COMPUTING AND THE HUMANITIES:
Summary of a Roundtable Meeting

Computer Science and Telecommunications Board,
National Research Council

in collaboration with

Coalition for Networked Information
National Initiative for a Networked Cultural Heritage
Two Ravens Institute

American Council of Learned Societies

ACLS OCCASIONAL PAPER, No. 41

ISSN 1041-536X

COMPUTING AND THE HUMANITIES:
Summary of a Roundtable Meeting

Computer Science and Telecommunications Board,
National Research Council

in collaboration with

Coalition for Networked Information
National Initiative for a Networked Cultural Heritage
Two Ravens Institute

American Council of Learned Societies

ACLS OCCASIONAL PAPER, No. 41

Contents

Note: This document reports the proceedings of a brainstorming discussion. It has been reviewed only by participants of the meeting and is not an official report of the National Research Council.

By the end of the day, we would like to have an answer to the question of what the Computer Science and Telecommunications Board can do to promote the application of information technology to the understanding of the human record.

—William A. Wulf

I. Introduction and Background

The preceding quotation is taken from opening remarks by William Wulf during an unprecedented meeting at the National Academy of Sciences building (Washington, DC) on March 28, 1997. The Computer Science and Telecommunications Board (CSTB) of the National Research Council sponsored a gathering of prominent individuals in the fields of computing and communications science, and arts and humanities research, in an attempt to explore the complexities of cross-disciplinary collaboration. The CSTB was joined by the Coalition for Networked Information, the National Initiative for a Networked Cultural Heritage, and the Two Ravens Institute (see Appendix One for information on the organizers) in planning a program that began with an exploration of the interests and methodologies of computer scientists and humanists, proceeded with a consideration of economic and institutional factors that shape how different communities pursue their interests, and culminated in brainstorming about the prospects for additional and more sustained interaction among the communities.

Among the key assumptions informing this meeting was an awareness that:

- facilitating the preservation of and access to humanistic information resources across a range of systems and media was of national importance;
- understanding of the human record would be significantly enhanced by developing cross-disciplinary and cross-media compatibility of systems to enable researchers and the general public to search a wide selection of humanities material in disparate locations, and to do so easily and creatively;

- identifying mutually satisfying mechanisms that would enable humanists to work more effectively, productively, and creatively with industry and academic technologists and generate new applications would be of enormous value to humanists and challenge computer scientists into new ways of thinking;
- understanding the intrinsic qualities of the material to be studied to enable appropriate electronic conversion and to foster generation of new material that may be possible only via electronic media would greatly enrich our cultural legacy.

Until very recently, computing and communications science was not seen as a partner to the arts and humanities (hereafter "computer science" and "humanities" for brevity). This perception was due in part to the exceptionally varied nature of humanistic study, an often loose coalition of disciplines that can include philosophy and some schools of sociology, anthropology and music, fine arts and the history of technology, and even some aspects of environmental studies (see Appendix Two). Cultural institutions such as museums and libraries are also included in the humanities purview. This realm of intellectual pursuit was deemed antithetical to precise, mathematical computing applications.

Yet all of these varied areas of study focus on the preservation, transmission, and interpretation of the human record, and all of these humanistic fields are emerging as implementers of computer-based technologies in scholarship and teaching applications. Despite an initial lower penetration of information technology, funding constraints, and inherently conservative methods, the humanities, like other groups in society, are making greater use of computing and communications for efficiency, long-distance interaction, and new forms of activity made possible by technology.

Early applications of computer science in the humanities involved statistical analysis of texts (digital concordances and determination of authorship), and computing of statistics relating to historical studies (census results, shifting wartime populations). More recent technological advances have facilitated progress in symbolic and textual storage and analysis, and newer technology enhances the storage and use of audio, image, and video material. Projects such as American Memory, ARTFL (American and French Research on the Treasury of the French Language), Perseus, the Making of America Project, and the Museum Educational and Site Licensing Project testify to increasingly sophisticated applications in the humanities encompassing text, image, sound, and multimedia combinations.[1] They also demonstrate how technology can strengthen links between the creation of new knowledge—research—and its dissemination—teaching and public presentations.

Computer science, in turn, has often focused on scientific and business applications, involving processing of numerical data and simply formatted text. Trends in computing and communications illuminate new opportunities associated with electronic "content": visualization of all kinds of information, digitization of music and other audio material, electronic publishing with increasingly elaborate formats, and more complex paradigms for finding, selecting, displaying, and making use of greater quantities and varieties of information.

These applications align the interests of computer scientists and humanists, and they spurred the convening of this group of practitioners. Recent progress in information technologies in the humanities and computer science also encourages more programmatic collaboration. Going into the March (1997) meeting, it was understood that effective partnerships among computing and humanities professionals call for an understanding of the opportunities in both sets of disciplines: what are shared intellectual interests; which problems in one domain provide challenges for the other; where can collaborations prove particularly fruitful. That understanding is uneven at best today. It also points to a range of important process questions: how do computing and humanities professionals do their work; what are their respective expectations for project size, cost, structure, and outcome; what barriers or constraints tend to shape projects; what incentives and disincentives can affect collaborations?

The March meeting took these questions as its point of departure. (See Appendix Three for agenda and participants.) The ad hoc group framed these issues by characterizing the needs of humanities scholars, delineating the systems now available and the economic and institutional context in which they work, and exploring the processes by which future progress can take place.

A digest of those discussions follows; the roundtable members are quoted as often as possible to convey the tenor and interactive nature of the meeting. Among the most important themes of the day's discussions is *collaboration*, which depends on language, methodology, and institutional context. Another is *structure*—the technical means (e.g., standards) for humanists to integrate and use computing systems in their research and applications. The third broad theme is *resources*, including money, time, and talent. These themes interact to the extent that shared intellectual excitement can bring humanists and computer scientists together and inspire new tools for humanities activities—but resource issues may determine ultimate outcomes.

II. Toward a Common Language: Methods and Context

Visions of the Future

Computer scientists and humanists often start with differing views of how computing should be developed for humanities applications. This contrast can be illustrated at the desktop, where most humanists work. Noting the fragmentation of information resources across physical and electronic sources, Jerome Saltzer proposed a goal

> *. . . of a digital library world, where every document in the library is something that I can inspect from my desktop. I am not looking for something elaborate. All I want is the ability to look at documents. The other goal for that future system that I think is within reach is that whenever you spot something in a document that makes a reference to something else, it ought to be possible to click on it and have that appear in the next window.*

Joseph Busch argued that not all information may be appropriate for desktop delivery.

> *I am a great believer that this is not something that everybody should do on their desktop. Not everybody should be an information server or is a good information server. I think we need to have service providers: not necessarily commercial service providers, but we have to understand the need for an institutional role of being a service provider.*

What services are provided by whom and how services are provided can have a fundamental impact on both system design and institutional roles.

Effective system design can be an elusive goal. Michael Brodie pointed to systems failures in banking, telecommunications, and other commercial sectors as evidence of the difficulty of moving from a theoretical vision to workable systems design and implementation. A first step is to understand the "humanities process," analogous to various business processes.

> *What is the method by which one establishes a premise and comes to a conclusion by interacting with other people? You study that process and translate it into the appropriate information technology requirements and build it. If you build a vague thing to support a general notion, it is very unlikely that you will succeed unless by some wonderful happenstance. . . . Run rampant in creating the vision but do justice to the resources that you request in order to achieve it.*

Collaboration, Brodie suggested, should extend to development of the ontologies that drive automation processes.

Much discussion of information technology focuses on information elements. But many also emphasize a vision of the future that more fully engages users as social animals, as individuals with distinct personalities. Computing and communications technologies have different merits, and considering them as a set brings with it the risk of missing important differences in their value to people. Mary Shaw suggested that contrasting the newer Web with older newsgroups can help to understand some of those differences.

> *The World Wide Web is one of the most sterile ways for people to interact with each other because the central human interaction has been washed out. As originally created, it was a way of pointing at some document out in cyberspace and saying, "I wish to see that document now." It has been larded about with mechanisms that let the documents animate themselves, but there is still not very much that lets you interact with real people. A much older use of the Internet involves electronic mail and newsgroups, which allow communities to develop around a shared interest in a shared topic.*
>
> *An important piece of technology differentiates the newsgroups from the World Wide Web and makes the information genuinely accessible to people in forms that they understand. Historically we have concentrated on stringing wires and making bits go faster. We had to do that to get started. But we now need to devote energy to making the information capabilities accessible to real people in the terms that they can understand and control rather than just making things go faster with more connectivity, more people, more kinds of representations. Technologists need to rebalance the investment in the underlying network and the investment in packaging the capability and the form that reaches the intended audience.*

The idea of active use and involvement with a medium as the key to its potential was amplified by Willard McCarty and Richard Liebhaber. Liebhaber argued that such potential implies a radical change in what people do with evolving technology.

> *Contextually, we are still focused on the written and statistical, whereas the movement is toward visual and virtual. We are working methodologically in a world of "store and forward," and we are moving toward a world of "forward and store." Our methodologies and our view of the technology in question is*

based on storing material somewhere and forwarding it to a user or a researcher. I believe technology and price, performance, and some of the other issues we discussed are leading us to a world of "forward and store." That is, instead of making it narrow and real-time, I am going to make it fatter and less than real-time, put it into a device and deal with it visually and virtually. I will not be connected to anything that is physical. What troubles me even more is that I believe the research and the real contextual thought that is going on to make those two changes happen is not being done in an academic environment, but in the commercial and industrial environment.

Liebhaber explained that progress in computing technology performance and associated reductions in cost will allow the movement of functionality from work space into play space and, ultimately, households. Trends also support a repositorial view of information: together with the economics supporting greater connectivity, this may broaden the potential for human studies and human interaction. Intellectual property becomes more important as the barriers from costly transport and storage and easy control of ideas and materials fall.

Imagine a world in which each of us had the opportunity to create for ourselves our own special interest magazines, which change by time of year or phase of life or whims of interest. Imagine a world in which instead of reading about the world of Thomas Aquinas, we download a series of images of Thomas Aquinas.

Liebhaber urged humanists to plan for novel system capabilities and designs, referring to his shift from store/forward to forward/store. Part of this trend involves recognizing and planning for growth in the consuming public, for whom the expansion of do-it-yourself options with automation extends to the finding, publishing, and repositorial handling of humanities material.

The coupling of technology and process is mediated by economics: people choose what they can afford. Responding to Edward Ayers' discussion of the potential of electronic media to support "democratic" access to humanities content, Charles Henry cautioned that humanists may have to work against the rhetoric of the National Information Infrastructure,[2] which tends to be commercial in its metaphors—pipelines, conduits, highways, and toll booths as "things that you can direct and meter." Humanists find little comfort in computer scientists' observations about emerging technology for fine-tuning charging schemes. Their experiences will, in any event, be shaped by the larger set of non-technical factors that govern intellectual property rights and associated public policy.

Jerome Saltzer observed that computer scientists themselves may not understand all of the potential of information technology.

> *Things that cost $100,000 in 1983 we can do for $100 today. . . . Most computer scientists are still reacting to this factor of a thousand that has come upon them. . . . They are still trying to figure out how to unwind all of the decisions that were made ten years ago based on things that cost 100 times as much, and get into step with today's costs. They are not looking ahead to see that the things that cost $100 today are soon going to cost $1.00. So when humanists talk to computer scientists, you have to realize that most computer scientists are, in a funny way, tied up in the past.*

Even scientists have difficulty coping with the rapid change in computing technology.

Universes of Discourse

Overcoming language differences was a major theme of the roundtable: all agreed that it is an issue for any cross-disciplinary collaboration. Pragmatically, Edward Ayers pointed to the rise of a pidgin language that spans computing and other disciplines. Michael Joyce captured the challenge eloquently in an opening statement using unusual words to illustrate the challenge of building a common language (See Appendix Four).

Willard McCarty observed that the homogenizing influence of information technology on methodology provides a basis for a common language.

> *What jumps immediately into focus after five years of teaching humanities computing to graduate students at Toronto, and now undergraduates and postgraduates in London, is the importance of methodologies. When you teach humanities computing what immediately becomes obvious is that the only subject you have to talk about is the methodology. You find that this is a great deal to talk about. There is a huge and very rich interdisciplinary common ground of techniques relating to the data that people in the humanities deal with, whatever their particular disciplinary orientation. There are common tools and techniques. One can easily take, for example, relational database management and highlight examples in various disciplines and have people recognize the utility of these things. But much more important*

> *than that is the new perspective on these disciplines that these tools bring about. Here I have the very strong suspicion that this transcends the humanities and has to do with what has been going on in the sciences. We do not know very much about this yet.... We can get to this common language by seeing what "falls out" from the application of computing to the humanities during the fifty years since Roberto Busa started his project and in the sciences.*

Methodology can provide one of the spaces where humanists and computer scientists meet. Summarizing his description of a variety of "potentiated spaces" created by humanists using computers, Joyce suggested that humanist and computer science methodologies have a reciprocal relationship: "our presence as human persons in real places continues as a value not despite but because of the ubiquity of virtual spaces."

Collaboration: "Synching Up"

Effective collaborations may be serendipitous, but they are often not accidental. A critical element of any collaboration is mutual respect, and that element is especially important in collaboration across disciplines. Individual disciplines tend to be chauvinistic, and disciplines that generate useful tools—including computer science—are concerned to have their contributions appreciated as such. Willard McCarty addressed that issue:

> *"Computing in the humanities" is not quite right, because the humanities do not own computing and should not swallow it up. We have to think much more in terms of a common language. The phrase "common languages," as George Steiner recognized, refers to the state of bliss before battle when everyone speaks one tongue. But we have suffered this fall into disciplinary specialization, and we have these separate city-states. . . . Perhaps a better analogy is the perspective of the Phoenician merchants who invented the alphabet—of people moving between civilizations needing to invent a meta-device, a meta-language, if you will, to represent the goods and services that they were trading to people speaking incompatible languages. . . . "Computing in the humanities" is not right because it suggests that computing moves into the disciplines and becomes absorbed in them. So you have a kind of Marxist theory of the withering away of the state and finally you have all of these*

disciplines with each of their computing experts in them. This is not the way to do things, because it ignores the fundamental contribution of the computer to the interdisciplinary dialogue, the fact that there is all of this material held in common in these techniques and approaches.

Although the roundtable was labeled "computing *and* the humanities" in recognition of the concerns McCarty raised, he cautioned against inferring from that wording that computing and the humanities are fundamentally separate. That inference, he explained, is "an illusion caused by a lack of historical perspective and perpetuated in the discipline-based structure of our institutions."

Stanley N. Katz underscored the challenge of transcending discipline-based structures, noting the mix of strength and constraint that they imply:

> *Probably everyone around this table would raise the banner of interdisciplinarity or of multidisciplinarity—or I like to talk about it as "nondisciplinarity." But probably none of us, if we are thoughtful, wants to give up disciplinarity. Method is important. It is enormously important. We do not know how to maintain a structure in which we have the virtues of both method and nondisciplinarity. It seems to me to be an enormously important challenge and very relevant to the kind of collaboration that we are talking about here.*

McCarty proposed looking more broadly at "humane learning, which includes the sciences." He suggested that computing from a humanistic point of view addresses how people think.

> *The computer, from a humanist's perspective, and I think from several others, is essentially a modeling device. That is, through it we determine or we play with how scholars think. And the interesting question that arises is the discrepancy between how the computer manages to do things and how human beings manage to do things. This generates many questions of interest to humanists because it has intimately to do with how the research perspective changes once you begin computing your texts, images, sounds, and other forms of humanities data.*

McCarty noted that scholars who shaped computing—for example, Alan Turing and Vannevar Bush—used models from the cognitive science branches of philosophy, neuropsychiatry, and other fields.

Many pioneering applications in the humanities have been documented, and several are accessible on the World Wide Web. (See Appendix Five.) To illustrate emerging possibilities, some of which

relate directly to the challenge of understanding and exploiting how people think, Mary Shaw presented examples of cross-disciplinary collaboration among computing, other sciences, and the arts. The Journey into the Living Cell project of the Carnegie Mellon University Studio for Creative Inquiry, which draws on the Fine Arts Department and Carnegie Science Center, was motivated by artists, who saw an opportunity to reach a new audience with a new medium. An interactive, visual system developed by biologists and artists, it involves the audience as collaborators in a performance. Another interactive performance project, associated with CMU's digital library project, involves an interactive representation of Albert Einstein using a script performed in different orders depending on the audience. Elements of this project include speech recognition and synthesis, information retrieval, and digital video. And Shaw's own work includes an education application that includes simulations, record keeping, and documents. She is collaborating with cognitive psychologists to understand how learning takes place.

Shaw derived from her and others' experiences some principles for effective collaboration that were echoed throughout the roundtable discussions. Foremost is for people to have some common cause, such as working on a particular project. This can generate positive feedback (drawing in students, colleagues, and follow-on activities). Participants must be full partners in order to reconcile different points of view along with different research styles, languages, and cultures. And institutions must find ways to remove administrative barriers (allocation of contract overhead, turf battles over income) in order to create the conditions for collaborations and for opportunities to be seized spontaneously when they arise. Edward Fox reinforced that point, noting that collaboration "cannot be marginalized," with humanists outside of a science project. He suggested that this had been an issue with the federal Digital Library Initiative (DLI).[3] Shaw commented more broadly that from the top down, institutions can only enhance or interfere.

Thomas DeFanti echoed Shaw by describing how partnerships between computer scientists and artists created the subdiscipline of computer graphics. He noted that SIGGRAPH conferences are distinguished from other computer meetings by the participation of artists, illustrators, and other visual workers, male and female.[4] Part of this mixture reflects the aesthetic imperative: "early on during the growth period it was clear that if we were going to promote this technology it had to look good." DeFanti's current lab engages artists as project managers because they surpass scientists and engineers in driving the software to commercialization. These artists have scientific training; they

are "a very cross-cultural bunch of people who are actually generating the technology."

Sandria Freitag suggested that the seemingly trivial frustrations about converting to different software may express a more fundamental concern about the match of technology to process: "Many people were right: WordPerfect is a much better package to work with than Word if you are writing." Humanists often perceive limited ability to shape the technology they use. Concern grows "from that very simple level to the much more complicated level of how do you harness the technology that can deal with the visual and the virtual." She remarked on the tensions revealed in the roundtable discussions—"between the notion that there is a technological expertise out there that can be leveraged on behalf of the humanities; and the other side of the issue, which is that the communities themselves have to structure these things, have to create them to address their own needs."

Freitag and others noted the emergence of a new mindset: a more subtle and profound transformation of what humanists do, and how that results from merely incremental improvements to the usability of any particular system. Michael Neuman cautioned against generalizing that phenomenon, at least in the near future. Attitudes about technology vary across the humanities. It may not be simply a matter of time before all humanists embrace computing tools and methodologies; it may not be possible to reach the entire community with any single project. But Willard McCarty suggested that broad forces shaping education impel accelerated attention to computing in the humanities.

> *Students come to us with the computer in their mental vocabulary as a kind of model for how thinking takes place. It is extremely important that humanists deal with computing in order to carry out their mission. We have no choice about this whatsoever. We also need to communicate with the public since we are no longer swabbed in superstitious reverence for what we do, as was true in my parents' generation.*

Humanists, explained Edward Ayers, have been trained in "how to deal with one kind of technology, the book," although it has taken centuries to master that technology. The more diverse and heterogeneous world of electronic information poses major challenges for instructors at all levels. The strong links between different kinds of humanities activities and education motivate interest in information technology as amplifying agents as well as interest in options for disseminating information about projects and systems with educational value.

Bruce Schatz noted that even in scientific contexts, professors and graduate students show different preferences for software and systems. Drawing from his experience with genome analysis systems, Schatz asked:

> *Would you like to have something that is used by professionals in the field now? You need to find something that they really want, that is so important to their lives that they will spend time using the systems that do not actually work. Or do you want something that is used by graduate students in scientific areas or undergraduates in teaching areas that does not really do anything useful but illustrates to them what the world will be like when they are professionals? Most of the high-end research systems are very, very popular among the graduate students because they know that that is what the world will be like five or ten years from now. But professors and professionals tend not to care because they can see that their immediate problem can be done more easily by some simpler mechanism. So you have to think about what kind of project you want to have before you start; that determines almost all of the other choices.*

Technology holds the promise of alleviating some of the problems of change that it creates. For example, Edward Ayers described how his center uses a model in which "an individual project has to be the carrier of its own set of instructions, somewhat like DNA; it has to be able to tell people how to use it at the same time." More generally, Tom DeFanti noted the impetus from distance learning:

> *If distance learning is going to happen, it will need everything that you are talking about in order for it to be successful. I think that it is going to happen because the technology will be there. How long it takes, I have no idea. How long it will be stalled, I have no idea.*

III. Software and Standards Development

Michael Joyce contrasted conventional and computer-based interactions: "Here we are visible within an invisible technology. Elsewhere we are invisible within a visible technology." Such characterizations help to frame questions about why humanists want software, what they want it to do, and how they want it to operate.

Computing and communications can affect what humanists work with—the presentation, representation, and cumulation of content—and how they do their work—the tools and methodologies. In humanities as elsewhere, software and standards drive available capabilities. Bruce Schatz noted that humanities applications benefit from the propensity of humanists to read, illustrated by a history of efforts to build electronic corpora and concordances and the building of systems to support analysis across multiple sources. He pointed to the emergence of more potent capabilities: "The era of comparative literature being technically possible for large bodies of humanistic material is coming very fast. It is not being pushed by humanities at all, but they are likely to be the greatest users."

Roundtable participants agreed on the need for more and better tools for creating, browsing, storing, finding, and retrieving complex objects with varying degrees of structure, from document-like to database and combinations. For example, Joseph Busch remarked that "databases are not necessarily oriented to functioning in new ways, and software is needed to make databases interoperate more with document-like objects." Databases, networks, and storage systems can enable wider overall participation in humanities activities.

Michael Neuman pointed to the need for item-level access to works under examination.

> *My background is in English literature. In Othello, as the play moves to its tragic conclusion, there is a stage direction that says "enter Othello and Desdemona in her bed." Well, a scholar asks how, on a stage without curtains, does a character enter in a bed? In what other plays can we find beds that appear in stage directions? So we would need not only access to a rich corpus of dramatic data, but also the ability to search for the word "bed" within the context of a structural component, namely the stage direction. Simple text searching will not be sufficient. The number of hits on the word "bed" would be monumental. So we need something greater than item-level access. We need component-level access.*

Visual and audio elements round out the potential represented by textual material and may drive new directions in scholarship. Mused Stephen Franklin:

> *If you go back 150 years, you can argue that the predominant forms of literature are textual: plays and poetry. I think that when the infrastructure gets there, Ed Ayers is going to start having oral material directly, rather than its representation in text. There is an enormous emotional impact. [There have been problems in ethnomusicology, for example, where] people could not sell enough CDs to justify even private pressing. So I do see, along with Chuck Henry, an enormous opportunity for democratization in terms of the publication.*

Beyond consideration of tools and treatment for specific kinds of media is telepresence, which promises to eliminate costs of transport and storage, as well as the ability to control ideas and material. Other complex systems will shape perceptions of what is possible, moving expectations beyond those that can be characterized as "automation." Willard McCarty acknowledged the difficulty of conceptualizing new possibilities.

> *The idea that a humanist should be able to get in the car and drive from point A to point B, as he or she has always done, rather than first fix the engine and perhaps manufacture a part so that it can work before it can drive from A to B, is fundamentally mistaken. What happens when you have a new form of transportation is you do not drive from A to B anymore because all of a sudden you can go to C, D, E, and F. So the whole world changes around you. While we still have people who only want to go from A to B, there must also be people who are losing time in that sense of struggling, of observing, of musing about and experiencing the most important thing of all, how the world has been changed by the technology. The derivation of a common language is from a group of people who march into the open spaces that Michael Joyce talked about and set up camp there—not from people who sit in one armed camp or another and receive envoys and go back and forth between these different realms, but from people who set up in the new realm and begin to operate in the common language.*

At issue in this discussion is evolving thinking about digital libraries and the nature of information systems generally. The digital library concept has broad appeal; in part because of the practical realities that

digital library projects represent, they are the focus of a major United States government research and demonstration initiative, and they represent a logical, comprehensible advance over conventional libraries.[6] Thus Joseph Busch declared that "operating in the information society means having access to a content-rich and context-rich distributed digital library." Library research today illustrates the difficulties in bridging electronic and paper worlds. As Jerome Saltzer noted, "if half of your material is on the computer and the other half is in the form of paper, that does not work nearly as well as working entirely on paper or entirely on a computer system." More sophisticated tools, such as those for story telling, pose even greater challenges. Beyond finding and reading the bits, problems of *interpreting* them are common across different kinds of data, remarked Jerome Saltzer, and are emblematic of the historian's challenge.

Shared artifacts and information were part of the rationale for building digital libraries. Future projects may involve spaces for collaboration as well as for artifacts, and evolving perspectives about what people want to do and how this will affect the design and architecture of systems.

Edward Fox proposed a framework for assessing system and software possibilities:

> *Fundamentally, we deal with four constructs here: streams, structures, spaces, and scenarios. In the world of streams, we have spent a long time talking about text. Now we are moving to multimedia, oral, and other kinds of representations. In both the "real" world and virtual new worlds, we have analog streams and bit streams. Structures is another area. We sometimes talk about syntax as an example, and we have spent a lot of time focusing on databases. In recent years there has been a focus on what is called unstructured information which is really not unstructured: an example is the book, which along with other kinds of literary works has tremendous structure. We call them unstructured because we do not know how to deal with them properly. The third aspect here is spaces. We are shifting, in part, from physical spaces to virtual spaces—in part, from term spaces to concept spaces. We have also heard today about methodological spaces, emotional spaces, collaborational spaces, and cultural spaces. The fourth aspect is scenarios. This gets into the shift from passive to active. On the computer side, we talk about functions and tasks. But in the humanities we speak of stories and of interactions, and these are other kinds of scenarios.*

The very concept of structure implies decision-making—*what* kind of structure will be designed—which in turn raises questions about *who* makes what decisions and *who* interprets what for whom. Humanists underscored the importance of finding ways to engage users of humanities information in building structures for content and analysis, observing that this did not happen often enough in the past. Today technology is making it easier for users to populate, update, and maintain databases, and infrastructures are needed to support such user roles.

Pros, Cons, and Experience with Standards

Standards can define functions, processes, or services. Humanist applications have advanced with the use of encoding and cataloguing/classification standards (Text Encoding Initiative [TEI], Standard Generalized Markup Language [SGML], and Hypertext Markup Language [HTML], and Z39.50[7]), but these same standards have limitations recognized by both humanists and technologists. Standards can extend to a variety of procedural or administrative functions, including the management of rights and permissions associated with intellectual property protection and more generally the use of content and services for which there are restrictions, fees, or other conditions.

Standards are a necessary evil or a mixed blessing, depending on one's particular perspective. Humanists, like other users of information technology, asked for standards in the hope of lessening the distractions of figuring out how to work with the technology. Technologists, mindful of how fast information technologies evolve (and, possibly, more tolerant of such change), argue for less emphasis on standards. Standards have the obvious appeal of facilitating the sharing of information. They also minimize the learning burden. At the same time, standards pose the risk of confining users to an obsolete technology or a lowest common denominator because of the difficulties and delays associated with setting standards.

Waves of innovation and ensuing system obsolescence in consumer and professional electronics make clear the challenge of preserving access to humanities material over time. Part of the problem relates to the changing base of support for digital formats. Michael Lesk, for example, reported that only half of the ten word-processing programs advertised in *Byte* magazine in 1985 remained on the market in 1995. In the field of music, Camilla Cai related the demise of reel-to-reel recording equipment to the problem of maintaining access to recorded information

over time after those recorded die or become otherwise unavailable. The problem is compounded, noted Joseph Busch, by the intrinsic concern of the humanities with history: the age of documents, for example, has been a long-standing area of concern. William Wulf explained that some scientific data also have properties of uniqueness—certain satellite data, for example, are not reproducible—and tend to be more concentrated than historical data.

The prospect of an evolving technology base implies a need to prepare to move beyond a given standard. In the library and archiving context, observed Lesk, the task of preserving information and content over time implies planning to refresh formats and other elements subject to standards. It is unreasonable to expect a single standard to endure. In the context of technological support for working with digital content, Mary Shaw suggested planning for technology to reconcile conflicting standards as an accompaniment to technology for helping to find content. "There are going to be different standards whether we like it or not. While it is worth investing in shared standards because that reduces the mayhem, we must, at the same time, learn to live with different standards because that is clearly unavoidable."

Edward Fox suggested considering actual experiences:

> *The TEI [Text Encoding Initiative] was a noble effort. What has happened to it? Where is it going to go? That has to be something that this group debates. It was the most rigorous investigation and representation of information that we have had in the last decade. It has not filtered into the computer science world with thinking and knowledge representation.*

Michael Neuman offered a more positive characterization of the TEI process but acknowledged a difficulty: "[the] determination of the components of text-based genres and how they could be represented using Standard Generalized Mark-up Language so that those details, given a large database, could be ferreted out, combined, juxtaposed." Neuman indicated that regardless of how well TEI served to encode humanities material, its value would not be realized so long as the tools currently available to deal with it remain intractable. He seemed to question whether the TEI as an application of SGML had any useful application to computer science work with knowledge representation, and what prospects were for scientists and humanists working together to produce more usable tools for making TEI more widely available.

Usefulness: Connect Capabilities to Real People

How much of a given content is available in a system is a critical element of its usefulness. As Joseph Busch observed about the objects of humanist activity,

> *. . . the most important condition necessary for humanists to work is a real critical mass of digital content, [which] consists not just of text but of pictures, sounds, and multimedia source materials, and the links between them.*

Computer scientists and humanists alike talk of getting at the "good stuff," but appraising and deciding on what that includes may be subjective and variable over time. The selection challenge is compounded: not only content, but format, tools, and approaches for finding, filtering, and indexing all involve selection. Especially subjective may be selection of what gets preserved in a context of prolific production of content made easier via computer technology. Information technology offers humanists and others with support for more inclusive collections as well as the threat of overload. Although technological support for information finding and filtering is growing, there are philosophical issues about the design of useful information repositories that affect both the resources available to humanists and the design and architecture of systems by computer scientists.

Bruce Schatz noted that

> *. . . the hardest thing about doing information-style projects like digital libraries is getting enough data coverage. . . . The problem with most of the electronic projects is that the information they can get electronically in an appropriate form is such a narrow segment of the actual information that people use that the system is illustrative but it is not actually useful.*

Other elements of usefulness pertain to how a system can be used. There are lessons to be learned from other domains. Bruce Schatz, for example, described how an innovative system he developed for molecular biologists studying nematode worms generated elements echoed in the larger human genome projects. "But the system itself withered and died because the things it did were illustrative of what the future was like but not immediately useful in the present." Schatz noted that objectives and activities may differ depending on the target users: practitioners, professionals, students, the general public.

Camilla Cai lamented the practical problems of learning how to use computer-based systems. Both the initial learning of a new system and the periodic need to learn additional systems add up to time lost.

Cai observed that some of her colleagues avoided using computers in order to avoid "wasting time." Other humanists echoed this concern with anecdotes about personal or departmental frustrations in adopting or upgrading information technology. These observations attest, in part, to a broad need for better user interfaces, which would facilitate adoption of computer-based technology by humanists for scholarship and for outreach to the general public (for example, systems aimed at museum patrons).

IV. Economic and Institutional Issues

Opportunities for computing and the humanities will come about in a broader context of institutional and economic realities, many of which are changing for a variety of reasons. Humanities have in common with computer science disciplinary dependence on colleges and universities; humanities tend to make particularly heavy use of libraries, as well. Educational institutions and libraries are evolving in response to the combined influences of constrained budgets, demographic change, intellectual and sponsor motivations for interdisciplinary collaborations, and the influence of information technology on costs and pedagogy.[8]

Michael Lesk pointed out that achieving critical mass in content is dependent upon capturing content, which can be expensive.

> *Running a digital library in the humanities is a problem. There is not a lot of money for these projects. There is vastly more material being produced now because of the Internet, but these resources are certainly not being catalogued. We really do not have a good answer to the question of what to do with material that is unlikely to be used.*

Lesk described how libraries serve the need for information that individuals might not be able to afford, and suggested that returning to individual acquisitions might diminish the diversity of information supported by the pooling function of a library.

Notwithstanding expectations for substantial change, scholars and leaders in these institutions work within social structures for academic recognition—tenure, promotion, access to research facilities—that seem to resist change. The result is that cross-disciplinary work is viewed with skepticism by many in academia, who fear that such work will not be rewarded comparably to intradisciplinary work, and that work involving development and implementation of systems may be stunted because of insufficient support for specialized implementation staff, documentation, evaluation, and other essential complements to the core intellectual work. Mary Shaw, for example, called for rewards for producers of editorial services, indexes, meta-indexes, and information evaluation and certification services.[9]

The structure and composition of typical humanities departments may impose additional constraints. For example, less faculty turnover as well as significant dependence on adjunct faculty who are not integrated into the university community tend to work against the introduction of new ideas and methodologies that might otherwise foster more activity in computing and the humanities.

Against this backdrop, two sets of concerns emerge as particularly strong among humanists: intellectual property rights as a driver of costs and constraints, and the infrastructure for computing and communications in the humanities. The vision articulated by Jerome Saltzer of comprehensive inspection of electronic content from desktops (and the ability to jump from citation to source) cannot currently be realized for reasons relating to access to copyrighted material in addition to the continued scarcity of access to computing and communications systems for many humanists.

Publishing

Whether the emphasis is on scholarship or on applications aimed at a broader public, humanities work is intended to be published and disseminated. In Willard McCarty's words, the Web "is a mechanism to publish documents to be used actively by people, certainly by humanists who are compulsive publishers." An associated need is increased awareness of digital content, which can itself be supported with information technology.

Information technology tantalizes by reducing reproduction costs, but dissemination also depends on the technology available to users. Drawing on frustrations associated with a large-scale Web publishing activity relating to the Civil War, Edward Ayers remarked that humanists are much stronger in the production than in the dissemination of electronic content. Infrastructure—computer equipment and communications bandwidth—remains expensive enough to be limited to institutions and scholars otherwise interested in electronic resources. Pessimistically, he suggested that

> *. . . if we cannot get this into the hands of people who want to touch the cultural heritage of the past, then I do not think it is worth the trouble. But when you have people out there with thirty students clustered around one computer on a 14.4 modem waiting for these pieces of the past to materialize before them, we are running into serious problems. And when it does materialize, if we have not equipped the teacher to point out what it might mean, then we are actually creating problems at the same time that we are supposedly creating potential.*

The economics of access are complex and outside the scope of this discussion, but they should be recognized as motivating alternative approaches to format, medium, and other aspects of information technology.[10] The importance of affordable innovations is underscored by humanists.

Costs depend on media choices and institutional relationships, all of which are in flux at this time. Edward Ayers described his center's ventures with CD-ROM publishing, explaining that CD-ROM provides higher bandwidth more easily in classrooms. He acknowledged, though, that while CD-ROM meets immediate needs, it may not have a long-term future.

It may be appropriate to consider separately material that must be captured electronically and newer material created directly in electronic forms. The opportunities to broaden access and use of newer material require planning and support to realize their potential, explained Edward Fox, who shared experiences in proselytizing, negotiating, and balancing scholarly and economic considerations through a project relating to the electronic publishing of dissertations.

Underlying the discussion of dissemination and access is what Stephen Franklin called "the gorilla behind the door": copyright. Humanists, like other members of the academic community, increasingly are concerned about the intellectual property quagmire. Many copyright holders have been keen to reassert their rights at the expense of fair use and other exemptions in the digital age, and many creators have been discouraged from publishing complex material by the difficulty of traversing the thicket of permissions. Franklin spoke of his own institution's fear of being sued for copyright infringement, which induces a chilling effect on the use of digital properties. On the other hand, pointing out that humanists are rediscovering the importance of being creators and owners of intellectual property as well as users in the digital environment, Stanley N. Katz briefly referred to the suit that the American Council of Learned Societies had brought against Viacom for infringing its rights.

Douglas Bennett noted that copyright discussions have tended to take place in an esoteric world with little connection to a group like that assembled for the roundtable: "Hardly anybody who knows what we want to do, and why we want to do it, is in the copyright discussion and vice versa." Participants agreed on the need to move beyond this situation. Humanists emphasized the need to preserve the "complicated fabric of balances" that copyright law has provided in the past. That balance has supported free circulation of information (for educational purposes) as well as recovery of costs by information aggregators and distributors. Accepting that people should be paid for good work and that money should change hands, Bennett asserted the need to preserve the basic principle of "limited sharing," analogous to the "first sale" principle, which allows one to pass on to others a book one has bought: "If we do not have limited sharing again, we are in big trouble."

Bennett also raised the issue of a potential parting of the ways between the commercial and educational worlds.

> *We are dangerously close to a situation where copyright in the future is shaped entirely around the needs of mass "infotainment" companies: Viacom, Disney, Warner, etc. What will work for them and work well for them and be good for the purposes they serve is probably not good for education and research. Underneath virtually every business plan being developed by that world is a pay-per-view strategy. Pay-per-view is deadly to the world of scholarship and research.*

Size and Wealth: Perception and Reality

Budget constraints affect all consumers, but humanists tend to find their resources more constrained than computer scientists and other members of the academic community. This concern is compounded at smaller institutions such as four-year colleges. Roundtable participants noted that access to computing and communications systems at colleges and universities is less automatic for humanists than for computer scientists. If scholars and other users lack desktop, building, and/or campus computing and communications resources, the nature and resource allocation of a project must encompass such infrastructure development or it must compensate for such a lack. Yet new technology may alter expectations and economics. As Jerome Saltzer joked, humanists may have benefits similar to those of developing nations, in the sense that they may leapfrog older and less effective technologies installed elsewhere.

The humanities are typically far less "capital-intensive" than other fields. They also tend to be less able to leverage a commercial or industrial base. William Wulf pointed out that, as a result, it is customary to speak of "the computer industry" but not a "humanities industry." Participants also noted that the dependence of museums and similar institutions on charity and other grants combines uncertainty with limited resources. This leads some to call for broader advocacy, as Joseph Busch put it, to "elevate the strategic importance of the humanities and arts in the public eye." In short, as the University of Pennsylvania's James O'Donnell has been known to quip, the humanities have traditionally been a "cheap date."

Mary Shaw asserted a need for seed funding to support "spontaneity" in scholarship, and humanists and technologists pointed to the need for

funding for computing and humanities to flow over significant periods of time. To date, many projects have received only start-up funds, which fail to guarantee continuity. Michael Lesk cited the example of a library obtaining funding to initiate or install new technology without support for long-term maintenance. There appears to be inadequate recognition of how technology creates obligations for the future. The NetLib program for distributing mathematical software continues to depend on NSF support; it is not self-sustaining. Edward Fox summarized the temporal element in humanities programs:

> *We have to support projects over time, and we have to support people over their careers. We have to deal with the rhythms of life and of people in different parts of the world and how they can collaborate across time. We have to deal with process and how we change and evolve and how students mature over their careers.*

The inherent difficulties perceived by humanists tend to vary with their response to circumstances. Joan Shigekawa described a trend toward entrepreneurialism among younger artists, who expect less from charitable sources and are creating new approaches to fund-raising.

> *We had a group of artists organized by the National Association of Artists' Organizations, based in Washington, D.C., come to the Rockefeller Foundation. All of them were under thirty. And they had quite a different point of view, a different outlook about even the making of art. It was much more entrepreneurial, like a throwback to the 1950s in not expecting so many foundation grants and not being dependent on them. These young artists are in a pioneering, self-sufficient, and entrepreneurial mode across the board, whether they are dancers or painters. Something is evolving, and it is in its earliest stages. We have been watching it with great interest, while at the same time looking for the areas where a small investment could be helpful.*

Along similar lines, Stephen Griffin estimated that thanks to cost-sharing, the Digital Library Initiative (DLI) achieves leverage of two or five to one (recognizing the difficulty of valuing collections of intellectual material). DLI partners provide staff, indexing services, and other non-monetary resources. Current projects in the four-year initiative receive $1 million annually in federal funds. Griffin acknowledged limitations in coordination among federal agencies not directly involved in the DLI as well as difficulty in creating mechanisms for co-sponsorship of nonscientific research activities. He speculated about future options: "More than anything else, the structure of the new program will have to ensure that

any type of activity, whatever its domain, can be funded at any time during the course of the effort. We will have to create a much more flexible program than the one currently available."[11]

Mary Shaw also noted that a real opportunity lurks within the problem:

> *The normal expected financial scale of computer science and the normal expected financial scale of the humanities, including the arts, are very, very different. This presents both a problem and an opportunity. The problem is fairly obvious—that it is very hard to collaborate across a gap of that size. Things that look like a big deal on one side look like a little deal on the other side, even though they turn out to be a big deal when you try to go through institutional procedures. However, this may also be an opportunity, in that there is a scale of funding in technology that, if we thought carefully and creatively, might admit an incremental funding for collaborative projects that could leverage the technology developments. Assuming genuine collaborations rather than shotgun weddings, this could provide a quite respectable level of funding for the humanities partner as an add-on to a large project. Not only funding, but also a test to the technology that is being provided as part of the larger project, would thereby be ensured.*

Questions about the lowest common denominator, and about where practical compromises are possible, are likely to become increasingly important as humanists and computer scientists work toward grander ends.

V. Next Steps: Talk First to Select Actions Better

Michael Joyce summed up roundtable sentiment by emphasizing the need to maintain dialogue while also constructing a knowledge representation that could perpetuate the roundtable's deliberations. While there might always be uncertainty about what to discuss regarding the collaboration of the humanities and computing sciences, such a representation could help frame future discourse, rendering the issues more coherent and tractable. It could be complemented by information tracking the propensity of humanists to assimilate different kinds of technology—avoiding the risk (noted by Michael Neuman) of overgeneralizing, given the differences across the humanities in terms of willingness and ability to experiment with new technology.

Willard McCarty asserted that further discussion is needed for its own sake and to support two additional objectives. First is the need to "clarify and communicate the nature of what is going on in humanities," to disseminate examples of the collection of data and the creation of access (such as an international effort to search out, catalog, and maintain a database). Second is the need for institutional models to illustrate current and future possibilities for collaboration among computer scientists and humanists.

Several participants urged additional work on existing projects as well as broader sharing of knowledge about ongoing projects. That broader sharing should be extended in particular to humanists fairly new to computing. Bruce Schatz argued that "[t]he main thing to do is to undertake a project or a series of projects that make it clear that humanities computing really works. . . . The best project would be one that the general public really cares about."

In summary, the roundtable participants offered the following suggestions in response to Dr. Wulf's original question about what CSTB might do to help foster a better understanding of the human record. They highlighted a series of activities that might perpetuate the rich and complex dialogue on the convergence of humanities and computing science:

- Publish a summary of the day's proceedings. One summary might be published by CSTB; another might appear as an American Council of Learned Societies' Occasional Paper; a third venue might offer a summary on the Web, with appropriate links to the projects mentioned.

- Organize focus groups to explore over time: methods and methodologies in the humanities and computing sciences, areas where collaboration between humanists and computing scientists might be most effective, and the potential influence of technology upon democracy.
- Organize and help fund a conference or conferences that would cover multidisciplinary demonstrations of technology in teaching and research; presentations of existing institutional models for teaching and applying technology; exploration of the uses of technology in the humanities, and how these applications benefit society as a whole; discussions of current standards (what has proven successful, what is needed for greater interoperability); exploration of the ways the business sector could be more integrated in supporting technology in the humanities, and means by which greater advocacy on the part of the business community might be achieved; presentations of case studies that highlight interdisciplinarity and shared taxonomies, including projects in the K-12 area; and a review of current debate on intellectual property, copyright, and the determining legal and economic phenomena relating to these areas.
- Organize and help fund a one-day meeting of humanities scholars to identify important collection material appropriate for digitization, and to foster coordination of digitizing projects in the United States.
- Sponsor a longer-term project to coordinate digital library initiatives in the United States with those of foreign libraries and nations.
- Establish discussion groups that would continue deliberations on themes relating to accessibility, including: bandwidth, rural access to digital resources, and wireless technology.
- Establish a body to continue discussions on the organizing themes of the roundtable: methodologies in the humanities and computing sciences, institutional and economic issues, and standards.

Other observations and recommendations included:

- Promulgating the roundtable's themes at existing conferences and venues, such as the CNI, Association for Literary and Linguistic Computing, Association for Computers and the Humanities, the various humanist professional society meetings, the digital library conferences, and SIGGRAPH and other professional computer science conferences that focus on approaches to presenting, representing, finding, organizing, and sharing information.

- Encouraging greater awareness of the importance of humanities computing, and more integration of humanities applications in future Digital Library Initiative competitions. That the Digital Library Initiative was singled out is symptomatic of a lack of frameworks for "pushing" the frontiers of computing and the humanities. Whether that particular initiative may be expanded to embrace more humanities projects is less important than that some initiative or program exist to foster computing and the humanities.
- Exploring ways to reward more programmatically the creation of meta-information.
- A formal endorsement by CSTB of the importance of humanities applications, and publication of a report on recommended outcomes of the collaboration of the humanities and computing sciences.

Notes

[1] American Memory <http://lcweb2.loc.gov/ammem/> is the online resource compiled by the Library of Congress' National Digital Library Program that provides access both to the Library's own collections of American history as well as to those of other libraries and archives across the country (through the National Digital Library Competition).

The ARTFL Project <http://humanities.uchicago.edu/ARTFL/> presents American and French Research on the Treasury of the French Language—a cooperative project established in 1981 by the French Government's Centre National de la Recherche Scientific and the University of Chicago to produce a new dictionary of the French language. In the process, a database was assembled of some 2,000 French texts from the thirteenth to twentieth centuries in literature, philosophy, the arts and sciences. ARTFL is available on the Web by subscription, although some reference works have unrestricted access.

The Perseus Project <http://www.perseus.tufts.edu/> is a constantly growing "digital library of resources for studying the ancient world," including texts, translations, philological tools, maps, essays, and 24,000 photographs. Based at Tufts University, Perseus was initiated in 1985 and now consists of a four-CD-ROM set, a large proportion of which is available without charge on the Web.

The Making of America <http://www.umdl.umich.edu/moa/> and <http://moa.cit.cornell.edu/MOA/moa-mission.html> is a collaborative digital library of primary documents of mid-nineteenth-century U.S. social history. Started in 1995 as a Mellon-funded project between Cornell University and the University of Michigan, the project currently has page images and searchable text of 5,000 volumes between 1850 and 1877. The project's ambition is to involve research institutions and national consortia to develop common protocols and consensus for the selection, conversion, storage, retrieval, and use of digitized materials.

The Museum Educational Site Licensing (MESL) Project <http://www.gii.getty.edu/mesl/> was launched in 1995 by the Getty Art History Information Program (now the Getty Information Institute) and MUSE Educational Media. This two-year collaboration brought together seven museums and seven universities to develop and test a model of licensing visual material across closed campus networks.

[2] The National Information Infrastructure refers to the body of communications and information resources and services that is becoming increasingly integrated; a variant is the Global Information Infrastruc-

ture, since communications services transcend national borders. NII components include the telephone system, cable and broadcast television, satellite and other wireless communications systems, and data bases, libraries, publications, and other information repositories. Some define the NII to include only those resources accessible electronically, others are more inclusive. The Internet is considered by many as a microcosm of the NII because it is composed of heterogeneous elements and integrates many information and communications resources.

[3] The digital library concept involves the integration of collections, people, and services across the cycle of information creation, dissemination, organization, finding, use, and preservation. See, for example, Paul Duguid and Daniel E. Atkins, *Digital Libraries: Report of the Santa Fe Planning Workshop on Distributed Knowledge Work Environments (March 9-11)* (Ann Arbor, MI: University of Michigan School of Information, 1997).

[4] With the Association for Computing Machinery, a major computer science professional organization, SIGGRAPH is the Special Interest Group (SIG) on Computer Graphics. It blends science and art to advance computer graphics and related human-computer interactions.

[5] See Note Three.

[6] Bruce R. Schatz, "Information Retrieval in Digital Libraries: Bringing Search to the Net," *Science* January 1997: 327-334.

[7] See Clifford A. Lynch, "The Z39.50 Information Retrieval Standard." *D-Lib Magazine* April 1997 <http://www.dlib.org/>. It explains that "Z39.50—properly 'Information Retrieval (Z39.50); Application Service Definition and Protocol Specification, ANSI/ISO Z39.50-1995'—is a protocol which specifies data structures and interchange rules that allow a client machine (called an 'origin' in the standard) to search databases on a server machine (called a 'target' in the standard) and retrieve records that are identified as a result of such a search."

[8] See, for example, Computer Science and Telecommunications Board, *Realizing the Information Future: The Internet and Beyond* (Washington, DC: National Academy Press, 1994) for commentary and references on trends in educational institutions and libraries.

[9] Note that the humanities differ from computer science in rewarding professional organization and evaluation of the work of others. The implications of Shaw's suggestion could range from the practical to the scholarly.

[10] See, for example: Association of American Universities Research Libraries, *Reports of the AAU Task Forces* (Washington, DC: Association of Research Libraries, 1994); Computer Science and Telecommunications Board, *Realizing the Information Future: The Internet and Beyond* (Washington, DC: National Academy Press, 1994); Anthony M. Cummings, Marcia L. Wit, William G. Bowen, Laura O. Lazarus, and Richard Ekman, *University Libraries and Scholarly Communication* (Washington, DC: Association of Research Libraries, 1992); and Lawrence Dowler, ed., *Gateways to Knowledge: The Role of Academic Libraries in Teaching, Learning, and Research* (Cambridge, MA: MIT Press, 1997).

[11] See Duguid and Atkins, *Digital Libraries*.

Appendix One

Computing and Humanities Roundtable Organizers

The Computer Science and Telecommunication Board (CSTB)

The Computer Science and Telecommunications Board's charter is broad: to ensure that the United States makes every effort to develop and use the major national resources represented in computer science, computer technology, and telecommunications. CSTB considers technical and policy issues pertaining to computer science, telecommunications, and associated technologies. The functions of the Board include:

- monitoring and promoting the health of the computer science, computing technology, and telecommunications fields, including attention as appropriate to the issues of human resources and information infrastructure;
- initiating studies involving computer science, computing technology, and telecommunications as critical resources and sources of national economic strength;
- responding to requests from the government, non-profit organizations, and private industry for expert input on computer science, computing technology, and telecommunications issues; and to requests from the government for expert input on computer and telecommunications systems planning, utilization, and modernization.

CSTB actively disseminates the results of its completed projects to those in a position to help implement their recommendations or otherwise make use of their insights. It provides a forum for the exchange of information on computer science, computing technology, and telecommunications.

Further information may be found at the CSTB Website at <www2.nas.edu/cstbweb>.

The Coalition for Networked Information (CNI)

The Coalition for Networked Information is an organization for institutions concerned with realizing the promise of high performance networks and computers for the advancement of scholarship and the enrichment of intellectual productivity. The Coalition was formed in

1990 by the Association of Research Libraries (ARL), Educom, and CAUSE. CNI pursues its mission through the aid of its membership; a 200-member task force made up of higher education institutions; publishers; network service providers; computer hardware, software, and systems companies; library networks and organizations; and public and state libraries.

Further information may be found at the CNI Website at <www.cni.org> .

The National Initiative for a Networked Cultural Heritage (NINCH)

The National Initiative For a Networked Cultural Heritage is a diverse coalition of cultural organizations dedicated to ensuring the greatest participation of all parts of the cultural community in the digital environment. Its vision of networked cultural heritage is of an integrated, distributed body of cultural material, seamlessly interoperable across many media, of the highest possible quality and fidelity, and easily usable and searchable by creators, scholars, the general public and by teachers and learners of all ages. NINCH's mission is to advocate for the inclusion of the cultural sector in all policy deliberations on the future of the information infrastructure and to educate policymakers, coalition members, and the general public about the critical importance of translating the vision of a connected, distributed, and accessible collection of cultural knowledge into a working reality.

Further information may be found at the NINCH Website at <www.ninch.org/> .

The Two Ravens Institute

The Two Ravens Institute provides a forum of convergence for scholars, teachers, students, writers, and others representing a number of academic and cultural perspectives to explore the transformational changes of networked technology on the contemporary social fabric. The perspectives the Institute adopts represent a merging of humanistic, social science, and scientific methodologies in order to better understand, and therefore predict, the effects of the digital revolution. Fundamental assumptions of the Institute include the belief that the growing digital networks will best serve teaching and research only if those networks are ultimately susceptible to human choice, experimentation, and creative application. The Two Ravens Institute undertakes:

- to refocus currently polarized and simplistic discussions about technology as it relates to culture, education, and the individual in terms that recognize the complexity and ambiguity of these issues;
- to invigorate these discussions with perspectives normally associated with the humanities perspectives largely absent from current discourse; and
- to foster an intellectual environment wherein individuals can assume greater responsibility for a strong and continuing democracy.

Appendix Two

What are the Humanities?

In our discussion we speak broadly of the humanities. We mean to include those areas normally understood as the arts, the humanities, and the social sciences. While the arts connote the arena of active creation or performance of works of human expression, the humanities comprise those areas of intellectual investigation that focus on the preservation, transmission, and interpretation of the human record. Their cousins, the social sciences, focus on the analysis of social forces and the modeling of laws underlying the formation of society and of the individual's relationship to society. The social sciences—economics, political science, sociology, psychology—rely more on observation of current social forces and analysis of data than on historical document and cultural object.

One might broadly categorize the humanities as follows:

- the history and analysis of society and culture—its currents, events, and forms as measured and observed through objects and documents. Specialization is broadly by the type of activity studied or the means of analysis (history of law, labor, politics, science, economics), period (archaeology, various period studies, e.g. Medieval Studies, 18th-century studies), or geographical area (by country or other area, e.g. English History; Oriental Studies). Methodologies vary: narrative and thematic analysis, comparative studies, detailed data analysis.
- history and analysis of forms of thought and creative expression: philosophy; spoken and written language (literature and linguistics, drama); visual forms (visual art and architecture, cinema, theater); music; and dance. Approaches vary: biographical narrative, detailed internal and comparative analysis of works, broader studies of aesthetics and their historical development.

Appendix Three

Roundtable Agenda and Participants, March 28, 1997

Meeting Co-Chairs:

Stanley N. Katz, Woodrow Wilson School, Princeton University; formerly President, American Council of Learned Societies
William A. Wulf, President, National Academy of Engineering

I. Toward a common language: methodologies and assumptions in humanities and computer sciences work

Michael Joyce, Vassar College
Bruce Schatz, University of Illinois

Discussants:
Edward Ayers, University of Virginia
Michael Brodie, GTE Laboratories, Inc.
Michael Neuman, Georgetown University and the Association of Computers and Humanities
Jerome Saltzer, Massachusetts Institute of Technology

II. Software and standards development: conditions and mechanisms for productive collaboration among humanists and industry and academic technologists

Joseph Busch, Getty Information Institute
Mary Shaw, Carnegie Mellon University

Discussants:
Camilla Cai, Kenyon College
Thomas DeFanti, University of Illinois, Chicago
Edward Fox, Virginia Polytechnic Institute and State University
Michael Lesk, Bellcore

III. Economic and institutional realities, assumptions, and infrastructures

Willard McCarty, King's College, London
Richard Liebhaber, Veronis, Suhler & Associates

Discussants:
Douglas Bennett, President, Earlham College; formerly
Vice President, American Council of Learned Societies
Stephen Franklin, University of California, Irvine
Sandria Freitag, American Historical Association
Michael McGrath, Colorado School of Mines
Joan Shigekawa, Rockefeller Foundation
Arthur Tsuchiya, National Endowment for the Arts

Appendix Four

Methodologies for Computing and the Humanities

Michael Joyce, Vassar College

John Unsworth of the Institute for Advanced Technologies in the Humanities has posted on the Web a wonderful talk about methodologies in the humanities, "Not Your Average Fool: The Humanist on the Internet." In many ways—and unlike John—I am your average fool, and so may seem the least likely person to perform this morning's task. Among humanists, I am considered to be more an artist, a writer, than a humanist. (This distinction may strike computer scientists as curious.) Among writers, as a hypertext novelist, I am considered a technologist. Among technologists, as the developer of a microcomputer hypertext system and structure editor, I am considered hopeless.

My state is interstitial, betwixt rather than between. And in that state I am neither truly hopeless (in fact, I am full of hope); nor, as you might expect, am I truly the least likely person to perform this task.

In the special anniversary issue of *Communications of the ACM* [February 1997], William Wulf draws upon Asian art to urge us to undertake an interstitial dialogue of exactly the sort we contemplate today. "As in oriental art and gardens," Wulf says, "a full appreciation requires a look at these spaces to discover the relationships." Wulf offers as examples the work of IATH scholars like Jerry McGann and Ed Ayers. It is both inevitable and important that we hear today about the methodologies of hypertextual historiographers like Ed, textual theorists like Jerry and my Vassar colleague Nancy Ide, and Willard McCarty. We will need to consider specific instances of scholarly work with data sets and data bases, text encoding and texts, image annotation and recognition, electronic publication and communication, global information systems, synchronous and asynchronous virtual communities, search engines and software agents, immersive computer environments and simulations, and the several other areas of humanities computing which are represented here today and whose methodologies I would not presume to characterize in detail.

Instead, I wish to touch on broad methodological perspectives and interstitial relationships. The American poet Charles Olson defines methodology as "how to use yourself and on what." I want to talk about the uses to which we humanists put ourselves in computing environments, and what spaces we inhabit and increasingly create. Humanists

create potentiated spaces in computer environments. By "potentiated spaces" I mean the kinds of morphogenetic computational objects that H. Van Dyke Parunak sees as a variety of artificial intelligence—where human beings supply the inferences, connections, and natural language processing which computers cannot provide. I want to characterize a few of these methodological spaces or perspectives for you.

The first such space is what we might call aporetic space, the space of doubt, skepticism, and consideration which eventually yields possibility, valorization, persistence, and meaning. In his ACM essay, Wulf says that looking at computational objects rather than spaces "can be distracting and misleading, focusing on the 'now' rather than the 'can be.'" While humanists might agree with him about the distraction, I hope that in the spirit of today's dialogue I can suggest a potentiated space of difference. For a humanist might amend Wulf's comment to say that we must focus on both the "can be" and the "continues."

We are surely not the first, but we are surely the most self-conscious age in terms of seeing ourselves living before a constantly impending future. In our technologies, our cultures, our entertainments and, increasingly, the way we constitute our communities and families, we live in an anticipatory state of constant "nextness." There is, of course, a branch of philosophy, eschatology—the study of ultimate destinies—which concerns those who see themselves as inhabiting the time before the future. Eschatological ages have both their virtues and their particular vices. The chief virtue is hope, that constant anticipation of the next which keeps us poised, unsettled and open to change. The chief vice is, paradoxically, inaction—a self-satisfied belief that there is no need to act in the face of a decisive and imminent history.

Most of us humanists see our task as encouraging virtues and discouraging vices insofar as we can recognize the difference between them. As a humanist and writer deeply involved with technology, I have for some time been concerned with the passivity that electronic media may encourage. The Web, for instance, too often seems a lonely pursuit, something which douses the crispness of difference and community in a salsa of shifting screens. Our culture has slipped the Web over our heads like a lonely guy slips on a T-shirt from the Hard Rock Cafe. The Web privileges the culture of brand names and corporate logos over the weave of our own multi-threaded culture and history. Like the lonely guy in the Hard Rock T-shirt, we haven't anyplace to go. The old places seem either deserted or dull and the infoscape seems paradoxically crowded and lonely, already mapped by somebody else, and often without a clear place for us. We are caught between a Hard Rock and Melrose Place.

The humanist's aporetic space, the space of considered doubt and possibility, of what continues and what persists, is meant to bookmark our passive acceptance of nextness with a space for reflection and action. And so when my Vassar history colleague and Ed Ayers protégé, Rebecca Edwards, builds a Website which considers women's political cartoons of the golden age, she highlights the continuing differences among us which are the fabric of what we have in common. More importantly, as Wulf notes about McGann's and Ayers' projects, Edwards creates "an environment in which students participate in the process of scholarship rather than its product." To the person in the Hard Rock T-shirt, such a process offers not only somewhere to go but somewhere to be, an arena for action.

In fact, each kind of space which humanists create or inhabit within computer environments is meant as an arena for action. A second kind of potentiated space is what I would call metaleptic space. Metalepsis is an obscure word even for many humanists, a rhetorical term with varying usage and not much currency. Yet the term describes a common human experience as well as an objective of certain domains of computational science. For metalepsis is a haunting, emergent meaning, "autopoesis," the way one form seems to anticipate, echo, transform, or transform into another.

Computer scientists will recognize how what at first glance might seem the dryly computational pursuits of the best known and longest established areas of humanistic technology—text encoding initiatives and the development of textual corpora—are at their best active efforts to create metaleptic spaces. To use Wulf's terms again, in such spaces the "now" of the text oscillates with both the "can be" and the "continues." Thus, through textual analysis my Vassar colleague Don Foster finds both the active making and emergent form of either a Shakespearean elegy or a contemporary political satire in the structured evidences of haunted language which computer scientists would recognize as knowledge representations. Other humanists are concerned with how the body is haunted by the machine and vice versa. Following the example of feminist scholars, "Queer Theorists," and others, new humanists like N. Katherine Hayles, Sherry Turkle, and Jay David Bolter seek to locate and affirm the importance of embodiment, of our physical humanity, in the face of increasingly seductive and polymorphous computer environments. They do so not as Luddites but as Harawayian co-evolutionists, conspirators, and cyborgs, participants in the continuous and transcendent *techne* of human civilization which haunts current and future technologies.

The last space I want to characterize is one which we are inhabiting even at this moment, since we are as much within a technology as we sit together in this room as when we interact in a computer environment. Here we are visible within an invisible technology; elsewhere we are invisible within a visible technology. Morphogenetic space is the space of constant potentiation and oscillation, where form dissipates in Prigogine's sense into form. All of us in this room, computer scientists and humanists alike, have experienced the human longing to touch upon, if for only a moment, the swirling and transformative stuff of human possibility and nature itself. All of us in this room, whether informed by scientific or humanistic scepticism, know that such moments exist only in the retrospection and reflection from which we fashion our representations. Humanists increasingly use computers as such spaces— whether creating VRML models for metaphors and metalepsis, or creating or analyzing multiple narratives of a changing world in hypertexts, or experiencing first hand through MOOs and more sophisticated virtual spaces the shifting perspectivalization which characterizes the transcendent and historical present tense which marks us as mortal beings.

Everywhere I speak or write I argue the same thing: that the value of our presence as human persons in real place continues as a value not despite but because of the ubiquity of virtual spaces. In this we humanists and scientists share reciprocal methodologies. Our embodiment graces actual and virtual space alike with the occasion for value and action which brings us together today.

Appendix Five

Illustrative World Wide Web Pointers for Computing and the Humanities

Organizational

American Arts and Letters Network: www.aaln.org
American Council of Learned Societies: www.acls.org
American Historical Association: web.gmu./chnm/aha.
American Studies Crossroads Project:
www.georgetown.edu/crossroads/crossroads.html
Argos Limited Area Search of the Ancient and Medieval Internet:
argos.evansville.edu/
Arts and Humanities Data Service:
www.kcl.ac.uk/projects/ahds/public/guides.htm
Association for Computers and the Humanities: www.ach.org
Carnegie Mellon University: www.cs.cmu.edu
Coalition for Networked Information: www.cni.org
Comp Lit 50A: Homer to Renaissance: eee.uci.edu/96f/21260
Computer Science and Telecommunications Board:
www2.nas.edu/cstbweb
CTI Textual Studies: info.ox.ac.uk/ctitext
Decameron Web:
www.brown.edu/Departments/Italian_Studies/dweb/dweb.shtml
Digital Library Initiative: www.cise.nsf.gov.iris/dlhome.html
Duke Papyrus Archive: scriptorium.lib.duke.edu/papyrus/
Earlham College: www.earlham.edu
Electronic Text Center at the University of Virginia:
etext.lib.virginia.edu/
Electronic Visualization Laboratory, University of Illinois, Chicago:
www.evl.uic.edu and www.eecs.uic.edu
English Language and Literature URLs, Yale University:
www.library.yale.edu/Internet/litenglish.html
Getty Information Institute: www.getty.edu
GTE: www.gte.com
Humanities Text Initiative, University of Michigan: www.hti.umich.edu/
Institute for Advanced Technology in the Humanities at the
University of Virginia, Charlottesville: jefferson.village.virginia.edu
Internet Movie DataBase: us.imdb.com/
Internet Shakespeare Editions Home Page: web.uvic.ca/shakespeare/

King's College London: www.kcl.ac.uk/kis/schools/hums/ruhc/wlm
The Labyrinth: Resources for Medieval Studies, Georgetown University: www.georgetown.edu/labyrinth/
National Academy of Engineering: www.nae.edu
National Endowment for the Arts: arts.endow.gov
National Endowment for the Humanities: www.neh.fed.us/
National Initiative for a Networked Cultural Heritage: www.ninch.org/
Orlando Project: www.ualberta.ca/ORLANDO
Perseus Project: www.perseus.tufts.edu
Shakespeare and the Globe: www.reading.ac.uk:80/globe/
Text Encoding Initiative: www.uic.edu/orgs.tei
Thesaurus Linguae Graecae: www.uci.edu:80/~tlg/
University of California, Irvine: www.oac.uci.edu/indiv/franklin
University of Illinois at Urbana-Champaign: www.cso.uiuc.edu
Vassar College: iberia.vassar.edu
Virginia Tech University: fox.cs.vt.edu
Voice of the Shuttle: humanitas.ucsb.edu/

Individual

Stephen Franklin, University of California: www.oac.uci.edu/indiv/franklin
Michael Joyce, Vassar College: iberia.vassar.edu/~mijoyce
Willard McCarty, King's College London www.kcl.ac.uk/kis/schools/hums/ruhc/wlm
Jerome Saltzer: web.mit.edu/saltzer/
Mary Shaw, Carnegie Mellon University: www.cs.cmu.edu/~shaw/

ACLS Occasional Papers

1. *A Life of Learning* (1987 Charles Homer Haskins Lecture) by Carl E. Schorske
2. *Perplexing Dreams: Is There a Core Tradition in the Humanities?* by Roger Shattuck
3. *R.M. Lumiansky: Scholar, Teacher, Spokesman for the Humanities*
4. *A Life of Learning* (1988 Charles Homer Haskins Lecture) by John Hope Franklin
5. *Learned Societies and the Evolution of the Disciplines* by Saul B. Cohen, David Bromwich, and George W. Stocking, Jr.
6. *The Humanities in the University: Strategies for the 1990's* by W.R. Connor et al.
7. *Speaking for the Humanities* by George Levine et al.
8. *The Agenda for the Humanities and Higher Education for the 21st Century* by Stephen Graubard
9. *A Life of Learning* (1989 Charles Homer Haskins Lecture) by Judith N. Shklar
10. *Viewpoints: Excerpts from the ACLS Conference on The Humanities in the 1990's* by Peter Conn et al.
11. *National Task Force on Scholarship and the Public Humanities*
12. *A Life of Learning* (1990 Charles Homer Haskins Lecture) by Paul Oskar Kristeller
13. *The ACLS Comparative Constitutionalism Project: Final Report*
14. *Scholars and Research Libraries in the 21st Century*
15. *Culture's New Frontier: Staking a Common Ground* by Naomi F. Collins
16. *The Improvement of Teaching* by Derek Bok; responses by Sylvia Grider, Francis Oakley, and George Rupp
17. *A Life of Learning* (1991 Charles Homer Haskins Lecture) by Milton Babbitt
18. *Fellowships in the Humanities, 1983-1991* by Douglas Greenberg
19. *A Life of Learning* (1992 Charles Homer Haskins Lecture) by D.W. Meinig
20. *The Humanities in the Schools*
21. *A Life of Learning* (1993 Charles Homer Haskins Lecture) by Annemarie Schimmel
22. *The Limits of Expression in American Intellectual Life* by Kathryn Abrams et al.
23. *Teaching the Humanities: Essays from the ACLS Elementary and Secondary Schools Teacher Curriculum Development Project*

24. *Perspectives on the Humanities and School-Based Curriculum Development* by Sandra Blackman et al.
25. *A Life of Learning* (1994 Charles Homer Haskins Lecture) by Robert K. Merton
26. *Changes in the Context for Creating Knowledge* by George Keller, Dennis O'Brien, and Susanne Hoeber Rudolph
27. *Rethinking Literary History—Comparatively* by Mario J. Valdés and Linda Hutcheon
28. *The Internationalization of Scholarship and Scholarly Societies*
29. *Poetry In and Out of the Classroom: Essays from the ACLS Elementary and Secondary Schools Teacher Curriculum Development Project*
30. *A Life of Learning* (1995 Charles Homer Haskins Lecture) by Phyllis Pray Bober
31. *Beyond the Academy: A Scholar's Obligations* by George R. Garrison et al.
32. *Scholarship and Teaching: A Matter of Mutual Support* by Francis Oakley
33. *The Professional Evaluation of Teaching* by James England, Pat Hutchings, and Wilbert J. McKeachie
34. *A Life of Learning* (1996 Charles Homer Haskins Lecture) by Robert William Fogel
35. *Collaborative Historiography: A Comparative Literary History of Latin America* by Linda Hutcheon, Djelal Kadir, and Mario J. Valdés
36. *New Connections for Scholars: The Changing Missions of a Learned Society in an Era of Digital Networks* by Douglas C. Bennett
37. *Information Technology in Humanities Scholarship: Achievements, Prospects, and Challenges—The United States Focus* by Pamela Pavliscak, Seamus Ross, and Charles Henry
38. *Report of the President, 1986-1997* by Stanley N. Katz
39. *A Life of Learning* (1997 Charles Homer Haskins Lecture) by Natalie Zemon Davis
40. *The Transformation of Humanistic Studies in the Twenty-first Century: Opportunities and Perils* by Thomas Bender, Stanley Chodorow, and Pauline Yu
41. *Computing in the Humanities: Summary of a Roundtable Meeting*