Why Technology Matters: the Humanities in the 21ˢᵗ Century

A Wisbey Lecture

King's College, University of London

16 October 2003

Stanley N. Katz

Woodrow Wilson School, Princeton University

Thank you so much for the invitation to present this first of the Wisbey Lectures. It is a

great honor. For me it is also a very special honor, since I was a graduate student in this College

during my Fulbright year in London, 1959-60. That was a very long time ago, but I still

remember very warmly my supervisor, then the Rhodes Professor of Imperial History at King's

College, the late Gerald S. Graham -- a perfectly wonderful man and a fine historian. I dedicate

this lecture to his memory.

Presumably everyone in this room believes that there is such a thing as the "digital

humanities," and that the digital environment is a promising new frontier for scholarship in the

humanities. Why else would you come to such a lecture? Well, I think so too, or I would not

have crossed the Atlantic to speak to you this evening. I fear, however, that there are still a good

many people who not only do not share our commitment. When I told a friend who presides over a large American philanthropic foundation that I was traveling to London to give a lecture on the digital humanities, I received (by instant return e-mail) the following response: "Only you could make a whole speech out of an oxymoron!" A second e-mail assured me that this was a compliment. I am uncertain.

In preparing for this evening, I reviewed some fairly recent library scholarship on use of computing and the Internet. It seems that several knowledgeable American librarians, surveying usage by their university faculty and graduate student patrons, find that academic humanists are not much interested in the broader digital humanities or even in the use of routine online scholarly resources.

For instance, Virginia Massey-Burzio, a librarian at Johns Hopkins University, writing in 1999, warned digital librarians that "not all users feel the same way or have the same attitude toward technology as our more vocal users." She means scientists and "younger students," but, and let's face it, she also means you and me. She found that a "small minority" of humanists used the Internet and only a "tiny" minority used it "for a specific purpose." She reports that her focus group participant humanists felt that "technology did not work well for the humanities," for "there is a book culture that is different from a technology culture." The distinction they drew to explain their attitude was that science and social science are "object-oriented," while humanists are "browsers." The humanists' major complaint seems to be about OPACs, which do not permit physical browsing – although so far as I can tell this is really a complaint about the

increasing prevalence of closed stack systems in major American libraries such as the Library of Congress.

But the humanists also complained about electronic journals (since they do not want to read on-screen), the lack of tech support for IT in humanities departments and even the World Wide Web, which they find "too busy, cluttered and game-like to take seriously." Massey-Burzio reports that humanists prefer to have librarians intermediate between themselves and digital information – "They have a low tolerance for anything complicated or time-consuming." And, of course, they demand made-to-order humanities software.[1]

Perhaps less surprisingly, since they were writing in 1994, Stephen Wiberly, Jr. and William G. Jones reported that humanists were "adopting new technologies slowly." But their research is worth attending to, since Wiberly has been studying humanist library usage for many years in a highly sophisticated fashion. They observe that, apart from word processing and OPAC searches, humanists use technology "less than scientists and social scientists for communication (e-mail), bibliographic searching, and storage, transmittal, and analysis of primary evidence." They attribute the difference to the sorts of evidence humanists use: "documents and artifacts created by persons whose activities and accomplishments the humanists seek to reconstruct, describe, and interpret -- as opposed to scientists and social scientists who create much of their own data

---

[1] Virginia Massey-Burzio, "The Rush to Technology: A View from the Humanists," *Library Trends*, 47/4, Spring, 1999, pp. 620-1, 627-8, 635-7.

. . . it is useful to look at all scholarship as a continuum from the physical sciences to the quantitative social sciences to the qualitative social sciences to the humanities. Moving along this continuum from the physical sciences to the humanities, one can say roughly that the scholar exercises decreasing control over the primary evidence that is analyzed. We suggest the proposition that the less control over primary evidence the scholar has, the harder it is to utilize information technology.

Wiberly and Jones stress the differences between the humanities and the social sciences, despite their relative proximity on the continuum.

Comparison of the work of humanists with that of social scientists shows that fundamentally humanists use sources created by the subjects of their research, while social scientists initiate and, much more than humanists, participate in the creation of their sources. This is a fundamental difference . . . because it points to what predominates in each area of scholarship.

The basic problem, they contend is that "humanistic sources are multifarious, often incongruous and diffuse, and harder to coordinate and manipulate than survey research data." And, finally, they agree with Massey-Burzio that humanists suffer from having "no generally accepted software package that can analyze such evidence according to the interpretive viewpoints that are evolving in their minds."[2]

---

[2] Stephen E. Wiberly, Jr. and William G. Jones, "Humanities Revisited: A Longitudinal Look at the Adoption of Information Technology," *College and Research Libraries*, November, 1994, 55/6, pp. 503-505.

I acknowledge that even 1999 is a long time ago in the rapidly changing world of humanities computing. Each of us knows some scholar who is still using pen and pad in order to process words (last week's <u>New York Times</u> obit of Edward Said put him in this category) and we also know a few scholars who do not have e-mail accounts. But, let's face it, these people are recognizably quaint these days, and I would guess that almost all of them are more than 50 years old. I have not seen a current survey of humanities usage, but it simply must be the case that more and more humanists use digital texts, since the quantity and quality of such information is expanding so rapidly. And it is hard to imagine the serious humanities scholar who does not regularly use OPACs and other primary source data available on the World Wide Web.

But perhaps the most dated observation in these two librarians' articles is their lament that the humanities do not have a customized software package. – that, and the failure of either article to mention humanities databases. So much has changed in our field that the analytical language of the 1990s is no longer satisfactory to describe it. So what exactly has changed?

Although some years ago I gave a lecture making fun of the notion that humanists thought of their desktop machines as overpriced typewriters[3], I believe it is the case that even computerized word processing and its digital accoutrements have made a major impact on the work habits of humanities scholars. We take these things for granted, but they make a difference. The capacity to manipulate text digitally is very important to us. We can revise more freely, up to the moment before we send off a text, even when the revisions would have necessitated completely retyping the manuscript. Younger scholars who have grown up with computers will not appreciate how much this capacity, just as the freedom to move text from

---

[3] http://www.princeton.edu/~artspol/workpap18.html

place to place, has changed our writing habits and, yes, probably improved our prose. But think also of the capacity to move from research notes and bibliography to the text we are composing, and the ease of creating and manipulating footnotes. We have become textually more accurate, since we type sources into our hard drives only once, and then cut and past them into text and footnotes rather than retyping them, often several times – though of course we had better take care not to keep track of what it is we are moving around, and especially whether or not it is our own words or quotations. Do you remember the tedium of such work in the analog era, and the number of errors you introduced in retyping? And, surely some of you are old enough to remember wondering whether it was worth adding a new footnote shortly before completing a manuscript, since to do so would necessitate the manual renumbering of every footnote in the text. We usually chose not to add the note, at least if I was typical.

Our work has been profoundly changed by the whole range of developments in digitization and telecommunications, developments that students and younger scholars take for granted. Consider what my work was like when I studied here forty-four years ago. I worked almost every day in the Manuscripts Room of the British Library, in its elegant former quarters in the British Museum. Some days I went to work in the Public Record Office, then in Chancery Lane. I took with me only a 5 x 8 note pad and several lead pencils, since fountain pens were not permitted. Each evening I filed my notes in a shoe box, arranging them according to the single concept or word that seemed most significant for my research. And on the weekends, when the archives were closed, I took out my Olivetti portable and typed what I had learned, using carbon paper to make copies – and fearing to make corrections, since that necessitated removing the paper and the carbon from the machine, and erasing both the original and the carbon separately,

before replacing both in the machine. Then, when I came to compose the actual chapters for my dissertation, I had to refile the 5 x 8 pages, since most of them had at least a second thought or bit of information that was needed elsewhere in my text. I cannot even imagine how much information I lost because I could not remember where it was filed. And then, of course, I had to find the original note pages to check the accuracy of my typing into the text, and again, later, to compose footnotes. The point is clear, though only the oldest members of this audience will fully appreciate what I am describing. Surely no one of us regrets leaving that part of the analog environment behind us!

But think just for the moment of the other new technologies, equally taken for granted (though only in the past decade or so), that enable us to do our scholarly work. The Internet, for instance. We use e-mail daily in order to communicate with our students, mentors and other scholars. We become virtual colleagues and sometimes mentors for graduate students and scholars whom we have never previously met, in parts of the world we have never visited. We e-mail drafts of papers to colleagues for review, and read theirs. We also use "Track Changes" and comparable technologies to distribute the editing of our writings, and thus use electronic communication to facilitate collaboration in ways that used to be largely unknown to humanists.

Or take another Internet-supported technology, the World Wide Web. We use the web to search distant library and archival catalogs, to exploit digital libraries and archives of all kinds and, generally, to search for information. New portals with advanced searching capacities, such as Google.com, enable us to identify and access information that was virtually inaccessible to us twenty years ago. Although it is still true that scholars working in the elite academy have

privileged access even to widely distributed digital information resources (due to the increasingly

propriety character of licensed databases), the individual scholar, working entirely

independently, now has remote access to formerly unimaginable information resources on his

own computer display. He or she also has access, I am happy to say to your colleague (and my

friend) Willard McCarty's Humanist e-list, which he describes as "an electronic seminar for

humanities computing."[4] McCarty notes on the website that that "[w]e in the academy have not

done a good job communicating our raison d'être to the rest of the world- arguably because so

many of us do not ourselves know what it is. Within the university, as outside it, fundamental

questions are seldom asked, but our fault is more serious because asking such questions is our

principal justification. The profound impact of computing on all aspects of modern life provides

therefore a great opportunity to engage in a long-overdue re-examination of what universities do

for the society of which they are a part."

All these things are obvious to everyone in this audience, as are the impact of other new

technologies. For the humanist, computation itself has created whole new scholarly disciplines,

computational linguistics being only the most obvious example. GIS enables us to use spatial

representation in ways that were inconceivable less than a generation ago. Much more recently

visualization software has revolutionized our capacity to imagine the physical character of the

past. The mark-up of text, as in the Text Encoding Initiative, enables us to use text accurately

and flexibly, through such technologies as HTML, SGML and XML. The emerging digital

environment is inherently multi-media, and formerly text-bound humanists are now exploring

---

4 http://www.princeton.edu/~mccarty/humanist

and exploiting sound and image. Historians, for instance are now studying the roles of sound and silence in the past.

For the humanist, perhaps nothing is more important than the capacity to organize and search large bodies of information. Historically, some of the greatest humanistic projects have been concordances, indices and bibliographies. We have always been committed to the compilation of vast bodies of related information, whether the complete works of single authors or the historical records of particular institutions. Famous scholars have devoted their lifetimes to such projects, frequently without being able to see them to completion. But the capacity to scan such material into relational databases and to mark it up in diverse and sophisticated ways enables us to do in decades what previously took many years, but it also enables us to retrieve information more effectively than even the best analog systems permitted. I have had some small contact of one such project in recent years, <u>The Dictionary of American Regional English</u>.[5] I can still remember, when I was an assistant professor at the University of Wisconsin, walking by the editor, Fred Cassidy's, library office and seeing the thousands of shoeboxes filled with small pieces of paper, each devoted to one observation of the use of a specific English word at a particular place and time. <u>DARE</u>, as it is now called, might never have left the shoeboxes of the editor's office if the advent of computers had not permitted the digitization of the data. Now the volumes are coming out regularly, with the wonderful enhancement of a digital map for each word, showing the frequency of its use geographically.

---

[5] Frederic G. Cassidy, Chief Editor, *The Dictionary of American Regional English* (Cambridge, MA., Harvard University Press, 1985-2002) 4 vols.

Most of these remarkable humanities projects take the form of digital databases, and it might be said that the new era of humanities scholarship is, really, the Era of the Database. Now of course a codex is also a database, but it is a fixed and unchangeable database, searchable only through a fixed and rigid indexing system. Digital databases are, however, dynamic rather than static, and they can be searched or otherwise manipulated in myriad ways. Today each of the humanities fields is experiencing the creation of large numbers of significant new databases that differ dramatically from one another, ranging from numeric to full-text to multi-media.

At least two things can be said about digital databases (and henceforth I will drop the modifier "digital"). On the one hand, they create challenges and opportunities for entirely new methods and subject matters for research. On the other hand, they are costly, administratively awkward and hard to assimilate into traditional academic structures. I will mostly be talking about the positive side of this equation, but let me here mention at least a few of the problematic aspects of the Era of the Database.

Sometime ago I proposed what I modestly dubbed "Katz's Law": *it costs more to maintain a database than to create it.* And it costs huge amounts to create databases. This is partly a function of the inherent labor and technology costs of the database, partly a function of the difficulty (and uncertainty) of the preservation of digital information, but mainly a function of the fact that databases are normally dynamic, and when they cease to grow (or change), they are likely to die (which is to say that they are likely to become uninteresting to users). Where does funding for academic databases come from? Normally it comes from "soft" money, either from state or philanthropic resources, but typically such funders are primarily interested in the

initial creation of the resource, rather than its continued development or its preservation. My colleagues at Princeton and I are now confronting precisely this problem with a database called CPANDA (Cultural Policy and the Arts National Data Archive), a large collection of cross-searchable numeric data sets for American arts activity. [6] We received a very generous philanthropic grant to build the Web-based site, with the promise of a follow-up grant to develop a business plan for the continuing expansion and maintenance of the site. But now the funder has changed its priorities, and $2 million later we are wondering what to do when the current CPANDA funding cycle expires in a year.

Where such projects should live in the university is another problem. CPANDA is a project of our university library, and frequently that is the case. But financially hard-pressed university libraries can rarely fund these projects, nor can other university units. What unit will advocate for the creation and maintenance of such information? Who will pay for and provide the technology support (primarily human resources, by the way) that make the creation and improvement of databases possible? The normal computer support unit ordinarily has neither the technical skills nor the incentive to play this role. What is needed are specialized units such as Harold Short, Willard McCarty and their colleagues are training personnel for here at King's College.[7] We have almost no comparable programs in the United States, and only a handful of humanities research support units. How will we pay for the time of humanities scholars to participate in digital humanities projects that do not necessarily lead to traditional analog academic publications? And how will we guarantee younger academic humanists that they will be rewarded rather than punished for their participation in the creation of databases and other

---

[6] http://www.cpanda.org
[7] http://www.kcl.ac.uk/humanities/cch/

digital products? I do not know the answers to these questions, and my purpose this evening is primarily to sketch what I see as the promise of databases for the digital humanities. But as an experienced academic humanities administrator I must acknowledge that there are fundamental financial, institutional and academic cultural barriers that stand firmly astride to the road to a brighter digital humanities future.

But allow me to return to my more optimistic and positive mode, and to use my own field, history, to illustrate what I mean about the promise of databases for the humanities. In many ways the field of history, and especially American history, has become one of the most interesting testbeds for the digital humanities. At least two major academic publishers, the Oxford and Cambridge University Presses, have decided to experiment with digitizing their history backlists. These have both been mounted experimentally at the University of Pennsylvania Library so that the intensity and quality of their usage, including the extent to which the texts are downloaded and printed out, can be systematically monitored.[8] The American Council of the Learned Societies has a similar initiative, the History E-Book Project, that has put online 750 history books.[9] The foregoing projects are all funded by the Andrew W. Mellon Foundation, which has been the leading funder of humanities digital innovation in the United States. This is part of the larger movement toward the publication of e-books, but it is focused on one field, and on academic usage, in the humanities. The result is that there is already more digital monographic information available for history than for any other field, and the Penn experiment will tell us a good deal about how it is used.

---

[8] http://verity.library.upenn.edu/search97cgi/s97_cgi?Action=Search&collection=ebooks&access_type=penn&resultcount=5000&OupQueryType=title&ResultTemplate=ebooks/home3books.hts
[9] http://www.historyebook.org/intro.html

History has also moved forward aggressively to digitize some of its major journals, through the creation of the History Cooperative, a consortium of the University of Illinois Press, the National Academies Press and 12 historical journals to publish online versions.[10] The group includes the The American Historical Review, The Journal of American History and The William and Mary Quarterly, and others of the preeminent journals in the field. The project thus provides an ongoing digital publication opportunity in the field, and, together with the remarkable Mellon Foundation-initiated JSTOR digital archive of the back files of many of the major historical journals, it means that there is also a tremendous, accessible and searchable body of digital journal information in the field. I will comment shortly on the implications of digital publication, but I want here simply to note the extent to which the content of historical information is currently available. It is supplemented at the present by the recent online version of the American National Biography, a living publication, adding more than 500 biographies a year and soon to add both image and sound to its biographies.[11]

The digital history field has recently been stimulated by the remarkable Center for History and the New Media at George Mason University, founded in 1994. Under the leadership of American historian Roy Rosenzweig (and with a superb staff of technologically adept historians as well as technologists), the Center aims to use:

. . . digital media and computer technology to change the ways that people- scholars, students, and the general public- learn about and use the past. We do that by bringing

---

[10] http://www.historycooperative.org/journals.html

[11] <http://www.anb.org/articles/home.html>

together the most exciting and innovative digital media with the latest and best historical

scholarship. We believe that serious scholarship and cutting edge multimedia can be

combined to promote an inclusive and democratic understanding of the past as well as a

broad historical literacy that fosters deep understanding of the most complex issues about

the past and present."

The Center is especially interested in working with teachers of history "to make effective use of

the new technology," especially through the use of information technology to improve student

learning. The Center has also built several interesting digital archives, including one on the 1965

and 1977 electrical blackouts in New York City[12], the recently announced *September 11 Digital

History Archive*[13] and *History Matters*[14], a digital American history survey course.[15]

The other major American digital humanities unit is the University of Virginia's Institute

for Advanced Technology in the Humanities, led until recently by the remarkable John

Unsworth, who has just become the Dean of the Library School at the University of Illinois.[16]

This is, to my way of thinking, the model general digital humanities operation in my country, a

wonderful collaboration between scholars and technologists across many of the fields of the

humanities. But here I want to stress its preeminent project in history, *Valley of the Shadow*,

now maintained by the IATH spin-off, the University of Virginia's Center for Digital History,

---

[12] http://blackout.gmu.edu/
[13] http://911digitalarchive.org/
[14] http://historymatters.gmu.edu/
[15] http://chnm.gmu.edu/index1.html

[16] http://www.iath.virginia.edu/

founded by Edward Ayers and now directed by William Thomas, III. The Center describes its

mission as:

> . . . to develop high-quality, well-researched, and
>
> reliable history resources in digital format and deliver them to
>
> schools, colleges, libraries, historical societies, and the general public
>
> via the Internet. By doing this we hope to transform how American
>
> history is taught, learned, understood, and accessed.
>
> To achieve this mission VCDH has committed itself to collaborative
>
> efforts with other institutions, both within and outside of the
>
> University of Virginia system, and to using innovative technologies
>
> which allow information to be freely accessed by a general audience.[17]

The most notable achievement of this Center, one that originated in IATH, is this Ayers-Thomas

database, a huge archive of material that, in the words of its promotional material, documents:

> . . . the lives of people in Augusta County, Virginia, and Franklin County, Pennsylvania,
>
> during the era of the American Civil War. Here you may explore thousands of original
>
> documents that allow you to see what life was like during the Civil War for the men and
>
> women of Augusta and Franklin.
>
> The *Valley of the Shadow* is different than many other history websites. It is more like a
>
> library than a single book. There is no "one" story in the Valley Project. Rather, what

---

[17] http://www.vcdh.virginia.edu/mission.html

you'll find are thousands of letters and diaries, census and government records,

newspapers and speeches, all of which record different aspects of daily life in these two

counties at the time of the Civil War. As you explore the extensive archive and you'll find

that you can flip through a Valley resident's Civil War diary, read what the county

newspapers reported about the battle of Gettysburg, or even search the census records to

see how much the average citizen owned in 1860 or 1870.[18]

The Valley of the Shadow, to which I will return, is equally adapted to teaching the history of the

American Civil War, and to serving as a research archive – in part because of the turn to the use

of research as a primary technique in undergraduate teaching.

The American Heritage Project at the University of California, Berkeley is a very

different sort of digital resource, a virtual archive designed to provide access to distributed

digital library information. The project's creators argue that "one of the most serious problems

facing knowledge seekers everywhere [is] the geographic distribution of both collections of

primary source material and the written guides describing and providing access to them." Their

virtual archive is the solution to this problem, "integrating into a single source, hundreds of

archival finding aids describing and providing access to a large body of primary source materials

from collections documenting American culture and history held by four major academic

research libraries." The Project is intended to "demonstrate the feasibility of providing both

scholars and average American citizens with user-friendly, universal Internet access to the

research collections of the world."[19] But in this instance the hype exceeds the promise, since it is

---

[18] http://valley.vcdh.virginia.edu/usingvalley/valleyguide.html

[19] http://sunsite.berkeley.edu/amher

an archive of finding aids to collections (quite valuable in itself) that does not permit access to the digital archives themselves.

Our community has also benefited enormously from the presence of H-Net, which began life as History-Net and whose formal name now is Humanities and Social Sciences Online.[20] Shrewdly administered by Mark Kornbluh of Michigan State University, H-Net is an online clearing house of moderated discussion lists in all the humanities and social sciences fields, as well as on special topics. Its coverage is very strong in history, and it has become the intellectual common ground and point of contact for specialists in my field. More recently it has been joined by the History Network Newsletter, a George Mason spin-off nicely described by its name.[21]

Perhaps this description of the emerging American institutions in digital history strikes you as a bit dry, but I ask you to consider what a remarkable technological infrastructure and body of digital source material and scholarly analysis has been created in almost exactly a decade. To me it indicates that the field has truly entered the digital age, although the sad truth is that at this point the technological infrastructure is more pervasive and extensive than the awareness of the scholarly historical community or even the competence of those who are aware to be able to use the infrastructure fruitfully. But out of small beginnings grow large and important things, and I am confident that now that we have built it, they will come.

---

[20] http://www.h-net.org

[21] http://HNN.us

The interesting question is to what end are we building a digital humanities world? Some of the answers are obvious, and I have already touched on them. At the very least, digitization makes more information more readily available to more scholars. Digital finding aids make it much easier to identify relevant source and secondary material. Digital word searching techniques not only facilitate the identification of specific information, but enable the researcher to compare and make connections across long periods of time and vast bodies of material. For most contemporary scholars, this level of technology suffices, and it enables them to do what they have traditionally done in a much more efficient and cost effective manner.

But moving up a step on the scale of complexity, the digital multimedia environment enables us to visualize and hear cultural phenomena in ways that even the most advance analog printing does not permit. The scholar can now manipulate information ranging from text to image to sound in ways that recreate old worlds and suggest worlds that never "really" existed. We are beginning to be able to search images and sounds in ways that were impossible before. We can specify links between image and text. We have remarkable mapping capacity. We can doubtless do many things that we have not yet discovered. This is to say that humanities scholars now have at the very least a radically expanded body of accessible information and repertoire of methodologies for identifying and manipulating that information than we have ever had before.

But another aspect of the digital humanities environment that we have only just begun to explore is its potential for interactivity. I do not want to argue that the analog environment lacks interaction between author and reader – we have a generation of scholarship on reader

responsiveness to text, image and sound. But interactivity takes on a new range of possibilities in the digital humanities world, and it is frequently at the core of the digital humanities project. Hyperlinks are the obvious case in point. In a digital "text" the "reader" is frequently invited to click on hot links that transport him to a different body of information. Once there, he is invited to click on other links, potentially ad infinitum. He may never get back to the original site, or he may reenter it at a point different from that at which he originally exited. This process is the fundamental structure of the cybernovels of writers such as Michael Joyce[22], but it is also at the core of most of the new humanities databases.

The most commonly cited example in my country is the Ayers-Thomas *Valley of the Shadow* database. You will remember that the site announces that "there is no one story" presented. The user of the site can enter through one of three chronological portals – "The Eve of War," "The War Years," or "The Aftermath." Each of these portals has a series of pages with different types, so that the user can click on 1860 statistics, newspapers, maps and images, letters and diaries, soldiers' records, census and tax records and so forth. The choice is his, and what he learns is a function of the ways in which he chooses to navigate the many sites in the several portals. Consider that The *Valley of the Shadow* is equally useful for teaching the history of the Civil War (and other things) and for scholarly research. The student can be set questions and projects to be addressed by navigating the site, and presented for evaluation in either analog or digital form, or both. The researcher has access to an incredible amount of source information hitherto not accessible in a single place. He, too, can present his results in a variety of formats.

---

[22] Joyce, Michael, Twilight, a symphony [electronic resource] : a hyperfiction (Watertown, MA : Eastgate Systems, c1996) computer laser optical disc

Furthermore, the site is thus both interactive and non-linear, producing both new relationships between "author" and user, and a non-narrative experience for the user.

And the interactivity does not stop with the website. Ayers and Thomas are about to publish an "article" in the December, 2003 American Historical Review. Like many established scholarly journals, the AHR currently publishes both analog and online versions. The online publication is only about three years old, and until now the e-journal has largely consisted of digital versions of the printed texts. But no longer. The Ayers-Thomas piece does not really have a central narrative, and it is full of choices for the user to make. The interactive experiences of users will be different from one another, so that there is no way to replicate this phenomenon in print. In effect, the AHR will now be publishing two different journals, and this will make a profound difference to the distribution and archiving of historical knowledge as the trend continues.

A final point about *Valley of the Shadow*. As I have already said, the site is equally well-adapted for research and for teaching.[23] I think this is a central and little-noted phenomenon of the digital humanities world. Almost all of the significant humanities databases are at least potentially dual-functional in this way, although of course some, like the Ayers-Thomas site, are specifically structured to facilitate teaching from them. John Unsworth has argued that the research function should guide the construction of humanities databases, since many sites built specifically for teaching have been weak analytically and from the point of view of content.

---

[23] But see Michael O'Malley and Roy Rosnezweig's argument that the Valley website is better suited to teaching than to research. Michael O'Malley and Roy Rosnezweig, "Brave New World or Blind Alley? American History on the World Wide Web," *The Journal of American History*, June 1997, 84/1, pp. 145-6.

Certainly the IATH experience give strong support for John's position. But nevertheless many

of the major scholarly sites are used for teaching, and I think that as we learn more about how to

teach digitally, this will become the norm. I suppose we should urge the creators of research

databases to think hard and early about how their web designs can subsequently be used to

support web-based teaching. And in this regard, the research being done by Mills Kelly of

George Mason University comparing the effectiveness of online with traditional classroom

teaching is extremely important.[24] But to the extent that we can make the case for the utility of

the use of digital technology in teaching, it will have a major impact on how universities think

about both databases and their creators – they may be more willing to fund their creation, and

more willing to recognize their creators for the original and productive scholars that they are.[25]

Let me now turn to a discussion of one of the most profound ways in which humanities

scholarship is being transformed by the introduction of digital technology – publication. We

used to think that defining "publication" was straightforward. By and large we meant the

process of submitting text, book or article length, to a institution that assessed it for quality and

reproduced the suitable submissions in some sort of bound print format – normally a codex –

offered for sale to the public. Legally, of course, the question of whether a text had been

published was and is quite complex, but I want to set that consideration aside. The transaction

was one between the author, the creator of the text, and the publisher, a specialized printing

organization. To some extent, this model still describes born-digital publication – an author can

submit a text to a publisher who will convert it into a digital file and offer it for sale as an e-

book, for instance. We have only just begun to go down this road for scholarly humanities

---

[24] http://chnm.gmu.edu/history/faculty/kelly/wciv/webography/webography.htm
[25] RichardLatner,"TeaSearch:ResearchThatCanAlsoTeach"http://www.h-net.org/teaching/essays/latner.html

publications, but this is precisely the main objective of the ACLS History E-Book Project and also of the American Historical Association's Gutenberg-e project[26], through which the Columbia University Press is producing prize-wining doctoral dissertations as born-digital books. The process is straightforward, but it leaves two questions unanswered. The first is whether humanities e-books will simply be digital traditional humanities texts, or whether they will be databases that take full advantage of the technology to create forms of knowledge that could not be presented in analog form? The second is whether the scholarly academic community will accept born-digital books as "real books" for purposes of promotion and tenure.

But if we think harder about e-publication, I think we will see that its manifestations are much broader and more complex. After all, one could (and, I think, should) say that an e-publication is among other things anything that is posted on the Internet, quite possibly including group e-mail (of the H-Net sort). After all, anything we post on our personal websites (such as this talk, as soon as I return home) is functionally published – if by publication we mean made generally available to the public. Moreover, most of what is personally published in this way is literally free to users. We also publish in only a slightly more formal way by posting material created for institutions. My Center at Princeton, the Center for Arts and Cultural Policy Studies, has recently posted all of its Working Papers on our website. Formerly we printed them out and sold them for a small fee, but we are more interested in having them read than in trying to recover our small analog publishing costs. There is a good deal less self-e-publication in the humanities than in the natural and life sciences, where the electronic pre-print has become the dominant form of publication, but the humanists are moving inexorably in the same direction.

---

[26] http://www.theaha.org/prizes/gutenberg/rdarnton2.cfm

And the irony is that improved search portals such as Google.com now normally make it easier for most people to find e-publications than print publications.

But of course posting on the Web is not at all the same as publishing a book or article with an established academic book publisher or journal. Among other things, traditional publishers provided the gate-keeping functions that ensured quality control. They increased quality through professional editing. They, in conjunction with academic libraries, ensured the long-term preservation of publications. Self-publication and non-peer reviewed Web publication raise serious questions of authority (who can vouch for authenticity?), quality and preservation of web-based material. Nevertheless, admitting these serious difficulties, scholarship is distributed more quickly and broadly than ever before. Even in the humanities, where time-value is not as important as in the sciences, this is a significant development.

But of course, as I have already noted, much electronic publication *is* peer-reviewed. There are now huge projects to retrospectively convert earlier humanities publications to digital form – the Oxford-Cambridge history project mentioned earlier is only one example. The number of campus-based e-text centers along the lines of the excellent unit created by David Seaman at the University of Virginia[27] is an example of our capacity to digitize large quantities of public domain text at very reasonable cost. And of course the rapidly growing number of humanities research/teaching databases are in themselves a new scholarly publishing industry. The earliest of these projects were produced for sale and distribution on CD-ROMs, but now they are mostly "published" online.

---

[27] http://etext.lib.virginia.edu/

E-publication thus has enormous promise for the humanities. It can be much more than digitized text. It can be interactive. And it normally needs to be produced in a collaborative fashion, given both its technological requirements and the ambitious intellectual scale of the new projects. Humanists have traditionally worked alone, but in the digital environment we need to interact with technologists, who themselves have to be knowledgeable about the content of the material they are manipulating. Digital publications have already opened new intellectual frontiers for the humanities and challenged our traditional methods and structures. That trend will continue, and I welcome it.

But there are problems in the digital paradise. Let me mention a few, briefly. The first is preservation of digital publications, since they are produced in an inherently unstable medium. It is not even clear who has the responsibility to look after the preservation of digital objects in the manner that libraries have traditionally taken responsibility for the preservation of books and journals. The second is access. While from one point of view the Internet provides the potential for global access to Web-based information, the digital divide creates the reality of tremendous asymmetries in access, even within the wealthier nations. And a potentially greater obstacle to access is the increasing privatization of digital information, given recent trends in the international law of intellectual property and the economic transition from the sale of information, especially to libraries, to the licensing of information, the legal transformation from a world of ownership to a world of mere rights to use. If the scholar is lucky enough to work in a wealthy university whose library can afford to purchase access to the huge quantity of proprietary information available only by license, he is very fortunate indeed. But many other scholars and the general public face the threat of being shut off from access to the same sources.

As I have mentioned, digital publication creates severe dilemmas in the specification of authority of (especially) Web-based publications. It creates difficulties of maintaining the uniformity of publication standards necessary to ensure interoperability. It creates challenges to peer review, for even if digital products are reviewed before dissemination, we need new modalities for assessing complex, interactive and multimedia products. It is less clear than it used to be who is the appropriate reviewer, since technology is now part of scholarly digital methodology, and most senior scholars at this point in time will have difficulty in assessing the full implications of digital scholarship. And all of these dilemmas also exist for post-publication peer-review for promotion, tenure and compensation in the university.

But perhaps the most interesting challenge of digital publication is how to deal with the question of finality. Roy Rosenzweig has argued that in many online publications it is hard to tell when an essay actually "ends.

> In effect, these essays undercut the unwritten social contract that exists between readers and writers of scholarly essays -- the social contract in which the author agrees to follow conventions of argumentation, organization, and documentation and the reader agrees to devote a certain amount of time to give the article a fair reading. They also undercut the sense in which a scholarly article can be seen as a very clearly defined and bounded "product," with a niche into which it cleanly slots. (Not coincidentally, they also unsettle

our standards for what "counts" as scholarly publication -- are these one article or many?

Publications of work in progress?)[28]

When a book is published, it is in some sense "done." It may appear in a subsequent edition, but

serial editions are easily tracked. The text cannot be altered except by being republished. Not so

for digital publications. They can be altered after being posted by the author, the publisher – and

even, under some circumstances, by the user community. Which is the real, or published

version? How do we authorize and cite such subject-to-alteration information? How do we

preserve the changes? Is a digital publication ever done? To the extent that digital texts are

interactive, is the real text simply one's person user-experience of the text? Even the formats are

unstable. Is the forthcoming Ayers-Thomas <u>AHR</u> article best described as an "article" in its

digital form? If not, what is it? Inevitably, as with any new technology, there are problems.

Nevertheless the possibilities opening up for the humanities in the Era of Information

Technology are immense. For me, the glass is half-full. And yet I recognize that for all that we

have accomplished, most humanities scholars have at best a passive relationship to the new

electronic world. This was the theme with which I began my lecture. On any university campus

or in any humanities disciplinary association, it is only the hardy few who knowledgeably

engage with more than the availability of digital and Internet-based information. The scholarship

of most humanists has been little affected by the more sophisticated aspects of the new

technology. They still teach in traditional ways, though they are all too likely to employ the

---

[28] Roy Rosnezweig, "Crashing the System? Hypertext and Scholarship on American Culture," *American Quarterly*, 51/2, 1999, pp. 244-5

most idiotic and counterproductive technologies, such as Power Point. We have a long way to go, then, but we are headed in the right direction.

I believe that our efforts now need to be institutionalized. As university scholars, our comparative advantage is in the magnificent way that the academic library community has embraced the new technology, and in the collaborative opportunities presented by libraries. This is why, for instance, we sought a grant to Firestone, the Princeton University library, in order to build the CPANDA database. My center contributed the subject matter expertise, but the library contributed both the technical skills (mainly provided by a medieval historian turned computer librarian, by the way) and the physical technology. And the library is committed to maintaining the database. This is precisely what has happened at the University of Virginia, where IATH resides in the Alderman Library.

But, whether in the library or not, each university needs to provide something like the IATH functionality. We need skilled technologists, ideally well-informed about or at least sensitive to humanities knowledge, to work with scholars as they create humanities databases. Few scholars can go it alone in this technology era. And if I am correct about the tight fit between improvements in teaching and digital teaching resources, campus humanities IT centers will also be or be related to the teaching/learning center. This center will require, in an ideal world, humanities computer specialists of the sort being trained at the Centre for Computing in the Humanities.

We will also require transformed university presses to produce and disseminate born-digital products. They, too, will need humanities computer specialists. They will need to collaborate systematically with both libraries (as the campus technology center) and scholars. And they will need to devise both new methods of peer review suitable to the digital environment and new modalities of "editing" suitable to the enhancement of electronic texts.

All of these institutions need to be more interactive and collaborative with the scholarly community. The old scholarly communications model of "we research in the library, we write at home and the publishers publish at their presses" simply cannot either describe the world we are already living in, nor can it provide the skills, collaboration and knowledge necessary to produce scholarship in the digital era. Some years ago I participated in a project of the Welch Medical Library at Johns Hopkins University which aimed to create what it called the "Knowledge Management" model for the creation and dissemination of knowledge. The idea, to oversimplify, was to have a press-library collaboration with bio-medical scholars to create an online genetics textbook that would be a living work of scholarship, constantly being updated by its authors and its users, and constantly being maintained by the library and press. Not a bad model for the humanities, as well.

So I think the brave new world of the digital humanities is coming into existence, though we have far to go in order to bring it to maturity. We need much more money, many more trained humanities technologists, new institutions both on-campus and off. We need to educate our scholarly colleagues and raise their consciousness of the promise of the new world. But

above all we need an articulated vision of where we are headed, and what we need to do in order

to get there.[29] My hope is that this series of lectures will be an important step in that direction.

Thank you.

---

[29] See the discussions of the January 2003 conference, "Transforming Disciplines, Computer Science and the Humanities," http://carnegie.rice.edu/toc.cfm