter Meirs (Meirs [n.d.]). Some remarkable statistics on the failure of digital newspapers are given by Vin Crosbie (Crosbie 2004a). For example, the Washington Post has an audited print circulation of 732,904. Its digital circulation is 424. In a followup, Crosbie pointed out that there are in fact many successful instances of digital publishing in other categories of business publication—just not digital newspapers (Crosbie 2004b).

[3]. This configuration is derived from a paper by Matt Zimmerman (Zimmerman 2004). This is actually (as Zimmerman intended) one of the simpler and more accessible attempts to put together an XML publishing system from Open Source tools. An instance of a more complex and powerful scheme is that used by STOA (see http://doxa.stoa.org/): this involves installing Fedora Core 2, tomcat, cocoon, the tomcat connector, MySQL and postgres (with the postgresql driver), modifying many configuration files and running various pieces of start-up
software. I note that the latest entry (as of March 3, 2005) in the installation blog begins "6 Sept 2004: we are now rebuilding this machine as a dedicated cocoon server", while the previous entry shows they "installed Fedora Core 2 on 3 September 2004,"—suggesting it took highly experienced system engineers several days to do basic installation.

Even systems which try to simplify matters by drawing all the pieces of software together so as to reduce the management of complex interactions can be little (if at all) less formidable to set up and run. The "Quick Start" instructions for installing the EXIST database system, for instance, run for pages: see http://exist.sourceforge.net/quickstart.html; a look at the wiki for this system gives a sense of the complications involved in running this software (http://wiki.exist-db.org/space/start). It also must be said that even after all the effort one might put into setting up and learning how to use these systems, they still have great difficulty carrying out certain tasks one might consider essential in a digital scholarly edition—showing a manuscript transcription page by page or manuscript line by manuscript line, presenting search results in a "key word in context" format, or showing individual hits highlighted within the source text.

[4]. I here distinguish the ability to make a complex scholarly edition in electronic form, such as those named earlier, from the ability to present relatively simple XML documents over the web. As the success of the digital library movement, especially in the US shows, many more institutions do have the ability to present simply-structured document (and image) collections (hence the proliferation of finding aids and electronic archives). However, the lack of the rich structure in these limits their ability to represent the wealth of interconnections within a scholarly edition, and thus limits their power as tools for research and reading.

[5]. The link to the Karasch site on the EDMACs index page is broken. The correct URL as of March 3, 2005 is http://karas.ch/cet/cetmikte.htm.

Works cited


Current issues in making digital editions of medieval texts – Or, do electronic scholarly editions have a future? Article. Apr 2005. An adequate digital text representation must therefore be compatible with the application of the formal procedures of information processing which give algorithmic form to current methods and practices of textual criticism and interpretation. The practice of markup, which became widespread with scholars who apply computational procedures to the study of a text, has revealed some difficulties which derive from fundamental theoretical options. These difficulties concern the conditions of adequacy of both kinds. On the one hand, a complete awareness of the theoretical View.

A digital text representation must therefore be compatible with the application of the formal procedures of information processing which give algorithmic form to current methods and practices of textual criticism and interpretation. The practice of markup, which became widespread with scholars who apply computational procedures to the study of a text, has revealed some difficulties which derive from fundamental theoretical options. These difficulties concern the conditions of adequacy of both kinds. On the one hand, a complete awareness of the theoretical View. Show abstract. Peter Robinson, "Current Issues in Making Digital Editions of Medieval Texts — or, Do Electronic Scholarly Editions Have a Future?" Digital Medievalist 1.1 (2005). Peter Robinson, "Where We Are with Electronic Scholarly Editions, and Where We Want to Be?" Jahrbuch für Computerphilologie Online 1.1 (2005). In print in Jahrbuch für Computerphilologie 2004, 123–143. The Electronic Library Foundation’s edition of the Canterbury Tales is available in a variety of format: in Middle English, Modern English, and facing page versions. Very good for student reading. Seems to be down at present.