

Roy Rosenzweig, Director  
703-993-4532 (w) • 703-522-2334 (h)  
425-984-9046 • [roy@gmu.edu](mailto:roy@gmu.edu)

July 1, 2004

Jesse Ausubel  
Program Director  
Alfred P. Sloan Foundation  
630 Fifth Avenue  
Suite 2550  
New York, NY 10111

Dear Jesse:

On behalf of CHNM and the ASHP, I am submitting the final report of the September 11 project. Delia said that you preferred something brief and so we have provided a two-page summary that reviews the four major areas of the original proposal and our work in each of those areas. We also include some notes on a fifth area—preservation—that was only gestured toward in the original proposal, but has absorbed a considerable amount of our time and efforts.

Appended to that summary report are a few documents that provide additional details—a narrative of the project that Jim Sparrow wrote for a collection on public history, a more detailed accounting of the different areas of work, an inventory of the collection, and an article by Dan Cohen for a history journal that discusses some of the implications of the work.

We are deeply grateful not just for your support of the project but even more for the original idea behind it. Your vision has led to what I think has been an extraordinarily successful project—a landmark in online digital collecting and digital preservation as well as a collection of enormous importance on the events of September 11, 2001.

Best wishes,



Roy Rosenzweig  
Mark and Barbara Fried Professor of History & New Media

***Final Report to The Alfred P. Sloan Foundation on the September 11 Digital Archive:  
Collecting, Preserving, and Presenting the History of the Twenty-First Century***

**SUMMARY REPORT**

With the generous support of the Alfred P. Sloan Foundation, the *September 11 Digital Archive* has established itself as the most important collection of digital materials related to September 11, 2001. It currently contains more than 140,000 digital items, and, with continuing cataloging and contributions, we anticipate that the number will increase by tens of thousands before the end of the year. Since the website was launched on January 11, 2002, it has received nearly 120 million hits and more than two million unique visitors, and it ranks first among more than 7.5 million “September 11” websites indexed by Google. The Archive has been featured by many international media outlets and is the repository of choice for many of the most important September 11 collections. Recognizing its importance, the Library of Congress has agreed to permanently preserve the Archive, making it the Library’s first major digital acquisition. Below we briefly review the four major sections of our original proposal to the Sloan Foundation. We have accomplished or exceeded almost all our goals in each of the areas. In addition, we have made substantial progress on a fifth area – preservation – that was not part of the original proposal. (In Appendix A, we provide a narrative overview of the project by Jim Sparrow; Appendix B offers more details on the topics covered below.)

**Organizing and collecting the electronic record**

Our efforts to organize and collect the electronic record have helped make the Archive the most popular portal to September 11 on the Internet. Our “Guide to Websites” contains 350 fully annotated sites, which, though fewer than the originally proposed 400, represents the full range of exemplary sites. In addition, each of these websites was archived in collaboration with the Library of Congress’s *September11.archive.org*, and we archived many others directly. As promised, we have interviewed the producers of the best of the September 11 websites, and those interviews influenced the development of the Archive.

**Collecting digital narratives and evidence of the September 11 attack and its aftermath**

This was both the most important and the most successful part of the project plan that resulted in a collection of more than 140,000 digital items, including 45,000 personal narratives, 62,000 e-mails and electronic communications, 14,000 images, 6,000 documents, and 4,500 audio and video files as summarized in Appendix C. We were so successful that the proposed collecting structure—witness and observer stories, participant historians, emails, and images—was overwhelmed by the tens of thousands of stories, emails, documents, and images contributed from around the world. In addition to our general collection efforts, we also collaborated on projects involving many different organizations, including, for example, Library of Congress, Smithsonian Institution National Museum of American History, the Red Cross, the Sonic Memorial Project, American Association of Museums, Museum of the City of New York, the Council on American Islamic Relations, the New-York Historical Society, Columbia University, New York University, and the Museum of Chinese in the Americas. Our aggressive outreach and our increasing prominence have made us the digital repository for materials from a wide range of groups, including for example, the Service Employees International Union, the SEPT11INFO discussion group, the Madison Area Peace Coalition, Global Kid’s E.A.9.11, the Hash House Harriers, Thank You Rescuers, and the Asociación Tepeyac de New York. Along

these same lines, we are in the final stages of protracted negotiations to become the digital repository for the highly successful *Here is New York* project and website. Their collection will add several terabytes of data to the Archive, including 6,000 digital photos, 500 hours of digital video interviews, and hundreds of thousands of email messages.

### **Contextualizing and teaching about the WTC and Pentagon attacks**

This section of the original grant proposal was scaled back in the final funding because less money was available than initially anticipated. In that context, we decided that it would be more effective to collect the lesson plans of others and focus our attention on making the Archive manageable for public visitors and future researchers. We have put substantial resources into developing a sophisticated and easy-to-use researcher's interface that will be launched in September 2004, on the third anniversary of the attacks. Incorporating the Archive into secondary and college classrooms required fewer of our resources, as it happened spontaneously. Archive staff also gave presentations and workshops to numerous secondary school and college instructors and classes to broadcast the availability of the Archive as a teaching tool.

### **Developing tools and models for collecting and preserving the past electronically**

We have given considerable attention to reporting and analyzing "the lessons for historians and archivists from the events of September 11." We have made numerous presentations at public and scholarly meetings (including a major symposium at the Library of Congress) and published articles in organizational newsletters and journals. In particular, Jim Sparrow wrote an article on the lessons of the project for a forthcoming public history text (see Narrative Report), and Dan Cohen wrote an article on this topic for the journal *Rethinking History* (see Appendix D). Other publications, including possibly a pamphlet co-published by the Library of Congress and CLIR, are in the works. We have also significantly expanded the scope and ambition of the sourcebook *Doing Digital History* promised in the proposal. It has become a book-length manuscript entitled *Making Online History* that will be published by the University of Pennsylvania Press and available free on the Echo website. (We have now completed a draft of the entire manuscript, and we hope to have it to the publisher in the fall.) We have fully developed most of the tools listed in the proposal, including an expanded version of Survey Builder and the electronic clipping tool, now packaged as part of the Web Scrapbook. A few tools projected in the proposal such as the email and image archive builder have been delayed because we have folded them into the next phase of the Echo project and because we have had to devote so many programming resources to the preservation activities described below.

### **Ensuring the permanence of the Archive**

In the original proposal we only indicated that we would "complete a plan for maintaining the project website in the years and decades beyond the end of 2003." Fortunately, we have gone considerably beyond this initial goal. After much work and negotiation, we made an agreement with the Library of Congress to become the permanent repository for the Archive. The successful conclusion of that agreement has of course entailed considerable additional work on our part, particularly work processing the Archive for accession by the Library. This means we have not only secured a permanent home for the Archive, but we are also making a major contribution to digital preservation. It bodes well for the future of Sloan's vision of online collecting that the Library is using the Archive as a test bed to establish standards in the management of digital materials.

## APPENDIX A: NARRATIVE REPORT

[Published as James T. Sparrow, “On the Web: The September 11 Digital Archive,” in James B. Gardner and Peter S. LaPaglia, *Public History: Essays from the Field* (Melbourne, Fl: Krieger, 2004).]

On September 11, 2002, more than 5,000 people gathered to tell their personal accounts of the tragic day that had rocked their lives exactly a year earlier. Hundreds more shared photographs, urgent messages, and other personal effects they had saved from that fateful day. An audience of more than 100,000 joined them, avidly attending to the minutest details of the witness they bore. As early as lunchtime, the number of those stopping by just to watch video clips presented by other participants approached 10,000. At the end of the day, the total number of objects of all types viewed by visitors ran to well over a million. Once through the entrance, each visitor discovered a space with a view of the past that fit their own circumstances and meandering predilections. As they wandered from exhibit to exhibit, visitors did not so much as brush the shoulders of the more than 700 other people who attempted to join in at the very same instant, in a constant flood of humanity that persisted unabated for over 18 hours. Indeed, they may not have had any inkling of just how large an assembled audience they had joined that night. Yet no visitor could have missed the presence of so many others who were willing to share their small part of history with the public. Indeed, it was precisely this sort of sharing that had drawn them all there in the first place, whether as participants or onlookers.

The place where they crossed paths was no place at all—or rather, it was a virtual space. In the “real” world, their steps fell within the confines of a web server’s overloaded circuitry and jammed hard drive, secreted away in a stuffy closet in Fairfax, Virginia. The website they visited—and built in the process, in many cases—was the September 11 Digital Archive, an online archive built by ordinary people who chose to document their small parts of this globe-

shaking event by going online and typing or uploading the digital remnants of their personal pasts.

In the months following the anniversary, thousands of visitors would return to the Archive, some on a daily basis to see the new entries as they mounted, others just once or twice to add their voice (or their photographic eye, or witnesses' ear) to the historical record and leave their memories there to rest. As images from digital cameras, forwarded email messages, voicemail narrations, and other digital testaments arrived by wire and mail, the Archive evolved into something wonderfully unanticipated. The team of historians who brought it into being (myself included) are still trying to decide where, exactly, it fits within the spectrum of historical practice. The one thing on which we can agree is that it is clearly and inescapably a work of public history in many senses of the term—some of them very new and still protean. This chapter will explore these new valences of practice and meaning, holding up the Archive as a case study of the broader movement to embrace the online environment. The Archive's brief history suggests that the web is indeed a place where meaningful public history can take a new form, and perhaps become more democratic as it "gets wired."

### **Public History in the Digital Age**

In recent years there has been an explosion of hundreds, if not thousands of web sites aiming to democratize public access to the raw materials of the past by placing them online. These range from the humblest personal web page put up by a Civil War enthusiast, to the efforts of local historical societies and museums to develop an online presence, all the way up to the formidable online presence of national museums and historical organizations. It seems that everyone must have a web site. The ever-expanding domain of what historian Roy Rosenzweig

calls “the history web” is already too massive to catalog or even define.<sup>1</sup> Almost overnight the web has enabled a new and expansive form of public history to emerge—although precisely where and how we should draw boundaries around either the “history” or the “public” we find online may never be determined.

To the skeptical, the great rush to the web may appear to be no more than a technomystical way to pour old wine into new bottles. Or worse, it may seem to augur a dystopian nightmare of technological determinism in which the machines decide what history can and should be, forcing the rich complexity of the past and its artifactual legacy into the deracinated cultural shallows of a medium that reduces everything to electrified ones and zeroes. Those inconvenient artifacts and modes of thinking and presentation which fail to translate well into the digital realm may be obscured or destroyed outright.<sup>2</sup> Web-based history may also be especially susceptible to the uncritical “progress talk” that historian Mike Wallace has observed dominating the narratives and modes of exhibition in museums of technology. Just as such museums can reflect the class biases of the engineers and corporate managers who patronize them, so too might the digital medium be prone to a comparably technophilic self-congratulation. And while museums of technology can suffer from the “tyranny of objects, the lure of the authentic,” historical web sites may likewise privilege digital objects over interpretation and historical relationships. There is a real danger that the magic of the internet may lead us to reify digital objects, with all their slick production values, even more easily than is the case with “genuine”

---

<sup>1</sup> Roy Rosenzweig, “The Road to Xanadu: Public and Private Pathways on the History Web,” *Journal of American History* 88, no. 2 (September 2001): 548-79.

<sup>2</sup> Nicholson Baker, *Double Fold: Libraries and the Assault on Paper* (New York: Random House, 2001); Sven Birkerts, *The Gutenberg Elegies: The Fate of Reading in an Electronic Culture* (Winchester: Faber and Faber, 1994).

(physical) artifacts.<sup>3</sup> Finally, there is a fear that the web may threaten the carefully constructed layers of authenticity, context and documentation that curators, archivists and other public historians have devoted their careers to establishing and preserving for our precious and fragile historical heritage. This concern is not surprising (or unreasonable) given the “wild west” atmosphere of the web today, and the ease with which digital documents may be altered without a trace. In such a climate, experienced web travelers cannot help wondering, “what am I really looking at, and is it genuine?”

Such concerns have their merits, and we should not allow history to fall prey to the same booster mentality that recently devastated the stock market when the “dot com” bubble burst. But even if we adopt an extremely cautious and pessimistic stance toward the new technology, it is clear that the embrace of the web is not a simple matter of technological determinism. As cultural studies scholar Randy Bass has observed in the context of American Studies, changing ideas about both technology and the practice of history have heavily shaped the specific practices we have adopted to take advantage of the web.<sup>4</sup> In general, recent years have seen a growing attention to historical agency, marginal or otherwise overlooked voices, multiple perspectives, and the broader experiences of ordinary people, all of which have cast a democratic light on recent developments in information technology. The rise of popular history on the web also dovetails with an evolving interest among curators, documentarists, and other practitioners of public history to more fully involve their audiences. As public historians have democratized and rethought the nature of their discipline, the siren call of the web has become increasingly hard to ignore. To those with eyes to see, it is apparent that the web has already unleashed an enthusiasm

---

<sup>3</sup> Mike Wallace, *Mickey Mouse History and Other Essays on American Memory* (Philadelphia: Temple University Press, 1996), ch. 4, esp. 77-9.

<sup>4</sup> Randy Bass, “Garden in the Machine: the Impact of American Studies on New Technologies,” working paper, <<http://www.georgetown.edu/faculty/bassr/garden.html>>.

and participatory ethos into the public experience of history that can be harnessed to salutary ends/

The web is built for a high degree of interactivity, which can be brought into physical exhibit space through the use of terminals, or extended out in time and space through a web presence that persists beyond an exhibit. Devices promoting popular engagement in historical thinking, such as quizzes, discussion forums, and opportunities for visitor feedback, are also relatively easy to set up on a web site.<sup>5</sup> Web sites provide a degree of flexibility impossible to attain in physical space, allowing for a multiplicity of views and vantage points. Online exhibits permit contextualization of extraordinary depth, extending well beyond the few hundred words that can fit on placards and in brochures or catalogues. They can also provide multiple views of artifacts as they were created or unearthed, an effective way to expose the process of curation and preservation while also increasing the historian's power to contextualize the raw materials of history. Just as multiple voices and perspectives enrich a fictional narrative, so too do multiple media and the diverse vantages they provide amplify the analytical and evocative power of historical accounts.<sup>6</sup>

The web makes possible "the novice in the archive," as Randy Bass has observed: it has the potential to radically democratize access to historical materials, thereby decentralizing historical practice as ordinary people gain access to the record of their own past. When researchers have asked museum goers and other ordinary Americans about their involvement in the past, they have discovered that the public's interests differ from what academics, curators and

---

<sup>5</sup> O'Malley and Rosenzweig, "Brave New World or Blind Alley? American History on the World Wide Web", *Journal of American History* (June 1997): 132-55; 147. For an example of how puzzles can be used to increase historical understanding, see History Matters' valuable resource, "Puzzled By the Past," <<http://historymatters.gmu.edu/browse/puzzled/>>.

<sup>6</sup> Roy Rosenzweig, "So, What's Next for Clio: CD-ROM and Historians," *Journal of American History* 88, no. 1 (March 1995): 1621-40. As Rosenzweig observes, even the personal shallowness of a figure such as Richard Nixon can be reproduced in all its historical specificity, with audio recordings and scanned memoranda providing rich windows into his psyche.

other experts anticipated. Many crave a direct connection to the past, with as little (apparent) mediation as possible.<sup>7</sup> This popular demand for greater access to the past, unobstructed by layers of expert control or by the numbing blandishments of textbooks, intersects powerfully with the decentralizing tendencies of the internet. To envision how thoroughly this trend might alter ordinary Americans' understanding of the past, imagine what would happen if anyone could conduct a full-text, Google-like search on all government documents under the Freedom of Information Act (FOIA). When anyone can search, sort, and follow links between archival materials and digital artifacts, the hierarchical relationships between experts and laypersons are destabilized. The very basis for expert knowledge loses a powerful set of real-world foundations.<sup>8</sup>

Perhaps for this reason, historians have been slow to respond to popular passions for participating in history—especially when they emanate from the internet. Indeed, some public history sites purposely restrict the open-ended possibilities of their web pages in an attempt to sustain their control over the visitor's experience. They limit links to outside sites, and treat their own sites' content as if it were simply another caption that happened to be mounted on a web site.<sup>9</sup> Consequently, although there is a proliferation of public history on the web, many “virtual exhibits” are really little more than posters for “real” installations in museums. In such cases, creating an historical website involves no new historical work, but rather becomes simply a chore akin to advertising.

Outside the confines of professional history the situation is somewhat different, with a number of sites devoted to online commemoration and history-making. Flourishing sites devoted

---

<sup>7</sup> Roy Rosenzweig and David Thelen, *The Presence of the Past: Popular Uses of History in American Life* (New York: Columbia University Press, 1998), chs. 1, 4, and *passim*.

<sup>8</sup> Bass, “Garden in the Machine.”

<sup>9</sup> O'Malley and Rosenzweig, “Brave New World or Blind Alley?”, 147-8.

to “real-world” communities, like the one devoted to Brainerd, Kansas, or the Rowville-Lysterfield community center in Australia, craft online spaces where local people build their own historical record, contributing family histories, reminiscences, folklore, and personal artifacts such as photos and scanned documents.<sup>10</sup> “Talkback” sites for television broadcasts such as PBS’s POV site, “Re: Vietnam, Stories Since the War” feature very active newsgroup-style discussions driven by viewers who, like many Vietnam veterans, are personally invested in both the presentation of history (“getting it right”), and recording their place within it.<sup>11</sup> Online communities, including SeniorNet’s World War II Living Memorial and the History Channel’s Veteran’s Forum, host discussion “threads” with thousands of historical recollections and conversations about the past (albeit mixed in with less historical commentary).<sup>12</sup> Some new media art installations also focus on memory in ways that extend the historical record. One such installation, the Collected Visions project at NYU, solicits photographs—both on line and during inaugural exhibits—which are annotated by the contributors with explanatory excerpts about their family history.<sup>13</sup> The powerful Atomic Memories virtual exhibit at San Francisco’s Exploratorium also began as an installation tied to a Web forum, juxtaposing personal memories of how visitors first learned about the bombing of Hiroshima and Nagasaki with reflections on the meaning of nuclear warfare.<sup>14</sup>

---

<sup>10</sup> Brainerd, Kansas: Time, Place and Memory on the Prairie Plains <<http://www.rootinaround.com/brainerd/>>, Rowville-Lysterfield History Project <<http://www.rlcnews.org.au/>>.

<sup>11</sup> <<http://www.pbs.org/pov/stories/vietnam/story.html>>.

<sup>12</sup> SeniorNet’s World War II Living Memorial <<http://www.seniornet.org/ww2/>>, HistoryChannel.com’s Veteran’s Forum <<http://www.veteransforum.historychannel.com/>>. As of April, 2001, the Veteran’s Forum contained nearly 20,000 postings, while SeniorNet’s World War II Memories Discussions section held over 6,000 entries. Veterans appear to be especially active online practitioners: e.g., see also the Atomic Veterans History Project <<http://www.aracnet.com/~pdxavets/>> and the Drop Zone Virtual Museum Oral History Project <<http://www.thedropzone.org/misc/mission.html>>.

<sup>13</sup> Collected Visions Gallery <<http://www.cvisions.cat.nyu.edu/gallery/index.html>>.

<sup>14</sup> Remembering Nagasaki: Atomic Memories <<http://www.exploratorium.edu/nagasaki/memories/amemory.html>>.

The following case study of the September 11 Digital Archive provides a glimpse into the workings of a project that, in a similarly participatory vein, seeks to take full advantage of the flexibility and interactivity of the web. Wedding popular participation to scholarly practice, the Archive aims to bring the public closer to the center of public history, thereby allowing ordinary Americans to literally make their own history.

### **The September 11 Digital Archive <<http://911digitalarchive.org/>>**

In the late summer of 2001, the Center for History and New Media (CHNM) at George Mason University was uniquely positioned to respond to the historical challenge of documenting popular experiences of the disaster about to unfold. Since the early 1990s, it had been a leader in the emerging field of computing and the humanities, producing CD-ROM textbooks, websites and other projects that sought to move beyond traditional historical practice. In 1998, I joined the Center with an experimental pilot, the Blackout History Project, which sought to use the web not simply for historical presentation, but to develop a methodology to collect, preserve and creating new, digital types of primary historical materials via the internet.<sup>15</sup> One of a score of projects funded by the Alfred P. Sloan Foundation between 1997 and 2000 to improve the preservation and documentation of the recent history of science and technology,<sup>16</sup> the Blackout project began from a basic premise. Modern life had “gone digital,” leaving traces of the past on hard disks and in other electronic storage media; it was time for history to catch up. Since January of 2001 a second, broader CHNM project named ECHO (“Exploring and Collecting History Online”) has continued to experiment with new methods in online collecting, producing a series of small

---

<sup>15</sup> <<http://blackout.gmu.edu/>>.

<sup>16</sup> The Blackout project was part of the first wave of Sloan projects, known as STIM: Science and Technology in the Making <<http://sloan.stanford.edu/>>. For a list of the other Sloan projects, consult the links in the right-hand sidebar of the ECHO Virtual Center <<http://echo.gmu.edu/center/>>.

projects on the recent history of science and technology. ECHO seeks to learn from its projects, the other collecting projects funded by Sloan, and from other early ventures in this new field, distilling their lessons into a set of best practices for online research with historical subjects. After five years of experimentation, we now have a collection of software tools, techniques, and case studies to guide historians in the ways that the web and other new media may be used to extend and enrich the historical record.<sup>17</sup>

As shock gave way to reflection in the days and weeks following the airplane attacks on the World Trade Center and the Pentagon on September 11, 2001, it became clear that ECHO's approach to documenting contemporary history was especially well-suited to the task of capturing a broad canvas of Americans' experiences for the historical record of "9/11" (as it has come to be known). We viewed this challenge with a mixture of hope and trepidation. Here was an event that demanded extraordinary delicacy, considering the sacrosanct wreath of commemoration which most Americans placed around the tragedy and its victims. Like the curators at the National Museum of American History, the Museum of the City of New York, and the other museums that banded together to respond to 9/11, we faced the dilemma of how to handle the collection of contemporary historical materials when popular feelings ran so high, with little emotional or temporal distance to lend perspective. Would our professional obligation to maintain a critical stance toward historical materials make us appear inappropriately aloof, thus driving away potential contributors? Or would emotionally gripping stories and artifacts,

---

<sup>17</sup> <<http://echo.gmu.edu/>>. Online collecting projects can be found in the Memory Bank <<http://echo.gmu.edu/memory/>>. For a précis of best practices, consult the Practical Guide <<http://echo.gmu.edu/guide/>>. The Guide will be expanded and accompanied by a forthcoming book authored by Roy Rosenzweig, the Director of CHNM, and Dan Cohen and Tom Scheinfeldt, the other co-directors of ECHO. For an effort that coincided with and shared many of the broad objectives of ECHO, see the History of Recent Science and Technology <<http://hrst.mit.edu/>>.

such as missing persons' posters, overshadow other important but less dramatic aspects?<sup>18</sup>

Despite these worries, we felt a great concern that something valuable would be lost if we didn't act quickly. Memories would fade, email would be deleted from inboxes, faxes and word processed diary entries would vanish in the ordinary cycle of hard drive crashes and software upgrades. The digital substratum that has come to line the foundations of daily life would corrode and slip away due to the relative fragility of magnetic and other types of electronic storage media.

Turning to historical precedent, we inquired whether an event of comparable magnitude had been preserved in fine-grained detail at the grassroots. Immediately following Pearl Harbor, government-employed social scientists had rushed out to capture the reactions of ordinary Americans in quick polls and "man on the street" interviews.<sup>19</sup> Although we lacked the funding or the staff of these efforts, we enjoyed the advantage of living in a time when millions of Americans self-document their experiences and views by composing weblogs (online diaries), sending email, joining chat sessions (real-time discussion forums), and otherwise bending the personal computer to their own devices. We saw an opportunity to use the web tools and methods we had already developed for ECHO to allow hundreds, perhaps thousands, to contribute their stories, email, images, voicemail, chat sessions, pager messages, home movies, digital artwork, word processed diaries and letters, and other electronic documents to a permanent historical record. Our web server could accept scores of thousands of entries and visitors almost as easily as it could accommodate a handful.

---

<sup>18</sup> James B. Gardner and Sarah M. Henry, "September 11 and the Mourning After: Reflections on Collecting and Interpreting the History of Tragedy," *The Public Historian* 24, no. 3 (Summer 2002): 37-52; 39-42, 50.

<sup>19</sup> Folklorist Alan Lomax, "assistant in charge" of the Archive of American Folksong in the Library of Congress when Pearl Harbor occurred, instructed ten fieldworkers (by telegram) to interview Americans on their reactions to the event the day after it happened. The tape recordings of those interviews are available in the Library of Congress, as described at <<http://memory.loc.gov/ammem/afcphtml/>>. Pollster Rensis Likert also conducted an opinion survey on the night of 12/7/1943. See Jean Converse, *Survey Research in the United States: Roots and Emergence, 1890-1960* (Berkeley: University of California Press, 1987), 160.

The September 11 Digital Archive began, fittingly enough, in an act of partnership. As with previous collecting efforts, it was through partnership that the most effective outreach was possible. Realizing the need for a broad foundation of historical expertise based in the two major metro areas affected by the attacks, the ECHO team and others at CHNM joined forces with a long-time partner, the American Social History Project (ASHP) at the City University of New York (CUNY) Graduate Center.<sup>20</sup> The ASHP had deep roots, both academic and otherwise, in the affected communities of New York City, as well as considerable expertise in history and new media. Armed with a fresh grant from the Sloan Foundation, which saw both a good cause and a chance to put its vision of online collecting to the test, we got to work in late December of 2001.

Our first challenges were practical and organizational. We had to quickly design a site that would collect the broadest possible range of digital materials from an almost infinitely wide spectrum of contributors, while avoiding the opposing dangers of either censoring or appearing to condone the extreme views that such a charged topic inevitably would attract. A related challenge involved concentrating our limited resources and staff time on the potentially unbounded task of outreach and recruiting, since this was an event that literally the whole world had witnessed in one form or another.

Right away we set up teams to handle critical tasks, including: database and interface programming; web design for look, feel and accessibility; outreach targeted at specific communities near the crash sites in New York City, Washington DC, and Shanksville, Pennsylvania; public relations aimed at local, regional and national media outlets; editors and vetting reviewers to provide quality control and legal protection for materials displayed on the

---

<sup>20</sup> Collaborations include History Matters <<http://historymatters.gmu.edu/>>, a resource for U.S. history teachers and students, and *Who Built America? From the Centennial Celebration of 1876 to the Great War of 1914*, and *Who Built America? From the Great War of 1914 to the Dawn of the Atomic Age in 1946*, 2 vols. CD-ROM (New York: Worth, 2000).

site; and partnerships with key organizations and institutions. Every few months or so we met in person at CUNY or GMU to firm up those intangible but essential connections that make teamwork possible across multiple sites. In between visits, we collaborated via email, phone and videoconference.

The first task, building the website, was easier than might otherwise have been the case thanks to our shelf of existing code and templates from ECHO. From our earlier efforts, and from the literature on web usability, we knew that web users' greatest frustrations centered around speed and ease of use, rather than the aesthetics of design, and that these concerns were driven by an overriding desire to locate some specific item or piece of information, rather than browse aimlessly.<sup>21</sup> Within a few weeks we were able to implement an interface in PHP, tied to a back-end MySQL database, that was very quick to respond to users' requests thanks to the efficiency of that particular software configuration. A spare home page featured our logo, some introductory text, and a navigation bar along the left side of the screen that allowed visitors to go directly to what interested them (stories, images, email, etc.). When the visitor clicked on a section, or paged through listings of collected objects within sections, they were constantly given clear navigational indicators from the sidebar (which had a consistent and almost rigidly hierarchical appearance on all pages), to help them orient themselves within the site. Throughout the site we took pains to employ the elements of user-centered web design: consistent (if spare) design elements such as font color and size, background color, logo appearance and placement; careful use of many small tables for layout (rather than larger tables with complex structures); an explicit (sometimes blatant) set of textual signposts; and search engines which allowed the

---

<sup>21</sup> See, e.g., GVU 10<sup>th</sup> Survey, Group G29, "Problems Using the Web," at <[http://www.gvu.gatech.edu/user\\_surveys/survey-1998-10/graphs/use/q11.htm](http://www.gvu.gatech.edu/user_surveys/survey-1998-10/graphs/use/q11.htm)>. Although bandwidth has improved for many web users since the survey was taken in 1998, speed is still a major consideration even for those enjoying fast connections, while clear, direct navigation has become even more prized as the web grows larger and the need to cut through the thicket of information increases apace.

impatient (or the simply curious) to bypass our carefully crafted navigational structure altogether.

In addition to clarity we added simplicity, which served both to assist our visitors and to reinforce the sense of place and permanence provided by the Archive. The domain name, 911digitalarchive.org, was preferable to a longer url such as chnm.gmu.edu/911 because it was easy to remember and immediately conveyed to our potential contributors who they were dealing with. Instead of the long, complex query strings that often fill the address bar at database-driven sites, making it hard for visitors to bookmark the url or even decipher it as a navigational clue, we used a configurable feature of our apache web server, “url rewriting,” to transform the appearance of queries into simple, logical directories. So, a visitor looking at the 847<sup>th</sup> email contributed to the site would see the url path /email/details/847 that clearly indicated where s/he was, and what s/he was looking at.<sup>22</sup> While this may seem a picayune detail, it is precisely the sort of usability feature that enables visitors to orient themselves. By clarifying the underlying function of the site, the url provides a definitive and unambiguous declaration that has the added advantage of suggesting that others who contribute will be adding something that will have a lasting location and identity, rather than being absorbed into an arbitrary morass of code that might change from one site revision to the next.

The use of images posed a special dilemma for us, both during the design and once visitors began uploading their own photographs and artwork. Image files can be used to create precise design elements that will not vary from browser to browser. This offers a source of control dearly sought by designers, who often feel vexed by the impossible variety of browsers, operating systems and hardware platforms on which their sites eventually may appear.

---

<sup>22</sup> i.e., the full url would be <http://911digitalarchive.org/email/details/847>.

Unfortunately, even the smallest and most unobtrusive of images can consume more bandwidth than a full page of simple text. Since visitors generally crave information, there is a corresponding need to minimize the ratio of images to text. However, effective design is nearly impossible without using some carefully selected images. For an event that was experienced by most as an awful vision of fire and destruction forever seared into the mind's eye, it was inadvisable to severely restrict the display of images—especially once contributors began uploading their own images to the Archive!

Our solution was to keep most images off the home page and out of other sections not directly concerned with visual materials. On those pages that did feature images, such as the extremely popular still images section which received more than 2,000 uploads by April of 2003, we allowed visitors to browse through miniatures, or “thumbnails” of contributions, so that when they viewed the full-size version it would be at their discretion, and not something that was forced on them by our design. After using the thumbnail-generating software for a while, however, we discovered that under rare circumstances it generated images that would crash certain browsers in subtly maddening ways.<sup>23</sup> This was just one example of the extreme difficulty of controlling the behavior of our site on others' computers. Our visitors might browse the site with extremely high or low screen resolutions, fast or slow processors and graphics cards, millions of colors or just a handful, images enabled or turned off by default, to mention only a few of the variables. Rather than trying to control every last detail of visual display through extensive use of graphics and tables, we opted for a flexible, simple design that would look elegant for most visitors and “degrade gracefully” in sub-optimal conditions.

---

<sup>23</sup> e.g., the browser would refuse to load images, “hanging” for no apparent reason but still allowing the display of new text.

Before opening the site to visitors and contributors, we also had to prepare a set of basic procedures and protocols. These were needed to ensure the substantive and legal integrity of our collections, protect the privacy and wishes of our contributors without reducing our collections (or the historical record) to their own personal billboard, and allow us to exercise a general level of quality control without enslaving ourselves to our own standards. The lack of authentication on the internet, and its fundamental statelessness, worked against these goals. How could we know that our contributors were who they claimed to be, and that what they contributed was genuine and belonged to them? A dark specter came to haunt our discussions, personified as an imaginary 13-year-old mischief maker residing in New Jersey and luxuriating in the late-night license afforded by his parents' cable modem hookup. The human subjects protocol required by CUNY's review board prohibited contributions from minors, while fear of violating our most dearly held axioms of professional ethics provided the other specifics of the phantasm.

To address these and related concerns, we adopted some common-sense practices that greatly reduced the chance of problem entries. Our software permitted us to record the unique IP address of our contributors, who were also required to provide a name, email address and age, before their contributions would be accepted. Immediately upon receiving an entry our software sent an email message to the address provided, thus sending the contributor an official record of his or her entry, while informing us of invented or expired addresses when these messages bounced back to us. Then, before the entry was added to the official Archive collections and displayed publicly on the site, we vetted each entry. Among other things, we checked the contributor's age, tried to establish that the contributed item was not clearly something produced by someone else (e.g. a news agency), and quickly confirmed that it was relevant to September 11 and not wildly inappropriate in its use of language or imagery. After an entry was fully

processed, it remained in the collections exactly as it had been entered, and was displayed on the website according to the contributor's specifications. In a few cases we did make changes or corrections for individuals who later made the case that their entries should be suppressed or altered in some acceptable manner.

Although these procedures would prove time consuming (particularly when thousands of entries flooded in overnight on the one-year anniversary), they helped us to maintain some basic level of control over the content. The need to exercise this sort of control was especially important given the highly charged climate after 9/11, and the inescapably sensitive nature of the contributions that would come in. In some infrequent cases where we realized that a photo was too professionally composed, or too familiar, to have been produced by the contributor, we withheld it from the public site for legal reasons. We could not conduct a copyright check on each image uploaded (even if such a thing were possible), but rather had to use our judgment as to what was an original. Outside of these limited scenarios where contributions were legally or ethically impermissible, the vetting team made every effort to avoid suppressing an entry unless it was absolutely required, our own personal qualms about the contributor's sensibilities notwithstanding.

Despite the pains we took to craft an effective web site and establish quality control procedures, the entire project would have been an exercise in entropy if we had failed to attract contributing visitors. This was no small consideration, as our previous experience with ECHO had shown. With the site in place (although lacking contributions) by January of 2002, we spent three months beta-testing our design and accumulating a critical minimal mass of entries by contacting friends, family, colleagues and acquaintances. We hoped these contacts would

snowball into further entries, although after a hundred or so entries the contacts had been exhausted.

Then, for the six-month anniversary in March, we worked with public relations professionals from CUNY, GMU and the private sector to develop a basic media plan targeting regional media coverage in the New York and DC metro areas. This plan involved considerable planning to get coverage for a press release, and to attract interviews with team members. Human interest stories ran in the *Wall Street Journal* (thanks to a little serendipity and our canvassing the financial district of Manhattan with specially made cards) and in the local county inserts for the metro section of the *Washington Post*. Other outlets picked up the story for the computer angle—including *InformationWeek*, a widely read journal for the information technology industry. A regional distribution network for news radio stations in Virginia also conducted a brief informational interview, thus reaching a broad and active segment of the public we might have missed if we had focused exclusively on internet audiences.

Due to the high profile and wide circulation of the media outlets in which the Archive was featured, this coverage would have driven traffic to the site no matter what the content. But we needed contributions, not just traffic (although we hoped the former would flow from the latter eventually). Accordingly, we focused our interviews and the wording of our press release very carefully to appeal to the sensibilities of the communities that had been most directly affected by the disaster. Paying due attention to communities close at hand also helped, since both CUNY and GMU had student populations and surrounding neighborhoods with the potential to yield thousands of contributors, located as they were near the two largest crash sites. In the weeks following the six month commemoration our traffic jumped significantly, and with it the contributions rose by a few hundred. Later in the summer of 2002, when our efforts had

resulted in a broader range of partnerships with allied institutions, our exposure in the media yielded a dramatic rise in the number and quality of our contributions. Features which ran on CNN (both the broadcast and the website), in the Associated Press newswire, and on New York City local television, were followed by surges of contributions that immediately ran into the low thousands by the eve of the first anniversary of the attacks.

At the same time that we pursued media coverage, we targeted specific groups and communities in the three crash sites whose contributions we felt were important but potentially hard to attract without special efforts. Chinatown, in New York City, had suffered an enormous blow due to its proximity to the World Trade Center. An outreach team at CUNY worked closely with community leaders there to ensure that Chinese-speakers and others would learn of the site, while also conducting interviews to increase the Archive's relevance for residents of that neighborhood. For similar reasons, we also provided a Spanish-language version of the site to reach the large Latino communities in Washington, DC and New York City. Some of our other outreach efforts focused on Arab high school students in the suburbs of Washington; undocumented immigrant workers and unionized American workers whose jobs were located near the crash sites in Washington and New York; residents of the community near the crash site in Somerset County, Pennsylvania; and volunteer medical workers. These outreach efforts were not guaranteed to produce more contributions directly to our site, but their cumulative effect was to establish a presence in important communities that were likely to fall beneath the radar of media coverage, and thus remain absent from the historical record (as has so often been the case with poor, marginalized or transient communities).

While national media coverage and local community outreach were critical in prodding contributors to help build the Archive, they were not, by themselves, sufficient to sustain a

momentum of ongoing contribution after the publicity had faded and the funding for community involvement dried up. The critical factor in getting contributions to sustain themselves over time was our good fortune in establishing strategic partnerships with institutions whom our potential contributors both trusted and visited or interacted with in some fashion. Early in the project we had been animated by the vision of a vast, popular base of “participant historians” who would write their own history. We did eventually attract visitors whose self-documentation approximated this description, but only by identifying and working with what might best be called “participant-centered historical institutions.” Our efforts to make the Archive visible and accessible benefited in myriad intangible ways from small exhibits or other modest collaborations with the Museum of the City of New York (MCNY), the Columbia Oral History Project, the amateur historians’ collective DocumentNY, the Middle East and Middle East American Center (MEMEAC), the public libraries of Somerset County, Pennsylvania, the Women in Military Service for America Memorial in Arlington, Virginia, and select public school systems in the affected communities. These efforts also served to connect the Archive with local communities of interest in concrete and persuasive ways. Slowly these efforts helped to transform what to most people was no more than a name and a url into an enterprise whose value they could understand, relate to, and eventually engage in a highly personal fashion.

If there was one catalyst that could be isolated as the key to transforming the Archive into a large-scale participatory exercise in popular history-making that now continues to grow of its own momentum, it would have to be our partnership with the National Museum of American History, and the carefully focused publicity surrounding it. The NMAH’s exhibit on 9/11, “Bearing Witness,” opened to wide acclaim on the first anniversary of the tragedy—although even before it opened there was a recognition that this would be a major and highly successful

installation.<sup>24</sup> In the spring and summer before the opening we worked closely with the NMAH to build the a website that would allow visitors to the museum, and others arriving only by the internet, to record their responses to the exhibit, register their reflections on their lives during the year since the attacks, and share their memories of the fateful day, all by adding entries to a specially designed portal to the Archive.<sup>25</sup>

Two important features of the exhibit were designed to encourage visitors to the museum to literally bear witness. In the last room, visitors were invited to reflect on the installation, and on 9/11 generally, by recording their thoughts either in writing, with pencils on index cards, or orally, by telling their story in telephone booths where a voicemail system could record their accounts. The voicemail system, which could also be accessed from outside the museum by dialing a toll-free number, was a particularly promising innovation, since many people who would shy away from recording their experiences on paper or on the web might find the conversational format of a voicemail message less intimidating and more appropriate to the intimate quality of their personal recollections. Voicemail messages held the added advantage that they could be converted on the fly into sound files and sent, as email attachments, to the Archive. Others might prefer the reflective mode encouraged by putting pencil to paper, so we arranged to have the cards scanned in bulk and later imported into the Archive's database.

The direct connection at the exhibit, reinforced by the accessibility of the website and toll-free telephone numbers, proved to be a powerful combination. On September 11, 2002, more than 3,500 new stories came to the Archive through the Bearing Witness portal alone. In the first week of the exhibit, the number of first-person narratives in the non-NMAH portion of the Archive also increased dramatically, jumping from around 2,100 to just under 8,000. Stories

---

<sup>24</sup> For a fuller description of this exhibit, see Gardner and Henry, "September 11 and the Mourning After."

<sup>25</sup> <<http://911digitalarchive.org/smithsonian/>>; linked from  
<<http://americanhistory.si.edu/september11/tellyourstory/index.asp>>.

were not the only things visitors added to the Archive. Visitors uploaded hundreds of images and email messages from their digital cameras and inboxes in the same week. And the contributions continued long after the anniversary. In the four months following the exhibit's opening, both the Bearing Witness and the regular Archive story banks received almost 2,000 additional entries apiece, while more than 1,000 voicemail messages and over 4,000 index cards arrived for inclusion in the Archive. (It should be added that old media has played a valuable role, too; thousands of photographs, email messages and other electronic files have arrived on CD-ROM in the mail, providing the vast bulk of our collections, which now number over 140,000 objects.) Visitors to the September 11 Digital Archive continue to make contributions to this day, long after the most active outreach attempts have concluded.

We are now in the process of securing long-term preservation and archival custody for our collections. In September of 2003 the Library of Congress will accession the Archive as one of its first major digital acquisitions. In preparation for that day, we are working to inventory and catalog our collections using open standards such as XML to ensure that future researchers will be able to access what our contributors have preserved. In this respect, our web design choices have provided yet another benefit. All objects are assigned a permanent object identifier (reflected in the url), which allows for stable and precise references, regardless of how those objects will be made available in the future. For similar reasons, the software we have used to store and present these objects is all "open-source," which means that future programmers should stand a decent chance of reproducing our server environment long after specifications for operating system, file structure, and other crucial functionality have evolved into unrecognizable formats. (Open source software is also free or very inexpensive, which allowed us to devote more resources to recruiting contributors. Hopefully it will also allow future archivists to work

with our materials without encountering the costly encumbrances of intellectual property law.)<sup>26</sup>

Saving these materials for the future thus presents yet another argument for keeping everything as direct and simple as possible—whether we’re thinking about middleware code, HTML and web design, electronic file formats for storing contributions (e.g. jpg, pdf, doc), navigational hierarchy, or even the internal documentation and formatting procedures we hide within the code that runs the site.

While long-term archiving and preservation of digital media are tasks which still pose considerable challenges to digital librarians, we are committed to ensuring that the Archive will endure. After all, we have an ethical obligation to our contributors that lies at the core of our guiding vision of the Archive; namely, that their good-faith efforts to help make a truly public history of 9/11 will live on in a permanent and accessible digital historical record.

## **Conclusion**

The September 11 Digital Archive has been strikingly dependent on the needs of its visitors, and on its ability to foster a sense of community online. In this regard, it resembles efforts by museums in recent decades to reach out more effectively to the audiences and communities they serve. As Andrea Hauenschild has written of the “new museology” that has begun to shape local and community-centered museums around the world, “what produces acceptance of the museum is its usefulness as recognized by the population in question.” Indeed, this requirement is even more fundamental on the web, where community is especially tenuous.

And unlike the “new” community-centered museums such as the Museum of Chinese in the

---

<sup>26</sup> On the challenge that changes in file formats and software pose to future historians, see Roy Rosenzweig, “Can We Save the Present for the Future?”, *American Historical Review* 108, no. 3 (June 2003): 735-62, esp. 741-2, 744-8. For a more general treatment of the pressing need to preserve ideas and other aspects of cultural heritage in open, non-proprietary formats, see Lawrence Lessig, *The Future of Ideas: The Fate of the Commons in a Connected World* (New York: Random House, 2001).

Americas centered on New York City's Chinatown, our sites, like the internet in general, have not generally been able to reach hard-to-reach and underserved communities.<sup>27</sup>

The key problematic for online historians is the lack of a sense of place, identity and authenticity on the web—something all web authors try very consciously to overcome through “branding” to foster a sense of community through shared design. Where are you, exactly, when you are visiting a web site? This question, which all but the most naïve web users must ask themselves at one time or another, undercuts the authenticity, authority and sense of permanence that even the most expertly crafted web site must take great pains to establish. Another challenge to authenticity is a byproduct of one of the internet's great strengths. The web allows us to “explode the narrative,” taking advantage of its “convergence of distribution” to enable a “polyvocal” history that captures all voices and perspectives without privileging any one over the other.<sup>28</sup> This gives rise to what might be called the “amateur perplex,” in which the enthusiasm and untrained bias of amateurs can foster historical knowledge and a zeal for primary materials, at the same time that it gives free rein to prejudices, distortions, falsehoods and inaccuracies (if not worse). As the multitude of Civil War sites built by enthusiasts and re-enactors attests, the populist energy unleashed by the web can cut both ways.<sup>29</sup>

To be credible, sites must either be tied to trusted real-world institutions, or go through the long and demanding process of establishing themselves as trustworthy. Here is an opportunity for museums and other institutions of public history to adopt a special role in the emergence of online history. In *The Presence of the Past*, Rosenzweig and Thelen found that museums were considered by most people to be the most trustworthy of institutional sources for

---

<sup>27</sup> <<http://www.moca-nyc.org/>>; Andrea Hauenschild, “Claims and Reality of the New Museology: Case Studies in Canada, the United States and Mexico” (Ph.D. dissertation in Ethnology, Hamburg University, 1988), section 4.1, para. 4-6. Available online at <<http://museumstudies.si.edu/claims2000.htm>>.

<sup>28</sup> Bass, “Garden in the Machine.”

<sup>29</sup> Rosenzweig, “Road to Xanadu,” para. 17, 23, 27.

historical information (according to their respondents), holding a special trust due to the immediacy of the past represented by the artifacts they hold.<sup>30</sup> Museums must work to make the most of this advantage, using it as leverage to transmit their high standards of historical practice to online practitioners, whether experts or amateurs. A window of opportunity for exercising this influence remains open to public historians so long as the website continues to evolve and gain definition as a genre of historical writing, presentation, and preservation. It is not clear how long the opportunity will remain open, however, given the extremely short time it takes for technologies and trends to “lock in” on the web.

---

<sup>30</sup> Rosenzweig and Thelen, *The Presence of the Past*, 20-21 (Table 1.3), 105-8, 235 (Table 2), 244-7 (Tables 9-12).

## **APPENDIX B: DETAILED REPORT**

The four main components of the original *September 11 Digital Archive* proposal – organizing and collecting the electronic record; collecting digital narratives and evidence of the September 11 attack and its aftermath; contextualizing and teaching about the WTC and Pentagon attacks; and developing tools and models for collecting and preserving the past electronically – have all been successfully accomplished. As with any highly interactive and highly public technology project, the final results do not precisely match our original intentions. The Archive greatly exceeded its initial collection targets in terms of the sheer amount of information collected, and now includes a much greater diversity of information than was outlined in the original proposal. Furthermore, over the course of the project, the Archive evolved into more than just a stand-alone operation, benefiting from close collaboration with many individuals, organizations, and institutions. Finally, as the Archive grew to include tens of thousands of objects, the management and long-term preservation of its 12 gigabytes of data became an especially pressing and challenging concern, necessitating the addition of a fifth major component to the Archive’s mission. This additional work is discussed here in the final section of this report, “Managing and ensuring the permanence of the Archive.”

### **Organizing and collecting the electronic record**

Organizing the electronic record involved surveying online materials and creating a portal to the best websites related to September 11. Our guide to more than 350 websites is fully annotated, contextualized, and searchable. While this number is lower than the 400 initially proposed, it is only because there were found to be fewer websites that merited annotation. In addition to the more general survey, Archive staff conducted interviews with the creators of ten of the best sites. These interviews influenced further planning of the Archive, and the results of this research will be included in a forthcoming CLIR pamphlet. Finally, in preparation for the conclusion of the official grant period, all annotations were fully updated as of June 1, 2004.

Collecting the electronic record was accomplished in collaboration with the Library of Congress’s *September11.archive.org* project. Though not suggested in the original project plan, we arranged for the Library and its partners at the Internet Archive to “spider” our entire Guide to Websites, allowing us to focus our efforts and reserve our server space for other aspects of the electronic record. Blogs, in particular, were shown to be especially susceptible to disappearing, and the Archive collected and preserved more than thirty of the best of these fragile websites.

### **Collecting digital narratives and evidence of the September 11 attack and its aftermath**

Collecting was intended to be the largest component of the project, and it has been an overwhelming success, resulting in an archive of more than 140,000 digital objects in more than 10 categories and more than 50 file formats. All text files are fully searchable as well as available for browsing, making the Archive an extremely useful and usable resource for researchers. The broad appeal of the Archive’s collections is evidenced by its popularity. To date its website has received more than 2,000,000 visitors and nearly 120,000,000 hits from more than 3,500 different servers. Feature articles on the Archive have been published or broadcast by the

Associated Press, Reuters, the *New York Times*, the *Wall Street Journal*, the *Washington Post*, the BBC, CNN and MSNBC, and Google, Yahoo!, and MSN all rank the Archive first in searches on the term “September 11.”

The original project plan divided our collecting efforts into five categories—witness and observer stories, email, images, and training participant historians. As the project grew, however, these initial categorizations proved to be incongruent with a rapidly expanding collection. For example, with ubiquitous images of the attacks and twenty-four hour news and media coverage, the distinction between “witness” and “observer” became unclear, as many people both “witnessed” and “observed” the events. Likewise, the original proposal envisioned a series of workshops to train participant historians in online collecting. Yet after surveying the full range activities of other mission-aligned organizations, partnerships with these organizations emerged as a better way to involve others in the work of the Archive. Among the Archive’s collecting partners are large organizations such as the New York Historical Society, the Smithsonian National Museum of American History, the Library of Congress, the American Association of Museums, the Service Employees International Union, and the American Red Cross, as well as smaller institutions and private collectors including Document NY, an independent online documentary project; “Everything After September 11,” a children’s website; Michael Ragsdale, a participant historian in N YC, whose flyer collection has become a major asset to the Archive; and *Wherewereyou.org*, a website created by college students just days after September 11 to collect stories of individuals’ experiences of the attacks. For these reasons, the following accounting will speak in terms ultimately more representative of the Archive as it stands today, dividing its collecting work into five new categories – *personal narratives*, *electronic communications*, *images*, *documents*, and *audio/video* collections.

Personal narratives. More than 45,000 personal narratives have been received from all quarters and by many different methods of outreach, no fewer than 16,000 of them directly through the Archive’s online web interface. Though an initial critical mass of contributions was provided by Archive staff, colleagues, students, friends, and family, and though early efforts at digital outreach via e-mail, bulletin board postings, and chat room conversations yielded a modest number of narratives, most of the personal narratives in the Archive were obtained as a result of direct community outreach efforts and institutional collaborations.

The bulk of the Archive’s direct outreach was conducted in New York City, with a focus on groups that were less likely to be included in the historical record. These efforts included posters in community centers, community newsletter articles, radio interviews, toll-free telephone numbers, and in-person collecting. The resulting collections are largely comprised of recorded and transcribed interviews from victims, eyewitnesses, and observers. Among the many groups with which we worked are Asociacion Tepeyac de NY; the Puerto Rican Family Institute; the Institute for Puerto Rican Elderly; Latino political leaders in the World Trade Center area; Muslim and Arab Americans through CUNY’s Middle East and Middle Eastern American Center (MEMEAC); minority and at-risk populations of students in the Bronx; the Service Employees International Union; Document NY; and Cornell University’s extension program for non-managerial white-collar workers, many of whom were in the WTC at the time of the attacks. Particularly notable are our ongoing efforts in cooperation with the Rockefeller Foundation, Columbia University, the Museum of Chinese in the Americas, and New York University’s

Asia/Pacific/American Studies Institute to collect personal narratives from members of Manhattan's Chinatown community and present them in a variety of languages over the internet as a resource for the community's defining of its future in the coming redevelopment efforts.

Outside of New York City, direct outreach efforts targeted Somerset County, Pennsylvania, Flight 93 victims' families, and Pentagon workers, victims' families, and rescue workers. Specifically, outreach in Somerset County included travel to the temporary memorial in Shanksville; setting up kiosks in the local libraries; working with local emergency services; contacting area photographers; publishing announcements in local papers; and giving presentations to local library science students and heritage professionals. In Northern Virginia, outreach included announcements in the *Pentagon Workforce News*; a partnership with Operation Solace, a victims' families' support group; contact with a Pentagon counseling team; direct appeals to victims' families; announcements in the Association of Flight Attendants' newsletter; work with capital-area musicians; presentations to Northern Virginia school teachers; and on-campus outreach to George Mason University students. Finally, outreach to Northern Virginia fire and rescue teams produced a significant number of images and stories. In most cases, however, institutional collaborations proved more effective than direct outreach in reaching groups in suburban Washington. These partnerships included the Library of Congress, the American Red Cross, and most notably, the Smithsonian Institution's National Museum of American History, which helped us collect more than 6,500 stories through our website and more than 25,000 fully digitized and indexed response cards from visitors to the Museum's "September 11: Bearing Witness" exhibition.

Electronic communications. The same outreach efforts and collaborations that produced the Archive's rich collection of personal narratives also produced other kinds of materials, most notably electronic communications. The Archive's collections of e-mails, instant messages, and other forms of electronic communications fall into two categories: individual e-mails submitted through the web interface, and larger collections contributed by groups or institutions. More than 1,000 e-mails have been contributed directly through the Archive's web interface, and more than 60,000 more have been obtained through direct personal and institutional outreach efforts. Among the larger collections of electronic communications are 12,000 e-mails from the Department of Justice, 3,500 postings to the SEPT11INFO discussion group, and the entirety of Global Kid's "Everything After September 11" collection of more than 3,000 e-mails from kids from around the world. Finally, we are also very hopeful that our long negotiations to become the digital repository for the hundreds of thousands of e-mails received by the *Here is New York: A Democracy in Photographs* project will be concluded shortly.

Images. The Archive's more than 14,000 digital images likewise fall into two categories. In the first category are more than 2,500 annotated photographs and pieces of digital folk art submitted through the website by nearly as many individual contributors. The second category contains dozens of small collections from individual contributors and several large institutional collections of many hundreds of images. Prominent among these are nearly 3,000 images contributed by the Army National Guard and more than 3,000 from Mike Garcia, a private collector. *Here is New York* should also yield more than 6,000 digital photographs taken by professional and amateur photographers in the coming months.

Documents. Though unanticipated in the original proposal, the more than 6,000 documents and reports collected by the Archive comprise some of its most interesting collections. Among these are such important institutional collections as the New York Fire Department's nearly 200 daily Incident Action Plans for the cleanup of Ground Zero and the National Guard Bureau's *Comprehensive Reference Resource* of more than 3000 documents. The Archive also contains some very personal records, such as a collection of nearly 1,000 flyers posted around New York in the immediate aftermath of the attacks documenting the search for missing loved ones, protests against the INS detentions, expressions of grief, and calls for retribution. Indeed the range of contributors to the Archive's document collections spans the cultural and political spectrum from peace activist organizations to hawkish groups like *deadarab.com* to satirists like the National Lampoon.

Audio/Video. Audio and video files form the final major category of material collected by the Archive. Most notable among the more than 4,000 audio files in the Archive are those obtained through our partnership with the Sonic Memorial Project, a collection that includes approximately 900 digital recordings of the life of the World Trade Center, its neighborhood, and the events of September 11. Another important audio collection stemmed from our collaboration with the Smithsonian's "Bearing Witness" exhibition. In cooperation with museum staff, the Archive installed two toll-free telephone lines in the exhibition gallery and encouraged visitors to leave two- to three-minute voice messages describing their experience of September 11. These messages were then e-mailed directly to our servers and archived as digital audio files, resulting in a collection of approximately 3,000 voicemails. Film, video, and Flash animation collections include more than 300 files from the United States, Italy, China, and the Middle-East – translated where necessary – produced by well-known documentarians, professional and amateur animators, and the general public, and represent some of the Archive's most popular resources. Finally the anticipated acquisition of *Here is New York* will significantly increase the Archive's collections, adding more than 500 hours of digital video interviews to our holdings.

### **Contextualizing and teaching about the WTC and Pentagon attacks**

This section of the original grant proposal was scaled back in the final funding, and in that context, we decided that it would be more effective to collect the lesson plans of others and focus our attention on making the Archive manageable for public visitors and future researchers.

The first of these activities resulted in a rich collection of materials collected from school groups and teachers. For example, outreach efforts with the Puerto Rican Family Institute and P.S. 721 yielded a rich collection of lessons plans and student work stemming from the "Super Rico" project, a school wide effort aiming to help students grapple with "their own experiences and fears of terrorism, violence, and family abuse". In other cases we leveraged the efforts of groups more educationally focused than us, archiving educational materials produced by the Social Science Research Council and the U.S. Department of Education, among others. Finally, our general collection contains hundreds of personal narratives, digital artworks, and other materials assigned as class work to students from around the country.

The Archive also has built a much-used "Frequently Asked Questions" page with links to general information about the attacks from the *New York Times* and *Washington Post*. Visitors

can find here detailed timelines of the events of September 11, information about the World Trade Center before, during, and after the attacks, and details about the hijacked planes, the victims of the attacks, and rescue efforts. The annotated Guide to Websites also provides a map to other noteworthy Internet resources. Finally, where possible, Archive staff have worked with teachers to help them better use the collections in their classrooms. These consultancies have included presentations to social studies teachers in Fairfax County, Virginia and to graduate history and information science courses at George Mason University and the University of Pittsburgh. In the years ahead, moreover, the Archive is scheduled to be used by teachers throughout Northern and Central Virginia as part of a set of Teaching American History grants coordinated by CHNM.

Proper collections management became more pressing as the Archive outgrew the original proposal's scope and vision. As the number of objects in the Archive increased, the need for an effective interface for researchers became a leading priority, and substantial resources were devoted to developing a sophisticated collections management system and easy-to-use researcher's interface. This new interface and a new design will be launched to the public in September 2004, on the third anniversary of the attacks. For the first time, it will provide easy public access to the Archive's entire collection.

### **Developing tools and models for collecting and preserving the past electronically**

The final proposed aim of the *September 11 Digital Archive* – to develop tools and models for collecting and preserving the past electronically – has also been fulfilled, and lessons learned from the development of the Archive have been widely shared through presentations and publications and reified as freely-available and easy-to-use software packages.

Presentations have included talks at scholarly meetings of the National Council on Public History, the American Historical Association, the Association of Mid-Atlantic Archivists, and the Southwest Oral History Association. More public presentations have included colloquia at the Smithsonian Institution, the Library of Congress, and the University of Pittsburgh School of Information Sciences, as well as media appearances on CNN, MSNBC, NPR, the BBC, Westwood One Radio, NY1 Television, NBC4 Television in Washington, and dozens of other broadcast and print outlets.

Publications range from the smallest newsletter notice to a forthcoming scholarly book. The former include abbreviated reports and case studies in journals and newsletters of the National Council on Public History, the Council on Library and Internet Resources, Organization of American Historians, and History News Network. The latter include published articles by Sparrow and Dan Cohen (Appendix A and Appendix D). Other publications, including possibly a pamphlet co-published by the Library of Congress and CLIR, are in the works. We have also significantly expanded the scope and ambition of the sourcebook *Doing Digital History* promised in the proposal. It has become a book-length manuscript entitled *Making Online History* that will be published by the University of Pennsylvania Press and made available free on the Echo website. In particular chapters on "Building an Audience," and "Collecting History Online" draw directly upon our experience with the *September 11 Digital Archive*. *Making Online History* will be the first book-length guide to digital history; we believe it will foster high

quality work in history and new media and will serve as a crucial resource in the next phase of the Echo project.

Finally, two and one-half years after the start of the grant period, an extensive and well-used suite of free tools for online collecting have emerged. An updated Survey Builder allows practitioners without any knowledge of programming languages to create and present text and image contributions to online forms. A new Web Scrapbook serves as an electronic clippings tool, allowing users to annotate, organize, and store text, emails, images and whole web pages in folders and share their work with others. These tools are constantly being improved and better integrated to meet our ongoing efforts to extend the practice of online collecting. As noted in the summary report, we postponed work on a couple of the tools projected in the original proposal, in part because we decided that they would be developed better as part of the next phase of work on Echo and in part because we had to devote substantial programming resources to building the researcher's interface and on preparing the collection for accession by the Library of Congress.

### **Managing and ensuring the permanence of the Archive**

In the original proposal we indicated only that we would “complete a plan for maintaining the project website in the years and decades beyond the end of 2003.” As the project grew more and more successful, however, receiving heavy use and continued growth, we felt obliged to expand upon our vague promise and see the Archive through to its permanent preservation. A meeting funded by the Sloan Foundation and hosted by CLIR in April 2002 particularly focused our attention on issues of collections management and digital preservation and provided an entrée to communities who could help us meet this important challenge. After considerable work and negotiation, the Library of Congress has agreed to become the permanent repository for the Archive, marking the first time that institution had accepted a major digital collection into its holdings. In fact, in addition to insuring the long-term preservation of the Archive, the Library will use the project as a test bed for management of digital materials. The National Digital Information Infrastructure Program (NDIIP), the Library's national digital initiative, is using the Archive to create and test software to effectively accession, manage, preserve, and disseminate heterogeneous digital materials. In addition, the Library is working with four universities on the Archive Ingest and Handling Test, using the Archive as a test bed to explore strategies for the ingest and preservation of digital archives.

The successful conclusion of this agreement has of course entailed considerable additional work on our part, particularly in working out the legalities of the gift and in processing the Archive for transfer to the Library. Long hours were spent negotiating the MOU and contacting contributors for permission. Similarly, as no formal guidelines were available to govern the preparation of the materials in the Archive for transfer to the Library, work by the Archive had to overcome roadblocks in creating standards for digital preservation. Our staff has spent many long hours working with the Library to make the collection ready for accession. The formal acceptance of the Archive and recognition of this work was marked by a full-day symposium entitled “September 11 as History: Collecting Today for Tomorrow” held on the eve of the second anniversary of the attacks.

In conclusion, the *September 11 Digital Archive*'s success in gathering varied materials from diverse communities and organizations underscores the value of institutional collaboration; the partnership of the American Social History Project and the Center for History and New Media allowed the Archive to draw upon each organizations' established "on-the-ground" networks and complementary new media strengths. The Archive has conscientiously collected and organized the electronic record of personal experiences of the September 11 attack and its aftermath. The public at large donated thousands of digital narratives, images, and sounds of their encounters with the enormity of September 11. The Archive provided a guide to online materials and information about the attacks, and created a framework to access the wealth of information held by the Archive as well. Developing tools and procedures for the Archive simultaneously created a model and methodology for collecting and preserving the past electronically, and the Archive itself will be used as a resource to set standards for digital preservation. Case studies and guidance are to be published in printed form and online, and in-person presentations broadcast the work of the Archive and possibilities of combining history and new media. Free Web-based tools are now available to aid and encourage students, teachers, researchers, and the historical profession to practice digital history. The *September 11 Digital Archive* truly is a project created of the moment, in response to a major contemporary event and using the best capabilities of current technology. The September 11 Digital Archive is also a project for the future, established to develop new avenues of historical preservation and research, and provide a window into the personal experiences of the world on September 11, 2001.

## APPENDIX C: INVENTORY

*Note: This is not a final inventory of the collection. Some numbers are estimated downward because of remaining processing work. Final totals will be higher even without further accessions such as the Here is New York collection.*

|                                  |              |                       |             |
|----------------------------------|--------------|-----------------------|-------------|
| <b>Personal Narratives</b>       | <b>45934</b> | Sonicmedia            | 272         |
| including:                       |              | *tyr_images*          | 70          |
|                                  |              | Webcontent            | 338         |
| *cair*                           | 10           | [dave_mccall]         | 550         |
| *lc_stories*                     | 258          | [louis_brendel]       | 559         |
| *nmah*                           | 6533         | [Mike_Garcia]         | 3480        |
| *rc_stories*                     | 34           | [rudymthompson]       | 371         |
| *seiu*                           | 55           | [saraschwitteck]      | 599         |
| [thank_you_rescuers]             | 479          | [book]                | 50          |
| smithsoniancards                 | 27899        | [Fritz]               | 500         |
|                                  |              | *lc_art*              | 7           |
| <b>Electronic Communications</b> | <b>61893</b> | [server]              | 1676        |
| including:                       |              |                       |             |
|                                  |              | <b>Documents</b>      | <b>6184</b> |
| doj                              | 12617        | including:            |             |
| hashers                          | 29           |                       |             |
| websites/weblogs                 | 16765        | Crr                   | 3119        |
| *coyle*                          | 439          | Documents             | 1           |
| *email*                          | 1184         | fdny_iap              | 130         |
| *lc_email*                       | 10           | Flyercollection       | 925         |
| [server]                         | 835          | lap                   | 79          |
| [andrew_manore]                  | 10500        | lowresflyers          | 925         |
| [andrew_mcmichael]               | 73           | objects               | 140         |
| [buz_grover]                     | 268          | sonicmedia            | 2           |
| [chad_blosser]                   | 96           | webcontent            | 145         |
| [cynthia_estrella]               | 750          | *articles*            | 486         |
| [jason_scott_luck]               | 135          | *tompaine*            | 230         |
| [jonathan_abbey]                 | 9743         | [server]              | 2           |
| [mills_kelly]                    | 107          |                       |             |
| [Fritz]                          | 300          | <b>Audio/Video</b>    | <b>4608</b> |
| [global_kids]                    | 3500         | including:            |             |
| [patrick_hirsch]                 | 100          |                       |             |
| [david_bendory]                  | 159          | documentnywav         | 2           |
| [andy_carvin]                    | 3560         | sonicmedia            | 897         |
| 911russian                       | 39           | [server]              | 108         |
| [satan_image]                    | 684          | [smithsonian_audio]   | 3000        |
| *stories*                        | 10666        | [sonicmemorial_audio] | 88          |
|                                  |              | [voicemail_audio]     | 164         |
| <b>Images</b>                    | <b>14359</b> | diganimbackup         | 118         |
| including:                       |              | imagineny             | 4           |
|                                  |              | objects               | 1           |
| crr ([national_guard_bureau])    | 2984         | sonic media           | 2           |
| diganimbackup                    | 6            | video                 | 1           |
| Documents                        | 52           | [server]              | 84          |
| ea911                            | 42           | [Mark_Fiore]          | 11          |
| embrace_the_memory               | 31           | [Pentagon Kiosk DVD]  | 61          |
| *images*                         | 2550         | [VHS]                 | 3           |
| *lc_images*                      | 7            | [Fritz]               | 17          |
| Objects                          | 46           | [911_videos]          | 47          |
| Sonic                            | 169          |                       |             |

## APPENDIX D:

[Published as Daniel J. Cohen, "History and the Second Decade of the Web," *Rethinking History* Vol. 8, No. 2, June 2004, pp. 293–301.]

*Over ten years of experience with the web has allowed us to understand what the medium does well and what it does poorly, and how we may be able to improve online historical efforts so that they capitalize on the web's strengths while avoiding its weaknesses. This essay explores three possible ways to advance digital history: interaction between historians and their subjects, interoperation of dispersed historical archives, and the analysis of online resources using computational methods. Thinking about such possibilities raises important, age-old questions about how we should preserve and chronicle the past.*

With the World Wide Web now in its second decade of existence, and with the euphoria of the dotcom era now well behind us, it is a good time to reflect on what historians have been able to do in this still immature medium and how in the next decade of its existence we can better make use of it. Generally a conservative bunch in terms of the adoption of new technology (if not in political inclination), historians have, mostly for better but surely on occasion for worse, incorporated the medium into their work over the past ten years. For the most part using the web—it seems appropriate in our post-bubble sobriety to drop the grandiose alliterative phrase as well as the capitalization—has meant the posting of materials for courses, exhibitions, and independent and collaborative work, as well as the communication of news and views from all corners of the discipline. Websites have flourished on almost every conceivable historical topic, created by historians from within and beyond the academy.

To be sure, this represents a tremendous step forward. As many have noted, the marginal cost of reaching another person with these electronic materials is almost zero, which represents a great advance over the physical world of paper (Negroponte 1996). With the web a historian can reach a hundred interested colleagues with an online essay as cheaply and simply as a dozen. Moreover, the structure of the web allows interested parties to access that essay from any

internet-connected computer, at any time, and to search the text for phrases or keywords. Creator and audience alike may link this piece to others on the web, catalogue and copy it, and even print it if so desired. By now most of us are familiar with these significant advantages that come with digital, rather than physical, historical resources.

One senses, however, that the medium of the web has not been exploited to its fullest if the best we can say about historians' use of this highly advanced computer network is that it has become a giant, global fax machine, faithfully reproducing and distributing copies of historical documents (primary and secondary), and related commentaries and professional missives. And because of the openness of the medium—and the always tenuous relationship between the professoriate and the large population of lay historians and the general public interested in history—many historians have found the web to be a mixed blessing: prolific but unmediated, powerful but untamed, open to all but taken seriously by few. When my colleagues Roy Rosenzweig, Michael O'Malley and Andrew McMichael first surveyed the historical web in 1995 they were impressed with the high number of websites on FDR: 49. Today that number stands at 628,000—and one would venture to guess that not all of them are compelling, accurate, and the product of the kind of deliberative thought and research that our profession honours (McMichael *et al.* 1996). Online history has not—and indeed may never—rival the gold standard of the book, replete with its physical and (hopefully) intellectual heft, peer review, and centuries of technical improvements such as the footnote, table of contents and index (Grafton 1997; O'Malley 2000).

Yet historians will never know if this apparent inequality of media is inherent in the natures of the web and print unless we work harder to capitalize on the advantages of the web and lessen its disadvantages, in the process creating new forms of history that can only exist

online. Surely it is ridiculous to ask this new medium to reproduce the exact successes of paper, just as it would be silly to denigrate Ken Burns' *Huey Long* for being less comprehensive than the biography by T. Harry Williams. Both worthy examinations of a crucial figure in American history, these two efforts had different goals that were commensurate with the strengths of their intended medium. Furthermore, just because the majority of websites on FDR are not worth the electrons they vibrate should not condemn by association the thoughtfully or creatively produced minority.

So what have we learned about the strengths and weaknesses of this young medium, and how can we maximize the web's advantages and minimize its disadvantages to create the best forms of online history? It is clear that the web is very good at connecting large numbers of people regardless of physical location; allowing for synchronous and asynchronous communications between those people and in general facilitating collaboration; enabling the storage and transmission of huge amounts of information; encoding information in such a way as to make it rapidly searchable by machines; removing barriers to publication and providing space for unbounded publications and parallel publications of a single entity (such as translations); and allowing for revision and updating. The web's disadvantages include its lack of stability and persistence; its difficulty in displaying human-friendly, legible formats for text, especially for the long-form narratives that have been the hallmark of history for generations; its failure to provide reliable clues that help one discern the real from the fake and the good from the bad; and its lack of the kinds of checks that in the physical world lead to an often beneficial winnowing of available information.

Examining this list of pluses and minuses, it should not be surprising that many in our profession have regarded at the web with suspicion. We spend much of our time as readers, and

we like our documents stable, authentic, persistent and legible. Moreover, when even web designers who have come out of academia such as Sarah Horton and Patrick J. Lynch note that it is best to break text on the web into ‘small chunks’, defenders of the traditional historical essay and book have had a right to be worried (Lynch & Horton 2002). In addition, high expectations for both ‘e-books’ and new, nonlinear forms of digital writing (‘hypertext’) have thus far proven unwarranted. Most of us still like our text linear and coherent, and in crisp black text on a white page.

Although technologists constantly praise computers as ‘universal machines’, and while it is true that they are indeed remarkably flexible and powerful, this does not mean that we should use them for every kind of historical effort or production. We can truly begin to think about history in the second decade of the web, and to postulate new, more appropriate and ideally more enlightening forms of online history when we begin to think beyond the distribution of documents (including essays, articles, exhibits, news and messages) and think instead about the collection, interrelation and exploration of those documents. Much of the problem here is conceptual: from the beginning of the web historians have largely discussed *nouns* such as web pages and websites rather than *verbs* such as searching, sorting, gathering and communicating.

A reorientation can occur when we remember that the web is a subset of the *internet*, its very name representing the way this computer network shuttles information *between* and *among* people, rather than just a publishing medium that goes from point A (historians) to point B (an interested audience). This character of the web is already being exploited well by websites that have furthered professional communication. For instance, H-Net (<http://www.h-net.msu.edu>) has online discussion groups on over a hundred topics that one can access via the website or through email. Every day historians ask each other for assistance with their research and analyze

new work in their fields. In contrast to paper media the internet seems ideally suited for this kind of vibrant, daily exchange.

There are other forms of interactivity, of two-way communication, on the web that are less developed but that have the potential to create new kinds of history on the web in this decade. What about communications not just among historians but between historians and their subjects? For historians working on topics in the post-Second World War era the web can be a tremendous resource for reaching those across the globe who might have recollections or materials that could further their investigations in the present and provide archives for future historians.

Two web projects I have worked on, the ECHO Project ([http:// echo.gmu.edu](http://echo.gmu.edu) ) and the September 11 Digital Archive ([http:// 911digitalarchive.org](http://911digitalarchive.org) ), have explored this possibility of using the web not only to present the past but also to collect it. Both sites use flexible databases and the upload capacity of the web (in addition to its vast capacity to download, or distribute, documents) to acquire materials from far-flung historical subjects such as scientists and engineers (in the case of ECHO) and those who witnessed and were affected by 9/11 (in the case of the September 11 Digital Archive). It is so far an imperfect science, to be sure, with pitfalls and insecurities, but the pay-off can be tremendous when a project that collects history online is successful. The nature and extent of what one can gather, while certainly different from a traditional oral history project or museum effort, seems just as enlightening and important as a future historical resource, and likely will grow more so as an increasing percentage of our communications and expressions occur in digital media.

That has certainly been the case with the September 11 Digital Archive, which the Library of Congress recently accessioned—the first major digital acquisition by the institution.

Started in January 2002 with a rudimentary website by myself and colleagues at the Center for History and New Media at George Mason University and our associates at the American Social History Project/Center for Media and Learning at the Graduate Center of the City University of New York, and funded by the Alfred P. Sloan Foundation, the archive now contains over 130,000 digital objects from the tragic events of that day and its aftermath, including stories, email, photos, artwork, poetry, audio recordings and digital video. The materials have been donated via the website and through less sophisticated means such as postal mail and the telephone (digitally connected to our servers) by over 30,000 individual contributors. The result is an extremely varied and vital archive of the experiences, thoughts and emotions of a broad spectrum of contributors from every state and from many nations. The archive contains the recollections and images of emergency workers and those who survived Ground Zero in New York and the Pentagon, the sounds of voicemail and the chatter of instant messages, as well as more impressionistic responses and subsequent reflections. While the September 11 Digital Archive was by far the largest and most comprehensive online effort to record this history, we were not alone in this endeavour; at the same time that we were permitting submissions through our website, dozens of other individuals and institutions were engaging in similar projects on the web, creating parallel collections that historians in decades to come will hopefully find useful when they try to understand the true meanings of 11 September 2001.

This concurrence also suggests a second, relatively unexplored possibility for history in the next decade of the web: the potential for ‘interoperability’. A word most frequently used by ‘web services’ software developers who work on integrating websites and databases for decentralized businesses, and associated with opaque acronyms and new information architectures such as the international standard XML and Microsoft’s.NET initiative,

interoperability holds promise for historical researchers interested in creating and accessing vast archives of documents that exist in different places on the web. For example, for my research on the history of nineteenth-century mathematics I have been following with great interest the virtual integration of several major archives of digitized mathematical texts by the Cornell University Library, the University of Michigan Library and the State and University Library Göttingen, called the Distributed Digital Library of Mathematical Monographs (<http://www.library.cornell.edu/mathbooks/>). This novel effort, facilitated by these new technologies that allow dispersed web servers to respond in tandem to requests for information, shows the power of interoperability. Certainly one could have gone to each library's website and conducted a separate search for books or phrases. However, the combined archive, created through seamless, hidden communications between these websites, is greater than the sum of its parts, and not only as a time saver. The Distributed Digital Library makes it far easier to get an overview of passages on a specific topic, such as non-Euclidean geometry, and to spot trends in the development of that subject during the nineteenth century. It represents both a quantitative and a qualitative advance over prior resources in the history of mathematics.

My own effort to explore the potential of interoperability has been an attempt to virtually combine the often disregarded educational products of tens of thousands of academics. By using an algorithm to analyze documents found through behind-the-scenes communications with Google's enormous database, my Syllabus Finder (<http://chnm.gmu.edu/tools/syllabi>) searches the web for pages that look like syllabi and, when successful, saves them in a special format. It can then present those syllabi to interested seekers, figure out which school, college or university a course is being taught at, and, more experimentally, attempt to extract assigned books and other notable features of a syllabus. In a sense the Syllabus Finder creates a single, overarching course

directory (though obviously it includes only syllabi that are posted to the web), with the potential to show how many courses are being taught on a specific topic, which books or articles are popular or influential, and which types of assignments faculty like to assign. My hope is that ultimately it will provide a unique window into the state of higher education. With over 250,000 syllabi collected so far (including over 10,000 history syllabi) it is not unreasonable to say that it is, by several orders of magnitude, the largest archive of course materials ever collected—and all of this was done in an automated fashion following my initial programming.

This collection of a quarter of a million syllabi, which could only have been acquired in a medium where text is machine-readable, searchable and easily reproduced, suggests another possibility for history in the second decade of the web: a digital cognate of ‘close reading’ that computer scientists call ‘data-mining’. Data-mining involves complex analyses of digital materials to find meaningful patterns. In a way it is a more advanced version of what some classicists and literary scholars have been doing for years when they have manually counted the frequency of certain words, or compared the various uses of those words, in a text or set of texts.

On the web the speed with which one is able to do this sort of textual analysis can enable both quick assessments of historical collections as well as more substantive investigations. When Michael Kazin used search tools to scan the September 11 Digital Archive for the frequency of words such as ‘patriotic’ and ‘freedom’, he came to some important, if initial, conclusions about the American reaction to the terrorist attacks. Kazin discovered that fewer Americans than one might imagine saw 9/11 in terms of nationalism, radical Islam versus the values of the West, or any other abstract framework. Instead, most saw the events in far more personal and local terms: the loss of a friend, the effect on a town or community, the impact on their family or jobs (Kazin 2003). Using similar electronic techniques, researchers are already probing the Syllabus Finder

database to investigate a variety of themes, such as the changing role of race and ethnicity in undergraduate education, the growth of distance learning, how course materials on the web differ from offline materials, and how globalization is affecting the teaching of history and international relations. Future enquiries may use an extremely powerful computer science method called ‘regular expressions’, which one can use to search documents for all kinds of text patterns, not just keyword or phrase matches.

As with interactivity and interoperability, early efforts at data-mining hint at a future of doing and using history on the web that may have more to do with providing materials and structures for historical research than with the presentation of a finished secondary source such as an electronic book or exhibit website. Done correctly, and taking advantage of the medium’s unparalleled ability to store, scan and interrelate documents, the web may enable new investigative tactics and resources. The 130,000 objects in the September 11 Digital Archive are not like any other 130,000 other objects in the Library of Congress; we have acquired them, and can explore them, in ways we are only beginning to think of. Such collections and their associated electronic methods have the potential (an appropriately cautious word at this point) to supplement, though not replace, traditional historical research in physical archives.

Finally, thinking about doing history on the web, just like thinking about history in any other medium, should ideally force us to revisit age-old questions about what history is and how we should engage in it. For some years I taught Herodotus and Thucydides to freshmen and perennially enjoyed (and felt professionally reinvigorated by) their sense of the differences between these two progenitors of the idea of history. Purveyor of a mostly tidied, marching chronology that aspired to the truth about what seemed to him (and to most historians since) to be the most devastating events to befall ancient Greece, and with a particular focus on

remarkable figures such as Pericles and Alcibiades, Thucydides' history of the Peloponnesian War captured the hearts and minds of most in my classes. Students readily identified his compelling narrative of warfare and diplomacy as 'history'. Other students found themselves drawn to Herodotus' idiosyncratic tangents and more expansive view of ancient culture beyond the battles between the Greeks and the Persians. This minority found Herodotus' sense of what belonged in the historical record refreshing: the strange rituals of non-Greek cultures, the sentiments of common people in addition to leading figures, competing and contradictory accounts. As Herodotus told his audience, he was saving and recounting it all because in the future people might have different notions of what or who is important: 'I will go forward in my account, covering alike the small and great cities of mankind. For of those that were great in earlier times most have now become small, and those that were great in my time were small in the time before. Since, then, I know that man's good fortune never abides in the same place, I will make mention of both alike' (Herodotus 1987, p. 35). Not having to worry about fitting his account into a limited number of pages, a rich description of what he had learned from a wide variety of sources, Greek and otherwise, seemed to Herodotus to be the sensible course to take.

Although the medium is still in its infancy it appears that the web may be more propitious for history in the inclusive and wide-ranging mode of Herodotus than in the resolute mode of Thucydides. On the web historians can supplement their narratives with virtually unbounded collections of sources, notes, graphs, charts, images and links that even a profligate publisher could not hope to fit into a book. In addition, given the open access of the web it seems appropriate to cast the widest possible net (so to speak) in projects like the September 11 Digital Archive, rather than focus on figures such as government leaders who will likely dominate coverage in print. The massive capacity of the web means that historians can push beyond the

selectivity of paper collections to create more comprehensive archives with multiple viewpoints and multiple formats (including audio and video as well as text). These archives, hopefully partially making up for their lack of a curator's touch with their size, scope and immediacy, will in turn require more sophisticated tools for future research. If carefully developed, such collections—ideally interoperable with others of their ilk—may provide historians not only of this generation but also of generations to come with the means to understand the past better and more deeply

#### **References:**

- Grafton, A. (1997) *The Footnote: A Curious History*, Harvard University Press, Cambridge, MA. Herodotus (1987) *The History*, trans. David Grene, University of Chicago Press, Chicago, IL.
- Kazin, M. (2003) '12/12 and 9/11: tales of power and tales of experience in contemporary history', *History News Network*, 11 September, <http://hnn.us/articles/1675.html>.
- Lynch, P. J. & Horton, S. (2002) *Web Style Guide* (2nd edn), Yale University Press, New Haven, CT.
- McMichael, A., O'Malley, M. & Rosenzweig, R. (1996) 'Historians and the web: a guide', *AHA Perspectives*, January, pp. 11–16.
- Negroponte, N. (1996) *Being Digital*, Vintage, New York.
- O'Malley, M. (2000) 'Building effective course sites: some thoughts on design for academic work', *Inventio*, Spring, [http://www.doit.gmu.edu/inventio/spring00/omalley/momalley\\_1.html](http://www.doit.gmu.edu/inventio/spring00/omalley/momalley_1.html)