

## **REPRESENTED CHRONOLOGICALLY:**

### **Charles Faulhaber**

**The Bancroft Library, University of California, Berkeley**

I would like to propose the **Encoded Archival Description**, particularly its instantiation in the Online Archive of California. This has been one of the more significant developments in making the resources of the nation's archives available to students and scholars.

(9/8/98)

---

### **Roy Rosenzweig**

**Prof. of History & Director, [Center for History & New Media](#)**

At the risk of self promotion, let me mention two current efforts in which the Center for History and New Media at GMU and my colleagues at the American Social History Project at CUNY are engaged, which I believe are making a difference.

One is our just-launched web site, "[History Matters: The US Survey on the Web](#)," which I believe is providing and will provide an invaluable set of resources for high school and college history teachers around the country. Our goal is to offer a gateway on to the Internet for teachers of American history and to give them a basic set of teaching resources and ideas. Without the Internet and the Web, there is simply no other way to get these primary sources materials, these teaching ideas, these interviews with distinguished teachers, these forums with leading scholars into the hands of a diverse audience of community college teachers in Tulsa, high school teachers in Eastern Washington, and university instructors in Kyoto. The ability of the dedicated, but often isolated, instructor to teach exciting material in new and exciting ways is instantly and dramatically enhanced through networking.

Related to this effort is our New Media Classroom seminars, which have energized first two groups of people who came to NYC for summer seminars in 1996 and 1997 and has now energized a much wider circle of people in regional workshops around the country.

Randy Bass's efforts through the Crossroads (and the New Media Classroom in which he is a partner) have had a similar catalyzing effect on the teaching of American studies.

I can say more about these and other efforts but I assume you want us to be brief. The work of the Center for History and New Media is summarized at <http://chnm.gmu.edu/desc.html>

To the extent that I have written anything that talks about the networking future (and I'm not sure that I am inclined toward "visionary" statements), it is contained in the essays that are on line at <http://chnm.gmu.edu/chnm/essays.html>. The essay on **US History on the Web** that I did with Mike O'Malley is probably my best effort at a statement of the potential of this medium.

(9/8/98)

---

## **William Wulf**

### **President, National Academy of Engineering**

I've attached an article, "University Alert: The Information Railroad Is Coming" -- an edited version of which appeared in *Issues in Science and Technology* about three years ago. It tries to look at the impact of IT on universities, but in the process, I hope, build an image of what the future might hold.

*Extract from linked article:*

...The easy examples are those that simply automate what was done manually: the reduction of data, the control of instruments, etc. The profound applications, however, are those that lead to whole new areas of research and new methods of investigation-- and thus to science that was not, and could not be done before: the final proof of the four color conjecture, analysis of molecules that have not been synthesized, measuring the properties of a single neuron by growing it on a silicon chip, watching a model of galaxies collide, and letting a scientist "feel" the forces as a drug docks in a protein. These applications have transformed the nature of scientific investigation; they led to questions that would not even have been asked before.

I don't think science, however, will be where we see the most dramatic impact. I say that despite a recent report from the National Research Council that I helped coauthor—a report that paints an expansive image of the transformation of scientific research. Instead, I believe that a more dramatic transformation is about to shake the foundations of scholarship in the liberal arts. Humanists more than scientists will lead the way to innovative applications of the technology in the university.

The comfortable stereotype of humanists as technophobic just doesn't apply anymore. The availability of both text and images in electronic form coupled with the processing power of modern computers allow the humanist to explore hypotheses and visualize relations that were previously lost in the mass of information sources. The presentation of humanists' scholarly results in electronic form is moving even faster. Precisely because of the complexities of the relationships they need to present, the subtle webs of relation and inference they need to express, electronic "hypertext" books and journals are emerging. Indeed, they are emerging faster, with more vigor, and with more effect on their disciplines than their counterparts in the sciences.

We all expect scientists and engineers to use computers in their research, but the notion that information technology could be central to humanistic scholarship is a bit more startling — at least to me. It was in large measure talking about the application of computers to historiography and the theory of text that opened my eyes to the larger issues that I am trying to raise here.

(9/8/98)

---

## **Robert Kolker**

### **Professor of English, University of Maryland**

The University of Virginia has perhaps the best record of humanities tech with the Institute for Advanced Technology in the Humanities and the E-Text Center. My place, University of Maryland, does not have great resources, but we have a galloping program:

Martha Nell Smith's Emily Dickinson project;

Neil Fraistat's Romantic Circles web site; and

my forthcoming CD-ROM, "Film, Form, and Culture."

Also, our Arts & Humanities self-support unit, The Center for Renaissance and Baroque Studies has been sponsoring K-12 outreach and hi-tech conferences for about 3 years: [www.inform.umd.edu/CRBS](http://www.inform.umd.edu/CRBS)

Bob Kolker

(9/8/98)

---

## **Peter Suber**

### **Earlham College**

Hippias, a search engine specializing in philosophy. While other search engines will pick up the same sites, they mix them with hundreds, sometimes thousands, of false positives. For example, just five minutes ago, an Alta Vista search for "Plato" returned 206,110 hits, including a town in Illinois, a software company, and many other non-philosophical sites. Hippias returned 929, all philosophical.

Hippias accomplishes this filtering through peer review. (I recruit and manage the editors, and serve as an editor myself.) Adding human, peer-review judgments to search engine technology is an exciting innovation that has made the web much more useful for philosophers. Most disciplines don't have a comparable tool yet.

Hippias has been online for 13 months and is very successful. Over 600 philosophy pages link to it. It processes between 600 and 1,000 search queries a day.

The successor to Hippias is called Noesis, and is already available in a beta version. Noesis is also a search engine for philosophy based on peer review, but it gives the user much more flexibility than Hippias does in searching for particular kinds of philosophical content. For example, users can limit their search for the word "Plato" to its occurrence in primary texts, journal articles, or book reviews. When it's out of beta, Noesis will offer far more options, which I could describe at length at any time.

(9/8/98)

---

**Janet H. Murray**

**Senior Research Scientist, Massachusetts Institute of Technology**

**Why we should fund centers of advanced humanities computing as part of a national initiative in digital technologies.**

Pioneering application projects in the humanities have played a crucial part so far in transmuting a medium of information transmission into a medium of communication. That is the continuing job at hand. We need the fastest wires and the smartest algorithms -- but we also need the appropriate organizational strategies to turn more and more data into greater and greater knowledge. To do this we have to develop the technology around specific knowledge-based applications. Humanists are particularly good at coming up with such applications because they have been so pointedly aware of the fact that they've outgrown the boundaries of the book and the library shelf. Pioneering humanities projects such as the **Brown Hypermedia Project, the Stanford Univ. Shakespeare Project, the Perseus Project now at Tufts, and MIT's Athena Language Learning Project, the Getty's controlled vocabulary work, and Berkeley's Digital Library applications** have significantly pushed forward our understanding of hypermedia, digital video, metadata, cross-collection searches, etc. by providing prototypes and test beds that demanded more than the cut-and-paste multi-media slide shows that were the norm for the corporate world. They helped to guide systems-level engineers toward developing the most important functionalities.

Furthermore, I would argue that the arts are absolutely crucial to the development of the medium because the nature of the artist is to pound the clay and find out what can be done with it. Playful applications like interactive fiction in the MUDs have become the basis for some of the most valued conferencing technologies. Narrative and visual artists can help us to develop "the communicative primitives" of the medium. We cannot count on existing entertainment interests to develop these any more than we could expect book publishers to invent the movies. If we want to move simulation games, for instance, out of the realm of the shoot-em-up and into the realm of educationally and humanly compelling virtual worlds, then we should fund ambitious digital systems to serve the needs of interactive creative artists.

Arts and humanities projects have the capacity to push the technology forward so that it becomes more humanly pliable as it becomes more powerful. They are also powerful antidotes to the growing terror of the machine. The President's Report has a bold vision of xerox parc-like "expedition" centers to chart the digital frontier. Let's

devote half of these to humanities and the arts -- perhaps focusing some of them on specific socially important themes like literacy education, multicultural history, democratic traditions, and violence prevention. The practicality and complexity of such domains, the passion for solution that they will engender, will focus the imaginative energy of the developers. I have always found that the best people are attracted to the hardest problems, especially when those problems are posed as part of a tangible human goal. Humanities projects can help focus the next stage of technological development, advancing our scientific mastery while deepening our understanding of who we are.

(9/10/98)

---

## **Gary Marchionini**

### **University of North Carolina at Chapel Hill**

Thank you for the opportunity to provide input. I offer a few observations, a few specific projects as exemplars, and some general pointers. I certainly am in no position to posit a vision statement for humanities computing but do look forward to reading what others envision. [Gary Marchionini, University of North Carolina at Chapel Hill]

#### **Observation 1.**

The humanities tradition has championed individuality, uniqueness, and personal expression. Many creative and useful resources are mounted by individuals with minimal technical support. Projects mounted by individuals are fragile enterprises and sustainable projects will require collaboration and long-term funding. The web is full of dead or incomplete individual efforts. Individuals re-discover organizational or presentational principles, struggle alone with larger challenges like intellectual property and privacy, and do not recognize interoperability issues early enough to influence basic design. There is a need for collaborative projects that disseminate best practices, sustain individual efforts, and coordinate such efforts. A related need is for scholars to value and reward collaborative efforts (promotion, tenure, etc.). This is a culture change that serious funding may accelerate.

#### **Observation 2.**

There are several good models for humanities collections but we have little evidence about their respective effectiveness for satisfying different needs. These models

include: deep collections with tools and scholarship attached for specific themes and content (e.g., Valley of the Shadow); general aggregations of primary materials with no integrative tools or commentaries (e.g., Project Gutenberg); digital archives of primary materials created exhaustively or opportunistically so that others can use them in new products (e.g., Library of Congress). It seems prudent to support research that investigates the efficacy of these models in different contexts.

### **Observation 3.**

There is a pressing need for more high quality digital content. Given the volume of humanities resources, only a microscopically small portion has been digitized. Moreover, the most popular items may be digitized multiple times. Funding for acquisition, digitization, markup, and added value commentaries and guides is needed. Note that the DLI2 initiative specifically excluded digitization efforts but such efforts are especially important to include in a humanities computing program.

### **Observation 4.**

To date, humanists have mainly adapted tools and design rubrics from science and engineering. Intermedia benefited from collaboration with computer science, IATH maintains strong ties to computer science, whereas some projects (e.g., Storyspace) emerged from special humanistic needs. Given that most of my emphasis here has been on content, it seems to me to also be a good idea to fund projects that aim to develop specialized tools, algorithms, and rubrics that flow from the special needs of different humanistic endeavors. It seems particularly opportune to encourage interactive and multimedia projects.

I offer four specific projects that I believe meet your request for extant and high-impact projects:

1. I see that Greg Crane was on your distribution list and he can surely provide more information but I must offer my impressions of why the [Perseus Project](#) has already demonstrated impact and can serve as one model for additional funding for computing in the humanities.

For the past 10 years I have been the external evaluator for the Perseus Project, a hypermedia corpus/digital library devoted to the ancient Greek world, and have watched it grow from a resource used in some specialized university classics courses to a publishing in the humanities center that is referenced by commercial online encyclopedias, used in a variety of humanities courses internationally (including distance education and other non-traditional venues), and accessed tens of thousands of times daily. I have presented evidence over

the past several years (through FIPSE funding, see <http://www.perseus.tufts.edu/FIPSE/>) that Perseus has contributed to systemic change in the field of classics. Funding for Perseus has come from a variety of public and private sources.

2. The Library of Congress National Digital Library Program has become an invaluable resource for scholarship and education at all levels. Millions of requests are served each month, many of which come from educational institutions that have come to rely on the availability of LC resources for online curricula and student assignments. Funding has been primarily through corporate sponsorship.
3. The Museum Site Licensing Project has provided a model for sharing museum treasures with the academic community. A project studying the economic factors related to MESL has recently been completed (Besser et.al.). These related projects provide realistic indicators for the complexities of sharing cultural artifacts, cataloging and delivering these objects, and determining the associated costs. Although these projects are now complete, the lessons learned bear dissemination, replication, and extension to other organizational settings and media. Similar projects should be encouraged. Funding has been mainly through private sources (Getty, Mellon). [See Bearman/Trant above on AMICO]
4. Documenting the American South is a full-text database of digitized and encoded resources on Southern history, literature and culture from the colonial period through the first decades of the twentieth century. Currently, it encompasses three projects: slave narratives, first-person narratives, and Southern literature. A fourth, based on Confederate imprints, is in development. The database, now in its third year, is administered by the Academic Affairs Library, University of North Carolina at Chapel Hill. The printed works and manuscripts come primarily from its Southern collections. An Editorial Board guides its development. DAS has received rave reviews from all segments of the population, and is noted as an exemplary site by several rating services. Funding has been through public and private funds (NEH, LC, Ameritech).

In addition to large-scale projects like the four above, various groups and centers have sprung up to address different aspects of computing in the humanities, e.g., [Institute for Advanced Technology in the Humanities](#) at the University of Virginia; the Text Encoding Initiative; and the Center for Electronic Texts in the Humanities. Surely, efforts like these should be expanded and supported by federal funding.

Finally, I offer two example gateway service that provides access to projects in the humanities. Such **gateway services** should also figure in any funding model NEH and other agencies develop.

The **National Endowment for the Humanities** [Edsitement](#) web page provides a vetted list of excellent humanities web sites. This gateway and others like it provide excellent starting points for teachers and scholars alike. One significant value added beyond the pointer service is quality control. Models for evaluation and summarizing (including surrogate creation and display) projects are needed.

A different, unvetted model is illustrated by the UNC Sunsite. It serves approximately 32,000,000 requests per month and is linked to by more than 160,000 external web pages (personal communication, Paul Jones). Although many of these accesses are for science and technology (many open source distributions such as Linux), large numbers of people depend on this and similar services for ongoing, specific humanities resources. We do not know how the vetting process affects usage and what Heisenberg-like effects vetting add. Supporting and studying usage in unvetted gateways and sites seems worthwhile. Here are a few examples that illustrate the broad mix of content and quality available from the UNC Sunsite.

- [Internet Poetry Archive](#) - some of the world's most distinguished poets reading their work, photos of the poets, critical papers, bibliography and texts - UNC Press, MetaLab, NC Arts Council, Sun, Cisco, Ventana participating in support. Poets so far: Czeslaw Milosz, Seamus Heaney, Margaret Walker, Philip Levine, Yusef Komunyakaa, and Robert Pinsky.
- Documenting the American South (see above)- giving access to important and often rare documents about the American South including first-person narratives, narratives of enslaved persons, literature, and more - UNC Libraries, MetaLab, NEH, LoC, Ameritech participating in support.
- Presentations from the Southern Folklife Collection - coherent presentations thematically developed from the SFC combined with new collecting in several areas including [Hollerin](#), [Moonshine](#), Medicine Shows, [Ghost Stories](#), musician Doc Watson and more.
- Folkden - musician Roger McGuinn's collection and performance of American folk music including annotations, tablature, lyrics and performance of over 30

songs with a new song added each month. Of special interest, as McGuinn was on the forefront of the 60s folk revival and the creator of folkrock as a genre.

- [Project Gutenberg](#) - a very long term text digitization project. In 1997, Project Gutenberg published its 1000th electronic text, Dante Alighieri's "Divine Comedy." Project founded and directed by Michael Hart (also assisting is UNC's Greg Newby).
- [Walker Percy Project](#) - The Vision of the Walker Percy Educational Project, Inc., is to found and develop an Internet Literary Center where students, scholars, and general readers from around the world may participate and collaborate in a range of ongoing discussion, research, and educational activities on the works of Walker Percy.
- [Expo](#) - Tim Berners-Lee's favorite WWW site (mentioned in several interviews through the years). Expo was build on collections from the Library of Congress including: "Rome Reborn," " Soviet Archives," "1492," "Dead Sea Scrolls," and "Spalato: Diocletian's palace at Split".
- [Doc Watson Exhibit](#) from the Southern Folklife Collection. Deep and detailed information about the legendary performer. Including photos, songs, historical documents and more.

(9/11/98)

---

## David Bearman, Jennifer Trant

### Archives & Museum Informatics

Thanks to Janet and Gary for getting us going here. We've a few things to add.

There is significant activity following up on the Museum Educational Site Licensing Project (that was mentioned by Gary Marchionini). The Art Museum Image Consortium (AMICO) is an open-membership, not-for-profit association of art collecting institutions in North America. This group - now at 25 - is creating a digital library of the documentation of their collections, and making it available for educational use. It involves three members of the MESL Management Committee (Maxwell Anderson - Director of the Whitney Museum, David Bearman and Jennifer Trant). The members have assembled a beta Library - about 20,000 works from 23 collections. You can see a thumbnail catalog, and much other background about AMICO at <http://www.amico.net> (**website no longer operational**).

The beta AMICO library is now being used in a University Testbed Project, by 18 campuses, and made available by the Research Libraries Group (RLG). Testbed

projects are addressing a research agenda that identified questions around the definition of "critical mass" and integration of digital resources into art and humanities teaching.

Other than a small planning grant from the Mellon (\$45K), the entire program has been self-funded. The Museums are paying a membership fee to belong to the consortium, and the universities are paying a subscription fee that covers both a license to the library and the costs of its delivery. All fees are based on cost-recovery for new expenses in distribution only. No money goes back to the museums for their digitization, or documentation costs. Archives & Museum Informatics have been staffing AMICO on a part-time consulting basis.

Needless to say, bootstrapping the consortium has been a very time consuming task. Maybe that's why people aren't as aware of it as they could be :) But now that the "proof of concept" for AMICO is completed, we're planning longer-term to address issues of scaling, pricing, and Library development, and to help identify needs for humanities knowledge models, process models and tools which exploit both.

Some excellent work has been done in **modeling scholarly knowledge** in some domains - the **Electronic Beowulf Project**, for example, has addressed these issues with respect to ways of using an ancient literary text - but best practices and examples of knowledge representations that can interoperate with standard methods for declaration of their schemas are not yet available. For inter-disciplinary research in the humanities, where all research is effectively interdisciplinary, these are critical. And humanities research depends on re-presentation of source materials that are largely not digital in the original.

When we were in York recently, at a conference of ethnomethodologists (a specific kind of sociology) who were studying museum work, we became painfully aware that we are without such studies as *Laboratory Life* by Woolgar and Latour, to help us understand how humanists do things. This means that we are often designing methods and systems in tandem. And not assuming that work processes in humanities are either shared across disciplines or simply the same as those of science.

**Tools to exploit knowledge representations** and fit into work processes are receiving little funding. We're investing a lot in digital content (without enough study of alternative methods of knowledge representation), but we don't have tools or the literacy in the humanities community to use them well. We don't have agreement on a set of analytical methodologies - thus content developed for one project cannot be exploited by others. We're seeing, for example, that project after project is discovering that when you work with images, you need to be able to compare and contrast more than one (this was the basic task in first year art history - outlined in all of the

disciplinary research handbooks). The implications of this for interoperable metadata and basic resource documentation at the point of digital capture are often quite simple (is the scale of the object built in to the metadata so that two objects could be moved into the same digital frame and their representations be correctly sized with respect to each other?), but typically overlooked. This may come back to Gary Marchionini's theme about working in isolation; perhaps it has more to do with an historic lack of focus on "humanistic method".

Editing *Archives and Museum Informatics* (a peer reviewed journal published by Kluwer) also brings home the lack of a literature in the methods of humanities computing. Many projects are conceived first as contributing to a research problem in their own disciplines and only subsequently as posing general computing problems. As a result, they are reported as discoveries in subject knowledge, not in computing knowledge. Common informatics issues that cross these projects are not drawn out, in part because rewards to the researchers are from within their own disciplines, if at all.

Ultimately, we need to find a way to characterise the computational problems of the Humanities in a generic way that is both comprehensible and attractive to the computing science field. Until we have an acknowledged problematic, proposals for humanities computing projects will be assessed and characterized based on their subject content rather than their contribution to the development of computing science.

Humanities computing projects are different and do have distinctive contributions to make. We'd argue for a way of presenting exemplary projects that highlighted this, and problems yet to be resolved: such as "Y1K - Issues of Time in Humanities Computing" - a classic case of knowledge representation that is crucial to decoding the record of the past and organizing it in a meaningful fashion so that new knowledge can be generated from it. In this case, we need to unlock many different systems for measuring and reporting time. A similar case can be made in unlocking the record of space - historical geography knowledge representation that will make it possible to collocate many information resources from the past.

(9/11/98)

---

**Marlene Manoff**

**Humanities Library, Massachusetts Institute of Technology**

Ann Wolpert, the Director of Libraries at MIT forwarded your message containing a request for already published articles about the possibilities of networked humanities resources. You might be interested in piece I originally wrote for a conference called Scholarly Publishing in the Next Millenium. A version of this paper was published in the *Canadian Journal of Communication*. A slightly different version is available on the web at

<http://libraries.mit.edu/humanities/manoff.html> **website no longer operational**

This piece looks at a number of humanities, and especially literary, applications of electronic technology and explores their strengths and weakness. The notes include live links to the web resources and projects that I discuss. The article is an attempt to explore the social and political implications of electronic tools for scholarly research. It draws distinctions between productive and less productive applications and tries to identify the more fruitful directions we might pursue.

(9/14/98)

---

## Steve Dietz

### Director, New Media Initiatives, Walker Art Center

The Walker Art Center and The Minneapolis Institute of Arts have been collaborating for over a year now on an extensive access and education joint project, which was initially funded by the State of Minnesota to digitize museum resources, provide Internet access to, and create educational programs.

The site was recently launched, with support from MCI, under the rubric ArtsConnectEd. Information about the project's genesis is at <http://www.walkerart.org/iaia>. **(website no longer operational)**

While both MIA and Walker are members of the **AMICO** consortium, what I think is interesting and uncommon about this project is the depth of the collaboration across knowledge domains (museum, library, archives) and heterogeneous data types (database, text, image, video, audio, hypermedia). The integration of resources points toward the beginning of a virtual, networked "museum" where the user is the focus, regardless of where information resides--who owns it.

Having just attended the ISEA and Ars Electronica conferences in Europe, what struck me in particular was the significant state funding for extra-institutional projects.

You are, of course, very familiar with many of the major EC-sponsored consortia. To these I would add virtual organizations such as the Hull Center for Time-Based Media, FACT (Liverpool), IDEA, and many other innovative organizations that sponsor projects and attempt to find the most appropriate venue for them. It seems to me that the US could benefit greatly from such "meta organizations," whether as sponsors of hard-core humanities research initiatives or educational programs or artistic/creative efforts.

Part of the long term goal of ArtsConnectEd is to become increasingly a "portal" for the sponsorship and identification of arts-oriented educational resources, regardless of the institutional owner. See, for example, ArtsNet Minnesota, which already has expanded beyond WAC and MIA to include the Weisman Art Museum, the Minnesota Museum of American Art and five community-based organizations around the state. Funding resources that encouraged such inter-institutional cooperation could help a great deal with this and similar projects.

(9/14/98)

---

## **Eleanor Fink**

### **Director, Getty Information Institute**

Your call for ideas is an opportunity for the cultural sector to become an innovator of technological change as opposed to serving simply as the content provider. There are several thoughts that come to mind.

Both a process and new tools are needed to provide subject access to the digital images and art information that are becoming accessible globally on the Web. For example, being able to search by image type will help overcome linguistic differences in a global market. Here more experiments like GII's and NEC's **Arthur** and **Amore** are needed.

As a whole we need to be able to catalogue text and images in new ways that are efficient and that provide the kind of generic access to knowledge that will serve the interests of consumers. Currently we are cataloguing based on traditional manual practices that are not efficient if we are to keep up with the demands of an information society. We need to develop and drive more efforts for rapidly encoding digital information.

**Sustainability of quality content** needs to be addressed. The private sector needs to understand that acquiring content is not in itself the answer. The problem is how to keep quality content available. Keeping it available will require participation of the cultural sector. We need to better understand the **sociology of cooperation**. More collaborative efforts around the formation of consortia such as **AMICO** are needed to better understand how we address sustainability of content.

Likewise, more research needs to be devoted to permanence of digital media. The **Time & Bits Conference** convened by the Long Now Foundation and the Getty Information and Conservation Institutes outlined several actions that need further development.

Forming digital libraries is only one part of the equation. More tools are also needed that will enable the interpretation, story telling, and packaging of information for educational curricula and virtual exhibitions.

All of the above is dependent on marketing research based on **demographics and how text and image information will be used**. We are witnessing the rise of two divergent consumer groups: the largest teenage and senior citizen populations. Each of these groups will generate needs for how information is designed and delivered. For children the internet is rapidly becoming the resource of first resort. These "clickerati" will have high expectations beyond current developments. Likewise, life-long learning for senior citizens will create new demands. We need to devote studies that inform both the cultural and private sectors on market needs.

(9/15/98)

---

## **John Unsworth**

**Institute for Advanced Technology in the Humanities, University of Virginia**

*The report, concluding that Federal support for research in information technology is "dangerously inadequate," calls for a major increase in funding of around a billion dollars over the next five years. It especially emphasized the importance of basic research with a focus on software development, the information infrastructure itself, high-end computing and on the social and economic impact of technology.*

I've been reading the examples that others have sent along, and their suggestions, and I wonder how many of those answer to the priorities expressed above (software development, infrastructure, high-end computing, the economic impact of technology). I would send you my own list of favorites, which would include the **Blake Archive** (a model of editorial standards and practices in electronic editions), the **Valley of the Shadow** (a fine demonstration of how technology can change the way a discipline conceives of itself), the **Perseus Project**, the **UVA Etext** and **Michigan Humanities Text Initiative collections**, and so on. Like the other lists, it would be full of great projects, with a certain prejudice for the locally produced, but I don't think it would have a lot to do with the expressed objectives of the report you describe.

If I were to try to address those objectives, I'd offer something different--and it would mostly be a list of things we ought to do, rather than things we've done.

- **Basic research with a focus on software development:** As we begin to develop significant electronic collections, the absence of analytical tools of all sorts is increasingly obvious, as is the need for them. All the electronic resources in the world will not transform research (or teaching, or the culture at large) if all we can do is search and browse them. We need **general-purpose, standards-based tools for visualization** (text-visualization for starters--see the **Dante project at IATH** for an example). We need to bring forward into the networked, XML-oriented world of distributed collections some earlier tools like Tact and Collate, and we need to extend them significantly to deal with more arbitrary data formats and structures. We need much better, and many more, tools for working with Unicode. We need--we especially need--an **institutional as well as a technical framework for coordinating, documenting, and preserving distributed software development**, where A builds the piece that she needs, and B can later find it and integrate it into something he needs.
- **The information infrastructure itself:** We already have the beginning of what we need here, in [Internet2](#), but we need to pump a lot more money into this if the next-generation internet is not going to replicate the current inequities of service and speed. Small colleges, poor colleges, remote universities, historically black institutions--all of these have something to contribute, but they won't, if they have to wait for five minutes for each web page to load.

- **High-end computing:** image analysis, analysis of time-dependent media like video and sound, very high resolution three-dimensional modeling based on real-world data, and distributed data modeling are all potentially of interest to those in the humanities who work with film, music, architecture, statistical data, etc. The current funding and administration of such projects does not usually bring humanists into the mix when it comes to designing the tools, so it's not much of a surprise if it's hard for them to see how the work of the molecular biologists and organic chemists might answer to their needs.
  
- **The social and economic impact of technology.** I think that the single most important thing I could propose in this category would be this: we need to provide funding that will reduce the risk in:
  - **experimenting with different economic models of scholarly publishing** (and documenting the empirical results), and
  - experimenting with methods for **interconnecting current scholarly research collections across commercial boundaries** (imagine if you had to go to the library on Wed. to read Oxford UP journals, and come back Saturday to read JHUP ones...).

As one who's been involved in electronic journal publishing for a long time, I can testify that it's come along a lot more slowly than anyone expected, and the perception of risk (on the part of publishers, and on the part of authors) is the only plausible explanation for that--as far as the intellectual goals of the research activity itself goes, the arguments are all on the side of ejournals.

If I had to choose one of the above as most important, in the long run, it would be the last, hands down. If you want to have a major impact on the production of knowledge, then research publication itself is the place to aim. Inasmuch as my earlier, praeterited examples go (Blake, the Valley of the Shadow, etc.), I'd sail them under this flag, rather than on their merits as collections of content. Change the terms on which people create and exchange and analyze research--in whatever discipline--and the rest will follow. Well, maybe not an end to injustice, but at least significant changes in the economy of knowledge.

(9/18/98)

---

Another idea. It is based in part on the kind of work that Ed Ayers has been doing, but it supposes an extension of that work both in disciplinary terms, and in local and institutional terms.

## **The Place Project:**

While I was at Lawrence Univ., talking with the faculty there, I stumbled on what I think is a great idea. Here's how it goes.

Small-town colleges and universities have a difficult time getting their faculty and their students trained and involved in electronic projects.

Why not put together a package of tools and training and documentation that would help such places do those things as a Place Project. Place projects would focus on the city or town in which the college or university in question was located: they would provide a framework for faculty and students in all disciplines to do primary research on the place in which that college or university is located, and combine the results, in electronic form, into a multi-faceted digital archive. Local history, art, film, television, literature, archaeology, architecture, sociology, geology, zoology, biology, etc. etc. etc. would be researched by students and faculty, and would be recorded in digital form under a single, sophisticated, standardized, and durable information structure.

A number of interesting results might be expected:

- within the university or college, one could expect students to have a keener sense of the practical applicability of the disciplines in which they received instruction.
- within the university or college, one could expect faculty to find new and very practical points of conjunction with colleagues in other disciplines.
- one could expect such a project to improve town-gown relations, by improving understanding of the local community on the part of the faculty and students, and by demonstrating to that community the commitment of faculty and students to understanding and preserving local culture, local history, and the local environment.
- if the concept of Place Projects caught hold and proliferated, the result would be a very rich and very durable information resource, like nothing we've ever had before.

Such a project would by its nature be eligible for funding from a variety of sources outside the college or university that initiated it--local and state historical societies, local and state government, individual and commercial donors from the local

community, national humanities and science foundations, and a broad array of private foundations as well.

Moreover, by its nature, this project would scale readily from one or two initial installations to a nationwide network of place-based digital information. Naturally, it would require substantial funding to create that national network, and it would require a collection of discipline-based and technologically sophisticated professionals to design the tools and documentation, and a cadre of advanced student participants to provide the training.

(9/19/98)

---

## Allen Renear

There are quite a number of wonderful projects being described and, if I don't include them below, I'll add a few of my own to the list in another piece of email later today, as well as a couple references to already published rationales for increased funding for humanities IT research projects. But in thinking over how I could best help with this effort, given the already impressive and fairly comprehensive contributions I've seen, I thought maybe I'd try to help in another way, by outlining the **two strategies** I think are most important in this sort of situation -- trying to shape federal R&D funding policy to make available to humanities IT research some small portion of resources primarily aimed at supporting economically important scientific research.

Nothing original here, these strategies are the obvious ones (even though we typically shrink from deploying them directly and without apology) -- but I think it is worth reminding ourselves of how they go, and then maybe considering whether we might want to step up to using them a little more boldly than we do.

But first I want to mention a third strategy, which, as important as it is, is too familiar and too simple to need discussing, that's "the Toto strategy", where one says, brightly, "...and the humanities too!". Nothing wrong with that, and, in fact, it regularly elicits pretty fair results for very little effort,. But there's no need to further elaborate on it here; it is easy to understand, easy to do, and familiar to all of us..

Now on to the two strategies I want to emphasize. They are also familiar, but unlike "Toto too!" worth thinking about a bit more.

**Strategy 1: Argue that support for humanities research can have a substantial economic benefit to the nation.**

The argument has two parts.

The first part is rehearsing the size and economic importance of the culture industry: popular music, television, movies, fashion, design, graphic arts, newspapers and magazines, trade publishing, libraries, museums, tourism, and so on. This part is easy. The culture industry is an absolutely massive part of our economy and an extremely important positive force in balance-of-trade; and, in addition, most of the service economy which is strictly speaking outside the culture industry is nevertheless thoroughly dependent upon many of the same methods, tools and techniques: as found in communication technologies, publishing, design and graphic arts, etc. In such a situation even small improvements and innovations in basic tools and techniques can have considerable larger economic effects.

The second part of the argument is to establish, or at least make it plausible, that **humanities IT research actually does have economic consequences**, that it makes difference. This part is harder, but it can be done; and it must be done. One begins by pointing out that humanities IT research has as its very subject developing theorized methods and techniques for creating, manipulating, analyzing, and using text, image, sound, etc.

At this point one could then go on to argue, in a detailed analytical discussion, that, therefore, if properly funded, humanities IT research will yield economically valuable results. But in fact the purely **analytical** argument, however worth doing, at least briefly, will not really be taken very seriously on its own unless one can also make the **empirical** case that humanities IT research has already had economic substantial results. That is, that innovative economically important theories, methods, and technologies for creating, managing, enhancing, and understanding cultural material have frequently originated in the sort of research we are promoting.

Can this be done? I think it can. To take just one example. Next year when you follow a WWW hypertext link in an online magazine, or internet game, or a virtual museum (or for that matter when buying some shoes), the underlying data structure for that link, XML-linking, will be based directly on work (TEI Extended Pointers) carried out by humanities scholars in the early 1990s. Let me say that again, within a few years most linking on the web -- and that's what the web is all about, linking, will be based on protocols developed by publicly funded (NEH) humanities scholars. And there are many more examples, particularly in preservation techniques, metadata, multimedia tools, and so on.. And if NLP and corpus linguistics counts as humanities, as we should at least argue, since it makes the case so well, there will be many more examples, in, e.g. areas of voice recognition, text understanding, etc. We will find that our examples actually play out in domains such popular entertainment, business communication, tourism, etc. But let's not shrink from that.

Of course this general pitch is not new to those in the humanities and arts who have been busy promoting the cause all along; and there is a fair amount of existing analysis and rhetoric already available that can be adapted. For a particularly **insightful UK analysis of the economic benefits of supporting the humanities and arts** see **John Laver's, "The Need to Invest in Research in Humanities & Arts"**. Although Laver (Chairman of the British Academy Arts and Humanities Research Board) does not directly address supporting humanities technology research in that report, he did address that exact issue later, at Kings College London last spring at a CCH event I attended. I'd summarize the theme of that presentation this way:

- don't be deferential, or defensive or subtle;
- humanities IT research has an extremely strong case and just needs to make it;
- it needs to say, boldly and emphatically and repeatedly, that it, and only humanities IT, is directly concerned with developing new methods, tools, and theories for dealing with the stuff our economy depends on: text, image, sound, cultural creations, cultural knowledge.

(One may also want to argue that when this work is adequately funded in the pre-competitive non-commercial environment, it is more likely, particularly where standards development is concerned, to achieve foundational interoperable technology based on best practices and current research rather than having development aborted prematurely by industry efforts to seize a short-term advantage by locking in a customer base. That advantage means a healthier competitive environment leading to more efficient and innovative development and obviating the need for legal measures to redress unfair practices)

## **2) Humanities IT research will be a source of deep research results**

Here the argument is typically in part empirical and part analytical.

The empirical part cites **examples of how the humanities IT community has already been the source of substantial research results**. Many of the examples used above will do just fine here as well, although with a different twist. **Why do we know as much as we do about documents, text structure, character sets, writing systems, classification systems, preservation issues, user behavior etc.** Because of the devotion of humanist scholars. Whether it is the intricacies of "overlapping hierarchies" in document structure, or data-driven typed hypertext links, or classification ontologies, some of the best and deepest and most consequential results come from this community. It is no accident that humanities scholars have been and

continue to play a major role in the most important areas of IT development: Michael Sperberg-McQueen's leadership in the development of XML, Steve DeRose's leadership in the development of XLL, etc.

The analytical part argues that the nature of humanities IT research is such that it is a **rich source of results**. This is because its problems are so hard, so complicated, so interdisciplinary, that solving them naturally elicit powerfully general and explanatory theories, theories that scale, that are scientifically productive and heuristic, and that connect with other established theories in other areas. E.g. solving the problem of document structure for ordinary, and ordinarily used, technical reports is fairly simple, even a bad theory will do. But bad theories don't scale, aren't productive, don't lead to new insights, don't support wide-ranging practices, aren't conducive to general tools, etc. But solve the problems presented by textual criticism and multidisciplinary textbases ... and you've got a document theory that just might be deep enough, and general enough, to be the framework for tools of genuinely new capabilities.

I have argued that humanities research IT in general, and textbase research in particular, is a very rich source of research results in: "**The Digital Libraries Research Agenda: What's Missing--and How Humanities Textbase Projects can Help,**" **D-LIB, July/August 1997.**

### *Part Two:*

#### **The TEI**

The **Text Encoding Initiative** is an absolutely stunning achievement. It must be at the top of the list. I'll let others characterize it in general (e.g. Chesnutt), but there is one thing, a very important thing, that they might miss. The achievement is not the guidelines -- however wonderful and timely they are. It is as *research program* that the TEI is so extraordinary. The TEI is **an approach to investigating text** that has been, and will continue to be, amazingly deep and productive; it will be spinning off important theories, techniques, practices, methodologies, tools and other results for many years to come. That's its real value. And that's why it is a funder's dream come true: this is exactly how funding is leveraged, by creating new research cultures, not just a result or two, or a resource, or a tool, but a whole new approach that turns out to be rich and productive of many theories, methods, and tools.

#### **The Women Writers Project**

I don't think anyone has mentioned the Women Writers Project yet. So I will. Though I'm not disinterested. Again I'm going to skip the obvious: how the WWP recovers women's writing, of many genres, and makes them available to everyone everywhere

in formats that provide fundamentally new ways of engaging the texts, and the consequences this has for research and teaching in history, literature, linguistics, anthropology, religion, etc. That case has been made, and made well, often enough before.

But let me generalize a bit. While many speak of digital libraries, metadata, explore the physical technology, exploit quick hits like image bases on the one hand, or distant visions like docuverses on the other, the WWP was *the* pioneer in the **development of a TEI-based encoding system for printed books considered as multi-disciplinary cultural objects**. That is, we really took on the fundamental question: how does a library represent real texts from real books, that are of substantial multidisciplinary interest as cultural objects, in ways that will genuinely and deeply exploit the potential of information technology.

The WWP has discovered a tremendous amount about about the intricacies of text structure -- as a knowledge representation system -- and the about the interplay between text structure and physical rendition in a printed book. This is what must be known if we are to build a real digital library, which would be a repository for real cultural objects and create fundamentally new and powerful techniques for engaging those objects. Moreover, this is what must be known if we are to have general robust theories of textuality, ones which support tools and methodologies that will truly scale and evolve.

(9/23/98)

---

## Morris Eaves

**Professor of English, University of Rochester, NY**

I wish I could bring a vision, but I bring only a project--the **William Blake Archive**, a free site on the Web for the past 3 years. We who work on it think it represents at least one significant aspect of the future of electronic humanities scholarship--both in itself and as a prototype for other image-oriented scholarly enterprises. We explain much of that in some materials we developed this past summer. Those will soon be available in final form in the About the Archive wing of our site. Meanwhile, for what it's worth, I'll attach a file that contains intermediate drafts of a substantial segment--in a form prepared for hard copy, however, rather than for the Web.

But I honestly don't mean to bury you in paper--or in information that is extraneous to your particular (very worthy) purposes. (IATH, as I'm sure you know, is a hotbed of humanities-computing development, so we are far from the only project worth mentioning--but we think we occupy an important place on the spectrum of possibilities.) I'm not sure what you really need--feel free to ignore any or all that I've sent. But if we can be of any further help, let us know. Morris Eaves (with Robert Essick and Joseph Viscomi, Editors, William Blake Archive, <http://www.iath.virginia.edu/blake>)

See: The William Blake Archive: Summary of Project (local file)

(9/26/98)

---

## Wendell Piez

### Mulberry Technologies

I feel I'm not really able to do justice, at this point, to your request. In part, this is because I haven't been in touch with the best current developments in the academy. I'm aware of some projects I myself consider exemplary -- prototypes at UVa like the William Blake project, or the "Documenting the American South" project at UNC-- but I haven't had a chance to develop the sense of range I think necessary to provide even a representative listing of the best examples available.

More particularly, however, this is because my current feeling is that we still have **much work to do to develop the technologies into something easier to use**. At present, the leaders are showing it is possible to develop scholarly electronic resources of tremendous flexibility and power in a platform- and software-independent way. This is a big advance, because projects that are platform- or software-dependent can never be more than proofs-of-concept. Markup languages based on SGML -- and now, on XML -- are making this possible, as the most readily available, externally specified (i.e., non-proprietary) form of data encoding flexible enough to meet the considerable requirements of Humanities applications. It is not particularly easy to do, however. Software is expensive and expertise is rare, not least because these technologies have not, themselves, been mainstream within computer science departments.

This leads me to conclude that for the moment, the most fruitful place to apply resources may be in testbed projects that take advantage, not just of the richness of materials and the wealth of interesting applications that can come out of Humanities departments, but of academic initiatives to provide **software** to implement these

projects. Currently, projects in the Humanities are more or less dependent on systems and software that Humanities departments simply cannot afford, even in the rare cases when they know of their existence. This "barrier to entry" has to be lowered. We must also keep in mind that the people who really learn from these projects are the *participants*, quite apart from any audience they may happen to reach.

Most of the work to do this is actually in **institutional bridge-building**. Relative to the kind of money routinely spent in science and technology programs, the projects need not be expensive. But they will require new kinds of cooperation between humanities and CS programs, cooperation that can only happen as each side gets better informed about the real contributions to common goals provided by the other. Funding, applied strategically, can be a great catalyst for the kinds of efforts necessary. Standards to support this work (XML and its related standards) are right now coming into place. It has to be stressed that the work has to be standards-based, both so that the encoding methods themselves can be open to study (would you teach carpentry by providing students with pre-fab kits?), and so that materials produced can have longevity beyond the four-or-five year lifespan of a software program.

As your survey is undoubtedly showing, proofs-of-concept exist -- although I also believe that new concepts will emerge as the work proceeds. To become widespread, the dependence of these projects on expensive software developed by vendors for other applications has to become a thing of the past. Within the academy itself are all the means and resources necessary. It's like a chemical suspension that only has to be shaken a bit for crystals to appear.

(9/26/98)

---

## Michael Joyce

### Vassar College

I am sorry to have taken so long to respond to your important call. The following comments expand on remarks prepared for the 1998 MLA forum, "Revolution or Evolution?: Electronic Resources in the Humanities." I explain my perhaps curious nominations in a headnote to that section.

All best wishes.

Michael

## STATEMENT

Computer spaces are spaces of becoming, their contours drifted over and reformed by what I have called elsewhere the blizzard of constant nextness. Computer spaces do continue to summon us to the traditional, which is to say evolutionary questions which we face as humanists- how do we understand who we have become, what links us to past accounts and actions, what binds us to hope and an imagined future. On the other hand the morphogenetic nature of the space of network culture inevitably summons us to new, which is to say revolutionary questions about self and past and hope alike: where can we be who we are, how can we know how others see themselves, how can we form our differences into shared perceptions of shifting experiences.

It is this latter summons, these more complex questions, which I mean to elicit in speaking about post-hypertextuality. The web itself is in a sense the first post-hypertextual medium, an impoverished one that has discarded many of the fundamental concerns of a fifty year history of hypertextuality. Hypertext throughout its foreshortened history has been characterized by a concern with finding new narratives to accommodate and commemorate shifting perspectives. The new media of post-hypertextuality truncate this process, substituting successive perspectives for narrative, novelty for memory.

An immediate task facing humanists in a post-hypertextual age is one we have claimed to find congenial throughout all ages: we must recover history. Attention to successive aftermaths in a place of constant nextness is the fundamental calling of humanists.. In creating new social spaces we are called to whole ways of making sense of the world rather than ways of making the world make sense as a whole. The hypertextual link held the reader responsible for contextualizing outcomes as well as predicting prospects-- both gestures which are familiar ones to humanists.

A post-hypertextuality attentive to its own history will ground itself in the succession of spaces where we understand our lives as taking place. Having fashioned fictional spaces of our individual links where can we gather? Where can we situate our linked responsibilities and shared differences so as to constitute real human community?

These concerns may sound like mere polemic but 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. Our embodiment graces actual and virtual space alike with the occasion for value. It is not enough merely to decenter the classroom in networked spaces if we do not insist upon holding ourselves, students and teachers alike, responsible for describing and

evaluating the constellations of succession and aftermath which we now share in lieu of centers. It is not enough to extend our inquiry into literature, history, art and philosophy through a web of linked discourses if we do not recognize and assert how what we have woven resituates and renews our expression of literary, historic, philosophic, scientific and artistic values. Stripped of transcendence, we embrace persistence; bereft of tradition, we celebrate recurrence; unmoored, we link ourselves through, rather than in, our technologies.

## **SUGGESTED SITES**

As regards exemplary projects, rather than suggest obvious choices (Ed Ayers "Valley of the Shadow" project at IATH or John Price-Wilkin's [American Verse Project](#) at the University of Michigan for instance ) I want to propose three quite different sites at varying stages of development, one maintained by an independent scholar, another by a university affiliated educational institute, and a third by a largely volunteer based social service and educational agency and NGO, each of which represent the humanities in action and suggest the potential for genuine networking. Two of these sites are local to the Mid-Hudson Valley where I live and so could be seen as representing a parochial interest were it not for the fact that they center upon historical figures, Franklin D. and Eleanor Roosevelt, whose much wider prominence may make this claim of locality seem disingenuous. Yet that to the degree network culture affirms our sense of place and confirms our understanding of how locality informs universality I would be comfortable with these recommendations solely on the basis of the scope and focus of their interests and their impact upon the communities they both serve and serve to widen.

### **1) The Nathaniel Hawthorne**

**Site <http://www.tiac.net/users/eldred/nh/hawthorne.html>**

is a model of what "a voluntary effort by a single amateur" (in the root sense) on the web should be but which unfortunately very few are. This is a rich and various set of texts, links and resources by an author attuned to issues of textual integrity, standard editions, mark-up and other scholarly and technological issues but not too stodgy to turn on the TV or talk about his love of Hawthorne with both neighbors and networked correspondents. The site's more idiosyncratic links prove the joys of the web and demonstrate the power of an engaged and engaging human being. Features like an essay contest for readers of the site or the "Search for the Great Carbuncle Source A transcript of some email messages, reprinted by permission of the authors"

create (or perhaps illuminate) a genuine community of readers. A list of college courses offers an expansive view of scholarly approaches to Hawthorne while photo essay on the Old Manse study is charming, intelligent, and utterly engaging. Other resources focus upon Hawthorne in popular media such as the Discovery Channel, Classic Comics and even (with appropriate warnings, god help us) Monarch Notes, general resources (encyclopedias and so on), and local (New England) or other idiosyncratic sites. This site sees learning as human, multiple, omnipresent and joyful. Hawthorne is presented as a living figure, in his time and our own, with an international audience whose network existed in the nineteenth century and is enhanced by the network of the late twentieth.

The site is cleanly designed with a surprising eye for graphics from time to time. Subheads and interspersed graphics are effective in a flea market sort of way. Texts are carefully edited and marked-up with attention to TEI conventions. Search engines are serviceable and fast. Text-based navigation tools are available at the top and bottom of pages; internal links move through pages and yet do not disorient. The author's access warnings deserve special mention for their inclusiveness and attention to human as well as legal issues. A good teacher could use this site to great advantage to show how the joy of learning leads to true human engagement. High school, college, and adult students alike could visit and revisit the site without losing interest, and in fact gaining in their interest in a still compelling major American author. The range of Hawthorne resources is liable to appeal to multiple audiences and the author offers compelling interactions through simple means. The textual database here is a generous contribution to our literary heritage and an instance of what it means to "develop active interest and mastery of the subject area." All in all the site is an instance of what the web once promised to become and a model of what citizen humanists might be able to create with appropriate funding and promotion.

## **2) The New Deal Network (<http://newdeal.feri.org/>)**

is a remarkable resource for text, photographs, and documents of FDR, his circle, and the New Deal geared to teachers and students. These "materials on the Great Depression and New Deal [are] from libraries and archives across the nation" and the current collection is described as "a small sampling of about 3,000 images from the National Archives in Washington and the FDR Library in Hyde Park, NY." The site includes "classroom lesson plans...related to [each] issue's feature stories....contributed by teachers who helped in the creation of this web site." Visiting teachers at the site are invited to join the network and share lesson plans. Organized around thematic issues (in both the periodical/magazine sense and the intellectual sense), one issue for instance is about work relief and welfare and focuses on WPA and CCC projects and how they enriched the lives of Americans. A rich collection of resources includes photographs, historical documents, oral histories, timelines and so on. Beautifully

designed with elegant and simple tab navigation (home, library, classroom and timeline), this site clearly has a student and general audience in mind and offers an engaging and multi-levelled set of resources and activities. An innovative "destinations" link on the home page leads students and visitors on a trip to a remote site (for instance the WPA murals at the University at Albany). While this resource has been growing steadily, it would profit mightily from dependable funding and offer a model for similar educational initiatives.

### **3) The Eleanor Roosevelt Center at Val-Kill (ERVK) website (<http://ervk.org/>)**

is currently celebrating the 50th anniversary of the Universal Declaration of Human Rights. ERVK describes itself as "a non-partisan, non-profit organization dedicated to Eleanor Roosevelt's belief that people can enhance the quality of their lives through purposeful action based on sensitive discourse among people of diverse perspectives focusing on the varied needs of society." Obviously this language anticipates the sorts of claims for polyvocality that have been made on behalf of the internet . This may explain why the site, barely the requisite year old, already seems something more than merely a "web presence" and much more a staging area and a place of reflection and action, not unlike the stone cottage of Eleanor Roosevelt which gives the organization its name and which served as a staging area for her own world-wide network of interests transcending socio-economic stratifications.

Of the three sites I am recommending this one could most obviously benefit from increased funding for network projects in the humanities. It is clear that the organization would like to expand its projects on-line and in the process both serve and stimulate distant and disparate communities through the network. Funding could help the organization serve as a repository for documents and resources for others interested in "the empowerment of women, racial and cultural diversity, the promotion of youth leadership, addressing the needs of children" as well as other projects "in the fields of human rights, democracy, international peace and the promotion of international understanding."

Here I wish to be careful to disclose my own relationship to this site. My wife, the hyperfiction writer Carolyn Guyer--whose *Hi\_Pitched Voices*, women's collaborative hypertext collective was an early and now legendary effort at truly constructive hypertext-- is an ERVK volunteer and designed its site. Her **Mother Millennium** project, co-sponsored by ERVK, has already begun to collect its proposed "2000 stories of mother by the year 2000." Mother Millennium's accumulating body of linked works on the subject of Mother is meant to be "told from as many different cultural perspectives as possible, world-wide" and to include "a variety of forms, including memoirs, graphical narratives, fiction, oral histories, poetry, essays, video, and sounds."

(9/27/98)

---

## Susan Hockey

**Professor and Director, Canadian Institute for Research Computing in Arts,  
University of Alberta**

I thought the response from Jennifer and David was excellent and so please add my name in support of everything they said.

I think there are two real issues to be noted here:

1. Those of us involved in this activity seem to be largely a community of producers. I am really interested to **know who the consumers are**, what they think of our products and for what purposes they will use those products. I don't mean here the "consumers" who are known to the producers and are lined up to write letters of support for grant applications or to sit on test panels, but the people out there who will use these products and create enough of a user community to make them sustainable.

2. "**Sustainability**" is my next point - I don't want to use the term "costs" as it raises so many hackles in academia. I think that we need to do a lot more fundamental research in how to make electronic products and how to ensure that they meet the needs of as many groups of users as possible without compromising the academic standards that we all expect to see in print publications and without eating up too many resources. This might mean less emphasis on the kind of "glitzy-interface" project which is usually trotted out for the Board of Governors and other dignitaries, but which also may well lack any serious academic depth. It ought to mean more emphasis on user studies, pilot projects that write up and publish their experiences in detail, and the development of general purpose tools that address user needs in the humanities.

I'm not really in favour of creating lists of exemplary projects unless it's absolutely clear by what criteria they are being judged. Things which might be exemplary for high school level are probably of much less use for advanced research.

(9/28/98)

---

## James Michalko

### President, Research Libraries Group, Inc.

Thanks for the invitation to contribute ideas and suggestions to bolster the argument for more funding for humanities technology support. Because the humanities are often overlooked when technology funds are handed out, there are areas where much-needed funding could provide quantum leaps forward in providing access to research resources and in turn, transform the kinds of humanities work that could be undertaken.

Despite a lack of funding, many efforts at using technology to support humanities research have been undertaken in academic departments, libraries, and by individuals, because the need exists and the value is recognized. Those early experimenters in the humanities world have amassed experience, developed methods, and learned some hard lessons. Increased funding can build on these early efforts, bringing those experienced pioneers into larger collaborative initiatives with their more inexperienced colleagues. In addition, **harnessing some of that local, individual experience can help us all move beyond "projects" to large scale production, and long term solutions** to the common problems that plague most academic institutions today.

The challenges of maintaining and sharing enormous amounts of information about research resources, forced the library world to lead in many areas of standards development and technical application. Libraries have established standard record formats, semantics and syntax; developed modes for document delivery; embraced international interoperability standards; and promoted dependable preservation techniques for conserving documents and reformatting information. These endeavors facilitate access, interchange, and preservation of information that supports all fields of research.

Initiatives in the library world over the past thirty years have focussed on describing, preserving, and sharing information within and among institutions. Current goals include those same functions, but we now have the promise of much greater return on investment, due to advancements in technology.

The Research Libraries Group for the past 25 years has been a leader in the library world's efforts to take best advantage of technology to support teaching and learning, especially in the areas of the humanities and social sciences. Promoting interinstitutional collaboration to solve common problems has been a hallmark of our success in developing data exchange standards like MARC and Z39.50; in promoting

standards for preservation and ILL protocols; and in finding better methods of enhancing access to unique research materials in archives, special collections and museums.

## **PROJECT EXAMPLES**

In addition to the development and implementation of standards mentioned above, some examples of RLG projects that have addressed specific issues in enhancing access to humanities resources include:

### Multi-institutional access projects

- "Seven States Project," brought together a group of US state archives to develop standardized description and a means for sharing government records appraisal and disposition information. It was subsequently extended to include 16 major government repositories, and resulted in an expanded use of RLIN for government records processing and reference, and the inclusion of government information in the mainstream of more traditional research resources.
- "Digital Image Access Project" was the first major multi-institutional effort to address access challenges and strategies in the digital environment for photographic materials. The project included a final workshop and published proceedings.
- The "Studies in Scarlet" project, a theme-based, multiple format, interdisciplinary, multi-institutional digital reformatting project brought together complementary, dispersed materials, demonstrating the power of a "virtual" collection - a sum greater than the total of all its parts. It also demonstrated the complex issues of content selection for digital collections, the problems in providing "seamless access" to widely varied formats of information, and the value of sharing information about what works and what doesn't in the process of a digital collection project.

### **Education**

- Symposia like "Selecting Library and Archive Collections for Digital Reformatting - An RLG Symposium," involved participants in the discussion of selection strategies for making local decisions about digital reformatting and engaging in digital projects.

- The publication of "Preserving Digital Information: Final Report and Recommendations of the CPA/RLG Task Force on Archiving of Digital Information," framed the key concepts in digital archiving and prompted an active and sustained international discussion about how to best address the issues. Many of the recommendations have been pursued while others have proved more challenging, especially the more technical problems surrounding long term digital archiving.
- Web-based information like "RLG DigiNews," helps practitioners stay abreast of current technological developments
- Workshops like Encoded Archival Description (EAD) sgml training sessions for the archival community, and "Managing Digital Imaging Projects" are used to widely promulgate the use of standards and best practices.

## **Collaboration**

- Developing best practices in such areas as digital image capture requires experimentation and collaboration
- Preventing duplication of effort by reporting digital preservation reformatting activity in MARC records will prove enormously cost-effective
- Assembling experts to address common issues like the Metadata Summit meeting of twenty-four representatives of the library, museum and archival research communities can be both timely and effective at identifying a common action agenda for information providers
- Shared tools like the "Worksheet for Estimating Digital Reformatting Costs", model RFPs, standards application guidelines,- all require resources to author, develop, test, and implement.
- Consortial pricing for services, like retrospective EAD conversion from a tested and trained service bureau can be an extremely helpful option for institutions that need assistance in taking their legacy finding aids into electronic form
- Integration of research materials from many repositories, like RLG's Archival Resources, provides integrated online access to finding aids, collection-level records, and digital versions of documents.

## **VISION**

### **The challenge of primary sources**

While technology is changing rapidly, our understanding of the need to use descriptive and technical standards and community based guidelines for both preservation and access increasingly assure us of the continuity of our efforts over time. Today, with the world wide web and the growing sense that we can influence and facilitate the development of information discovery and delivery, the goal continues to be providing the researcher an apparently seamless web of related resources that exist in different repositories around the country and increasingly around the world.

Primary materials are the mother lode of academic and cultural research, but are among the least accessible for users and often present special description and preservation needs. Primary sources come in many different forms: manuscripts, early editions, papers, archives, sound recordings, digital resources, and material objects. **These materials, so essential to researchers, have also been the most notoriously difficult to locate and access. They enable us to define ourselves as a civilization. If we can open up historically inaccessible special collections and link them to the existing base of scholarly publication, we will allow scholars to expand and refine our knowledge of every area of human endeavor.**

Unlike books and periodicals, which have long enjoyed detailed cataloging in unified formats that could be readily translated into the digital universe, primary sources have not been so easily labeled for access and navigation. If primary sources are not easily found in the digital universe, then the work of building our information infrastructure is incomplete.

In the existing large U.S. bibliographic databases, archival materials and special collections are often represented with a brief collection level entry that can point researchers to resources. By making more detailed descriptions found in archival finding aids accessible electronically, the researcher can gain much deeper understanding of the kinds of materials included in these collections. And by actually presenting selected documents or discrete series of records in a digital format, the user has the advantage of seeing and evaluating unique information without having to first travel to the collection.

Working together we can establish agreement on standards and practices, to test methods in real-world settings, to disseminate results, and finally to implement systems for making primary sources available to researchers digitally.

In building virtual collections, the entire spectrum of descriptive practice and access techniques - bibliographic records, encoded finding aids, and descriptive metadata, many linked to whole information objects - can be used to create a seamless web of research resources. And the technical infrastructures can provide discovery and

delivery of these resources in an internationally distributed fashion. Commitment to supporting long-term access to and preservation of this information will ensure that the fruits of the labor are widely accessible into the future.

## **Collaboration**

The use of digital technology has captivated the world of information professionals and providers. The number of projects and initiatives to take advantage of new, more powerful methods of access and research are myriad. Most importantly, the library and archival fields are concerned that the technical and intellectual developments be consistent with standards for information exchange and that they support long-term access to and preservation of information in a distributed network that is compatible among institutions and across geographic and political boundaries. We are no longer solely dependent on massive pools of information residing in central locations for information delivery. But the distributed model of information sharing requires many of the same information exchange standards and some new ones to ensure the long-term quality access to our research resources.

[The Research Libraries Group](#) (which is an international alliance of universities and colleges, national libraries, archives, historical societies, museums and independent research collections, and public libraries) helps institutions do collectively what it not possible to do individually. Together we envision a world where researchers in any field can gain digital access to the materials they need to support their research. Creating such a unified digital research infrastructure is obviously beyond the scope of any single institution's resources. Furthermore, any standards, practices, and technologies used to build this infrastructure must be mutually agreed upon or the efforts will be hampered by a vagary of methods and approaches.

Useful URLs for more information about RLG programs and initiatives:

Home site: <http://www.rlg.org/toc.html>

Archival Resources: <http://www.rlg.org/arrhome.html>

Museum Resources: <http://www.rlg.org/strat/projmusres.html>

European Library Data Initiative: <http://www.rlg.org/strat/projeuro.html>

Marriage, Women and the Law (a digital collection project): <http://www.rlg.org/scarlet.html>

(9/30/98)

---

## The Getty Information Institute

(see separate file)

(10/1/98)

---

## Guy Hermann

### Mystic Seaport

(10/13/98)

#### **EXAMPLE**

As there seems to be some own horn tooting happening, I'll nominate **Exploring Amistad** (XA) <http://amistad.mysticseaport.org> (**website no longer operational**) as an excellent, low-budget, effort to draw together a widely dispersed digital archive and place it in a historical context. Like some other projects, XA collects digital versions of documents which help tell an important story. XA extends the concept by providing contextual materials and secondary sources which link into and back out of the digital archive. The goal is to help mediate and make accessible the primary documents for both teachers and students who may be unfamiliar with either the story or with the nature and use of historic documents.

#### **VISION**

The web offers a significant opportunity to bring the real stuff of history out of the archives. But connecting that raw material to the needs and backgrounds of a broad audience will be a new challenge. How do we make primary documents meaningful? We have a chance to fundamentally reshape how we, as a people, understand and know our own history. How we do this will require many more **experiments with contextualization**. Funding these experiments is a critical next step.