Orietta Da Rold

Old and Middle English Texts on CD-ROM: An Overview

Our computer-mediated culture has revolutionised the way in which information is delivered and the computer has become an indispensable commodity in all aspects of everyday life. Everybody is now familiar with terms such as ‘hypertext’, ‘hypermedia’ and ‘hyperstructure’, which refer to non-linear, non-sequential information for an interactive environment to be delivered on a terminal. Material generated with this methodology uses a combination of media, is of interest to different disciplines and can be published on or off line. Although the academic community welcomed electronic publishing as a new medium for distributing and accessing information, scholars have also come to use computer technology for teaching and research. Multimedia publications could be exploited to give or search for alternative answers. Hypermedia products present readers with revisited, updated and additional material, encouraging textual scholars to explore texts in different ways; for example to ask the question ‘what is a text in the digital era?’ and ‘how can it be presented or re-presented?’

This paper aims to review electronic publications within the domain of Old and Middle English. It will investigate the off-line market of the CD-ROM, evaluating outcome, research potential and user-friendliness of these publications. This overview should establish the extent to which it is still profitable to favour the CD-ROM as means of publication and outline the reasons for this.

A CD-ROM (Compact Disk Read-Only Memory) is a data carrier, belonging to the family of optical media and “was designed as the storage format for general computer data - in addition to uncompressed audio data”. It has now developed into an ideal support for transmitting all kind of

---

1 I should like to thank Angelo Gambella, Karen Britland and Jennifer Marshall for reading a draft of this paper.
3 A medium can be described “as a means for distribution and presentation of information. Examples of a medium are text graphics, speech and music”. However, media “can [also] be classified with respect to different criteria… We classify media according to perception, representation, presentation, storage, transmission, and information exchange” Ralf Steinmetz and Klara Nahrstedt, Multimedia: Computing, Communications, and Applications, (Upper Saddle River, NJ and Prentice Hall PTR : Prentice-Hall International (UK), 1995), p. 10. Hence, for the purpose of this paper, medium refers both to the content and the physical support used to make such content available.
8 Steinmetz and Nahrstedt, Multimedia: Computing, Communications, and Applications, p. 188.; see also Nielsen, Multimedia and Hypertext : The Internet and Beyond, pp. 160-4.
media. It has a high storage capacity, low cost of production and it is physically robust; it is the ideal medium for electronic publishing.

Despite the expectations of the early 90s the market for CD-ROM never developed and the hopes for digital publications are now set on the DVD (Digital Versatile Disk). Certain disciplines more than others made use of CD-ROM. Print publishers enthusiastically embraced the new medium, as they were keen to explore electronic publishing. They were, however, inexperienced. They lacked technical expertise and the market had too few commercial prospects. As Wise explained: “The original promise of the medium, that it would create a new multimedia product that would lead to the death of the book, has not been fulfilled.” The failure of electronic publishing on CD-ROM was not only a consequence of a poor market, but it also depended on the technical support this publication required and its cost. In 1996, Susan Hockey explains:

CD-ROM is still the medium of choice for many electronic publications, largely because it is easier for publishers and librarians to handle. It fits in better with procedures for handling books and in some ways can be seen as an extension of them. However, for long term CD-ROM has serious limitations. There has been much discussion about the longevity of CD-ROMS. Since the medium has been around only a short time, no one knows whether they will be physically readable after a long period. Even if it is readable, the format of the files stored on it will most likely not be. CD-ROMs are also too small. They cannot hold sufficient data, especially when images and other non-textual media are used. Even with text only, six hundred megabytes is are not enough for a lot of text indexed in different ways.

Several years have passed and academic projects are still adopting CD-ROMs as a publishing medium, as I will discuss below. One can notice, however, that electronic publications of this kind have developed over the years. It is interesting to review Old and Middle English texts published on CD-ROM to test the use of this medium within a relatively restricted discipline. It is accepted that medieval studies are undergoing a revival, partially or mainly ignited by the implementation of technology within the discipline. In the following overview, each project or publication will be identified and considered in an approximate chronological order. They will not be treated extensively as some of them have already attracted much publicity, however, they will be assessed following some of the principles enunciated by Shillingsburg and Sperberg-McQueen. Shillingsburg considered eight main points to which electronic scholarly editions should comply: 1. Usability; 2. Transportability; 3. Archive specifications; 4. Security and Order; 5. Integrity; 6. Expandability; 7. Printability and 8. User friendliness. Sperberg-McQueen is skilchier about the requirements for electronic publications, nevertheless he argues that “accessibility without needless technical barriers to use; longevity; and intellectual integrity” should be essential requisites.

Although the two approaches are different they

---

both examine the roles of the editor and the audience of the publications, focusing on what an editor or author should consider when planning or preparing a successful electronic publication. To what extent were these publications on CD-ROM trying to fulfill these requirements? Were they pursuing a different scope? Should they be considered “book as a machine knowledge”.  

In 1996 the Wife of Bath’s Prologue of Chaucer’s Canterbury Tales appeared on CD-ROM edited by Peter Robinson. This was the first publication within the framework of The Canterbury Tales Project established in the 1990s. This publication was accepted with enthusiasm. For the first time the transcriptions and images, though in black and white, of the fifty-four manuscripts and four pre-1500 printed editions of one of Chaucer’s tales were fully accessible. This new approach to manuscripts allows editors and critics to have access to all witnesses of the poem to study the development of the text with all its evidence. The CD-ROM also collated readings and lines through all witnesses. The collation is prepared with software called Collate, especially developed by Robinson for electronic analysis. The CD-ROM provides a spelling database and a description of the witnesses by Dan Mosser. In 2000 The General Prologue was published following the framework of the former publication, but it included two additional essays by Peter Robinson: one on a textual variants workshop; the other a stemmatic commentary. These two publications do not aim to prepare a scholarly critical edition of Chaucer’s texts, they provide evidence and computerised tale by tale analysis of the Canterbury Tales readings to help to understand anew the text tradition of the poem. The main limitations of these CD-ROMs are the software used to read the data, which displays the data with several pop up windows and does not allow one to view images and transcriptions simultaneously. Information is delivered with DynaText. This system issues electronic books both for SGML and XML encoded files. During the 90s the software was considered an excellent tool which could handle texts very well, however it then became apparent that it had serious compatibility problems with other hypermedia material.

To overcome some difficulties of the earlier CD-ROMs, Peter Robinson has developed a new software: Anastasia (Analytical System tools and SGML/XML Integration Applications). It was used


22 For an overview of this system see Peter Robinson, The Transcription of Primary Textual Sources Using Sgml, vol. 6 (Oxford: Office for Humanities Communication Publications, 1994). See also: ‘Inso Announces DynaText® 4.0 and DynaWeb® 4.0, the Leading Standards-Based Publishing Solution for Delivering Documents on CD-ROM, LAN, or the Web”, <http://xml.coverages.org/dynaweb4-980317.html>.  


24 On this tool see: Anastasia: Analytical System Tools and SGML/XML Integration Applications at
in Estelle Stubbs’s *The Hengwrt Chaucer Digital Facsimile*, the last publication by the *Canterbury Tales* Project which appeared in 2000. This is a remarkable achievement among the electronic publications on CD-ROM. It contains full transcriptions and high resolution colour images of the Hengwrt Chaucer (Aberystwyth, National Library of Wales, MS Peniarth 392 D) of the *Canterbury Tales*, with collation against the Ellesmere manuscript (California, San Marino, Huntington Library MS El. 26 C 9) copied by the same scribe, diagrams of colour of inks, running heads and tale order. The codicological work that the editor carried out on the manuscript is reflected both in her observations and in the detailed comments, which appeared in the ‘about’ sections. Manuscripts and transcriptions can be viewed in different ways both side by side or individually, including opening and conjugate folios. The size of the image can be chosen at one’s convenience. The search tool is also quite advanced as it can be used as a concordance or a spelling database. Additional features are the ‘go to’ section, which allows one to jump to the textual part of interest. The publication is correlated by two essays: one by Dan Mosser on the description of the manuscript; and another by Simon Horobin on the language of the manuscript. This is a fine publication, which can be used both in the classroom and as a research tool. Its interdisciplinary strengths within the domain of medieval book production and textual studies can be exploited using the several features available on the CD-ROM.²⁵

An abridged version of this edition will be available on the internet and on CD-ROM with an introduction by Caridwen Loyd-Morgan.²⁶

Despite the short list of publications, the *Canterbury Tales* Project’s contribution to the development and understanding of electronic editions in the past ten years is unquestionable. The forthcoming program, however, is a significant one. *The Miller’s Tale* edited by Peter Robinson, *The Nun’s Priest Tale* edited by Paul Thomas and the *British Library Caxton’s Chaucer* edited by Barbara Bordalejo have been announced, and will be published by Scholarly Digital Editions. They promise interesting additional tools for study and research.²⁷

In 1997 there appeared *Geoffrey Chaucer’s Book of the Duchess: A Hypertext Edition* edited by Murray McGillivray.²⁸ The editor explains that his hypertext edition ‘is one that uses the technology of the interactive windowed screen in order to improve the print-form scholarly edition’. These improvements can be summarised in two main points: ‘the windowing that allows several different documents or parts of documents to be brought to the screen at one time, and hypertext linking, which allows the reader to choose his or her own path through the edition (guided by the editor’s present construction of links)’.²⁹ Following this strategy the editor devises a straight-forward list of links from the ‘Main Menu’, which serves as a table of content connecting the several sections of the editions. The CD contains an introduction, a reading edition (with links to glossary, notes, and sound files) and a critical edition with textual notes, transcriptions and images of manuscripts and Thynne’s 1532 edition. It provides transcripts from edited texts of ‘Chaucer’s Main Old French and Latin Sources for the Book of the Duchess’, which include abstracts from editions of Machaut, Froissart, Ovid and Statius with translations. The edition also comprises an ‘how to use’ section, an SGML versions of transcriptions, critical edition, and source texts, linking to a protected web site, which unfortunately provides only two SGML files: Thynne’s edition of *the Book of the Duchess* and Gaullaume de

---


²⁷ An overview of the new publication program and features of the forthcoming publications was presented in a paper ‘Where Next with the Canterbury Tales Project?’ by Peter Robinson and Barbara Bordalejo at the Early Book Society Conference, Durham, July 10-14, 2003.


Machaut’s *Dit de la Fonteinne Amoreuse*. There is also a link to a web site for updates and additional on-line material.\(^{30}\)

It is a user-friendly edition, which is easily accessible. Its HTML structure affords flexibility and can be viewed with any browser, although one should note that, for example, Netscape 4.5 does not read all the instructions for displaying the frame structure correctly. The frame structure is used throughout the edition to accomplish the editorial aim to view more features at the same time. A section of the edition has especially been created to compare two or more texts, images or editions. This is indeed useful, but daunting at times especially with a high number of frames to compare. It would also have been quite helpful to have a search feature for folios and transcripts, which would have minimised the use of hyperlinks and maximised the use of the edition as a research tool. The edition does have an emphasis on hypertext as the editor points out, but it is also a first step into incorporating multimedia features into an electronic edition. The sound files are an interesting teaching tool and a pleasant addition, although they can support the sound of only a few lines. Indeed, as Blake emphasised: “This CD-ROM presents a lot of useful material in a simple electronic form”\(^{31}\).

In 1993 the British Library began the Electronic Beowulf project “as one of a number of initiatives to increase access to its collections by the use of imaging and network technology”\(^{32}\), in collaboration with Kevin Kiernan of the University of Kentucky and Paul Szarmach of the Medieval Institute, Western Michigan University.\(^{33}\) This enterprise led to the publication of *The Electronic Beowulf* under the editorship of Kevin Kiernan in 1999. The edition is published on two CD-ROMs. CD 1 provides the digital colour images, transcriptions and edition of the Beowulf manuscript (British Library, Cotton Vitellius A. xv). A comprehensive glossary, eighteenth-century transcriptions and copies of the 1815 first edition with early nineteenth-century collations of the manuscript correlates the facsimile edition of the Beowulf manuscript. Users can choose the material to view and its arrangement, and can perform sophisticated searches of the edition and the transcriptions with a pop up window. An interesting feature of the edition is the close scrutiny of letters concealed by the paper frame protecting each folio after the restoration carried out in 1845, which can be done by clicking on a small blue square button on the toolbar. This feature also provides an explanation, appearing on the right frame of the screen, of the damaged letter.\(^{34}\) CD1 includes background information about the edition (Guide) and a section with the Project History and a list of online articles by Michael Ellis, Kevin Kiernan, Andrew Prescott, and Elizabeth Solopova. CD2 contains guides to the use of *The Electronic Beowulf*, and the Supplement provides access to the digital images of Vitellius A. xv both as ‘Facing pages’ and ‘Sheet Collation’. It includes diagrams of the reconstruction of the quires now lost after the fire, which almost destroyed the manuscript in 1731. This edition is an excellent tool for study and research, although users may encounter some difficulties in loading the edition on a terminal.\(^{35}\) There is also little information on the technology used for the interface. However, a search in the SGML files of the transcription is provided, and a discussion on how they were prepared is also included in an article by Solopova.\(^{36}\)
In 1999 another project published its first electronic publication: *The Piers Plowman Electronic Archive, Vol. 1: Corpus Christi College, Oxford MS 201 (F)*, edited by R. Adams, H. N. Duggan, E. Eliason, R. Hanna III, J. Price-Wilkin and T. Turville-Petre. This was followed the year after by *The Piers Plowman Electronic Archive, Vol. 2: Trinity College, Cambridge, Ms B. 15.17*, edited by T. Turville-Petre and H. N. Duggan. Both publications appeared on CD-ROM published by Ann Arbor: University of Michigan Press & SEENET. These are, to date, the first two electronic publications of the *Piers Plowman* Archive whose long-term goal ‘is the creation of a multi-level, hyper-textually linked electronic archive of the textual tradition of all three versions of the fourteenth-century allegorical dream vision Piers Plowman’. The publications of the project are devised as single electronic editions of the manuscripts of the Piers Plowman tradition. Volume 1 is the edition of Corpus Christi College, Oxford MS 201, which contains a preface and an introduction with the transcriptions and the colour images of the manuscript. It comprises, in appendix, the text of the manuscript prepared for TACT concordance programs, a scribal linguistic profile, essays by A. I. Doyle on the scribe of the manuscript and by Linda Lee on the Binder together with a ‘ReadMe’ section for help. The edition is organised with pop up windows which display the transcriptions of each *passus* with notes where relevant. The transcripts can be shown according to different styles: ‘Scribal, Critical, AllTags, Diplomatic’. Each display performs a different function and is tagged on a line by line basis. A search tool is provided and images are linked to the transcripts, but cannot be shown alongside the text, although zooming in and out is possible. The material available in the publication is very good, perhaps limited by a software which does not allow the flexibility of search and display that one may desire. All the text is tagged in SGML and is viewed with MultidocPro. The proliferation of pop up windows is clearly organised and offers very useful information indeed, but perhaps an interface which could provide a frame based access to various material would be less obstructive and would display the material more efficiently. Nevertheless, this is an interesting publication with a good emphasis on scribal practice and linguistic features.

The last CD-ROM, so far available within the Old and Middle English domain, though marginally, appeared in 2000. It contains a selection of forty colour images of British Library Cotton MS Nero D.iva, commonly known as The Lindisfarne Gospels, an eighth century manuscript of the gospels of Mathew, Mark, Luke, and John in Latin, copied at Lindisfarne, which was glossed in Old English and beautifully illuminated throughout. The CD-ROM is a macromedia product and can be viewed by clicking on the image and dragging the image across. It has a 2D feature named ‘turning the page’, which literally give the effect of the page being turned. Each page has a textual feature, explaining some of the main elements viewed in that page, correlated by an audio file if one wishes to hear this explanation. The CD-ROM includes a section with background information and an essay by Janet Backhouse on the artistic and historical value of the manuscript. This publication differs in many ways from the one described above. It does not include transcriptions of the selected folios or any search facilities, the emphasis is on the book and on the software which allows one to appreciate the manuscript in all its glory. The selection is now available on the web without the essay by Backhouse, alongside other e-books featuring the turning the page software.

Tito Orlandi emphasised the importance of the theoretical approach to the transfer of a text from one medium to another. He was calling for a better understanding of what an electronic edition could be. He explained: ‘We must also find a way of obtaining from the one dimension of the electronic support the same possibilities of the bidimensional environment, typical of manuscript and print. In

39 For information on this product see: “Multidoc Pro® Product Brief” at <http://xml.coverpages.org/multidocPro.html>.
41 See “Turning the Pages on the web” at <http://www.bl.uk/collections/treasures/digitisation.html>.
other words, we must make explicit the meaning of all the components of the printed object”.

The overview of the electronic publications discussed above encompasses the idea expressed by Orlandi. They are all remarkable attempts at thinking about a different way of looking at texts, images and books from the printed or the manuscript page. Some of these publications encountered the limitations of technology, but they thought and rethought the problem of the mise en page of information from one medium to another. They fulfil the requirements envisaged by Shillingsburg and Sperberg-McQueen, although the problem of longevity is still an issue. Scholars working on texts came to an agreement that SGML and XML should be a standard to be adopted for the flexibility of use and adaptation. The TEI consortium was founded to reach a compromise among scholars working with electronic texts and this should be the way forward. Another solution to longevity is the availability on the internet of editions which appeared on CD-ROM. Cases in point are the British Library turning-the-page edition and the edition of Hengwrt. However, publication on the internet means public access (unless the site is restricted), which implies copyright issues. Availability of images of manuscripts, for example, on websites, which do not belong to the copyright holder, could be a problem. Moreover, material usually duplicated on line does not have the same content of a publication on CD-ROM, for obvious commercial reasons. The fashion for so called ebooks started with the turn of the century and now ebooks are readily available to download from the internet at little cost. Recent research considered how the ebook market is increasing its potential and hence its business and a recent project at the AHDS (Arts and Humanities Data Service) Centre for Literature, Languages, and Linguistics based at the Oxford Text Archive is studying this phenomenon. Is this the way forward? As complex hypermedia publications, would ebooks be easily available and downloadable? Are CD-ROM publications still worth considering?

At present despite the limitations that an editors and a publishers can face when choosing this medium it is still one of the most practical ways of making information available. In an ideal world we would be able to access everything on the internet, be it in a research library with free internet access. As it is reality differs. In her conclusive chapter on Electronic texts in the Humanities, Susan Hockey writes: ‘Electronic texts should really be treated as a supplement to books, to use for those functions which are difficult to do with books, for example searching, analysis, and hypertext linking. It is easy to create books or printed material from properly structured electronic text, but the creation of books should be just one of many applications”. Following from here, one should also keep on experimenting with the mise en page of electronic publications, stretching technology even to the 3D if possible. I believe that at the moment the best way to exploit the data from a user view point is to keep on publishing on CD-ROM. This medium allows a publication to be both a tool for research and a guide on the recent findings. A publication on CD-ROM is thought to have a finite shelf-life, however, it can still be advantageous, as it can be accessed quickly, used readily in the classroom and it presents certain kinds of information (particularly images) effectively. A publication of this sort can be designed to accommodate textual and contextual information that can rapidly be searched and compared.

---

43 “Welcome to the TEI we site” at <http://www.tei-c.org/>.
46 For an overview on this project see: “An Investigation into Free Ebooks” at <http://ota.ahds.ac.uk/ebooks/JISC/>.
47 Hockey, Electronic Texts in the Humanities, p. 166.