The Future of Preserving the Past

by Daniel J. Cohen

Consider the effort expended to save a rich and representative historical record of perhaps the two most tragic days in American history in the past century: December 7, 1941, and September 11, 2001. The National Archives preserved military photographs of the chaos at Pearl Harbor on December 7 as well as communications and damage assessments. The Office of Naval Records and Library recorded the names of those who died or were wounded. Meanwhile, other government branches and institutions undertook more wide-ranging preservation activities. The Library of Congress acquired the annotated typescript of the National Broadcasting Corporation's breaking news account. In addition to saving military records, the National Archives catalogued the reactions of government officials in public announcements and private correspondence. The National Park Service administers the *USS Arizona* Memorial of Pearl Harbor in Hawaii to preserve the underwater remains of the ship, while providing visitors a sense of the day's events and repercussions.(Figure 1)

In a mode more active than reactive, others sought to save the character of the Pearl Harbor attack by seeking out the views of average Americans. Pioneering folklorist Alan Lomax, working at the Library of Congress's Archive of American Folk Song, sent out an urgent telegram on December 8 to like-minded colleagues around the country imploring them to record the sentiments of the American people. In the next three days these interviewers, using cutting-edge technologies such as direct-to-disc machines that recorded sound directly onto platters that could be played immediately like normal records, gathered commentary from dozens of people in 15 states—a total of 4 1/2 hours of powerful expression. In subsequent years, historians have mined other national and local archives, letters and diaries, and the memories of Americans and Japanese to create a comprehensive picture of this day of infamy.

Sixty years later, on and after another day of infamy, September 11, 2001, professional and amateur archivists and historians again sought to record the aftermath of a horrific event. Widely varying initiatives began almost immediately, engaging in selective acquisition and broad opportunism, active outreach to historical subjects and passive collecting of artifacts, short-term haphazard gathering and careful long-term preservation. Projects modeled on those of 1941 quickly arose. At Columbia University, the Oral History Research Office and the Institute for Social and Economic Research Policy created the

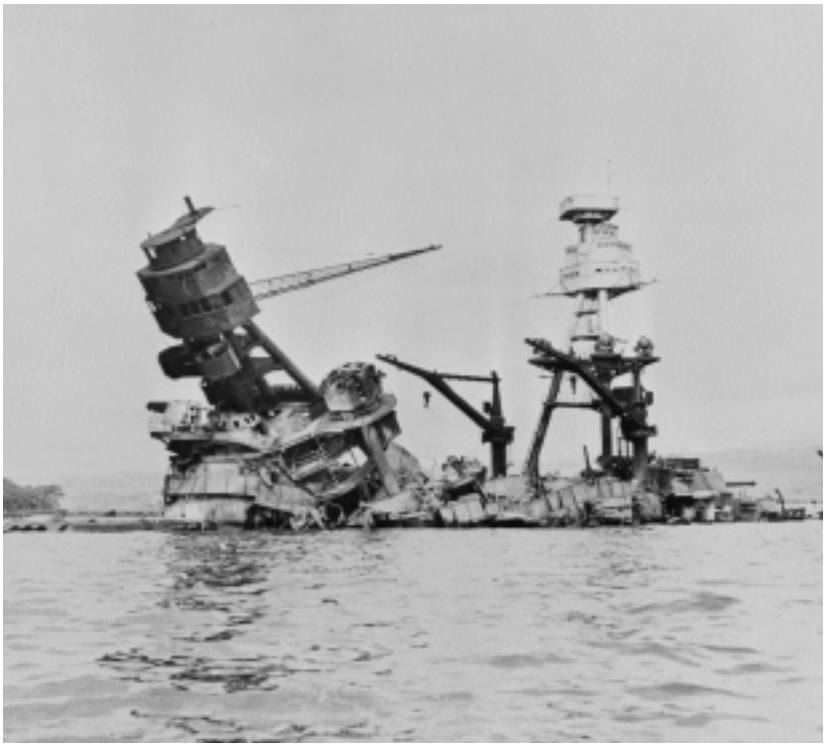

FIGURE 1
An official United States Navy photograph captures the USS Arizona *following the attack on Pearl Harbor on December 7, 1941. In 1944, the Office of War Information's Overseas Picture Division transferred the film of this image to the Library of Congress. (Courtesy of Prints and Photographs Division, Library of Congress)*

September 11, 2001 Oral History Narrative and Memory Project, which has conducted more than 300 interviews with people affected by the terrorist attacks in New York, New Jersey, and the Boston and Washington, DC, regions, including interviewees who escaped the World Trade Center or lived in its shadow and Afghan and Muslim immigrants.[1] As it had 60 years earlier, the American Folklife Center at the Library of Congress, the descendant of the Archive of American Folk Song, sent out a notice to folklorists across the United States to record the "thoughts and feelings expressed by average citizens." This distributed network of oral historians donated approximately 300 hours of audiotape to the library, collected in 19 states and a military base in Italy.[2] The library's September 11, 2001 Documentary Project also gathered a smaller number of video interviews, written narratives, drawings, and photographs.[3]

Despite the efforts following September 11, which were orders of magnitude larger than those of Lomax and his small band of colleagues, the nature of the historical record had changed in many ways. Media no longer meant the radio broadcasts of a few national networks but now meant hundreds of audio and video broadcasts. Far more expansively, the record of 9/11 was to be found in new media such as websites, email, and other forms of electronic communication and expression, forms that have become an increasingly significant part of America's and the industrialized world's cultural output.

To be sure, in the weeks and months after September 11, museums, libraries,

and archives began to address the changing nature and scope of the historical record. In doing so, however, they had to abandon, at least in part, well-established models drawn from oral history and archival science. The explosion of historical sources in a digital age necessitated this evolution in preservation tactics. For example, whereas photographs of the attack at Pearl Harbor number at most a few thousand—the largest collection, at the National Archives and Records Administration, comprises a mere 5 boxes with about 200 images in each box—the photographic record of September 11, 2001, likely numbers in the millions of images. Indeed, with the proliferation of personal cameras since 1941, and especially with the spread of digital cameras in the last decade, 9/11 may be among the most photographed events in history.

Given the enormous size of the photographic record of 9/11, a variety of organizations, not just those in the preservation business, have had little trouble building impressive archives. The United States National Institute of Standards and Technology (NIST), as part of its investigation into why structural elements in the twin towers failed, gathered more than 6,000 images from 185 professional and amateur photographers, from almost every conceivable angle and covering virtually every moment, and in some cases fractions of a second, of the towers' collapse.[4] As seen in the remarkable Here is New York "democracy of photographs" collection of 5,000 images from hundreds of contributors, each photographer literally as well as figuratively had his or her own perspective on the event.[5]

Some preservation institutions recognized the proliferation and importance of new digital media. Looking to supplement their standard accessions of the printed editions of newspapers after September 11, the Library of Congress, in partnership with the Internet Archive, WebArchivist.org, and the Pew Internet and American Life Project, archived 30,000 websites from September 11 to December 1, 2001. This massive collection of digital materials will undoubtedly be of great value to future researchers. But even this impressive undertaking saved less than *one-thousandth* of the roughly 32 million websites in existence in September 2001.[6]

Others took a more active role, soliciting a variety of digital reactions and artifacts through online projects not dissimilar from Alan Lomax's grassroots effort to capture a wide range of perspectives from across the country after the Pearl Harbor attack. The September 11 Digital Archive at the Center for History and New Media at George Mason University, co-produced by the American Social History Project/Center for Media and Learning at the City University of New York Graduate Center, which I co-directed, tried to capture digital sources from everyday people. We used a website, available at http://911digitalarchive.org, digital telephone lines, and less technologically sophisticated methods like note cards that were later scanned, to save personal stories, emails, photographs and works of art, instant messages, pager commu-

FIGURE 2
The remnants of 1 World Trade Center following the September 11, 2001, terrorist attacks are recorded in this digital photograph made by a Goldman Sachs employee who worked across the street. Now part of the September 11 Digital Archive, this image accompanied the photographer's vivid email recollections. (Photograph by David Bendory, courtesy of September 11 Digital Archive)

nications, and other forms of expression and communication from 9/11 and its aftermath. Thus far the archive has collected more than 150,000 items from thousands of individual contributors. In the fall of 2003, the collection was accessioned by the Library of Congress, one of the library's first major digital acquisitions.[7] (Figure 2)

Concerns About Digital Collections

The vast expansion of the historical record into new media between December 7, 1941 and September 11, 2001 presents serious challenges that will have to be surmounted in the coming years if future scholars and the public are to have access to an adequate record of the past. Yet despite the urgency of dealing with this mutating record, many in the cultural heritage community have major reservations about digital collecting, due in part to an understandable aversion to the complicated hardware and software involved, but more importantly because of some very real concerns about the nature of online work. At the same time that the web has enabled an exponential increase in cultural production, some argue that online collecting misses those older, less educated, or less well-to-do subjects who may not have access to the necessary technology.[8] Furthermore, the shift from analog to digital entails a change from well-known and relatively stable forms such as paper to forms for which the preservation path is unclear.

Digital collections are characterized as being shallow and less useful for research than traditional archives, for which provenance and selection criteria

are critical. It is unrealistic to expect that the Library of Congress could pre-screen 30,000 websites for quality or relevance to 9/11. The staggering numbers possible in digital collecting renders ineffectual some central tenets and time-honored procedures of archival and library science. Another common problem, encountered by many online collecting projects that actively solicit digital materials, is the opposite of this abundance: the failure to collect much at all because few people hear about or contribute to their websites. An inverse relationship between the quantity of digital artifacts gathered and the general quality of those artifacts may exist.[9]

Perhaps the most profound benefit of online collecting is an unparalleled opportunity to allow more varied perspectives in the historical record than ever before. Networked information technology can allow ordinary people and marginalized constituencies not only a larger presence in an online archive, but also generally a more important role in the dialogue of history.

Digital collections are indeed more susceptible to problems of quality because they often lack the helpful selection bias of a knowledgeable curator and the pressure to maintain strict criteria for inclusion engendered by limited physical storage space. Web collections formed around the submissions of scattered contributors or thousands of websites and blogs have a very different character from traditional archives. Digital collections tend to be less organized and more capricious in what they cover.

On a more positive note, digital archives can be far larger, more diverse, and more inclusive than traditional archives. Perhaps the most profound benefit of online collecting is an unparalleled opportunity to allow more varied perspectives in the historical record than ever before. Networked information technology can allow ordinary people and marginalized constituencies not only a larger presence in an online archive, but also generally a more important role in the dialogue of history. "The Net is a people's medium: the good, the bad and the ugly," Brewster Kahle, the founder of the Internet Archive, has said. "The interesting, the picayune and the profane. It's all there."[10]

A less obvious but perhaps more important measure of the "quality" of a digital historical collection becomes apparent when the collection is assessed as a whole. Like any collection, there will be a minority of striking contributions among a sea of mundane or seemingly irrelevant entries. Historians who have browsed box after box in a paper archive trying to find key pieces of evidence for their research will know this principle well. The propensity of digital collectors to save virtually everything given the low cost of digital storage and the difficulty of using selection criteria may make these percentages worse. Yet, a few well-written perspectives or telling archival images may form the basis of a

new interpretation, or help to buttress an existing but partial understanding of a historical moment. At the same time, the greater size and diversity of online collections allow more opportunities to look for patterns. Why do certain types of stories recur? What does that reveal about popular experience and the ways that experience is transformed into memory?

Because of a digital collection's superior ability to be searched, historians can plumb electronic documents in revealing and novel ways. The speed of analysis can enable quick assessments of historical collections and more substantive investigations. For instance, when historian 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 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 or another abstract framework. Instead, most saw the events in personal and local terms, the loss of a friend, the effect on a town or community, the impact on their family or job.[11]

Active Solicitation of Digital Materials

Reaching out to and interacting with historical subjects online, either in real time or asynchronously, is far more economical than traditional oral history. With subjects writing their own narratives, the cost of transcription is avoided. While live individual interviews are often quite thorough and invaluable resources, online initiatives to collect personal histories can capture a far greater number at lower cost and acquire associated digital materials, such as photographs, just as cheaply.

Of course, even if highly successful in the future, online interaction with historical subjects will not mean the end of traditional ways of gathering recent history. As oral historian Linda Shopes observes, newer technological methods will have a hard time competing with many aspects of the oral historian's craft: "the cultivation of rapport and…lengthy, in-depth narratives through intense face-to-face contact; the use of subtle paralinguistic cues as an aid to moving the conversation along; the talent of responding to a particular comment, in the moment, with the breakthrough question, the probe that gets underneath a narrator's words."[12] Instead, using the Internet will likely complement these older methods.

Acquiring historical materials and recollections online is more difficult than setting up a rudimentary website because it entails digital tools to receive, process, and store submissions. To adequately capture the past in this way, more technical hurdles must be surmounted to allow for historical documents and artifacts to flow inward rather than merely outward, as they do on the web pages of most museums, archives, and historical sites.

The good news is that online interactivity is becoming easier each year. The same digital technologies that have made the historical record proliferate into new forms give us the best hope to capture that record. Not everyone needs a custom-programmed archival system such as the one constructed for the September 11 Digital Archive. Much of the infrastructure and software required to do simple or even moderately complex online collecting is available and cultural institutions and independent scholars should take advantage of these technologies.

Probably the oldest and still quite useful technology for online collecting is email, the choice of some of the most successful projects. Keith Whittle's Atomic Veterans History Project, devoted to the community of veterans who participated in nuclear testing during the Cold War, has collected and posted more than 600 personal narratives from former soldiers, acquired solely through email. As Whittle discovered, emailers include attachments such as scanned photographs, many of which grace the website alongside the narratives. Email also allows for long-term interactions, follow-up, and detailed exchanges. An online collecting project can get started right away with a simple web design that uses email links to encourage and accept submissions.[13]

Blogs have given millions of Internet users a taste of what it is like not just to read and view the web, but also to post to it. Many ways of maintaining a blog also allow for more than one person to post and for contributors to add images and multimedia files, creating an ever-expanding and multifaceted discussion about topics of interest. The ease with which one can add materials makes blogs an attractive possibility for a basic collecting site. Blogs have built-in search features and the ability to export whatever is collected to other locations.[14]

New forms of instantaneous communication on the Internet will further expand the toolkit for collecting history online. Millions are now using instant messaging (IM) software that permits real-time communication with individuals around the globe. Although they do not have the tonal inflections of a spoken dialogue, these typed conversations have the advantage of being self-documenting, unlike oral history interviews, which require expensive transcriptions. More recent versions of these IM programs also allow rudimentary audio and video chats, which opens up the possibility of a future that is much like the past of traditional oral history. Technical concerns such as installing and configuring appropriate software and hardware for digital collecting should recede, ultimately, into the background.

What will remain in the foreground are the qualitative concerns, especially the question of provenance raised by the solicitation of historical materials from unseen contributors. Given the slippery character of digital materials, how can we ensure that what we receive over the Internet is authentic, or that historical narratives we receive really are from the people they say they are?

Some of these worries are relatively easy to address. Concern about the falsification of digital historical documents and metadata (information about such artifacts) has mostly turned out to be a phantom problem.[15] I am not alone in this assessment. Newspaper websites have found that relatively few people enter fake information. In one study, the *Philadelphia Inquirer* discovered that only 10 to 15 percent of their 300,000 registered users had entered bad email addresses, and some of those were merely by accident or due to technical difficulties. Zip codes and other less problematic bits of personal information are falsified at an even lower rate.[16]

The nonprofit mission of online historical archives should produce even higher rates of honesty. Most people who take the time to submit something to a digital project share a cultural institution's or dedicated researcher's goals and interest in creating an accurate historical record. In addition, some technical methods can help double-check online contributions. Every computer connected to the web has an Internet Protocol (IP) address. A small bit of programming code can capture this address. If a researcher is skeptical that a contribution has come from a specific person or location, a WHOIS search, which translates an IP address into a semi-readable format that often includes a contributor's Internet service provider and broad area of service, may result in helpful information.[17] Less cloak-and-dagger is a simple email or telephone follow-up with a contributor to thank them for their contribution; this presents an opportunity to ask contributors if they might have any other documents or recollections and whether they might know of other contacts.

The best defense against online fraud comes from traditional skills. Historians have always had to assess the reliability of their sources. Countless notable forgeries exist on paper. As Donald Ritchie has pointed out, written memoirs and traditional oral histories are filled with exaggerations and distortions.[18] Historians will have to continue to look for evidence of internal consistency and weigh them against other sources. In any media, new or old, solid research is the basis of sound scholarship.

Despite the challenges and insecurities surrounding digital collecting, it has become a burgeoning practice. Recently, for example, the British Library, the Victoria and Albert Museum, the Museum of London, and a number of other British museums and archives pooled their resources to display and collect stories of immigration to the United Kingdom in a project called Moving Here. Thus far the project has posted almost 400 stories and artifacts, mainly digitized versions of existing archive records but also new materials acquired via the site, ranging from a documentary video on Caribbean life to the reflections of recent African immigrants. The British Broadcasting Corporation's online project to gather the stories of Britain's World War II veterans and survivors of the London Blitz, entitled WW2 People's War, has been even more successful, with over 1,000 narratives gathered through the BBC's website after only 8 months, including dozens of harrowing accounts of D-Day.[19]

In the United States, the National Park Foundation, the National Park Service, and the Ford Motor Company are using the Internet to collect first-hand narratives of life during wartime for the Rosie the Riveter/World War II Home Front National Historical Park in Richmond, California. So far more than 6,000 former home front workers have contributed stories. *National Geographic's* Remembering Pearl Harbor site has received over 1,000 entries in its memory book. Over 500 people have recorded their personal stories and artifacts of the Civil Rights Movement on a site co-sponsored by the American Association for Retired Persons, the Leadership Conference on Civil Rights, and the Library of Congress. The Alfred P. Sloan Foundation has supported dozens of online collecting projects on science and technology in the belief that the history of these subjects is growing much faster than our ability to gather it through more conventional means.

Although there remains a healthy skepticism in the oral history community about the usefulness and reliability of narratives collected online, several new projects by major oral history centers demonstrate the benefits of online collecting. Even Columbia University, the home of the nation's first oral history program, is encouraging alumni to join in writing Columbia's history by contributing stories online.[20]

Saving Existing Digital Sources

The main challenge for those interested in a more passive form of digital collecting is how to preserve what is collected for the long term. This is a serious challenge faced by actively acquired digital collections as well. Electronic resources are profoundly unstable, far more so than physical objects like books. The foremost American authority on the longevity of various media, NIST, still cannot give a precise timeline for the deterioration of many of the formats we currently rely on to store precious digital resources.

A recent report by NIST researcher Fred R. Byers notes that estimates vary from 20 to 200 years for popular media such as the CD and DVD. Anecdotal evidence shows that the imperfect way most people and institutions store digital media leads to much faster losses. For example, a significant fraction of collections from the 1980s of audio CDs, one of the first digital formats to become widely available to the public, may already be unplayable. The Library of Congress, which holds roughly 150,000 audio CDs in conditions almost certainly far better than those of personal collections, estimates that between 1 and 10 percent of the discs in their collection already contain serious data errors.[21]

Moreover, nondigital materials are often usable following modest deterioration, while digital objects such as CDs frequently become unusable at the first sign of corruption. We have gleaned information from letters and photographs

discolored by exposure to decades of sunlight, from hieroglyphs worn away by centuries of wind-blown sand, and from papyri partially eaten by ancient insects. By contrast, a stray static charge or wayward magnetic field can wreak havoc on the media used to store digital sources.

Beyond the possibilities of data corruption, all digital objects also require a special set of eyes, often unique hardware, and an accompanying operating system and application software, to view or read them properly. The absence of these associated technologies can mean the effective loss of digital resources, even if those resources remain fully intact. There have already been several versions of HTML, the underlying language of the web, enough to cause many of the web pages created in the early 1990s to be partially unreadable. The University of Michigan's Margaret Hedstrom, a leading expert on digital archiving, bluntly wrote in a recent report on the state of the art, "No acceptable methods exist today to preserve complex digital objects that contain combinations of text, data, images, audio, and video and that require specific software applications for reuse." In short, historians, archivists, librarians, and museum curators, even those strongly committed to the long-term preservation of recent history, enter uncharted waters when they try to save the past digitally.[22]

Computer scientists and digitally savvy librarians and archivists are working on possible solutions to these challenges, from software like the Massachusetts Institute of Technology Libraries' and Hewlett-Packard's DSpace or the University of Virginia's and Cornell University's Fedora, and through broad initiatives like the Library of Congress's National Digital Information Infrastructure and Preservation Program. But we are still in the very early stages of the creation of these new digital archives, and many prototypes and methods will undoubtedly disappear. Most readers of this article will not become active participants in these complex projects, but they are worth keeping an eye on to understand when possible solutions might become available.[23]

Worrying too much about the long-term fate of digital materials in many ways puts the cart ahead of the horse. The average web page exists for a mere 44 days, after which it can never be reproduced. Instead of worrying about long-term preservation, most of us should focus on acquiring the materials in jeopardy in the first place and on shorter-term preservation horizons, 5 to 10 years, through well-known and effective techniques such as frequent backups stored in multiple locations and transferring files regularly to new storage media, such as from aging floppy discs to DVD-ROMs. If we do not have the artifacts to begin with, we will never be able to transfer them to one of the more permanent digital archives being created by the technologists.[24]

Taking First Steps

The importance of moving quickly to save extant digital materials is exceedingly evident in the case of 9/11. People turned to the Internet as a "commons"; it became a place to communicate and comment and share their feelings and perspectives. For example, nearly 20 million Americans used email to rekindle old friendships after 9/11. Thirteen percent of Internet users participated in online discussions after the attacks. People approached the Internet as a place to debate the United States government's response to terrorism (46 percent), to find or give consolation (22 percent), and to explore ways of dealing locally with the attacks and their aftermath (19 percent). Rather than in tangible diaries and letters, there was an outpouring of thoughts and emotions in thousands of blogs on September 11 and the following days, and in millions of emails and instant messages.

"For the first time," wrote one electronic newsletter editor, "the nation and the world could talk with itself, doing what humans do when the innocent suffer: cry, inform, and most important, tell the story together." Just four years later, many of these potent reactions already have been permanently lost in a discarded email or blog account, to willful or unconscious deletion, or on the unrecoverable magnetic surface of a crashed hard drive. Had the Library of Congress and its partners decided months later, instead of within mere hours, to save the web pages from 9/11 and immediately afterwards, many already would have vanished into the digital ether.[25]

Humans have always found ways to express their feelings and their history to each other and to a wide audience. Today this is being done increasingly in digital rather than analog forms, instantaneously to a vast global audience. In an age in which a significant segment of the record of modern life exists in digital form—a segment that will only grow in the years to come—ways will need to be found to capture digital documents, messages, images, audio, and video before they are altered or erased if our descendants are to understand how we lived. A future in which the cultural heritage community does not make extensive use of digital technologies as part of their mission is difficult, if not impossible, to imagine. Much more can, and must, be done if those interested in preserving a robust historical record are to fulfill their mission in the 21st century.[26]

Daniel J. Cohen is an assistant professor in the department of history and art history, and the director of research projects at the Center for History and New Media at George Mason University. He can be reached at dcohen@gmu.edu.

I would like to thank Roy Rosenzweig for his critical input and collaboration in the development of this article. Much of the article springs from work on

our book *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web* (Philadelphia: University of Pennsylvania Press, forthcoming in 2005). I would also like to thank Tom Scheinfeldt for equally valuable collaboration, for his thoughts and writings on this topic and for his incredibly nimble direction of the September 11 Digital Archive and the Echo Project. I have been helped greatly by the research assistance of a number of my other colleagues at the Center for History and New Media at George Mason University, especially Joan Fragaszy.

Notes

1. Oral History Research Office, Columbia University, *September 11th 2001 Oral History Narrative and Memory Project*, http://www.columbia.edu/cu/lweb/indiv/oral/sept11.html, accessed May 11, 2005; Mary Marshall Clark, "The September 11, 2001, Oral History Narrative and Memory Project: A First Report," *Journal of American History* 89 (September 2002): 569-579.

2. American Folklife Center, The Library of Congress, *September 11, 2001 Documentary Project*, http://www.loc.gov/folklife/nineeleven/index.html, accessed May 11, 2005.

3. The Library of Congress, *September 11, 2001 Documentary* Project in the American Memory Collection, http://memory.loc.gov/ammem/collections/911_archive/, accessed May 11, 2005.

4. National Institute of Standards and Technology, "Documentary Information Received by NIST," http://wtc.nist.gov/media/docs_info_received.htm, accessed May 11, 2005.

5. *Here is New York: A Democracy of Photographs*, http://www.hereisnewyork.org.

6. The Library of Congress, *The September 11 Web Archive*, http://www.loc.gov/minerva/collect/sept11/, accessed May 11, 2005; Netcraft, *Netcraft Web Server Survey*, September 2001, http://www.netcraft.com/Survey/index-200109.html, accessed May 11, 2005.

7. The Center for History and New Media at George Mason University and the American Social History Project/Center for Media and Learning at the Graduate Center at the City University of New York, *The September 11 Digital Archive*, http://911digitalarchive.org, accessed May 11, 2005.

8. According to the Netcraft web server survey in February 2005. See http://news.netcraft.com/archives/web_server_survey.html, accessed May 11, 2005, for the latest numbers.

9. On the vast digital corpus of the Clinton White House, see Adrienne M. Woods, "Building the Archives of the Future," *Quarterly* 2, no. 6 (December 2001), http://www.nasm.si.edu/research/arch/temp/marac/the_quarterly/Dec2001.html, accessed May 11, 2005. On the strange potential of future digital archives to contain either enormous numbers of documents or very few, see Roy Rosenzweig, "Scarcity or Abundance? Preserving the Past in a Digital Era" *American Historical Review* 108, no. 3 (June 2003): 735-762. For more on how to create a successful online collecting project, and how to promote it, see Daniel J. Cohen and Roy Rosenzweig, *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web* (Philadelphia: University of Pennsylvania Press, forthcoming in 2005), chapters 5-6.

10. As quoted in Lee Dembart, "Go Wayback," *International Herald Tribune*, March 4, 2002, http://www.iht.com/articles/2002/03/04/itend04_ed3_.php, accessed May 11, 2005.

11. Michael Kazin, "12/12 and 9/11: Tales of Power and Tales of Experience in Contemporary History," *History News Network*, September 11, 2003, http://hnn.us/articles/1675.html, accessed May 11, 2005.

12. Linda Shopes, "The Internet and Collecting the History of the Present," paper presented at *September 11 as History: Collecting Today for Tomorrow*, Washington, DC, September 10, 2003. For more on this "rapport" and the way rich historical accounts arise during the live interaction of interviewer and interviewee, see Alessandro Portelli, *The Battle of Valle Giulia: Oral History and the Art of Dialogue* (Madison: The University of Wisconsin Press, 1997) and Michael Frisch, *A Shared Authority: Essays on the Craft and Meaning of Oral and Public History* (Albany: State University of New York Press, 1991).

13. Keith Whittle, *Atomic Veterans History Project*, http://www.aracnet.com/~pdxavets/, accessed May 11, 2005.

14. There are two main types of blogging systems: those hosted on one's own server and those hosted on a blog company's server. Certain versions of both types are free, though there are also paid versions that have more features. By far the three most prevalent hosted blogging systems are Blogger, owned by Google, http://www.blogger.com, accessed May 11, 2005, LiveJournal, run by a small team of software developers and staff, http://www.livejournal.com, accessed May 11, 2005, and AOL Journals, owned by TimeWarner, http://hometown.aol.com, accessed May 11, 2005. Although it exists in a commercial version, LiveJournal can also be downloaded for free and installed on your server. LiveJournal and Six Apart's Movable Type, http://www.moveabletype.org, accessed May 11, 2005, are the predominant do-it-yourself blogging systems (Six Apart also runs a commercial hosting service for Movable Type blogs called TypePad, http://www.typepad.com, accessed May 11, 2005).

Many other free and commercial blog sites and programs (including the open source WordPress) exist for those who find the dominant software and hosts too basic, or who demand other features like message encryption or the automatic resizing of images for web display. A full list of software packages can be found at http://en.wikipedia.org/wiki/Weblog# Blogging_systems, accessed May 11, 2005, and hosts for a variety of blog packages can be found at http://directory.google.com/Top/Computers/Internet/On_the_Web/Weblogs/Tools/Hosts/, accessed May 11, 2005. In general, however, the top four systems will be suitable in most cases. Two of the more sophisticated blogging systems are the free *Nucleus CMS*, http://www.nucleuscms.org/, accessed May 11, 2005, and pMachine's *ExpressionEngine*, http://www.pmachine.com/expressionengine/, accessed May 11, 2005.

More information on blogs can be found in National Institute for Technology & Liberal Education, "Market Share," *NITLE Blog Census*, http://www.blogcensus.net/?page=tools, accessed May 11, 2005; Weblogs Compendium, "Blog Tools" and "Blog Hosting," http://www.lights.com/weblogs/tools.html, accessed May 11, 2005, and http://www.lights.com/weblogs/hosting.html, accessed May 11, 2005.

15. My rough estimate is that over the last 4 years less than 10 percent of the nearly 1,000 submissions to the Echo Project, a set of experimental online collecting efforts in the recent history of science, technology, and industry, have been off-topic or suspect. Center for History and New Media, *Echo Project*, http://echo.gmu.edu, accessed May 11, 2005.

16. San Jose Mercury News, "Web Newspaper Registration Stirs Debate," Mercurynews.com, June 13, 2004, http://www.mercurynews.com/mld/mercurynews/8915529.htm, accessed June 2, 2005. Online collecting projects that focus on sensitive topics obviously may encounter more resistance to revealing accurate personal information. See R. Coomber, "Using the Internet for Survey Research," *Sociological Research Online*, 2, no. 2 (1997), http://www.socresonline.org.uk/2/2/2.html, accessed May 11, 2005.

17. The American Registry for Internet Numbers has a free IP lookup service at http://www.arin.net/whois/, accessed May 11, 2005. Non-U.S. domains (those with two-letter country codes at the end) can be located through Uwhois.com, http://uwhois.com, accessed May 11, 2005. Domains that end in .aero, .arpa, .biz, .com, .coop, .edu, .info, .int, .museum, .net, and .org can be located through the governing body for the web, the Internet Corporation for Assigned Names and Numbers (ICANN), at http://www.internic.net/whois.html, accessed May 11, 2005. There are several commercial services that scan worldwide IP addresses, e.g., Network-tools.com, http://network-tools.com/, accessed May 11, 2005, and Network Solutions, http://www.networksolutions.com/en_US/whois/index.jhtml, accessed May 11, 2005.

18. Donald A. Ritchie, *Doing Oral History: A Practical Guide* (2nd ed.; Oxford: Oxford University Press, 2003), 32.

19. Moving Here, *Moving Here: 200 Years of Migration to England*, http://www.movinghere.org.uk/, accessed May 11, 2005; BBC, *WW2 People's War*, http://www.bbc.co.uk/dna/ww2/, accessed May 11, 2005.

20. National Park Foundation, "Rosie the Riveter Stories," *Ford Motor Company Sponsored Programs*, http://www.nationalparks.org/proudpartners/partner_ford_rtrs.shtml, accessed May 11, 2005; National Geographic, *Remembering Pearl Harbor*, http://plasma.nationalgeographic.com/pearlharbor/, accessed May 11, 2005; *Voices of Civil Rights*, http://www.voicesofcivilrights.org, accessed May 11, 2005; The Alfred P. Sloan Foundation, "History of Science and Technology," http://sloan.org/programs/scitech_historysci.shtml, accessed May 11, 2005; *C250 Perspectives: Write Columbia's History*, http://c250.columbia.edu/c250_perspectives/write_history/, accessed May 11, 2005; *Vietnam Archive Oral History Project* http://www.vietnam.ttu.edu/oralhistory/participation/index.htm, accessed May 11, 2005.

21. Fred R. Byers, *Care and Handling of CDs and DVDs: A Guide for Librarians and Archivists*, (Washington, DC: Council on Library and Information Resources, 2003), http://www.clir.org/pubs/reports/pub121/sec4.html, accessed May 11, 2005; Peter Svensson, "CDs and DVDs Not So Immortal After All," Associated Press, May 5, 2004, http://www.usatoday.com/tech/news/2004-05-05-disc-rot_x.htm, accessed May 11, 2005; Basil Manns and Chandru J. Shahani, *Longevity Of CD Media, Research At The Library Of Congress* (Washington, DC: Library of Congress, 2003), http://gort.ucsd.edu/mtdocs/archives/diglet/001522.html, accessed May 11, 2005; Eva Orbanz, Helen P. Harriso, and Henning Schou, eds., *Archiving the Audio-Visual Heritage: A Joint Technical Symposium* (Berlin: Stiftung Deutsche Kinemathek, 1988); Diane Vogt-O'Connor, "Care of Archival Compact Discs," *Conserve O Gram*, 19/19 (Washington, DC: National Park Service, 1996), http://www.cr.nps.gov/museum/publications/conserveogram/19-19.pdf, accessed May 11, 2005.

22. Margaret Hedstrom, "It's About Time: Research Challenges in Digital Archiving and Long-Term Preservation," white paper sponsored by the National Science Foundation and the Library of Congress, August 2003, 8. For more on the challenges libraries and archives currently face with the proliferation of digital artifacts and collections, see Daniel Greenstein, Bill Ivey, Anne R. Kenney, Brian Lavoie, and Abby Smith, *Access in the Future Tense* (Washington, DC: Council on Library and Information Resources, 2004), http://www.clir.org/pubs/reports/pub126/contents.html, accessed May 11, 2005, and *Building a National Strategy for Preservation: Issues in Digital Media Archiving* (Washington, DC: Council on Library and Information Resources and the Library of Congress, 2002), http://www.clir.org/pubs/reports/pub106/contents.html, accessed May 11, 2005.

23. *DSpace*, http://www.dspace.org; fedora, http://www.fedora.info, accessed May 11, 2005; National Digital Information Infrastructure and Preservation Program, http://www.digitalpreservation.gov, accessed May 11, 2005.

24. On the longevity of web pages, see Gail Fineberg, "Capturing the Web," *Library of Congress Information Bulletin* 62, no. 4 (April 2003), http://www.loc.gov/loc/lcib/0304/digital.html, accessed May 11, 2005.

25. Pew Internet and American Life Project, *The Commons of the Tragedy* and *How Americans Used the Internet After the Terror Attack* (Washington, DC: Pew Internet and American Life Project, 2001). Quotation from *The Commons of the Tragedy.* See also Amy Harmon, "The Toll: Real Solace in a Virtual World: Memorials Take Root on the Web," *New York Times*, (September 11, 2002): G39.

26. For more on the growth of Internet usage, especially as a place for communication, expression, and dialogue, see Deborah Fallows, *The Internet and Daily Life* (Washington, DC: Pew Internet and American Life Project, 2004).