The History of Scholarly Digital Editions, Plc.

Peter Robinson*

In this paper, I relate the history of a single small academic scholarly publishing house, specializing in the publication of scholarly editions in digital form: Scholarly Digital Editions.¹ While there have been many publications about the possibilities of scholarly digital editions, there have been very few studies of actual digital editions and almost none at all about how they come to be published or about the place of commercial enterprise in this endeavour. This paper attempts to fill that gap and offers some suggestions about how the digital publication of scholarly editions might evolve in the future.

It has been recognized, almost from the very first application of computing methods to the humanities, that scholarly editions are especially likely to benefit from, and be transformed by, the revolutions in information technology of the last century. The very first attempt to use computers in humanities research was Father Roberto Busa's work with IBM on the texts of Thomas Aquinas. Although not itself a scholarly edition (Busa's interests lay in the analysis of meaning in the text, and not on its transmission) his work established, first, that computer technology could deal with large amounts of text and, second, that useful results might derive from this employment of computer technology. Textual scholars were quick to see the potential, and Vinton Dearing, Petty and Gibson, and John Griffiths in the 1960s explored the creation of machine-readable representations of literary works, their collation by computer, and the analysis of the results of the collation. Their pioneering work, and

* Peter Robinson is Bateman Professor of English at the University of Saskatchewan. He is currently leading development of the Textual Communities project, which aims to provide a platform for collaborative scholarly editing. He is active in the development of standards for digital resources, formerly as a member of the Text Encoding Initiative and as leader of the European Union-funded MASTER project, and has published on Chaucer, scholarly editing, and digital humanities.

¹ At several points in the article, I use "we" when speaking of Scholarly Digital Editions. Commonly, the pronoun includes Barbara Bordalejo, for years the company's chief support and resource.

that of other scholars through the 1970s on, is usefully summarized by Susan M. Hockey in two monographs on electronic applications in the humanities.²

Up to the 1980s, while interesting work could be done with computing methods on some of the materials of scholarly editions, the editions themselves were not published in any electronic form. One might prepare electronic versions of texts, interrogate those electronic versions with computer tools, and then write an article which appeared in print (or, perhaps – and increasingly through the late 1980s – report on the work in an electronic discussion group); or one might use computer methods to prepare the text of the edition, and then publish in print (as with the many editions, almost all in Germany, produced with Wilhelm Ott using the TU-STEP program that he began developing in the late 1960s, or many others based on the TeX family of programs, first issued by Donald Knuth in 1978). While some scholarly editions indeed existed in electronic form, for most scholars there was no practical way of publishing them in any form but print. The example of the *Thesaurus Linguae Graecae* is instructive: the scholars behind this project began gathering Greek texts in the late 1960s, distributed them on magnetic tape in the mid-1970s, and then shifted to the dedicated Ibycus computer system from early in the 1980s. For scholars who were not David Packard, electronic publication was simply impossible, and even if they did somehow acquire an electronic text, there was little they could do with it. In 1986, I acquired an electronic copy of Andrea van Arkel's edition of the Icelandic manuscript *Modruvallabók* and made a concordance of it using the Oxford Concordance Program. I then printed the concordance onto a two-foot pile of paper and used that in my research. Indeed, the prospect of electronic publication appeared so distant that in 1980 Hockey described the final step of making a critical edition as "printing of the text and apparatus."³ Nowhere did she contemplate the possibility of distribution in electronic form.

² Susan M. Hockey, *A Guide to Computer Applications in the Humanities* (Baltimore: Johns Hopkins University Press, 1980), 144–67; and *Electronic Texts in the Humanities: Principles and Practice* (Oxford: Oxford University Press, 2000), 145.

³ Hockey, *Guide to Computer Applications*, 144. By 2000, however, matters had so changed that Hockey added the phrase "and Electronic Editions" to the title "Textual Criticism" of the equivalent chapter in her 1980 monograph and spent half the chapter discussing the first electronic editions.

By the mid-1980s there were already substantial quantities of scholarly text in electronic form. However, there was no way, then, to publish these texts. One could transfer the text to disc and mail it to the Oxford Text Archive, which might mail out copies to anyone who asked, but this was hardly publication. Two events changed this. The first event was the arrival of the CD-ROM form for publication of data. In those days, when a computer disc typically carried around a single megabyte of data, and many computers had even less memory than that, the prospect of fitting six hundred megabytes of data onto a single disc was breathtaking. However, it took some time for CD-ROM drives to become standard on computers. The technology to burn them was expensive and there was little software for the presentation of these large quantities of text and images. On the other hand, there was already a considerable appetite for such large collections, and with very few publishers able to supply that appetite, those who did, did well. Accordingly, the first CD-ROMs were produced as commercial enterprises, notably the collections published by Chadwyck-Healey (the FullText English Poetry Database and others) selling for some £30,000 as a CD-ROM set. Particularly active was Oxford University Press (OUP). The press had begun experiments with the use of computers in their dictionary division, focusing particularly on the *Oxford English Dictionary* itself, and it published the first electronic version of the complete *OED* on CD-ROM in 1988. Encouraged by this, OUP established an ambitious electronic publishing division, led for a time by Andrew Rosenheim (later managing director of Penguin, and now editor of the Amazon UK Kindle Singles store). In a short span of time, this new division accumulated an impressive list of prospective editions, though in fact it took many years for most to emerge (for example, the flagship Bergen Wittgenstein Nachlass was not published till 2000 though under contract with OUP from the early 1990s). OUP and Chadwyck-Healey were not alone: the University of Michigan Press committed itself to the publication of the SEENET series of Medieval English and Norse texts and Jerome McGann's Rosetti Archive, and University Microfilms International took on the publication of the University of Sheffield's Hartlib Papers, again on CD-ROM.

Accordingly, around 1992, when what became the Canterbury Tales Project was taking shape, the publication landscape for scholarly editions in digital form looked very different from the situation today. Then, a few publishers were committed to, or already publishing, electronic texts on CD-ROMs for sale. In essence, this model was

identical to the one that had prevailed for academic book publication for centuries. A scholar (or group of scholars) had an idea for an edition and took it to the publisher; the publisher sought scholarly opinion on the idea, agreed to publication, issued a contract, and then published and distributed the result. It was different in that the edition was a CD-ROM, not a book; and different too, in that the publisher's costs for the new electronic medium might be rather high and open-ended, in contrast with the established world of print publication. However, the promise of electronic publication, even in those pre-web days, was such that several publishers had already tested the waters.

It happened that Cambridge University Press, observing Oxford's foray into electronic publication, decided that it too wanted to explore this new medium. Andrew Brown, then responsible for humanities publishing at the press, came to Oxford and spoke with me and others about the impact of the new technologies. I went to Cambridge and met Kevin Taylor, who assumed direct responsibility for Cambridge's electronic publication experiments. Cambridge has a long history of publication of critical editions and Taylor and Brown shared both a personal commitment to continuing that work and an interest in exploring what could be done with scholarly editions in the electronic medium. One might also believe that they got a certain pleasure in taking on editions that a press in another British university had felt unable to support. (OUP had passed up the chance to publish both my work on the *Tales* and Anne McDermott's *Johnson's Dictionary*.) Accordingly, they were prepared to put substantial resources into the making of electronic editions, enabling publication in these years not merely of the *Wife of Bath* and *General Prologue* but also Anne McDermott's edition of *Johnson's Dictionary*, four iterations of Jim Harner's *World Shakespeare Bibliography*, Christie Carson and Jacky Bratton's *King Lear*, and the *History of Parliament on CD-ROM*. Among these was my edition, *The Wife of Bath's Prologue on CD-ROM* – the first major achievement in the series – and we were close to publishing a second, Elizabeth Solopova's *General Prologue on CD-ROM* (eventually published in June 2000). Both of these were single-tale editions, presenting all the witnesses for a single tale (fifty-eight for the *Wife of Bath* and fifty-four for the *General Prologue*).

In late 1999, however, the leaders and researchers of the *Canterbury Tales* Project, based at De Montfort University in Leicester, had a problem. We had been working for several years on the Hengwrt manuscript of *The Canterbury Tales*, probably the oldest and (by many

people's estimate) the most important of all manuscripts of the *Tales*. We had a superb transcription of the whole manuscript, which was initiated by Estelle Stubbs and Michael Pidd under the direction of Norman Blake and myself at the University of Sheffield from 1994 on, and then checked multiple times by project members in the years since. The National Library of Wales had agreed to create a full set of digital photographs of the manuscript for us. In the course of leading the transcription effort, Stubbs had examined the manuscript with an attention that it had likely never received before, and she was forming a series of arguments about the manuscript: how it was produced, its history, and its place in the copying of the *Tales*. The transcription, the images, and Stubbs's commentaries would constitute a substantial edition. We planned that the Hengwrt edition would be the first single-manuscript edition, offering the whole *Tales* as present in a single manuscript.

All seemed well until Kevin Taylor, our editor at Cambridge who had nurtured the project through the first two electronic editions and who had arranged for the press to give an extraordinary degree of support for us in those years, informed us (with polite understatement) that Cambridge could not commit to publishing our CD-ROMs after the General Prologue. This was not a definite statement of withdrawal. He and I had worked together since 1992 on a series of experiments with electronic publication. However, one can already sense in a 1998 article he and I co-wrote about the making of the Wife of Bath CD-ROM that Cambridge was having doubts about the enterprise. Taylor noted that, at the price of £150 (US \$250), the press would "be looking to sell about 250 copies to cover our direct costs."⁴ By October 1997, less than a year after publication, 180 copies had been sold, suggesting that for this edition, at least, 250 copies was a reasonable sales target. In the event, 348 individual copies and five local area network (LAN) copies were sold. However, other editions that Taylor and I worked on were rather less successful. Further, sales at this level met the direct costs only: the costs of actual production like software and other licensing, preparing the data and interface programming for the CD-ROMs, and physical manufacturing.⁵ Cambridge is a large organization and

⁴ Peter M.W. Robinson, and Kevin Taylor, "Publishing an Electronic Textual Edition: The Case of the *Wife of Bath's Prologue* on CD-ROM," *Computers and the Humanities* 32 (1998): 280.

⁵ Kevin Taylor has provided the following figures concerning Cambridge sales and costs (e-mails to the author of 15 and 16 November 2012): *Wife of Bath*,

sales are needed to pay for warehouses, salaries, and infrastructure. Taylor remarked to me at the time that merely to insert an item in the Cambridge worldwide sales catalogue cost £15,000, giving an idea of these overhead expenses. Further, because almost everything we were doing was unlike anything else happening at the press, the costs could not be spread across other sectors. A printed edition of the same work and of similar sophistication might have sold four times as many copies. One can also detect some hesitation in Taylor's remark in our article: "was it worth all this effort, or could we have dispensed with some of the more rarefied features?"⁶

To compound the situation, while the National Library of Wales was prepared to photograph the manuscript for us, they wanted £10,000 for the publication rights. We had come to an impasse. From the tenor of my frequent conversations with Kevin Taylor, there was no possibility that Cambridge would be willing to spend so large a sum; indeed, the answer was so clear that I never asked him. At the same time, he was hinting to me with extreme delicacy that Cambridge would be quite relieved if it ceased to be our publisher sometime in the future. Nor was there any other likely publisher on the horizon – rather the reverse. In a few years preceding the publication of our 1998 article, we had seen two major university presses that had invested heavily in electronic publications similar to ours close down their efforts (OUP – hence their lack of interest in Anne McDermott's and my work – and University of Michigan Press). Commercial publishers like Dorling Kindersley, HarperCollins, and Penguin had followed the same path: substantial investment and high hopes had given way to disillusionment and withdrawal.

348 + 5 LAN licences, direct production costs £21,202; *Johnson's Dictionary* 824 copies + 2 LAN licences + 2 WAN licences, direct production costs £33,355; *General Prologue* 258 copies + 1 WAN licence, direct production costs £12,191. On these figures, only *Johnson's Dictionary* would have made a significant return which could be set against indirect costs, with the two Chaucer editions managing only a small surplus over direct costs. Other editions did considerably less well. *The History of Parliament CD-ROM* was very expensive to produce, as the whole of the text of twenty-three published volumes had to be captured in SGML and then painstakingly corrected: thus the production costs were logged at £54,459 for the press, with only 220 copies, 8 LAN, and 1 WAN licence sold. Further, the revenues from this publication had to be shared with the History of Parliament Trust. As Taylor comments (e-mail to the author of 16 November 2012: "It wasn't a business model that worked too well, but I don't think it would today either. That shouldn't detract from the sheer brilliance of the products."

⁶ Robinson and Taylor, 283.

A factor in these discussions was the second event that made publication of electronic materials possible (the first being CD-ROMs): the advent of the World Wide Web from 1993 onwards. I remarked above on the closeness of the CD-ROM publication model to that of traditional print publication; by contrast, the web offered no such familiar model, while making it clear that the mode of CD-ROM publication, to which the press had committed, was likely to be superseded (if not completely supplanted) by web publication. But web publication was so different from traditional publication and there were (and still are) so many uncertainties about it that the press was not prepared at that time to devote yet more resources towards yet another new direction, with a still more uncertain outcome.

I spent much time in 1999 trying to find a way past these problems. For a period, it seemed that De Montfort University, where the project and I were then based, might be able to set up a publishing arm to carry on where Cambridge thought it could not. However, these discussions amounted to nothing. Finally, the idea came to me: I could publish the Hengwrt edition myself. Not only might we publish this and later project materials; the sudden disappearance of digital publishers able to cope with enterprises of our scale (as opposed to the mass publication strategies of the likes of Chadwyck-Healy, now ProQuest) meant that there might be other scholars with born-digital work looking for a publisher. We might be able to publish these, too. I was strongly advised that I should set up a private limited company as the publisher, rather than publish them myself as a personal enterprise. The company structure would provide me (and my family) with protection if things went badly wrong and also provide some confidence to institutions with which we were dealing. For this, I needed the permission of my employer, who also required safeguards against conflicts of interest. All this took some time. Permission from De Montfort did not come until the end of May 2000. I borrowed some money to pay the National Library of Wales and the University of Sheffield (who decided that it, too, needed payment for the transcripts on which project staff at Sheffield had worked), filled out various forms, and Scholarly Digital Editions was born.

On 25 October 2000, to coincide with the sixth hundredth anniversary of Chaucer's death, we published the *Hengwrt Chaucer Digital Facsimile*, edited by Estelle Stubbs. Since then we have brought out a further ten editions. I give the approximate sales figures for

each, with its publication date and the usual price charged for an individual copy (table 1).

Table 1. Works Published by Scholarly Digital Editions

Edition	Date	Price	Copies Sold (approx.)
<i>The Hengwrt Chaucer Digital Facsimile</i> , ed. Estelle Stubbs	25/10/2000	£82	320
<i>The Bayeux Tapestry Digital Edition</i> , ed. Martin Foys	31/10/2002	£30	2500
<i>The Hengwrt Chaucer Standard Edition</i> , ed. Ceridwen Lloyd-Morgan	3/11/2003	£20	80
<i>Caxton's Canterbury Tales: The British Library Copies</i> , ed. Barbara Bordalejo	8/10/2003	£20	150
<i>The Miller's Tale on CD-ROM</i> , ed. Elena Pierazzo	5/5/2004	£50	110
<i>The Parliament Rolls of Medieval England</i> , eds. Christopher Given-Wilson et al.	5/5/2005	£50	560
<i>The Nun's Priest's Tale on CD-ROM</i> , ed. Paul Thomas	3/5/2006	£50	70
<i>Dante: Monarchia</i> , ed. Prue Shaw	3/5/2006	£50	60
<i>Leiden Armenian Lexical Textbase</i> (Internet), ed. Jos Weitenberg	21/2/2008	£50	36
<i>A Digital Catalogue of the Pre-1500 Manuscripts and Incunables of the Canterbury Tales</i> , ed. Daniel W. Mosser	10/5/2010	£50	16
<i>Dante Alighieri: Commedia</i> , ed. Prue Shaw	1/12/2010	£100	16

Did we make money? So far as these publications are concerned, the answer is a simple, flat no; we lost money on these publications. Like Cambridge, we had to cope with the direct costs of production: licensing, interface programming, and packaging. As is common for new companies, we overspent on direct costs in the first years. We paid too much to the National Library of Wales for the image rights and too much to the University of Sheffield for rights to the transcriptions for the *Hengwrt Chaucer*. We paid too much to OUP for the rights to translations included in the *Bayeux Tapestry*. We created beautiful, but very expensive, packaging and publicity materials for the *Hengwrt*

Chaucer, and (to a lesser degree) for the *Bayeux Tapestry*. After those two publications, we became shrewder about how we packaged our CD-ROMs, and developed for the *Caxton's Canterbury Tales* and later publications a standard packaging which cost about one fifth of the *Hengwrt Chaucer* packaging. The excessive sums we paid in the first instance, however, were a weight against our revenues for years afterward.

Note too that our revenues were rather lower than one might presume from this table alone. One cannot calculate our revenue simply by multiplying price by copies sold. Many copies were sold by resellers, who kept between 40 to 80 percent of the sticker price for themselves. For many copies of the *Bayeux Tapestry*, we received a payment of less than US \$6 after every reseller along the sales chain had taken their cut. Indeed, the company only survived the first years because we also carried out a considerable amount of paid consultancy. This allowed us to pay a computer programmer (originally Andrew West, later Zeth Green) to work for us from 2000 to 2009, and to contribute to the necessary programming for the editions. We were able to charge for the consultancy on a "cost plus" basis: the "plus" subsidized our publication program. The simplest measure of profit, for a limited company like us, is tax paid. In the first years of the company we built up such a substantial deficit that up to now we have not reported any profit after allowances for losses for a single year. We expect this will change in 2013 as the company has finally been trading at an annual profit for several years, and profits are at last on the point of overtaking accrued losses. But we would have preferred to be making profits and paying taxes years ago.

This table also is striking because of the asymmetry of the sales. A few publications sold very well: the *Bayeux Tapestry* may claim to have sold more copies than any comparable scholarly digital edition ever. The *Parliament Rolls* also did exceptionally well, especially when one factors in institutional Internet licences (fifty-one sold to now), although our revenues from this publication had to be shared with the National Archives. One may note that we experimented too with cutting the price, marketing Barbara Bordalejo's excellent *Caxton's Canterbury Tales* for just £20. Sales did not increase. The *Hengwrt Chaucer* has still to recover its production costs, despite a healthy sales level, because of the excessive fees and costs we paid in the first year of publication. Other publications have failed to meet their production costs. Particularly notable are the poor sales of four publications: *The Miller's Tale* and *The Nun's Priest's Tale* (both Chaucer) and

the *Monarchia* and *Commedia* (both Dante, both edited by Prue Shaw, now Emeritus Reader in Italian Studies at University College London). All four are made to a model that is effectively unique to Scholarly Digital Editions. They contain full transcripts, images, collations, and analyses of multiple versions of one text – more than fifty for the *Miller* and *Nuns' Priest*, twenty-two for the *Monarchia*, and nine (including two modern print editions) for the *Commedia*.

The preparation and publication of editions made to this model is more demanding than for the presentation of a single version, such as the *Hengwrt Chaucer*, or other digital facsimile single-witness publications, the model to which nearly all digital scholarly editions conform.⁷ To make these publications, not only must one gather images of each page and transcribe them, but one must also compare all the different versions word-by-word, to create a complete and accurate record of the differences among the versions (a collation). Creating a collation, as these publications all do, is not simply a matter of running a program (such as the widely used “diff”) across the versions. All these publications use an interactive, computer-assisted collation system developed by myself called Collate in the early 1990s. This requires that the scholar look at the collation of every word, one at a time, and gives tools to adjust the differences found. The variants judged to be insignificant by the scholar can be removed (usually, variations in spelling or punctuation), while those judged to be significant can be optimized (additions or omissions, phrases, and the like), the aim being to represent as clearly as possible the differences among the witnesses at that point. This is an exacting and time-consuming process and requires so high a level of scholarly judgement that it is usually reserved for a senior editor.

These four publications have a further feature. One of the major innovations in textual scholarship in the last twenty years (though one still little known in Anglophone textual editing) has been the application of phylogenetic methods. Phylogenetic methods were developed in evolutionary biology for the creation of hypothetical “family trees” –relational groupings of populations of organisms through the analysis of the degree to which they share characteristics. Textual scholars can posit relations among manuscripts in the same

⁷ For more on models see Kevin Kiernan, “Digital Facsimiles in Editing,” in *Electronic Textual Editing*, ed. Lou Burnard et al. (New York: Modern Language Association of America, 2006), 262–68; and Elena Pierazzo, “A Rationale of Digital Documentary Editions,” *Literary and Linguistic Computing* 26 (2011): 463–77.

manner, by assessing the number of variants they share.⁸ All four of these publications (like the Cambridge *General Prologue*) submit the full data of the agreements and disagreements among the witnesses recorded in the collation to phylogenetic programs, which create family trees that hypothesize the way in which the manuscripts might relate to each other.

These four publications have a level of sophistication far beyond any of our other publications and no other publisher has produced anything like them.⁹ One could argue that they represent attempts to reach the pinnacle of textual scholarship, the making of an edition of a text existing in many versions. The *Commedia* edition is also significant because it contributes to (and indeed was occasioned by) a vigorous debate among Dante scholars concerning the editing of the *Commedia*, sparked by Federico Sanguineti's controversial 2001 edition. Yet we have sold only sixteen copies in two years (our Italian co-publisher, Sismel, may have done better). Kevin Taylor observed that what we were trying to do at Cambridge was ahead of its time.¹⁰ One may wonder if indeed there will ever be a time for editions such as these. Are scholars sufficiently interested in editions of this type to justify the immense effort and expense of their production? Or should we think of these efforts as editions for the ages, not to be consumed for a single school assignment (as so many of our purchasers of the *Bayeux Tapestry* have used that publication), but to wait years and years for the readers who will use them to ask questions we cannot imagine?

This returns to the question that faced us when we began Scholarly Digital Editions. Cambridge was not able to find a satisfactory model for electronic publication of born-digital materials such as these. Have we been able to find a model where Cambridge could not? Speaking commercially, we have not: witness our failure to make a profit after ten years, even excluding from our costs any of the time that I have

⁸ On the success of phylogenetics in stemmatics, see the list of publications at <http://www.textualscholarship.org/newstemmatics/bibliography/index.html>.

⁹ The nearest to these, in presenting a text in many versions, are the ongoing editions of the Greek New Testament at Birmingham and Münster (<http://www.iohannes.com>, and <http://nttranscripts.uni-muenster.de/AnaServer?NTtranscripts+o+start.anv>) as well as the Liverpool Cancioneros (<http://cancionerovirtual.liv.ac.uk/AnaServer?dutton+o+start.anv+ms=BC&sms=3>). All these use the same configurations of transcription and collation software as our four publications, but without the analytic materials described here.

¹⁰ E-mail from Kevin Taylor to the author, 15 November 2012.

spent over the years on these editions. (Barbara Bordalejo, too, has worked on all of them.) However, while continuing commercial success is the shortest way to maintain a long-term publication model, it is not the only way. One might choose to continue a publishing enterprise, as we have for the last twelve years, not for reasons of profit, but because it brings other benefits that sufficiently offset the expenses of effort and money. This has been the case with Scholarly Digital Editions – there have been such benefits. The company has continued through years even when it was losing money because these benefits more than compensated for the loss and the risk.

The first such benefit is the body of editions that we have been able to bring out and their value to the public. Several of these, notably the *Bayeux Tapestry*, the *Hengwrt Chaucer Digital Facsimile*, the *Parliament Rolls of Medieval England*, and the *Monarchia*, have won honours and extraordinary reviews. The *Bayeux Tapestry* has won almost every award available to it including a *Choice* Outstanding Title award and the 2005 prize of the International Society of Anglo-Saxonists for the best new edition in any medium. *History* magazine began its review of the *Parliament Rolls* with the declaration, “Nobody involved in any aspect of medieval research can afford to do without this publication.”¹¹ It is also the case that if we had not published these, no one else would have. As explained earlier, the original motive for setting up the company was to publish Stubbs’s edition of the *Hengwrt Chaucer* when there seemed no other way of achieving this publication. Following publication of the *Hengwrt Chaucer*, Martin Foys (then an early-career scholar still establishing his academic path) approached us at our stall at the Kalamazoo Medieval Congress in 2001. He had been working since 1996 on a digital edition of the *Bayeux Tapestry*, which was nearly complete and quite stunning, but he could not find anyone to publish it. Would we be interested? We said yes, and this became our best-seller. In 2004, a similar problem arose over the *Parliament Rolls*. This project began in 1997 with Leverhulme Trust funding and Cambridge agreeing to act as publisher. I worked (with funding from Cambridge) with the project team, headed by Chris Given-Wilson, to capture the data and create the whole edition, which was to be launched in 2005. However, by 2004 Cambridge was having doubts about this publication similar

¹¹ For these and other reviews, see Scholarly Digital Editions, “Scholarly Digital Editions: Reviews of Some of Our Publications,” accessed 10 June 2013, <http://www.sd-editions.com/reviews/index.html>.

to those that Kevin Taylor had expressed to me in 1999. During the summer of 2004, there were intense discussions between myself, the National Archives, and Cambridge. They culminated in an agreement that Scholarly Digital Editions, the National Archives, and the History of Parliament Trust would together publish the *Parliament Rolls*, with Scholarly Digital Editions leading the physical production and distribution. We were able to publish on schedule in May 2005. At eight thirty in the morning of the first Thursday of the Kalamazoo conference we had a line of twenty scholars waiting to purchase the newly minted CD-ROM. One could argue that this was a temporary problem only for the *Parliament Rolls*, which is now also published on the British History Online website, but I cannot see how any of our other editions would ever have been published if we had not been there.

We know, too, that these editions have been used. The 2,500 copies of the *Bayeux Tapestry* represent only a fraction of the readers this publication has reached. Teachers frequently use the CD-ROM as the basis for a lecture to a class and often as the basis for an entire course. With these mass uses, and allowing for a large number of illegal copies as the CD-ROM has no copy protection, we estimate that over a hundred thousand readers have seen this publication. One might make a similar calculation for the *Parliament Rolls*, where, in addition to the 560 CD-ROM sales, some fifty-one institutions have purchased licences permitting unlimited Internet access by all students, staff, and associates to the online version.

The second benefit, beyond commercial revenue, that Scholarly Digital Editions has brought, is that its publications have advanced the careers of the scholars whose work they embody. At the time of the *Hengwrt Chaucer*, Stubbs was on a short-term research contract. She has remained in university employment since, working on a variety of medieval-related research projects, an outcome likely to have been aided by the credit she earned as editor of the *Hengwrt Chaucer Digital Facsimile*. Foys is now an associate professor of English at Drew College, the executive director of the International Society of Anglo-Saxons, the co-director of several impressive digital humanities projects, and the author of two well-received monographs. To the extent that we were able to publish his *Bayeux Tapestry* at a crucial moment of his career, this edition helped him on his way.

The person whose academic career has been most affected by Scholarly Digital Editions is, of course, myself. At the time of writing, I hold a named chair in the Department of English at the University

The History of Scholarly Digital Editions, Plc.

Peter Robinson*

In this paper, I relate the history of a single small academic scholarly publishing house, specializing in the publication of scholarly editions in digital form: Scholarly Digital Editions.¹ While there have been many publications about the possibilities of scholarly digital editions, there have been very few studies of actual digital editions and almost none at all about how they come to be published or about the place of commercial enterprise in this endeavour. This paper attempts to fill that gap and offers some suggestions about how the digital publication of scholarly editions might evolve in the future.

It has been recognized, almost from the very first application of computing methods to the humanities, that scholarly editions are especially likely to benefit from, and be transformed by, the revolutions in information technology of the last century. The very first attempt to use computers in humanities research was Father Roberto Busa's work with IBM on the texts of Thomas Aquinas. Although not itself a scholarly edition (Busa's interests lay in the analysis of meaning in the text, and not on its transmission) his work established, first, that computer technology could deal with large amounts of text and, second, that useful results might derive from this employment of computer technology. Textual scholars were quick to see the potential, and Vinton Dearing, Petty and Gibson, and John Griffiths in the 1960s explored the creation of machine-readable representations of literary works, their collation by computer, and the analysis of the results of the collation. Their pioneering work, and

* Peter Robinson is Bateman Professor of English at the University of Saskatchewan. He is currently leading development of the Textual Communities project, which aims to provide a platform for collaborative scholarly editing. He is active in the development of standards for digital resources, formerly as a member of the Text Encoding Initiative and as leader of the European Union-funded MASTER project, and has published on Chaucer, scholarly editing, and digital humanities.

¹ At several points in the article, I use "we" when speaking of Scholarly Digital Editions. Commonly, the pronoun includes Barbara Bordalejo, for years the company's chief support and resource.

that of other scholars through the 1970s on, is usefully summarized by Susan M. Hockey in two monographs on electronic applications in the humanities.²

Up to the 1980s, while interesting work could be done with computing methods on some of the materials of scholarly editions, the editions themselves were not published in any electronic form. One might prepare electronic versions of texts, interrogate those electronic versions with computer tools, and then write an article which appeared in print (or, perhaps – and increasingly through the late 1980s – report on the work in an electronic discussion group); or one might use computer methods to prepare the text of the edition, and then publish in print (as with the many editions, almost all in Germany, produced with Wilhelm Ott using the TU-STEP program that he began developing in the late 1960s, or many others based on the TeX family of programs, first issued by Donald Knuth in 1978). While some scholarly editions indeed existed in electronic form, for most scholars there was no practical way of publishing them in any form but print. The example of the *Thesaurus Linguae Graecae* is instructive: the scholars behind this project began gathering Greek texts in the late 1960s, distributed them on magnetic tape in the mid-1970s, and then shifted to the dedicated Ibycus computer system from early in the 1980s. For scholars who were not David Packard, electronic publication was simply impossible, and even if they did somehow acquire an electronic text, there was little they could do with it. In 1986, I acquired an electronic copy of Andrea van Arkel's edition of the Icelandic manuscript *Modruvallabók* and made a concordance of it using the Oxford Concordance Program. I then printed the concordance onto a two-foot pile of paper and used that in my research. Indeed, the prospect of electronic publication appeared so distant that in 1980 Hockey described the final step of making a critical edition as "printing of the text and apparatus."³ Nowhere did she contemplate the possibility of distribution in electronic form.

² Susan M. Hockey, *A Guide to Computer Applications in the Humanities* (Baltimore: Johns Hopkins University Press, 1980), 144–67; and *Electronic Texts in the Humanities: Principles and Practice* (Oxford: Oxford University Press, 2000), 145.

³ Hockey, *Guide to Computer Applications*, 144. By 2000, however, matters had so changed that Hockey added the phrase "and Electronic Editions" to the title "Textual Criticism" of the equivalent chapter in her 1980 monograph and spent half the chapter discussing the first electronic editions.

By the mid-1980s there were already substantial quantities of scholarly text in electronic form. However, there was no way, then, to publish these texts. One could transfer the text to disc and mail it to the Oxford Text Archive, which might mail out copies to anyone who asked, but this was hardly publication. Two events changed this. The first event was the arrival of the CD-ROM form for publication of data. In those days, when a computer disc typically carried around a single megabyte of data, and many computers had even less memory than that, the prospect of fitting six hundred megabytes of data onto a single disc was breathtaking. However, it took some time for CD-ROM drives to become standard on computers. The technology to burn them was expensive and there was little software for the presentation of these large quantities of text and images. On the other hand, there was already a considerable appetite for such large collections, and with very few publishers able to supply that appetite, those who did, did well. Accordingly, the first CD-ROMs were produced as commercial enterprises, notably the collections published by Chadwyck-Healey (the FullText English Poetry Database and others) selling for some £30,000 as a CD-ROM set. Particularly active was Oxford University Press (OUP). The press had begun experiments with the use of computers in their dictionary division, focusing particularly on the *Oxford English Dictionary* itself, and it published the first electronic version of the complete *OED* on CD-ROM in 1988. Encouraged by this, OUP established an ambitious electronic publishing division, led for a time by Andrew Rosenheim (later managing director of Penguin, and now editor of the Amazon UK Kindle Singles store). In a short span of time, this new division accumulated an impressive list of prospective editions, though in fact it took many years for most to emerge (for example, the flagship Bergen Wittgenstein Nachlass was not published till 2000 though under contract with OUP from the early 1990s). OUP and Chadwyck-Healey were not alone: the University of Michigan Press committed itself to the publication of the SEENET series of Medieval English and Norse texts and Jerome McGann's Rosetti Archive, and University Microfilms International took on the publication of the University of Sheffield's Hartlib Papers, again on CD-ROM.

Accordingly, around 1992, when what became the Canterbury Tales Project was taking shape, the publication landscape for scholarly editions in digital form looked very different from the situation today. Then, a few publishers were committed to, or already publishing, electronic texts on CD-ROMs for sale. In essence, this model was

identical to the one that had prevailed for academic book publication for centuries. A scholar (or group of scholars) had an idea for an edition and took it to the publisher; the publisher sought scholarly opinion on the idea, agreed to publication, issued a contract, and then published and distributed the result. It was different in that the edition was a CD-ROM, not a book; and different too, in that the publisher's costs for the new electronic medium might be rather high and open-ended, in contrast with the established world of print publication. However, the promise of electronic publication, even in those pre-web days, was such that several publishers had already tested the waters.

It happened that Cambridge University Press, observing Oxford's foray into electronic publication, decided that it too wanted to explore this new medium. Andrew Brown, then responsible for humanities publishing at the press, came to Oxford and spoke with me and others about the impact of the new technologies. I went to Cambridge and met Kevin Taylor, who assumed direct responsibility for Cambridge's electronic publication experiments. Cambridge has a long history of publication of critical editions and Taylor and Brown shared both a personal commitment to continuing that work and an interest in exploring what could be done with scholarly editions in the electronic medium. One might also believe that they got a certain pleasure in taking on editions that a press in another British university had felt unable to support. (OUP had passed up the chance to publish both my work on the *Tales* and Anne McDermott's *Johnson's Dictionary*.) Accordingly, they were prepared to put substantial resources into the making of electronic editions, enabling publication in these years not merely of the *Wife of Bath* and *General Prologue* but also Anne McDermott's edition of *Johnson's Dictionary*, four iterations of Jim Harner's *World Shakespeare Bibliography*, Christie Carson and Jacky Bratton's *King Lear*, and the *History of Parliament on CD-ROM*. Among these was my edition, *The Wife of Bath's Prologue on CD-ROM* – the first major achievement in the series – and we were close to publishing a second, Elizabeth Solopova's *General Prologue on CD-ROM* (eventually published in June 2000). Both of these were single-tale editions, presenting all the witnesses for a single tale (fifty-eight for the *Wife of Bath* and fifty-four for the *General Prologue*).

In late 1999, however, the leaders and researchers of the *Canterbury Tales* Project, based at De Montfort University in Leicester, had a problem. We had been working for several years on the Hengwrt manuscript of *The Canterbury Tales*, probably the oldest and (by many

people's estimate) the most important of all manuscripts of the *Tales*. We had a superb transcription of the whole manuscript, which was initiated by Estelle Stubbs and Michael Pidd under the direction of Norman Blake and myself at the University of Sheffield from 1994 on, and then checked multiple times by project members in the years since. The National Library of Wales had agreed to create a full set of digital photographs of the manuscript for us. In the course of leading the transcription effort, Stubbs had examined the manuscript with an attention that it had likely never received before, and she was forming a series of arguments about the manuscript: how it was produced, its history, and its place in the copying of the *Tales*. The transcription, the images, and Stubbs's commentaries would constitute a substantial edition. We planned that the Hengwrt edition would be the first single-manuscript edition, offering the whole *Tales* as present in a single manuscript.

All seemed well until Kevin Taylor, our editor at Cambridge who had nurtured the project through the first two electronic editions and who had arranged for the press to give an extraordinary degree of support for us in those years, informed us (with polite understatement) that Cambridge could not commit to publishing our CD-ROMs after the General Prologue. This was not a definite statement of withdrawal. He and I had worked together since 1992 on a series of experiments with electronic publication. However, one can already sense in a 1998 article he and I co-wrote about the making of the Wife of Bath CD-ROM that Cambridge was having doubts about the enterprise. Taylor noted that, at the price of £150 (US \$250), the press would "be looking to sell about 250 copies to cover our direct costs."⁴ By October 1997, less than a year after publication, 180 copies had been sold, suggesting that for this edition, at least, 250 copies was a reasonable sales target. In the event, 348 individual copies and five local area network (LAN) copies were sold. However, other editions that Taylor and I worked on were rather less successful. Further, sales at this level met the direct costs only: the costs of actual production like software and other licensing, preparing the data and interface programming for the CD-ROMs, and physical manufacturing.⁵ Cambridge is a large organization and

⁴ Peter M.W. Robinson, and Kevin Taylor, "Publishing an Electronic Textual Edition: The Case of the *Wife of Bath's Prologue* on CD-ROM," *Computers and the Humanities* 32 (1998): 280.

⁵ Kevin Taylor has provided the following figures concerning Cambridge sales and costs (e-mails to the author of 15 and 16 November 2012): *Wife of Bath*,

sales are needed to pay for warehouses, salaries, and infrastructure. Taylor remarked to me at the time that merely to insert an item in the Cambridge worldwide sales catalogue cost £15,000, giving an idea of these overhead expenses. Further, because almost everything we were doing was unlike anything else happening at the press, the costs could not be spread across other sectors. A printed edition of the same work and of similar sophistication might have sold four times as many copies. One can also detect some hesitation in Taylor's remark in our article: "was it worth all this effort, or could we have dispensed with some of the more rarefied features?"⁶

To compound the situation, while the National Library of Wales was prepared to photograph the manuscript for us, they wanted £10,000 for the publication rights. We had come to an impasse. From the tenor of my frequent conversations with Kevin Taylor, there was no possibility that Cambridge would be willing to spend so large a sum; indeed, the answer was so clear that I never asked him. At the same time, he was hinting to me with extreme delicacy that Cambridge would be quite relieved if it ceased to be our publisher sometime in the future. Nor was there any other likely publisher on the horizon – rather the reverse. In a few years preceding the publication of our 1998 article, we had seen two major university presses that had invested heavily in electronic publications similar to ours close down their efforts (OUP – hence their lack of interest in Anne McDermott's and my work – and University of Michigan Press). Commercial publishers like Dorling Kindersley, HarperCollins, and Penguin had followed the same path: substantial investment and high hopes had given way to disillusionment and withdrawal.

348 + 5 LAN licences, direct production costs £21,202; *Johnson's Dictionary* 824 copies + 2 LAN licences + 2 WAN licences, direct production costs £33,355; *General Prologue* 258 copies + 1 WAN licence, direct production costs £12,191. On these figures, only *Johnson's Dictionary* would have made a significant return which could be set against indirect costs, with the two Chaucer editions managing only a small surplus over direct costs. Other editions did considerably less well. *The History of Parliament CD-ROM* was very expensive to produce, as the whole of the text of twenty-three published volumes had to be captured in SGML and then painstakingly corrected: thus the production costs were logged at £54,459 for the press, with only 220 copies, 8 LAN, and 1 WAN licence sold. Further, the revenues from this publication had to be shared with the History of Parliament Trust. As Taylor comments (e-mail to the author of 16 November 2012: "It wasn't a business model that worked too well, but I don't think it would today either. That shouldn't detract from the sheer brilliance of the products."

⁶ Robinson and Taylor, 283.

A factor in these discussions was the second event that made publication of electronic materials possible (the first being CD-ROMs): the advent of the World Wide Web from 1993 onwards. I remarked above on the closeness of the CD-ROM publication model to that of traditional print publication; by contrast, the web offered no such familiar model, while making it clear that the mode of CD-ROM publication, to which the press had committed, was likely to be superseded (if not completely supplanted) by web publication. But web publication was so different from traditional publication and there were (and still are) so many uncertainties about it that the press was not prepared at that time to devote yet more resources towards yet another new direction, with a still more uncertain outcome.

I spent much time in 1999 trying to find a way past these problems. For a period, it seemed that De Montfort University, where the project and I were then based, might be able to set up a publishing arm to carry on where Cambridge thought it could not. However, these discussions amounted to nothing. Finally, the idea came to me: I could publish the Hengwrt edition myself. Not only might we publish this and later project materials; the sudden disappearance of digital publishers able to cope with enterprises of our scale (as opposed to the mass publication strategies of the likes of Chadwyck-Healy, now ProQuest) meant that there might be other scholars with born-digital work looking for a publisher. We might be able to publish these, too. I was strongly advised that I should set up a private limited company as the publisher, rather than publish them myself as a personal enterprise. The company structure would provide me (and my family) with protection if things went badly wrong and also provide some confidence to institutions with which we were dealing. For this, I needed the permission of my employer, who also required safeguards against conflicts of interest. All this took some time. Permission from De Montfort did not come until the end of May 2000. I borrowed some money to pay the National Library of Wales and the University of Sheffield (who decided that it, too, needed payment for the transcripts on which project staff at Sheffield had worked), filled out various forms, and Scholarly Digital Editions was born.

On 25 October 2000, to coincide with the sixth hundredth anniversary of Chaucer's death, we published the *Hengwrt Chaucer Digital Facsimile*, edited by Estelle Stubbs. Since then we have brought out a further ten editions. I give the approximate sales figures for

each, with its publication date and the usual price charged for an individual copy (table 1).

Table 1. Works Published by Scholarly Digital Editions

Edition	Date	Price	Copies Sold (approx.)
<i>The Hengwrt Chaucer Digital Facsimile</i> , ed. Estelle Stubbs	25/10/2000	£82	320
<i>The Bayeux Tapestry Digital Edition</i> , ed. Martin Foys	31/10/2002	£30	2500
<i>The Hengwrt Chaucer Standard Edition</i> , ed. Ceridwen Lloyd-Morgan	3/11/2003	£20	80
<i>Caxton's Canterbury Tales: The British Library Copies</i> , ed. Barbara Bordalejo	8/10/2003	£20	150
<i>The Miller's Tale on CD-ROM</i> , ed. Elena Pierazzo	5/5/2004	£50	110
<i>The Parliament Rolls of Medieval England</i> , eds. Christopher Given-Wilson et al.	5/5/2005	£50	560
<i>The Nun's Priest's Tale on CD-ROM</i> , ed. Paul Thomas	3/5/2006	£50	70
<i>Dante: Monarchia</i> , ed. Prue Shaw	3/5/2006	£50	60
<i>Leiden Armenian Lexical Textbase</i> (Internet), ed. Jos Weitenberg	21/2/2008	£50	36
<i>A Digital Catalogue of the Pre-1500 Manuscripts and Incunables of the Canterbury Tales</i> , ed. Daniel W. Mosser	10/5/2010	£50	16
<i>Dante Alighieri: Commedia</i> , ed. Prue Shaw	1/12/2010	£100	16

Did we make money? So far as these publications are concerned, the answer is a simple, flat no; we lost money on these publications. Like Cambridge, we had to cope with the direct costs of production: licensing, interface programming, and packaging. As is common for new companies, we overspent on direct costs in the first years. We paid too much to the National Library of Wales for the image rights and too much to the University of Sheffield for rights to the transcriptions for the *Hengwrt Chaucer*. We paid too much to OUP for the rights to translations included in the *Bayeux Tapestry*. We created beautiful, but very expensive, packaging and publicity materials for the *Hengwrt*

Chaucer, and (to a lesser degree) for the *Bayeux Tapestry*. After those two publications, we became shrewder about how we packaged our CD-ROMs, and developed for the *Caxton's Canterbury Tales* and later publications a standard packaging which cost about one fifth of the *Hengwrt Chaucer* packaging. The excessive sums we paid in the first instance, however, were a weight against our revenues for years afterward.

Note too that our revenues were rather lower than one might presume from this table alone. One cannot calculate our revenue simply by multiplying price by copies sold. Many copies were sold by resellers, who kept between 40 to 80 percent of the sticker price for themselves. For many copies of the *Bayeux Tapestry*, we received a payment of less than US \$6 after every reseller along the sales chain had taken their cut. Indeed, the company only survived the first years because we also carried out a considerable amount of paid consultancy. This allowed us to pay a computer programmer (originally Andrew West, later Zeth Green) to work for us from 2000 to 2009, and to contribute to the necessary programming for the editions. We were able to charge for the consultancy on a "cost plus" basis: the "plus" subsidized our publication program. The simplest measure of profit, for a limited company like us, is tax paid. In the first years of the company we built up such a substantial deficit that up to now we have not reported any profit after allowances for losses for a single year. We expect this will change in 2013 as the company has finally been trading at an annual profit for several years, and profits are at last on the point of overtaking accrued losses. But we would have preferred to be making profits and paying taxes years ago.

This table also is striking because of the asymmetry of the sales. A few publications sold very well: the *Bayeux Tapestry* may claim to have sold more copies than any comparable scholarly digital edition ever. The *Parliament Rolls* also did exceptionally well, especially when one factors in institutional Internet licences (fifty-one sold to now), although our revenues from this publication had to be shared with the National Archives. One may note that we experimented too with cutting the price, marketing Barbara Bordalejo's excellent *Caxton's Canterbury Tales* for just £20. Sales did not increase. The *Hengwrt Chaucer* has still to recover its production costs, despite a healthy sales level, because of the excessive fees and costs we paid in the first year of publication. Other publications have failed to meet their production costs. Particularly notable are the poor sales of four publications: *The Miller's Tale* and *The Nun's Priest's Tale* (both Chaucer) and

the *Monarchia* and *Commedia* (both Dante, both edited by Prue Shaw, now Emeritus Reader in Italian Studies at University College London). All four are made to a model that is effectively unique to Scholarly Digital Editions. They contain full transcripts, images, collations, and analyses of multiple versions of one text – more than fifty for the *Miller* and *Nuns' Priest*, twenty-two for the *Monarchia*, and nine (including two modern print editions) for the *Commedia*.

The preparation and publication of editions made to this model is more demanding than for the presentation of a single version, such as the *Hengwrt Chaucer*, or other digital facsimile single-witness publications, the model to which nearly all digital scholarly editions conform.⁷ To make these publications, not only must one gather images of each page and transcribe them, but one must also compare all the different versions word-by-word, to create a complete and accurate record of the differences among the versions (a collation). Creating a collation, as these publications all do, is not simply a matter of running a program (such as the widely used “diff”) across the versions. All these publications use an interactive, computer-assisted collation system developed by myself called Collate in the early 1990s. This requires that the scholar look at the collation of every word, one at a time, and gives tools to adjust the differences found. The variants judged to be insignificant by the scholar can be removed (usually, variations in spelling or punctuation), while those judged to be significant can be optimized (additions or omissions, phrases, and the like), the aim being to represent as clearly as possible the differences among the witnesses at that point. This is an exacting and time-consuming process and requires so high a level of scholarly judgement that it is usually reserved for a senior editor.

These four publications have a further feature. One of the major innovations in textual scholarship in the last twenty years (though one still little known in Anglophone textual editing) has been the application of phylogenetic methods. Phylogenetic methods were developed in evolutionary biology for the creation of hypothetical “family trees” –relational groupings of populations of organisms through the analysis of the degree to which they share characteristics. Textual scholars can posit relations among manuscripts in the same

⁷ For more on models see Kevin Kiernan, “Digital Facsimiles in Editing,” in *Electronic Textual Editing*, ed. Lou Burnard et al. (New York: Modern Language Association of America, 2006), 262–68; and Elena Pierazzo, “A Rationale of Digital Documentary Editions,” *Literary and Linguistic Computing* 26 (2011): 463–77.

manner, by assessing the number of variants they share.⁸ All four of these publications (like the Cambridge *General Prologue*) submit the full data of the agreements and disagreements among the witnesses recorded in the collation to phylogenetic programs, which create family trees that hypothesize the way in which the manuscripts might relate to each other.

These four publications have a level of sophistication far beyond any of our other publications and no other publisher has produced anything like them.⁹ One could argue that they represent attempts to reach the pinnacle of textual scholarship, the making of an edition of a text existing in many versions. The *Commedia* edition is also significant because it contributes to (and indeed was occasioned by) a vigorous debate among Dante scholars concerning the editing of the *Commedia*, sparked by Federico Sanguineti's controversial 2001 edition. Yet we have sold only sixteen copies in two years (our Italian co-publisher, Sismel, may have done better). Kevin Taylor observed that what we were trying to do at Cambridge was ahead of its time.¹⁰ One may wonder if indeed there will ever be a time for editions such as these. Are scholars sufficiently interested in editions of this type to justify the immense effort and expense of their production? Or should we think of these efforts as editions for the ages, not to be consumed for a single school assignment (as so many of our purchasers of the *Bayeux Tapestry* have used that publication), but to wait years and years for the readers who will use them to ask questions we cannot imagine?

This returns to the question that faced us when we began Scholarly Digital Editions. Cambridge was not able to find a satisfactory model for electronic publication of born-digital materials such as these. Have we been able to find a model where Cambridge could not? Speaking commercially, we have not: witness our failure to make a profit after ten years, even excluding from our costs any of the time that I have

⁸ On the success of phylogenetics in stemmatics, see the list of publications at <http://www.textualscholarship.org/newstemmatics/bibliography/index.html>.

⁹ The nearest to these, in presenting a text in many versions, are the ongoing editions of the Greek New Testament at Birmingham and Münster (<http://www.iohannes.com>, and <http://nttranscripts.uni-muenster.de/AnaServer?NTtranscripts+o+start.anv>) as well as the Liverpool Cancioneros (<http://cancionerovirtual.liv.ac.uk/AnaServer?dutton+o+start.anv+ms=BC&sms=3>). All these use the same configurations of transcription and collation software as our four publications, but without the analytic materials described here.

¹⁰ E-mail from Kevin Taylor to the author, 15 November 2012.

spent over the years on these editions. (Barbara Bordalejo, too, has worked on all of them.) However, while continuing commercial success is the shortest way to maintain a long-term publication model, it is not the only way. One might choose to continue a publishing enterprise, as we have for the last twelve years, not for reasons of profit, but because it brings other benefits that sufficiently offset the expenses of effort and money. This has been the case with Scholarly Digital Editions – there have been such benefits. The company has continued through years even when it was losing money because these benefits more than compensated for the loss and the risk.

The first such benefit is the body of editions that we have been able to bring out and their value to the public. Several of these, notably the *Bayeux Tapestry*, the *Hengwrt Chaucer Digital Facsimile*, the *Parliament Rolls of Medieval England*, and the *Monarchia*, have won honours and extraordinary reviews. The *Bayeux Tapestry* has won almost every award available to it including a *Choice* Outstanding Title award and the 2005 prize of the International Society of Anglo-Saxonists for the best new edition in any medium. *History* magazine began its review of the *Parliament Rolls* with the declaration, “Nobody involved in any aspect of medieval research can afford to do without this publication.”¹¹ It is also the case that if we had not published these, no one else would have. As explained earlier, the original motive for setting up the company was to publish Stubbs’s edition of the *Hengwrt Chaucer* when there seemed no other way of achieving this publication. Following publication of the *Hengwrt Chaucer*, Martin Foys (then an early-career scholar still establishing his academic path) approached us at our stall at the Kalamazoo Medieval Congress in 2001. He had been working since 1996 on a digital edition of the *Bayeux Tapestry*, which was nearly complete and quite stunning, but he could not find anyone to publish it. Would we be interested? We said yes, and this became our best-seller. In 2004, a similar problem arose over the *Parliament Rolls*. This project began in 1997 with Leverhulme Trust funding and Cambridge agreeing to act as publisher. I worked (with funding from Cambridge) with the project team, headed by Chris Given-Wilson, to capture the data and create the whole edition, which was to be launched in 2005. However, by 2004 Cambridge was having doubts about this publication similar

¹¹ For these and other reviews, see Scholarly Digital Editions, “Scholarly Digital Editions: Reviews of Some of Our Publications,” accessed 10 June 2013, <http://www.sd-editions.com/reviews/index.html>.

to those that Kevin Taylor had expressed to me in 1999. During the summer of 2004, there were intense discussions between myself, the National Archives, and Cambridge. They culminated in an agreement that Scholarly Digital Editions, the National Archives, and the History of Parliament Trust would together publish the *Parliament Rolls*, with Scholarly Digital Editions leading the physical production and distribution. We were able to publish on schedule in May 2005. At eight thirty in the morning of the first Thursday of the Kalamazoo conference we had a line of twenty scholars waiting to purchase the newly minted CD-ROM. One could argue that this was a temporary problem only for the *Parliament Rolls*, which is now also published on the British History Online website, but I cannot see how any of our other editions would ever have been published if we had not been there.

We know, too, that these editions have been used. The 2,500 copies of the *Bayeux Tapestry* represent only a fraction of the readers this publication has reached. Teachers frequently use the CD-ROM as the basis for a lecture to a class and often as the basis for an entire course. With these mass uses, and allowing for a large number of illegal copies as the CD-ROM has no copy protection, we estimate that over a hundred thousand readers have seen this publication. One might make a similar calculation for the *Parliament Rolls*, where, in addition to the 560 CD-ROM sales, some fifty-one institutions have purchased licences permitting unlimited Internet access by all students, staff, and associates to the online version.

The second benefit, beyond commercial revenue, that Scholarly Digital Editions has brought, is that its publications have advanced the careers of the scholars whose work they embody. At the time of the *Hengwrt Chaucer*, Stubbs was on a short-term research contract. She has remained in university employment since, working on a variety of medieval-related research projects, an outcome likely to have been aided by the credit she earned as editor of the *Hengwrt Chaucer Digital Facsimile*. Foys is now an associate professor of English at Drew College, the executive director of the International Society of Anglo-Saxons, the co-director of several impressive digital humanities projects, and the author of two well-received monographs. To the extent that we were able to publish his *Bayeux Tapestry* at a crucial moment of his career, this edition helped him on his way.

The person whose academic career has been most affected by Scholarly Digital Editions is, of course, myself. At the time of writing, I hold a named chair in the Department of English at the University

of Saskatchewan. As a late starter in academic life (I did not begin graduate study until my mid-30s), this is more than I might ever have expected to achieve. I have been involved in Scholarly Digital Editions for over half my academic career. In some important respects, my work with the company helped me win this post and perhaps was crucial to it. The editions that I had achieved myself through the company and (even more) the editions that I had helped others achieve were very much on the minds of those who appointed me. However, what if, instead of devoting so much time, energy, and resources to creating editions for the company, I had followed the more usual career path for a digital humanist, concentrating on building a digital humanities centre, nurturing digital humanities initiatives, attending digital humanities conferences, and editing collections of essays on digital humanities? It is true that my preoccupation with making scholarly editions in digital form, and helping others make them, has perhaps taken my work out of the eye of digital humanists, but this does not much bother me. We made our editions for students and readers of Chaucer, of Dante, of Armenian, and of Medieval English history and literature, not for digital humanists alone. I created Scholarly Digital Editions because I value scholarly editing.

In one respect, however, it has to be said that involvement with the company did damage my career. Although it has made no profit, Scholarly Digital Editions is a for-profit company. Many academics are suspicious of commerce and profit.¹² This has created difficulties in two areas. The first is that several edition initiatives I have explored with other scholars have come to nothing because either those scholars, or the grant agencies from whom they were seeking funding, thought that publication should be on a free-to-all online site, and not by a commercial agency, albeit one selling CDs and online access at relatively cheap rates. Eventually, those scholars chose to publish with a digital humanities centre. Two examples of this were the Emily Dickinson Papers project, led by Martha Nell Smith, which went to the Institute for Advanced Technology in the Humanities in

¹² One may see this suspicion in the comment “\$\$\$\$\$\$” by Sebastian Rahtz on the Scholarly Digital Editions Facebook page (<https://www.facebook.com/ScholarlyDigitalEditions?fref=ts>) and in the anonymized reviewers’ comments on several grant submissions in which we were named as the project publisher, remarking unfavourably on the notion that the project output might be made available through paid-for access, even though the same reviewers saw no difficulty in the publication of other project outputs in paid-for print materials.

Virginia¹³ and the Fine Rolls Project, led by the National Archives, which went to King's College London.¹⁴ In both cases, we went a considerable distance towards a publication agreement and our inability to offer completely free access was certainly a factor in the project principals' choice to go elsewhere. Had I chosen the more usual digital humanities path of publishing these materials free to all from a university-hosted (and, ultimately, grant-funded) site, I might have been a partner in those projects.

The second area of difficulty was specific to a particular project, the Canterbury Tales Project. I explained earlier that Scholarly Digital Editions was founded specifically so that Estelle Stubbs's edition of the *Hengwrt Chaucer* could be published. The *Hengwrt Chaucer* arose from work done by Stubbs, Michael Pidd, and others at Sheffield in the early years of the Canterbury Tales Project, then directed at Sheffield by Norman Blake and supervised by me, transcribing the manuscripts of the *Tales*. On Norman Blake's retirement from Sheffield in 1998, the project moved to De Montfort University and I became formal director of the project, a move that coincided with the shift of publisher from Cambridge to Scholarly Digital Editions. The plan was to continue work on the transcripts begun at Sheffield and to publish these eventually in some suitable form. Sheffield was a partner in a project created to this plan and funded by the UK Arts and Humanities Research Board and led by myself at De Montfort from 1999 to 2004. An agreement between the two universities in 2000 appeared to ratify this plan, without achieving formal signature. Multiple assurances were given by staff at Sheffield in the years from 2000 that full agreement would be reached, and several times we seemed on the point of agreement. Accordingly, from 2000 right up to 2009, I and others at De Montfort and then at Birmingham continued to work on these transcripts, believing that eventually we would be able to publish them.

However, Stubbs and Pidd opposed this agreement. In time, they were able to persuade authorities at Sheffield to refuse permission to Scholarly Digital Editions to publish any materials arising from Sheffield. This constituted the greatest part of the work I had done since 2000 and also the most significant publications we had planned

¹³ Dickinson Electronic Archive, accessed 10 June 2013, <http://archive.emilydickinson.org/>.

¹⁴ Henry III Fine Rolls Project, last modified 21 August 2013, <http://www.finerollshenry3.org.uk/home.html>.

for the company. Our correspondence with Sheffield makes clear that the for-profit status of Scholarly Digital Editions was a factor in Sheffield's refusal. One letter from Sheffield expresses the belief that files provided to the project for academic purposes were used without permission in the company's for-profit commercial publications.¹⁵ This did not happen, and the company never made a profit from any publications related to the *Tales*, including Stubbs's *Hengwrt Chaucer*, for which the company was expressly founded.¹⁶ Ultimately, Stubbs and Pidd published the transcripts – that is, those produced at Sheffield between 1994 and 1999, but apparently without the later work done at De Montfort and Birmingham – through the Sheffield Humanities Research Institute.¹⁷

For me, what we were able to publish justifies the creation of the company and all the work we did over the years. We did not achieve all our aims. Our original intuition was that there was a place for small specialist publishing companies in the digital and scholarly editing realms: one needed only a computer and Internet access to set up such a publisher; many scholars are engaged in editing projects and are looking for a publisher; and we could reduce overheads to a minimum. Surely, others would follow and we would (after a decade or so) be one among many, or at least several, small and specialist companies busily jostling for space in a crowded market. A decade has followed, and not only are we still on our own, but we have been able to publish much less than we might have done. This is partly because of the impasse with the University of Sheffield, but it is also due to a factor that we did not anticipate in 1999: the rise of a few large digital humanities centres acting as publishers in their own right. Some nine of the nineteen projects listed as actively developing at King's College London¹⁸ have a significant textual editing component; any of these might have been published by Scholarly Digital Editions,

¹⁵ Letter from Sarah Fulton, University of Sheffield, to the author, 11 March 2008.

¹⁶ The Sheffield transcripts in the *Hengwrt Chaucer* were, as noted above, licensed, paid for, and fully acknowledged. The Sheffield transcripts in *The Miller's Tale* and *The Nun's Priest's Tale* were fully acknowledged, and licence agreements with Sheffield were sought by the company but never finalized. No part of any Sheffield transcription was used in Barbara Bordalejo's edition of *Caxton's Canterbury Tales*.

¹⁷ The Norman Blake Editions of the *Canterbury Tales*, "The Multitext Edition," University of Sheffield, last modified 2013, <http://www.chaucermss.org/>.

¹⁸ Department of Digital Humanities, "Current Projects A–Z," King's College London, accessed 10 June 2013, <http://www.kcl.ac.uk/artshums/depts/ddh/research/projects/current/index.aspx>.

as might many of the completed projects (such as the Henry III Fine Rolls project and the Swift Archive). One can find many similar projects on the websites of the Maryland Institute for Technology in the Humanities,¹⁹ the Institute for Advanced Technology in the Humanities,²⁰ and, of course, at the Humanities Research Institute in Sheffield, including the Canterbury Tales transcripts originally to be published by ourselves. All these sites can offer what Scholarly Digital Editions cannot: free access to these publications online to anyone. Given this, and the suspicion many scholars have of anything which might be commercial, we cannot compete. The nearest effective competitor to Scholarly Digital Editions, in that it publishes scholarly editions similar to those we publish and charges for them, is the University of Virginia's Rotunda Press, established with large grants from the university and the Mellon Foundation.

There is much to celebrate in the publishing achievements of digital humanities centres over the last decade. A few points should be made, however. First, no one involved in a digital humanities centre over the last decade has ever done what I and Barbara Bordalejo do every May at the Scholarly Digital Editions stall at the Medieval Congress in Kalamazoo. We show a scholar one or more of our publications; we demonstrate them; and then we wait to see if they take out their credit card and buy. Like Cambridge University Press and like every general-sale publisher that has ever existed, we live by individuals liking what we do enough to pay for it. Digital humanities centres have no such imperative. Their over-riding need is to produce a final report and a website that satisfy the funding agency. This can lead, rather easily, to neglect of the individual. A case in point is the disappearance of CD-ROMs. It is common to assert that CD-ROMs have gone the way of button-up boots, and yet we have sold nearly four thousand over a decade. Moreover, we have offered one publication both online and in CD-ROM form (*Parliament Rolls of Medieval England*) and individual buyers have chosen the CD-ROM by something like fifty to one.

How can this be? Libraries dislike CD-ROMs for obvious reasons, but individual scholars, who have to pay their own money or pay from a hoarded professional allowance, like CD-ROMs, much as they like books. Many digital publishers offer online subscriptions to libraries, institutions, and schools, but not to individuals. For

¹⁹ University of Maryland, accessed 10 June 2013, <http://mith.umd.edu/>.

²⁰ University of Virginia, accessed 10 June 2013, <http://www.iath.virginia.edu/>.

example, Harrassowitz does not offer individual online subscriptions to the Parker Library, nor does ProQuest for Early English Books Online (EEBO). Rotunda raises this bias towards the institution to a kind of art form and divides institutions into as many as six tiers ranging from “high research universities” down to high schools, and bundles “unaffiliated individuals” into the lowest tier while asking them to pay \$295 for access to the Dolly Madison Papers.

There is another result of the concentration of scholarly editing in digital humanities centres. It costs a considerable amount of money to create an edition, from data gathering to final publication, and it is not uncommon for the funding for a single edition to run to hundreds of thousands, even millions, of dollars. Because of the key roles played by digital humanities centres in the making of such editions, funding agencies prefer to give grants to partnerships between scholars and a digital humanities centre. Who, then, chooses what editions are made, and by whom? In this scenario, the choice is made by funding agencies working with digital humanities centres. There are clear dangers here. As digital editions have become more and more expensive to make, funding has flowed towards the safe option: the marquee project, with a famous author or collection.

One justification for Scholarly Digital Editions, I submit, is that we were able to publish Martin Foys’s *Bayeux Tapestry Digital Edition*. Suppose a young scholar now had a similar project, and was looking for a publisher. Where would she or he go? If he went to a digital humanities centre, the first question he would be asked is: do you have funding? This is unsatisfactory. For several years, I observed the efforts of two first-ranked and senior scholars to bring proposed digital editions to fruition: Murray McGillivray, for his edition of the Cotton Nero manuscript, and Daniel O’Donnell, for his Visionary Cross project. Finally, they have won funding, but there is something wrong with a model in which the making of a scholarly edition does not depend on a publisher or on the needs of scholars, but on the funding agency. A healthy academic ecosystem includes scholars, publishers, funding agencies, libraries, students, and readers. We have lost publishers, or rather they have been replaced by less-focussed digital humanities centres without a primary publication agenda, while funding agencies have excessive sway. Sometimes, it seems that we are heading for a world in which anonymous funding agencies decide which editions should be funded and anonymous librarians decide which editions should be available in their institutions in perfect exclusion of the individual scholar. We are not there, but we

are closer than we should be. One observes that the decisive factor in the existence of projects such as the Olive Shreiner Letters²¹ or Jane Austen's Fiction Manuscripts²² was not the needs of scholars or readers but the ability of the project proposers to persuade a funding agency to pay out a large sum of money.

Once more, I am writing of models. Where next with Scholarly Digital Editions? Can we find the model of publication which has so far eluded us, that will allow us to continue to publish challenging and outstanding editions? Consider Prue Shaw's edition of Dante's *Commedia*. We have sold just sixteen copies in two years, yet this edition may claim to be the most fully realized and valuable of all our publications. It contains transcripts and images of the whole *Commedia* in seven key early witnesses as well as a full collation and analysis of the text of these seven against the full text of two key modern editions. Most of all, it is Shaw's edition: she checked, many times over, every word of the transcripts and collation; she prepared a meticulous argument concerning the relations among the manuscripts, along the way demolishing the view of the manuscript relations offered by Federico Sanguineti. It is a tour de force of textual scholarship. If Scholarly Digital Editions had published nothing else at all, this alone would justify our efforts. We want to continue on and to add manuscripts to the seven included in Shaw's edition. How? The prevailing model for the making of a digital editions begins with the scholar's getting an agreement with a library, and proceeds to his or her securing of funding and approaching a digital humanities centre, which helps gather the materials and makes the edition. The problem here is money. To make an edition the way we did for the *Canterbury Tales* or for the *Commedia* costs so much money as to make grant funding essential. Accordingly, in the last few years we have been exploring ways to reduce radically the funding needed and hence to open up the field to many more scholars to make digital editions. We believe that if we can reduce the cost of making digital editions so much that in most cases scholars can fund the work through the resources usually available in universities, then far more digital editions will be made. The key to this is what we are calling "textual communities" – many people contributing to the gathering

²¹ University of Edinburgh, last updated July 2013, <http://www.oliveschreinerletters.ed.ac.uk/>.

²² University of Oxford and King's College London, last updated 31 July 2013, <http://www.janeausten.ac.uk/index.html>.

of the fundamental materials for an edition.²³ Our searing experience with the University of Sheffield had this fortunate result: it focussed our attention on the need for a different model of collaboration, in which people contribute to the making of editions not because of the project and what they might win from it, but because of their interest in the work itself. Practically, we express this by insisting that all contributions to the textual communities we will make (including for the Canterbury Tales) should be made available under the Creative Commons Attribution Share-Alike licence. This will avoid forever the situation where an entire project can be held hostage by a few people who worked on it.

The insistence on this licence has another important effect. Unlike many academic groups, we do not advocate the non-commercial form of the licence. Rather, we will encourage commercial entities to make use of the community-created materials of our edition. More than this, we specify that textual communities should make all collaboratively made materials available independent of any one interface. The combination of the two – the unrestricted licence and the access to all materials independent of the interface – has far-reaching consequences. It means that a commercial agent (or any agent at all, including any digital humanities grouping) can bring our materials into their own publication, re-arrange them, combine them with other materials, and create their own highly-tuned edition built on and around them. To return to the Shaw edition, we can envisage a senior textual scholar of Shaw's calibre creating a new edition of the *Commedia*, adding transcriptions and collations made by others to her edition but correcting and refining them, and then developing a new textual commentary. We see a role for Scholarly Digital Editions and publishers like ourselves in this process. We can create custom interfaces, giving better access to the materials than anyone else; we can include paid-for images unavailable on the open web; we can provide an important validating function, warranting the quality of the edition. Along the way, we think many more than sixteen people will find this useful. We assume here that editions will no longer be single print objects, but hybrid collections of multiple materials. Some of the materials will be developed by many people and will be open

²³ The Textual Communities project, based at the University of Saskatchewan, is supported by a Canada Foundation for Innovation grant; my co-investigators are Brent Nelson and Frank Klaassen. See <http://www.textualcommunities.usask.ca>.

to everyone; others will be developed by one or two and accessed through payment. The publication, moreover, will be ongoing, the editions evolving through multiple forms of multiple websites.

In 1999 we sought a model that might allow us to make the editions we wanted to make. Up to now, what we have made has conformed to a type that the librarians of Alexandria would have recognized: an authoritative scholar, supported perhaps by a tightly-controlled group of assistants, decides everything. We see in enterprises like the Transcribe Bentham project²⁴ and in analogous “citizen scholar” enterprises such as the biodiversity and bio-map projects at the Australian Museum,²⁵ pathways to a new type of partnership between scholars and readers, mixing elements of the new “social edition” with traditional authority-based editing.²⁶ Meanwhile, there will continue to be a place for commercial publishers, some of them small like Scholarly Digital Editions, to carry out their longstanding roles: to enable, to advise, to make the an edition as useful and efficient as possible, and to put it in the hands of readers.

SOMMAIRE

La brève histoire des livres en sciences humaines sur support électronique a d’abord débuté par des publications sur CD-ROM lancées par quelques éditeurs commerciaux avant de passer par des publications à compte d’auteur diffusées par des chercheurs sur le web ou par le truchement de centres d’humanités numériques. On relate l’expérience d’un petit éditeur spécialisé dans l’édition de livres électroniques en sciences humaines. On fournit des tableaux représentant des ventes de CD-ROM, DVD-ROM et sur Internet en vue de démontrer que la compagnie ne peut suivre un

²⁴ Transcribe Bentham: A Participatory Initiative, “Welcome to Transcribe Bentham,” University College London, last modified 21 August 2013, <http://blogs.ucl.ac.uk/transcribe-bentham/>.

²⁵ “DigiVol (Voluntary Digitisation Project),” Australian Museum, accessed 20 June 2013, <http://australianmuseum.net.au/Rapid-Digitisation-Project>.

²⁶ For more on the “social edition” see Ray Siemens et al., “Toward Modeling the Social Edition: An Approach to Understanding the Electronic Scholarly Edition in the Context of New and Emerging Social Media,” *Literary and Linguistic Computing* 27, no. 4 (2012): 445–61, accessed 30 May 30 2012, doi: 10.1093/llic/fqs013.

modèle commercial régulier dans le cadre de ses activités. D'autres considérations justifient cependant le travail de la compagnie et l'article prévoit qu'elle pourra continuer d'exister en offrant à un large public certaines de ses publications en ligne gratuitement tandis que d'autres portant sur des recherches plus pointues et réservés aux spécialistes seront accessibles moyennant finances.