Current Trends of Computing in the Humanities in England

Current Trends of Computing in the Humanities in England

Dr. Marilyn Deegan*

Abstract

Computerizing the materials used by humanistic scholars requires sophisticated and innovative use of the latest of new technologies as these are areas which pose particularly complex problems of mixed media data handling. The popular view of humanists is that they are either non-technological or that their needs are extremely simple, far simpler than those of scientists; perhaps limited to the use of word-processors and a little e-mail. This is far from being the truth: the key limiting factor on the use of computers in the humanities in the past has been the inability of the technology to handle many of the highly technical and abstruse problems of data handing which are dealt with daily in humanities scholarship. Now, however, the technology is rapidly catching up with the disciplines in its ability to handle the varieties of media: text, images, graphics, sound, and video.

Computers are now used by humanists around the world for many different applications in a large number of subject areas, for quantitative analysis of those topics which are amenable to numerical study, such as the analysis of datasets for historians, for the study of paintings and artefacts by art historians and archaeologists, for the sophisticated study of textuality by literary critics and textual editors, for the study of the moving image by film theorists, and many, many more. The topics which I should like to consider in this paper are: electronic publishing and the changes which are being wrought in the publishing industry; the uses of networks; image digitization; hypermedia and multimedia; and finally the uses of computers in humanities teaching. In particular, I will be looking at the current state of these in Britain.

* Centre for Humanities Computing, Oxford University   Fax/Phone: 44-865-273221
  Computing Services : 13 Banbury Road Oxford OX2 6NN UK   E-mail : marilyn@vax.ox.ac.uk

Introduction

'The humanities' is actually a large number of extremely diverse disciplines, which include text-based studies: literature in all languages, history, philosophy, theology, etc; image and artifact-based studies: including archaeology and art history, and film and media studies; and linguistic studies: the study of all languages, classical, medieval, and modern, linguistics; sound-based studies: music and spoken corpora. Computerizing the materials used by humanistic scholars in these subjects requires sophisticated and innovative use of the latest of new technologies as these are areas which pose particularly complex problems of mixed media data handling. The popular view of humanists is that they are either non-technological or that their needs are extremely simple, far simpler than those of scientists; perhaps limited to the use of word-processors and a little e-mail. This is far from being the truth: the key limiting factor on the use of computers in the humanities in the past has been the inability of the technology to handle many of the highly technical and abstruse problems of data handling which are dealt with daily in humanities scholarship. Now, however, the technology is rapidly catching up with the disciplines in its ability to handle the varieties of media: text, images, graphics, sound, and video. Thus, it is not the naiveté of the users and the relative simplicity of their needs which has prevented academics in these subject areas from engaging with computing in the past, but rather the opposite: many felt that there was no use for them in a machine which could perform ever larger calculations at ever greater speeds, and it is only with the ability of the computer now to handle mixed media almost as well as it handles calculations that it is now a machine which is seen to be a natural resource in these subjects.

Humanities computing has, in the scope of the history of computation, a relatively long history itself. Indeed, within a year of the development of the first stored program computer in 1948, Father Roberto Busa was making plans to produce a complete word index and concordance to the works of Thomas Aquinas. The first volume of this appeared in 1973[1]. The earliest uses of computers in the humanities generally involved the analysis of literary texts, and it is a great tribute to the diligence of the early pioneers of computing in the humanities that they persevered despite some of the difficult problems of input and machine representation of character sets. For instance, input was via punch cards or punch tape, and lower case letters could not always be represented. Only the Roman alphabet could be used with any ease, and non-Roman character sets remained one of the most long-standing and intractable problems which has recently been solved. It is also remarkable that humanists were even prepared to teach themselves the arcane rules of statistical

analysis in order to study their text. Anthony Kenny for instance, a classical scholar and philosopher, even wrote a book on the statistical analysis of style for the use of other textual scholars[2].

Computers are now used by humanists around the world for many different applications in a large number of subject areas, for quantitative analysis of those topics which are amenable to numerical study, such as the analysis of datasets for historians, for the study of paintings and artifacts by art historians and archaeologists, for the sophisticated study of textuality by literary critics and textual editors, for the study of the moving image by film theorists, and many, many more[3]. The topics which I should like to consider in this paper are: electronic publishing and the changes which are being wrought in the publishing industry; the uses of networks; image digitization; hypermedia and multimedia; and finally the uses of computers in humanities teaching.

Electronic Publishing

There has been a huge increase in the number of publishers wishing to engage in publishing materials in electronic form in the last year or two. Most of these are conventional academic and trade publishers, although there have also been some firms specifically established purely to create or to publish CD-ROMs. A recent report on electronic publishing presented to the Sir Stanley Unwin Trust states:

> 'The market for electronic books may be in its infancy, but it is growing fast, and while in some sectors it may complement the book trade, it is encroaching and may threaten to supersede it in other areas, where the advantages of compressed storage, sophisticated retrieval and cross-reference, or multimedia display are particularly relevant'.

The report also quotes Peter Kindersley of the publishers Dorling-Kindersley as saying that the book has a future, but he is not sure that this will be in paper form[4].

Electronic publishing has a number of advantages and disadvantages for the scholar, for the publisher, and for the custodians: those we charge with looking after, storing, and cataloguing our scholarly materials.

Advantages

One of the key advantages in the modern publishing industry is that CD-ROMs are much cheaper to produce than weighty and complex reference works produced in book form; the objects themselves are small, and smaller print runs can be contemplated, meaning that there is less need for expensive warehouse storage. Electronic publications allow easier

retrieval of complex information structures and can permit better integration of different media, for instance an encyclopaedia can include video clips and sound as well as text and still images. It is also useful that resources can be more easily shared: for example CD-ROM-based teaching materials can be mounted on a network for access by a whole class of students simultaneously, and information which changes rapidly can be updated more easily, quickly, and cheaply.

Disadvantages

As with anything in life, there are always disadvantages which must be offset against advantages, particularly with something new. For example, the rapid pace of change of hardware and software brings the potential of rapid obsolescence; and the constituent materials need to be permanently 'managed' which has never been necessary in quite the same way with books. Also, the integration of large volumes of diverse media brings problems of copyright and intellectual property rights on an unprecedented scale. It has to be assumed, too, that all potential users will have access to the relevant hardware; and also that they will require a level of training which is not needed for them to use books.

Some examples

A growing number of different kinds of electronic publications are now being produced for the humanities scholar, and so I will try to describe some of the major ones currently available or which are coming soon. First of all, there are a number of scholarly editing projects which are adopting electronic publication as the most effective means of presenting the complex materials which push the printed book to the limits of its organizational possibilities. These include the Canterbury Tales Project based at Oxford University which will be the first publication of the new Cambridge Electronic Editions series from Cambridge University Press. The Project intends to make available, over a ten-year period, full transcripts of the text of every manuscript and pre-1500 printed edition of the Canterbury Tales, together with computer images of every page of every manuscript and early edition, collations of all these texts, and analyses of the textual tradition based on the transcripts. Cambridge University Press is to publish these materials in CD-ROM form commencing in late 1994 with the fifty-five manuscripts and four pre-1500 printed editions of the Wife of Bath's Prologue[5]. Major electronic editions of Shakespeare, Johnson, Hardy, and Yeats are currently being discussed by the Press and scholars in these areas.

The Canterbury Tales Project is being published by a major academic press, but there are other routes to reputable publication. A group of scholars in North America and

Europe has recently established a society for the publication of electronic scholarly editions of early texts analogous to the Early English Text Society which has existed for more than a hundred years to publish scholarly editions in book form. The society is known as SEENET—Society for Early English and Norse Electronic Texts—and is being published by University of Michigan Press. The first proposed work in this series is the *Piers Plowman* Electronic Archive which is being constructed by a team of scholars in England and the USA and which plans to create a multi-level, hypertextually linked electronic archive of the textual tradition of all three versions of the fourteenth-century poem *Piers Plowman*, written by William Langland. Initially, the team plans to make documentary editions of eight important manuscripts in the B tradition and one manuscript each of the A and C versions, reconstruct and annotate the B archetype, establish a critical edition of the B version, and write appropriate textual, linguistic, and codicological notes[6].The case of *Piers Plowman* is an especially interesting one for the editor working in the electronic medium as it is a poem which itself has always been in a particularly fluid state. Langland worked on the poem extensively throughout his life and scholars recognize at least three distinct versions of it. The poet himself therefore, was transgressing the textual boundaries which define what a finished work of art should be, and he released into literary debate a cultural object of a peculiar unfixity. Printed editions of this work have always been unsatisfactory both as a medium for the editor to work in and as a means of presenting this textual unfixity to the reader. The electronic edition is clearly the best means of dealing with a work of this kind.

In contrast to the detailed study of single works or single authors which the above projects are carrying out, some publishers are producing vast archives of whole genres of works or of existing massive scholarly resources. Chadwyck-Healey, which is based in Cambridge, England and has a branch in the USA has produced the enormous work of Latin patristics known as the *Patrologia Latina* as a CD-ROM. The printed version of this is some 130 huge volumes, the electronic is held on four CD-ROMs. Chadwyck-Healey is also producing major resources for scholars and students of English Literature. They have just completed the issue of their English Poetry Full Text Database, which contains the work of over 1,000 poets from the earliest period of English writings up to 1900. This again is contained on four CD-ROMs and is enormously costly to purchase. However, the company has just announced that it will be producing a more popular version of this, with major canonical authors and more multimedia capability than the main full text database currently has. This will be considerably cheaper than the full version—£500 as opposed to £25,000. They have also embarked on a similar scholarly database of English Verse

Drama, and they have just announced two new products: *Editions and Adaptations of Shakespeare* and *The Bible in English.* The Shakespeare package will consist of a full-text database of major historical editions and theatre adaptations of the works of William Shakespeare. There will be the complete text of eleven major editions of Shakespeare's works, twenty-five contemporary quarto and octavo editions, and over a hundred adaptations, sequels, and burlesques. The Bible package will include the complete text of fifteen editions of the Bible, ranging from the earliest Anglo-Saxon translations through the Middle Ages and the Renaissance to the modern period[7].

Chadwyck-Healey aim for completeness and comprehensiveness, which comes with an expensive price tag. Oxford University Press, on the other hand, is producing materials at the more modest end of the price range, at costs which might appeal to the individual department, academic, or even, for some of their cheapest products, for the student. The Oxford English Dictionary on CD-ROM is their flagship publication, and one which has already had great success. They have now produced what they have called the Reference Shelf Series; this is a set of resources for writers in different subject areas: Business, Languages, Science, etc. Each Reference Shelf contains dictionaries and guides for the writer to have easy access to a range of information which can be called up from within standard word processors. For instance, the Language Shelf has the Oxford French, English, Italian, and German Minidictionaries in one package. As well as reference materials, OUP is producing a range of high-quality but low-priced electronic texts aimed at the student market. They are doing so in partnership with the WordCruncher Corporation in the USA: there are fifteen titles due in the first tranche which should be released by the end of 1994, thereafter they will be producing in further tranches of fifteen. The texts are from their World's Classics series, and deal with major authors of the western cultural tradition: Dickens, Chaucer, Austen, Voltaire, Homer, among others[8].

There are other sources of electronic textual resources than commercial publishers. There are for instance text archives such as the Oxford Text Archive which offer free distribution of public domain texts[9]. The On-Line Book Index, available through the Internet, lists sources of such texts, and the On-Line Book Initiative aims to make many of these available on-line[10]. The Initiative claims to be a project to make a large collection of freely redistributable text available in a common format for others to do with as they like. A word of warning about obtaining electronic texts from sources which distribute them free: these texts may not be particularly accurate, and those who make them available do not always understand the need for scholars to know the sources and provenances of the texts they use. Also they may be encoded in ways which make them less

than useful. Having said that, judicious selection can result in some useful finds.

Networking Humanities Resources

Most academics have now heard of the Internet, and many are adept at navigating around it, but for those who are unsure about it or even scared, it might be useful to explain what it is, what it can do, and why scholars might find it useful. The 'Internet' has in the UK and the USA had a great deal of publicity in the national press in the last year or two, in part because the interest being taken in it by the US government and the Information Super-highway which they are proposing to create. But what exactly is the Internet? Well, it is an integration of national networks throughout the world all of which can talk to each other because they use the same set of conventions or protocols—the Internet Protocol or IP—to allow the exchange of information. Estimates of the size of the Internet are bound to be inaccurate and quickly outdated, but just to give some idea of its reach and influence, one recent book claims that at the beginning of 1994, the Internet consisted of some 11,000 linked networks, with 1.7 million host computers, and, according to some estimates, between 10 and 25 million users. The growth rate is currently approximately fifteen percent monthly[11]. A large proportion of the usage of Internet services is academic, and a significant and growing proportion of these academics are humanists. The very size of the Internet community affects the way scholarship proceeds, for increases in the mass of participants inevitably change the mode of participation. Access to the range of resources and services provided through networks is as important a development in scholarship as word-processing, which liberated scholars from the fixity of words written or typed on a page and allowed them an intermediate stage between the thought and its fixed realization. Drafting and redrafting became different concepts from hitherto and gave place to an on-going fluidity of composition. This new mode of written production underlies much of what scholars do on networks: communication and interaction with resources are also liberated, less constrained. The potential effects of these new technologies on humanistic scholarship are enormous: the use of high speed computer networks is certain to revolutionize the way disciplines are studied by academics and students in the future. Present problems of moving high quality images, sound, film, and multimedia around will no longer exist and academics will have access to primary materials on an unprecedented scale. The changes which this will bring about in academic life will be at least as great as the changes wrought by the advent of print technology. Given that e-mail has been in use for less than two decades, while we have had printing for 500 years and the telephone for 100, its impact is already disproportionately large[12]. The ability to access primary materials from the desk is of

great value in a period when travel budgets for scholars are being continually eroded. The most important effect of this is scholarly rather than financial, however: the virtuality of the medium offers a possibility of simultaneity of interaction with these sources which could never be otherwise possible. For instance, a medieval palaeographer examining a manuscript in the Bodleian Library in Oxford will be able to call up a manuscript in Vienna or Prague or New York on her computer screen and compare the scripts directly instead of having to rely on memory or photographs which may take months to arrive. It will also be possible to compare the manuscripts interactively with the guidance of scholars in two or more locations simultaneously. An art historian will be able to assemble the entire oeuvre of an artist even though the individual works are in many galleries all over the world. Textual scholars will be able to search textbases without even being aware of where the individual texts reside in the virtual universe[13].

Another important influence on the disciplines is the ready access which scholars have to each other through the electronic networks. The human networks of scholarly debate have, of course, been in existence for as long as there have been scholars: scholarship and teaching are processes of exchange. The electronic networks map onto these human networks and allow a faster, more active rate of interchange on a global scale. Interestingly, some scholars see this as a welcome return to an earlier state of scholarly interaction: The American humanistic scholar Patrick Conner, for instance, suggests that 'Telecommunications allows humanities scholars to reunite the academy which has been split in the west since the days of Martin Luther and Henry Ⅷ. I can now write and talk to people around the world as easily as I write and talk to my co-workers in the English department at West Virginia University'[14]. This is truly McLuhan's 'global village' and it changes dramatically the nature of the interactions as well as the speed at which they happen. Now the scholarly campus is the network, not the institution or department; every individual scholar, even those based at remote geographical locations, is a part of the wider scholarly community.

It has been claimed that accessibility to data and data resources becomes more democratic when these are networked, as it is as easy then for students to access materials as for their teachers. It is possible in the networked environment for a sharing of resources to take place, with 'virtual courses' being offered between institutions which may be widely separated geographically. A note of caution must be sounded at claims of democratization, however. Is this really so, or is there a further subtle elitism of which we must be wary in the sharing of networked resources? What about scholars and students in ill-funded institutions or outside the rich Western industrial nations? Is it still cheaper and more

democratic for them to buy a book or visit a library? Or will they have ready access to the libraries of the world through the electronic networks and be able to use resources that were formerly unattainable?

Currently, humanists use networks for informal communication on both a public and private level through electronic mail and discussion lists; for accessing bibliographic records through local and remote library catalogues, and through on-line databases like BLAISE; for interacting with primary sources (predominantly textual at the moment, although there are growing numbers of images on the networks); for searching secondary sources; for collaborative work; and for publishing in and reading electronic journals. Use of remote resources through new network management tools is increasing rapidly—many of these tools are used daily by scholars for whom it is as routine to turn on the computer and search the resources of the world as it is to make themselves a cup of coffee.

Two particularly interesting projects which are in progress in Britain at the moment are the British Library's Beowulf manuscript digitization project and the Bodleian Library's Image Archive project. Under its Initiatives for Access programme, the British Library is making available high quality digitized colour images of the Beowulf manuscript which can be accessed through the computer networks. The Library has found throughout this project that it is not just increased ease of access to images of manuscripts which the new technologies offer, but they can also enhance the quality and usefulness of that access. While the Library was investigating a range of digitizing cameras, there were demonstrations of a number of the potential candidates. At one of these demonstrations a page from a manuscript which had been damaged in the tragic fire at Ashburnham House, home of the great collector Sir Robert Cotton was shown. This is virtually unreadable in the original, with dark brown ink on top of dark brown vellum. The application of an image enhancement technique known as contrast stretching lightened the background and darkened the ink to make the page as readable as if it had never been damaged. It seemed at the time quite miraculous.

In Oxford, the Bodleian Library is also planning a programme of digitization of some of its major holdings in the Bodleian Image Archive Project. This project is planning a large programme of conversion of manuscript images to digital form. It has been consulting with the British Library during the piloting of the Beowulf Project, and will be purchasing the same camera, the Kontron ProgRes 3012, which has shown its value to the British Library. This is not Oxford's first step in this direction: the Bodleian already has a project under way to produce an image database of 6,000 items of printed ephemera concerned with the history of the motor industry in cooperation with Toyota in Japan, and is planning to

digitize its collection of 30,000 slides of iconography from medieval manuscripts. The Ashmolean Museum is scanning 100,000 photographs of figure-decorated Greek vases from the Beazley Archive which it is making available over the international networks and the Institute of Archaeology has a number of projects which are digitizing images of original objects and networking them. These include the Tchalenko Archive of photographs of near eastern architecture, the Cresswell Archive of photographs of Islamic architecture, and the Index of Celtic Coins.

The main objectives of the Image Archive Project, which involves the Bodleian and other libraries in Oxford, are: to define the issues involved in the capture and storage of images of rare, fragile, and valuable materials in libraries and to produce guidelines resolving these issues for institutions considering such projects; these issues will include preservation, surrogate acceptability, and the durability of the digital images. To establish technical procedures and standards for the capture process; to integrate the archival and retrieval system with the online bibliographical system in use in Oxford University and accessible over the international networks; to adopt an effective indexing system which combines structured text entry at time of capture with ease of retrieval from the final archive. The problems to be solved include the management of such images over a long period of time: microfilm, the main surrogate storage medium at the moment, deteriorates over time, in particular if colour film is used. It is however accessible using relative simple technology. Digital images require highly sophisticated hardware for their display, and this hardware becomes rapidly obsolete. This is a problem which has also been found in the uses and storage of electronic text. Some key texts were captured in the 1960s, the early days of textual computing, and the media of capture (punch cards, punch tape, etc) are no longer available. If the texts are not moved up consciously through generations of hardware and software, then the life of an electronic text is potentially very short, shorter even than that of a text printed on the most fragile acid-based paper. Contrast this with a medieval manuscript written on parchment which will be readable after 1500 years, even (or perhaps especially) if it is never opened during those years. If we are to spend vast amounts of money capturing digital images, then we need to ensure that the images survive as long as possible. We also need to preserve these secondary images as a means of conserving the originals and controlling access to them. If the surrogates are sufficiently good, then use of the fragile originals can be restricted; but if we have to keep rescanning the originals the use of surrogates as a conservation method becomes pointless. There is a problem inherent in making images of fragile manuscripts widely available, it seems to me, and that is that the publicity might result in more requests to handle the originals, rather than fewer. While

these requests can always be refused by the libraries, dealing with them puts further loads on already busy staff.

Hypermedia and Multimedia[15]

A great deal of interest is now being shown in the UK in the possibilities of hypermedia and multimedia for academic research and teaching: the great attraction of hypermedia for academics in literary and other humanities subjects is that it models what they already do in their research and teaching. While many of the systems being developed have more form than useful academic content, this is an area which will probably eventually prove to be of great value in the humanities, and already there are some flagship projects which have released good material. One example of an exemplary product is the Perseus CDROM developed in a collaborative setting, by classicists and archaeologists at Harvard University, Bowdoin College, Pomona College, and St. Olaf's College, among others. Perseus is a multimedia database designed to facilitate the study of Archaic and Classical Greece and to expand the ways in which ancient Greek literature, history, art and archaeology can be examined. Perseus is intended to be both a teaching and a research tool, for the beginning student as well as for the specialist, and is in use in a number of educational institutions in the UK[16].

The technologies of hypermedia and multimedia, which most people (including myself) use as synonymous, have as their theoretical underpinning the notion of hypertext. Although the term hypertext was first coined in the 1960s the original concept, as most workers in hypertext and hypermedia are aware, had its origins in an idea propounded in the 1940s by Vannaver Bush, one of Roosevelt's scientific advisors. The human mind, reasoned Bush, works by association, but information retrieval mechanisms operate along fixed and linear pathways, rendering access slow and difficult. He designed a theoretical machine, the Memex or Memory Extender, intended to mimic the brain's powers of associative linking. The digital computer had not yet emerged, and so the Memex was based on analog technology: information was to be stored on microfilm and accessed using mechanical devices to locate particular items. Any two items in the Memex could be associated electronically with each other in some permanent way, allowing the user to leave 'trails' through the information space.

Hypertext, which uses the same notions of associative linking as Memex, is the computational manifestation of Bush's theories. It is a means of linking together textual materials using what have become known as 'nodes' and 'links'. A node is a piece of text of any length (word, sentence, paragraph, even a whole book) and a link is the mechanism

by which this node is connected to another node. Links can connect chunks of text within the same document or between different documents. These concepts are extremely useful in humanities teaching and research because humanities scholarship is a highly complex process which involves its practitioners in following trails of paper and other media through libraries, museums, and archives all around the world. Indeed, one can propose a non-electronic analogy of nodes and links drawn from the humanities area and this is the scholarly book or article where the links are, for example, footnote numbers. The reader follows the link in the text, usually a superscript numeral, and finds the other end marked by the same number at the foot of the page or at the end of the work. Links, too, can take the reader to another part of the same text, using perhaps a cross-reference of the form 'see page 45', or to another text by giving the bibliographic reference. In the electronic hypertext, links are followed by pointing at them using a mouse and clicking; the other end of the link is then displayed immediately. Nodes and links can also connect text with non-textual media—sound, images, video—hence the term hypermedia. Electronic hypertexts have become popular recently because they function particularly well when implemented through the graphical user interface, the standard screen configuration on Macintosh, and now PC, microcomputers.

The book, which many claim is already doomed, has served humanities scholars well for many centuries and strategies have evolved over those centuries which allow books to function as non-electronic hypertexts. Even in medieval manuscripts and early printed books, great complexity of information structures could be represented in ways which readers could make sense of, although they often had to be experienced in navigating their paths through the intricacies of page design, just as modern users of hypertexts need to have simply-designed systems or a great deal of experience to avoid being 'lost in hyper-space'. A well-designed hypertext system can offer modes of access to materials superior to those offered by the printed book. There are also many other advantages to publication of materials in electronic form besides this ease of access. While I believe (and hope) that the book printed on paper will always survive, hypermedia systems can excel in the presentation of highly structured mixed media information[17]. We have recently developed two very modest systems at the Centre for Humanities Computing in Oxford, a package on the poetry of the First World War, and a package to provide teaching materials for the Old English poem The Dream of the Rood. We found that clearing the copyright on the relatively restricted range of materials incorporated in these was a greater task than any other in the development of the package.

The First World War package is centred around one of the most accomplished poets of

the First World War, Isaac Rosenberg, and in particular his poem 'Break of Day in the Trenches' and after the opening title screen the first real card the user is presented with is the text of the poem complete with buttons to show the history of the amendments made to the text by Rosenberg throughout 1916. From here readers can go directly to information concerning the poet (i.e. his life, attitudes, and other works); to a section entitled Analogues (the work of contemporary poets both male and female, plus some of the less well-known poetry surviving only in ephemeral publications like the trench newspaper The Wipers Times); and to a larger module on the First World War itself. To aid navigation there is a background map facility. Finally, although students are free to browse around the various cards of information using the map facility or the built in hypertext links, they cannot exit from the program without returning to the initial poem. This was a deliberate design policy so that hopefully on re-reading the poem their perception of the work may have changed[18].

The Dream of the Rood package is a hypermedia edition of the Old English poem of the same name, a primary text for the study of Old English in British and North American universities. The hypermedia edition contains a full glossary and notes, an Old English grammar, translations, analogue material, background introductory essays, colour photo-graphic images, and a full bibliography. The poem is central to the package. Select any word in the poem and click on the Glossary button to retrieve glossary information for that word. A further click will reveal grammatical information, or all occurrences of the word and its variant forms throughout the poem, or will link to further notes. Click on the Analogue button to display a list of documents; on the Topics button to display a list of background essays; on the Images button to display the list of image titles[19].

There are of course many other examples of hypermedia packages for teaching and research in the humanities; some of these are produced commercially, others developed in the academic environment.

Computers in Humanities Teaching
The Possibilities
The use of computers in the humanities was originally the province of only the most dedicated researchers. Some of the tools and techniques which they have used or developed are now being made more widely available to students, and are in some areas transforming teaching.

When making the decision to introduce computers into the teaching of arts subjects it is most important to be realistic about the possible benefits. There are, for instance, some

areas of text-based study which are more amenable to the application of computers than others: for instance, subjects which are less accessible to the student because they deal with unfamiliar texts, languages, and cultures; classics, medieval languages and literatures, non-European languages and literatures, for example. There are also methodologies which are particularly suitable for computation: stylistics, philology, lexical analysis, and other linguistic manipulations. For accessible languages and cultures, and for textual studies which seek to elucidate the non-enumerable qualities of a text, the benefits of using computers in teaching are not so obvious. With some modes, genres, and works, there will never be an acceptable substitute for reading books and talking about them. In areas where it is appropriate to use computers in teaching, the benefits can be enormous. Textual analysis and retrieval using concordancing and free text retrieval software, can, for instance, allow the student or scholar to find all instances of a particular word or concept in a range of contexts and thus offer some informed critical judgements about a text.

In history, enabling students to access real datasets and manipulate them is an invaluable tool for teaching. Computers can also be used to display numeric information in a graphic form to make it easier for non-numerate students to understand and manipulate. Simulation packages also allow students to explore the implications of critical decisions and their potential effects upon the course of history. Access to a range of original materials is also made possible for the undergraduate, where formerly these were only available to the researcher.

For students working with languages of all periods and cultures, Computer-Assisted Language Learning (CALL) techniques and drill packages can be invaluable, and there are many suitable packages now available. These are in some ways even easier to utilize for classical and medieval languages than for modern languages, because current learning of these languages is generally passive acquisition so the intensive teaching needed to produce active competence is not required. With a little imagination, drill packages can be produced which instil the rudiments of grammar without boring or frustrating the learner. Given the lack of traditional grammar teaching in British schools, and no doubt schools elsewhere, these aids are needed if the students are to have any understanding of the texts they are expected to read at higher education level.

In archaeology and art history computers are being used increasingly as tools to handle large amounts of data about the objects of study as well as presenting images of the objects themselves. The advantages of computer presentation of these materials over book presentation can be considerable: much more information can be handled and accessed through multiple pathways, which allows the information to be viewed from different

perspectives. Perseus, described above, is a good example of a teaching tool for both archaeology and art history: it has detailed studies of all the major archaeological sites in Greece, with maps, plans, and images of objects found therein. The CD-ROM has some 5, 000 colour images and many more black and white and line drawings. The objects can be searched and sorted by type, date, provenance, site, etc.

The Problems

The problems of attempting to introduce computer-aided learning into an arts faculty can be enormous. Traditionally, teaching and research in humanities subjects have been extremely cheap in terms of equipment: teaching requires a teacher, students, a room, and a blackboard; research a library and something to write with. Many older academics have never even learnt to use a typewriter, having had access to good secretarial support for most of their working lives. The introduction of new technologies is not a matter of merely providing hardware and software, but of changing fundamentally the attitudes and working practices of a substantial number of academics, and of changing the administration and financing of their institutions. This comes at a time when these academics are under considerable pressure to publish ever more books and articles, to teach ever more students, and to handle the concomitant increase in their administrative burdens. They are rightly suspicious of and resistant to any proposed changes which seem, in the short term, to increase their work-loads even further. There are, too, a number of fundamental misconceptions about the use of computers in teaching. Teachers believe that they might be replaced by machines, government and funding bodies believe that using computers will substantially reduce unit costs. Both these views are partly, but not entirely, true. The use of computers in the teaching of literary subjects will not reduce costs in the short term, although it may in the long term. Teachers will not be replaced, although their roles may change. They may be able to make better use of resources with technological aids, but as yet we have only anecdotal evidence to suggest that this may be the case.

The Computers in Teaching Initiative: A National Structure for Encouraging Computer-Based Learning

In 1982 the British Computer Board for Universities and Research Councils set up a working party under the chairmanship of Dorothy Nelson to review computer facilities for teaching in British higher education. This working party was charged to consider and make recommendations to the Board on the type and level of facilities that should be provided for teaching and on the services that should be provided by computing centres to support

teaching facilities.

The report produced by the working party, now commonly known as the 'Nelson Report', made a number of firm recommendations to universities to encourage them to increase the provision of hardware and software for use in teaching, and to provide training and support for teachers who wished to introduce the use of computers into their teaching. As a direct result of the suggestions of the Nelson Report, therefore, the Computer Board decided in 1983 to establish the Computers in Teaching Initiative (CTI), and it funded the first phase of this with a total expenditure by 1987 of £9.5 million. The aims of the Initiative in its early stages were: to promote greater awareness in all subjects of higher education of the potential benefits of the use of computers in the teaching and learning process; to help individuals develop the requisite skills; to provide or produce appropriate courseware for use in university courses; to evaluate and assess the hardware, software, and organizational requirements for the successful introduction of information technology into university teaching.

The first phase of the Initiative attracted some 700 proposals for projects, covering the full range of disciplines taught in British higher education institutions. Of these 139 were funded for up to three years; most of them included some software development, alongside a serious attempt to incorporate computer-based learning into individual courses. There were a number of projects in the humanities, in language learning, textual studies, history, archaeology, among others.

Evaluation of these early projects showed that they had achieved a measure of success in a number of different ways: they had promoted awareness of the possibilities within the institutions which had housed projects; they had trained staff in the use of computer hardware and software, and, most importantly, they had had significant impact on the learning experience of students. An assessment was carried out at the end of this phase of the Initiative to consider just what it was that students had gained. Was it increased understanding and competence in their discipline-based studies, better use of computers, or both? Interestingly, the study carried out showed that the greatest effect was on the general levels of computer literacy among both staff and students involved in the projects, although the effects on academic competence were felt to be significant. As well as increased levels of competence, other important outputs from the initiative were the actual products of some projects. For instance, some 63% of projects produced software which could be used outside the originating institution.

As a direct result of phase one of the CTI, a rather different phase of the Initiative was begun in 1989. This involved the establishment of 20 subject-specific Centres to promote

and support the use of computers in university teaching. These cover most disciplines taught in British higher education, and while they do not themselves develop software or courseware, they disseminate information around university departments in their discipline areas by means of newsletters, resources guides, visits to institutions, and network delivery of information. These Centres have been so successful in fostering the permeation of computer-based learning throughout higher education in the five years since they were established, that the British Higher Education Funding Councils have recently agreed to fund them for a further five years from 1 August 1994. Four of these Centres are in the humanities disciplines: Textual Studies at Oxford; History with Archaeology and Art History at Glasgow; Music at Lancaster; Modern Languages at Hull[20].

The Future

One of the great driving forces of invention in the history of computation has been the formulation of theories which are far ahead of the technical possibilities: Babbage's difference engine, Bush's Memex, the theoretical but untestable algorithms of Ada Lovelace and others. Hypertext is now technically possible to implement, particularly on a relatively small scale, but the theories which are now being constructed extend it to enormous, even global dimensions. Ted Nelson who originally coined the term hypertext, is, for instance, working on a global library, a 'New Alexandria', called the Xanadu project, which plans a worldwide network of hypertextually-linked documents. Readers will traverse this hypertextual universe of documents (known as a 'docuverse') linking and reading where they will, paying through an automatic charging system for the texts they read. Fanciful though this may seem, we have seen in the last two years great strides forward in the development of such large scale projects through the Internet using the World-Wide Web.

CD-ROM publication of hypermedia systems for the humanities is the present transfer and access mechanism of choice, but it looks as if network publication will soon take over from this. The World Wide Web currently uses rather more primitive hyperlinking than can be supported on CD-ROM, but this will soon change: it will be irrelevant where in the virtual universe the hypertextually linked media reside, and developments in high band width connections will make the transfer of high quality digital images, film, and sound as easy as the transfer of text now is.

In the humanities, we may be able to establish a global community of scholars who teach students residing in different countries or even on different continents. We may be able to access library materials from all over the world at the touch of a button: we can already

find out the contents of these libraries, soon we will also be able to have access to them. These are exciting possibilities, but they are also very frightening. We find it very difficult to deal adequately with the information which we can currently access: how will we deal with the exponential growth in the information available, particularly in the situation where, as with the Internet, much of the information is of such poor quality as not to be usable by scholars? Also, the pace at which we are expected to do research, scholarship, and publication is accelerating so rapidly that there is little time for the considered reflection which our work used to be subject to before it was released for general view. In view of these possible problems, it is incumbent upon academics to engage with both using and critically evaluating the new technologies so that they become our servants, not our masters.

Footnotes

[1] See Susan Hockey, *A Guide to Computer Applications in the Humanities* (London: Duckworth, 1980), pp. 68-69 for a brief description of Busa's project.

[2] Anthony Kenny, *The Computation of Style* (Oxford: Pergamon, 1982).

[3] There are a number of sources of information for those who wish to learn more about the range of projects and possibilities in computing for the humanities. A good starting point is the free publication *Information Technology in Humanities Scholarship: British Achievements, Prospects, and Barriers* obtainable from the Office for Humanities Communication, Oxford University Computing Services, 13 Banbury Road, Oxford, OX2 6NN, UK. This is soon to be translated into Japanese. Also useful is Ian Lancashire (ed.), *The Humanities Computing Yearbook 1989-90: A Comprehensive Guide to Software and Resources* (Oxford: Clarendon Press, 1990). The journals *Literary and Linguistic Computing and Computers in the Humanities* should also be consulted.

[4] Josie Dixon, *Electronic Publishing: American Developments, British Implications.* A Report for the Sir Stanley Unwin Foundation (1994).

[5] See Norman Blake and Peter Robinson, *The* Canterbury Tales *Project Occasional Papers Volume I* (Oxford: Office for Humanities Communication Publications 5, 1993). The Project also publishes a Newsletter, available from Dr Peter Robinson, Computers and Variant Texts Project, Oxford University Computing Services, 13 Banbury Road, Oxford OX2 6NN; e-mail peterr@vax.ox.ac.uk.

[6] The most up-to-date information about this project, along with facilities to search the *Piers Plowman* archive is available on the World-Wide Web at http://jefferson.village.virginia.edu/piers/archive.goals.html

[7] Further information about all these products can be obtained from Chadwyck-Healey Ltd, Cambridge Place, Cambridge, CB2 1NR, UK.

[8] For further details about the range of electronic publications available, contact Electronic Publishing, Oxford University Press, Walton Street, Oxford OX2 6DP, UK.

[9] The Oxford Text Archive, Oxford University Computing Services, 13 Banbury Road, Oxford

OX2 6NN, UK. E-mail: archive@vax.ox.ac.uk

[10] The Online Book Initiative is available through ftp or gopher: ftp world.std.com; *login anonymous; cd obi; gopher world.std.com/OBI The Online Book Initiative.*

[11] Paul Gilster, *The Internet Navigator* (New York: John Wiley and Sons, 1993).

[12] See Willard McCarty, 'HUMANIST: Lesson from a Global Electronic Seminar', *Computers and the Humanities*, 26 (1992), 203-222, and Patrick W. Conner, 'Networking in the Humanities: Lessons from ANSAXNET', *Computers and the Humanities*, 26 (1992), 196-204, for some interesting ideas about the influence of electronic communications on scholarship.

[13] See Patrick W. Conner, 'Hypertext in the Last Days of the Book', *Bulletin of the John Rylands University Library of Manchester*, 74 (1992), 7-24, for a thoughtful and reflective examination of the virtual medium and access to scholarly materials.

[14] Conner, 'Networking in the Humanities', p.196.

[15] An introduction to the use of hypermedia systems in humanities scholarship and teaching can be found in Marilyn Deegan, Nicola Timbrell, and Lorraine Warren, *Hypermedia in the Humanities* (Sheffield: CVCP/UDSU, 1992). This costs £6 and can be obtained from Jean Burgan, (ITTI products), CVCP, Level Six, University House, Sheffield, S10 2TN; tel: 44-742-725248; fax: 44-742-728705; e-mail: j.burgan@pa.shef.ac.uk

[16] The Perseus CD ROM 1.0 is published by Yale University Press and can be ordered from John Wiley and Sons Ltd, Distribution Centre, Southern Cross Trading Estate, Shripney Road, Bognor Regis, West Sussex PO22 9SA; tel: 44-243-829121; fax: 44-243-820250. Version 2 is to be released shortly.

[17] Many of the political, intellectual, and economic issues around hypertext are discussed at length in Warren Chernaik, Caroline Davis, and Marilyn Deegan, *The Politics of the Electronic Text* (Oxford: Office for Humanities Communication Publications, 3, 1993).

[18] This package can be obtained from Dr Stuart Lee, Centre for Humanities Computing, Oxford University Computing Services, 13 Banbury Road, Oxford OX2 6NN, UK; e-mail: stuart @vax.ox.ac.uk. Dr Lee is currently working on version 2 of this, which should be available on a CD-ROM at modest cost to academic establishments.

[19] Copies of this can be obtained from Jean Burgan at the address given in note 15 above. The cost is £25.

[20] There are a number of useful articles on the individual CTI Centres and on the Initiative as a whole in *Computers and Education* 19 (1992) which is a special issue of the journal devoted to the CTI.