

Atminties komunikacija archyvuose, bibliotekose ir muziejuose

Tarptautinės mokslinės konferencijos medžiaga



Communication of memory in archives,
libraries and museums

Proceedings of the international scientific conference

VILNIUS UNIVERSITY
INSTITUTE OF LIBRARY AND INFORMATION SCIENCE

COMMUNICATION OF MEMORY IN ARCHIVES,
LIBRARIES AND MUSEUMS: THE INTERACTION
OF SCIENCE, POLICY AND PRACTICE

4–5 October 2007, Vilnius

Proceedings of the international scientific conference

Vilnius University Press
2008

CIRCULATING CULTURE FOR THE KNOWLEDGE CONTINUUM: LIVING HISTORY, DIGITAL HISTORY AND THE HISTORY WEB

DR MARIJA DALBELLO

Rutgers, The State University of New Jersey,
School of Communication, Information and Library Studies
Department of Library and Information Science, New Brunswick, USA,
e-mail: dalbello@scils.rutgers.edu

This paper surveys the cultural record in the digital environments and the current efforts to capture this record and circulate it as knowledge, documents, and collections in memory institutions, and provide a basis for the creation of new knowledge. The goals of digital preservation are interpreted in the light of recent arguments about the role of the humanities in providing access to the complete human experience, of the changing idea of the archive representing that experience, and of the roles of memory institutions in supporting the humanities project. Two sets of current preservation activities are identified and surveyed – web archiving (of national web spaces, web spheres) and curated collections of primary sources from the history web. The emerging forms of interpretive and point-of-view history, invented archives, and digital libraries capturing local history, everyday experience and community memory illustrate how digital media can support interpretive and multi-perspective historiography.

Keywords: digital history, digital preservation, memory institutions, history web

Introduction

Tradition is memory that has become historically aware of itself. [29]

What is the future in general of the humanities and arts when the former seems destined only for what information industries call “content” and the latter for “multimedia entertainment”? [25]

The humanities preserve our cultural legacy – not as a collection of static artifacts, but stimuli to acts of interpretation. [14]

In his famous essay on the history of art as a humanistic discipline, Erwin Panofsky states that all humanities are historical because they are involved in “salvaging records from the streams of time,” of managing culture and the construction of history [32]. More recently, Johanna Drucker [14] identifies preservation and interpretation as core programs for the humanities. More precisely, in Alan Liu’s words, “preservation conserves and curates, while interpretation – especially with its contemporary humanistic connotations of questioning and destabilizing (descended from the genealogy of Hegelian antithesis, the modernist avant-garde, postmodern punk, and so on) – critically reevaluates” [26, 251]. Thus, preservation and interpretation enable personal and collective identity to be formed through access to experiences of the self, others, and the past, as well as critical

engagement with such experiences articulated in the program for the humanities.

Organized memory collections (such as digital libraries) and the work of managing the emerging cultural record in the digital realm and its circulation in the knowledge continuum present the model for the humanities as a type of knowledge work. The problem of managing memory in large technical systems to support the humanistic endeavor of interpretation, is tied to the practice of information work and its traditional concern with content and information, rather than with culture. Bridging the gap between knowledge and cultural artifacts is a challenge for information workers as they deal with digital futures. Building meaningful collections calls for rethinking historiography in the digital realm, that is the purpose and uses of digital cultural records.

Let us start with the current content and the growing mass of cultural heritage dispersed through the public web, in addition to digitization efforts, web archiving, scholarly archives, and a range of digital collections in addition to commercial entities. The institutional forms and institutionally collected websites are discontinuous knowledge spaces because grassroots culture is only partially captured within institutional frameworks – as documents, records and collections. Thus, it is important to consider how an emerging heritage can provide the basis for creating new knowledge. This can be done by studying the efforts of the communities of researchers and practitioners in shaping the History Web, learning about it, and using it to create new knowledge. This essay examines some of the attempts at archiving the web as well as collections that exemplify the transmission of culture in the Knowledge Continuum*. They are considered as two approaches to capturing a digital record of culture, and the “creation of new knowledge ... [that] forms part of a continuum which has no beginning and no end” [39].

Archiving the web: national web spaces and web spheres

One of the key functions of archives is to enable historians to understand what happened in the past. The Internet is the most recent technology to create a record of the past. Although it has been recognized that collecting digital records falls within the purview of memory institutions, there are problems inherent to channeling web-based cultural artifacts into the Knowledge Continuum, as well as there are specific challenges for the practice of “a historian’s craft” [3] in the context of digital media. While the Internet preserves and facilitates the articulation of social spaces and new genres, it is sometimes perceived as a source of misleading information that is either problematic or irrelevant for the Knowledge Continuum.

* The term Knowledge Continuum in its original context:

The Knowledge Continuum assesses the various elements that contribute to scholarly communication and seeks to provide a continuum of support services whereby the identification, absorption, utilization, and manipulation of existing knowledge merge with the organization, creation, and dissemination of new knowledge. The concept sought to capitalize on, and extend, an organizational structure which was already formally in place. Thus, while the NIST [National Institute of Standards and Technology] research library and information support resources and services as well as the technical publications and information services were brought together organizationally within the Office of Information Services in 1983, it was not until 1993 that a major concerted effort was initiated to complete the unification process through the creation and adoption of the concept of a Knowledge Continuum. [39]

Additionally, web-based material is notoriously ephemeral and continually updated, two features that present a serious obstacle to building and preserving perpetual archives [17, 46]. While the work of preserving digital memory in the context of heritage (memory that becomes “historically aware of itself” [29]) is fraught by perpetual change resulting in an overflowing archive, there have been some notable efforts to capture and archive the cultural record of the web and impose unifying principles of provenance and authority in forming digital collections.

That work has been under way for some time. This paper will highlight three major types of initiatives for archiving the web and its content for access as public historical record and explore their implications for historiography. This overview focuses on the public web, excluding the commercial efforts that have also resulted in collections of significance. Although significant, such efforts that are outside the public domain remain poorly documented as pointed out by Paul Duguid in his evaluation of the Google Books project [15].

Archiving the International Public Web: The Internet Archive

The best-known project for capturing a record of the web since 1996 is the *Internet Archive*, “a digital library of Internet sites and other cultural artifacts in digital form” [20]. Developed and maintained as an independent effort through a philanthropic initiative of the Internet pioneer Brewster Kahle, the current collection of 85 billion historical web pages of the *Internet Archive* has become a multi-terabyte repository of web content and the most extensive existing effort to archive the indexable web. The major components of the *Internet Archive* are successive incarnations of published web pages (accessible to the public since October 24, 2001). The collection is carried out by *Alexa Internet*, a web navigation service gathering over 100 gigabytes of data a day from around the world (including publicly accessible World Wide Web pages, the Gopher hierarchy, the Netnews bulletin board system, and downloadable software). As stated at the *Internet Archive* website the content of the collection is by no means filtered or selected: it may contain information that might be deemed offensive, disturbing, pornographic, racist, or otherwise objectionable*.

Long-term preservation at multiple sites diminishes the risk of accidents and natural disasters destroying the data. The problem of obsolescence is countered by frequent migration of data to new storage tapes. This also includes the collection of software emulators to aid future researchers, historians, and scholars. The resource is free and open to the public.

Although it is still a largely under-utilized resource, the *Internet Archive* is becoming increasingly relevant for historians of the web spaces and has attracted institutional partners in joint projects (such as the Smithsonian Institution, and the Library of Congress). The *Internet Archive* has not grown to a prominence that makes it essential for using the international and long-term record of the Internet. Because it mirrors the structure of the web, researchers have been using the *Internet Archive* as a laboratory

* The collection consists of publicly available documents only, thus avoiding infringement of copyright laws. If there is any indication that their owners do not want them archived or employ robot exclusion mechanisms, these restrictions are obeyed, as is the possibility of manual removal upon request [20].

for research teams to conduct studies on retrieval methods since 1996, long before this resource became public in 2001. Although technical knowledge and programming skills are needed for creating research corpora, there are also benefits for general historical (and legal) research that involves verification, authentication and retrieval. In its public form, the *Internet Archive* allows for retrieval of multiple historical versions of web pages by means of the "Wayback Machine" search engine. The *Internet Archive* is also used to track the transformation of websites. For example, in a recent study of the transformation of publicly accessible pages of the National Digital Library Program at the Library of Congress from 1996 to 2002, it allowed a reconstruction of successive versions of the *American Memory* site and how it has been represented on the Library of Congress home page, to supplement interpretive work on the history of a national digital library [9; 10; also 6; 7].

The *Internet Archive* provides access to a global digital heritage while it preserves and mirrors its diversity. It is a representation of the international public web.

Other initiatives that use a harvesting approach to build collections of digital artifacts work within set boundaries of national web spaces, which are stored in large repositories that contain snapshots and excerpts of Web data. These projects are discussed next.

National Web Spaces

In her recent article on *PANDORA*, Australia's Web Archive [31], Margaret Phillips establishes two broad national approaches to collecting and preserving online publications that emerged in the mid-to late 1990s, "as a small number of national libraries began archiving programs and exploring different approaches to archiving national documentary heritage online ... the whole domain or comprehensive approach, and the selective approach" [33]. Two representative early systems for archiving national web spaces are represented by the Swedish and the Austrian initiatives. These initiatives are by no means the only ones. Complementing numerous studies focusing solely on the Internet Archive [37], these projects have been described in computer science literature on website archiving.

The Swedish Archive initiative *Kulturav3* pioneered the harvesting of national web spaces from 1996. *AOLAP Austrian On-Line Archive* went public in 2001 [1]. *AOLAP* presents a well-documented case study for understanding such a national initiative of creating an on-line archive [34; 35].

The components of a Web archive are the content, methods of harvesting and provision of access to this content. The types of materials collected include personal web pages. They are harvested using web crawlers (exemplified by *NEDLIB Project*, *Alexa* and *Combine* harvesters) to identify a particular domain and then to pick up all the linked documents (similar to a suction of interrelated networks). The technical challenges of acquisition, archiving, and long-term storage and preservation of data are several components discussed in the sizable literature accompanying the development of web archives. In that same literature that focuses on technical solutions and policies for archiving, it is not yet clear how these archives have been used, beyond their implied uses as the passive record of national web spaces.

Collaborative frameworks for sharing information and experience among the different national projects have been available since 1997. One such early collaborative initiative is *The Nordic Web Archive*, “an early Nordic national libraries’ forum for coordination and exchange of experience in the fields of harvesting and archiving web documents” [30]. This project started in 2000 and ended in 2002 [18]. The other significant initiative from that era was the *International Research on Permanent Authentic Records in Electronic Systems InterPARES 1* (1999–2001), followed by *InterPARES 2* (2001–2006), a Canada-based project for knowledge sharing [21]. The international collaborative 12-member consortium* *International Internet Preservation Consortium (IIPC)* established in 2003 (documented at *netpreserve.org*), has most recently (2007) opened its membership “to libraries, archives, museums and cultural heritage institutions everywhere that are interested in collecting and preserving Internet content” [28]. The early experiences in web archiving resulted in a sizeable literature of reports, global involvement, and evolving international efforts for the preservation of cultural records on the web.

The experimental stage in web archiving that established a few initial models of capturing national heritage space on the web from 1996–2006 has produced two distinct models – the comprehensive and selective approaches. In addition to these initiatives, national and international, there are efforts best described as the topical approach to identify relevant slices of the web, also known as web spheres.

Topical Approach to Web Archiving: Web Spheres of the Political Web

The web spheres approach for capturing the cultural record of the web is linked to collaborative initiatives of political scientists, technologists and librarians (in the United States). This was done through isolating web spheres of event-focused records. The projects resulting from such an approach are recorded on *MINERVA*, Library of Congress Web Archiving & Preservation site [27; see also 24]. Web sphere analysis originates from two social scientists, political science researchers Kirsten Foot at the University of Washington and Steven M. Schneider at SUNY Institute of Technology (Albany, New York), who founded *WebArchivist.org*. Through a collaborative effort with the *Internet Archive*, the Library of Congress, and the Pew Internet & American Life Project, they launched an initiative to record the sites that sprung up in response to the attacks on the World Trade Center in New York on September 11, 2001 [36]. This project resulted in the first topically organized web archive and is a combination of a commemorative archive and research corpus for political science research that utilized the technologies of web harvesting developed by Brewster Kahle (founder of the *Internet Archive*), and the partnership with the Library of Congress.

Ever since, the *WebArchivist.org* initiative has provided the impetus for diverse efforts “to identify, archive and catalogue born-digital materials relevant to Web spheres of interest ... [to] analyze web spheres for publications and presentations ... [and to measure] breadth, scope, emergence, features, structures for action and other aspects”

* The charter members of the consortium include the national libraries of Australia, Canada, Denmark, Finland, France, Iceland, Italy, Norway, Sweden, the British Library (UK), the Library of Congress (USA) and the Internet Archive (USA) [28].

[40]. This approach has defined a web sphere as an archival and historical record and is the foundation for considering web analysis as a humanities project as identified at the outset of this essay. The web sphere as a type of collection demonstrates the principle that preservation supports interpretation. It not only presents the multiplicity of naturally occurring perspectives but it also allows for access to the experiences of self, others, and the past, to critical engagement and personal expression experiences as articulated in Drucker's program for the humanities.

Because the web sphere approach aims to preserve official and unofficial sites and to explore all sides of an issue, it provides the basis for interpretation, research, and debate. An experimental collaboration of web archivists with the Library of Congress thus formed an official heritage collection, and initiated rethinking the archive as a space for political sphere analysis. In fact, this effort combined grassroots academic (*WebArchivist.org*), with philanthropic (*Internet Archive*), and public institutional resources (*Library of Congress*).

The first attempt at archiving the Political Web captured sites connected to 9/11 events. As stated on the occasion of its official launching on September 10, 2002, this web archive remains the Library of Congress' "first major digital acquisition of September 11, 2001, materials with the addition to its collections of *The September 11 Digital Archive*" [13]. It was seen as a heritage project, as reported in this USA Today news item from November 10, 2001:

The Web archive will include thousands of sites, ranging from news to charity efforts and memorials after the attacks. ... Diane Kresh of the Library of Congress says the archive will be invaluable, especially for future scholars: "It's a snapshot of what people were thinking at a very critical time in this nation's history. Reflections are coming from all around the globe. It's truly a global response to a single tragic event that's touching off a new way of life for Americans."... Kirsten Foot, an assistant communications professor at the University of Washington who is working on the Web archive, agrees. The attacks are "part of history now. It's important to go back and reflect on what went on, what people did," she says. [22]

Other web capture projects (listed on *MINERVA*, Library of Congress Web Archiving & Preservation site) include Election 2000, Election 2002, and 107th Congress Web archive (with several others in production). Together, these web archiving projects have grown to more than 70 terabytes of data (according to May 2007 statistic) [24]. These efforts are integrated in the *National Digital Information Infrastructure and Preservation Program (NDIIPP)* at the Library of Congress.

As noted earlier, the Political Web has an inherently dialectic structure that is transferred to archiving the web spheres and topical harvesting*. The national libraries through their preservation and heritage initiatives have been at the forefront of developments described so far. Yet let us not forget that the harvesting and archiving technology was pioneered outside of memory institutions (as exemplified by the *Internet Archive*), and that collecting cultural artifacts through web archiving depends on innovative tech-

* The 2004 report of the Center for Research Libraries examines American and European efforts at archiving web-based political communications. The report outlines the archiving policies and procedures, and the process for assessing, capturing and storing web sites for the Political Web Project [4].

nologies for data warehousing.

Despite claims that web sphere analysis has revolutionized cyber-culture studies, such studies have yet to be fully manifested. Web archiving applications are just beginning.

Curated collections of the history web

In contrast to web archiving focusing on the record of the most recent past of post-1996 era, digital heritage is increasingly being created through digitization programs of primary source material and historical commentary. They provide digital continuity for pre-World Wide Web era materials; the technologies of access allow for reinterpretation of primary sources and improve access for scholars and the public at large. Others are the collections of primary source materials and history websites produced by libraries and local history institutions and by scholars (literary and historical scholarly archives developed by researchers for their own use). Added here can also be historical websites developed and used by genealogists and enthusiasts. Although these collections often provide old materials in the new medium, the digital libraries feature innovative and novel presentation. In the text that follows, I will highlight a few of resources of the curated web. Some of them extend and even challenge the existing ideas of collections and the archive in the digital medium, while others conform to existing models for collections.

Research Collections Online: National Digital Libraries

Curated memory spaces in research and national digital libraries include full-text access or web exhibits and searchable databases created in the first wave of digital library development from 1996 to 2002. These digitized collections are for the most part providing more access to the existing research collections preserved in research libraries and other collections of research materials. Digitization has presented opportunities for smaller libraries and archives to get their materials online, thus allowing for often obscure materials to become more accessible. Primarily featuring historical treasures, visual and public domain material, these collections have increased access to research collections, opening them to educational purposes and to the general public. The *American Memory* collections at the Library of Congress present a case in point with the extension of its collections' user base to primary school students, who were previously not served by the Library of Congress collections [9; 10]. More radical uses of digital library technology by public libraries and local history collections to manage community memory are another recent development.

Living History and the Digital Library: Collecting Community Memory

Digital library projects that connect collections of documentary materials to communities and their memories represent a very special type of material but are relevant in this discussion about how to disseminate culture for the Knowledge Continuum. Although not limited to digital libraries, the ability to record "living memory" and

circulate it through the public web to intensify and stimulate the connection of collections to the everyday and historical experience of communities has been extraordinary. Some of the earliest community memory projects, from a sample of early digital library projects [8], are represented by Brandeis University's, "A People's History: 1999," and "Bridgeport Working: Voices from the 20th Century" (Bridgeport Public Library). Similar to these are: "American Life Histories: Manuscripts from the Federal Writers' Projects, 1936–1942" (Library of Congress), "African Voices" (Smithsonian Institution), "Talking History: Labor History Archive" (The University at Albany, State University of New York), and "Bioscience and Biotechnology in History" (University of California–Berkeley Bancroft Library). What they have in common is the extension of oral history to collect and record memory narratives. Others, such as "Benedicte Wrensted: An Idaho Photographer in Focus" (Idaho Museum of Natural History), have been instrumental in initiating a dialogue about memory objects between institutions and living memories transmitted in local communities:

One of the goals of this exhibition has been to demonstrate the ways in which photographs, even those a century old, can be placed in historical context. Only 1% of the Wrensted images at the NARA [National Archives and Records Administration] were identified at the onset of the project. Once they were shown to the descendants at the Fort Hall Indian Reservation, the families of origin were discovered. Individual names were recovered from written records, and today 84% of Wrensted subjects have been identified. (From original description of "Benedicte Wrensted: An Idaho Photographer in Focus" in the ARL Digital Initiatives Database at: <http://www.arl.org/did>; accessed June 31, 2001)*.

In this example, digitized photographs reconnected a Native American community in a cross-generational memory loop. The project was initiated at the Idaho Museum of Natural History, and is now preserved at the University of Virginia [2].

The emerging form of living library (interactive, cross-temporal digital collections) can stimulate the construction by shared storytelling of vernacular histories and the transfer of cross-generational knowledge about historical artifacts. It can supplement genealogy research and aid in the construction of community resource databases. Facilitating memory processes through interactive archives of historical documents is a way of preserving local knowledge and recognition of common heritage. The technologies and tools for digital storytelling support access to large digital oral history archives, and collaborative work in online communities.

The previous cases emphasize how memory institutions can become active participants in shaping historical records by offering ways for interaction with existing collections of knowledge bases of local subjects and neighborhoods. Memory institutions are central in conceptualizing locality and empowering communities. Digital libraries could become sites and agencies for such knowledge production by creating dynamic memory spaces for local heritage and access to primary source material without the customary protocols of research and access.

* The database is not archived in the Internet Archive but a version is available from ARL upon request. The description of the database in the Internet Archive is at: <http://web.archive.org/web/20060108004048/www.arl.org/did>.

Digital libraries provide the context for such knowledge production. In that sense, one may agree with Derrida that the archive is future oriented because it defines the context for transmission of texts.

The archive has always been a pledge, and like every pledge, a token of the future. To put it more trivially: what is no longer archived in the same way is no longer lived in the same way. Archivable meaning is also and in advance codetermined by the structure that archives. [12, 18]

Once again the notion of archive as “a pledge” to the future harks back to the formula that preservation and dialectic interpretation are the basis for humanistic engagement and Drucker’s program for the humanities. The need for the memory institutions to engage in social entrepreneurship is part of the pledge of the archive to support the identities of communities and social selves.

History Websites and Invented Archives

In their recent overview of digital history web, Dan Cohen and Roy Rosenzweig [5] surveyed the 32,599 history websites from Yahoo’s directory, identifying the following categories: organizational websites (museums, historical societies, historic sites), websites of genealogical societies and history departments, websites created by history enthusiasts, online syllabi, and online journals. Further, the ten-year record of the history web includes archives (primary sources), exhibits, films, scholarship, essays (secondary sources), teaching resources for students and teachers, discussion online (dialogue), and organizational information about historical groups.

They have noted that these resources had transformed the traditional idea of the archive and that many of these collections do not any more share provenance in the sense of “firm history of the custody of a coherent body of materials since their origin” [5, 25]. And they commonly mix published and unpublished sources. Expanded access to primary sources through searching is significant but it entails difficulty to find specific data and it involves cross-collection searching. The new user base comprises students and non-academics; the creators of these archives are likely to “think out of the box.” The implications for historiography are not only in the much larger accessibility of primary source material, but in the possibilities for building unconventional historical explanations, applying inventive design, and using unique approaches and unusual primary sources to provide new interpretations of history. There are also problems with grassroots digital history sites because they may present traditional historiography, old theories, and long discredited arguments, and some may even rehash old prejudices [5].

A striking example of innovation and the transformation of the idea of the archive is the invented archive [5, 29], an example of which is the *Valley of the Shadow* (from *The Institute for Advanced Technology in the Humanities*) [19]. This digital archive comprises 1400 letters and diaries, complete census records from 1860, 45 GIS maps, and 700 photographs and images, together with an implicit interpretation of materials [5; 38] that form a compendium of documents about two Civil War communities and their two viewpoints, one Northern and one Southern. The effect is unique, because the collection not only provides comprehensive access to primary source materials, but it

blurs the usual carefully maintained distinction between an archive and a historical argument. In this example, the archive integrates an interpretive dialectic of history. Thus, the *Valley of the Shadow* helps demonstrate how invented archives contribute to historical understanding and offer a point-of-view perspective.

The invented archives can provide material for dialectic interpretation, present parallel perspectives, and also incorporate features for interactivity and intervention in the archive through personal reflections and integration of pedagogical perspectives. The dialogue between the archive and inquiry learning may take form of solving mysteries and interrogating historical evidence more closely. History sites can remove the barrier between scholars and the public.

Implications for "a historian's craft"

Historians are engaged in interpretative work and their bibliographic methods are focused on text and primary sources. Improved access to text and expansion of the range of available texts and primary source materials provide opportunities for expanding the scope of web-based historical discourse. Primary sources are being collected in digital repositories in order to provide continuity for existing collections in curated digital collections. They aggregated to make searchable digital artifacts from "naturally occurring" historical record--the cultural objects found on the web. What then is the potential for historiography of these new types of collections?

Repositories of Web Pages and Web Bases

As shown in the examples presented earlier, web archive can be defined in terms of technologies that define their capabilities, and in terms of historiographic objectives. They are based on the technologies of data warehousing (DWH) that allow analysis and exploration from a different vantage point using OLAP (On-Line Archive Processing). The analytical capabilities of such systems are also defined by the capabilities of data warehousing technology⁵; thus, collecting is defined technologically. That archive⁶ is limited to web page documents.

The primary users of data warehouses could be web historians forging collaboration with technologists or seeking to acquire such skills as part of their training. Memory institutions could serve as facilitators of technologies, collections and their uses. Collecting vast amounts of data that provide snapshots of web spaces is a reality, although there are still legal and technical obstacles to enabling the use of these collections by the general public and the scholars.

⁵ The OLAP tools allow one to view and analyze the data from different perspectives and compute measures across many dimensions, for querying and analysis of web-linkage data storage. The information attributes of OLAP are facts and dimensions. They are variously organized for retrieval using large databases.

⁶ The comparison of examples of data repositories discussed here to the traditional archive is superficial because of problems surrounding selection, access, and the research purpose of these data repositories that have largely been developed around the genre of the web page.

Practicing "A Historian's Craft" from Point-of-View Perspective

This survey has highlighted examples of curated collections integrating interpretation and interactive components in their presentation of primary sources. This is in agreement with collection development in American librarianship that emphasizes commitment to provision of access to multiple points of view and to point-of-view evidence. With digital collections, such perspectival approach can be enhanced with interactivity. Interactivity is not only a technical component of the system; interactivity assumes interpretation of primary source material from multiple points of view.

Conclusion: digital memory in the knowledge continuum

The record of the first days of the Internet has been lost for the large part, and its full history needs to be reconstructed by other means [5]. Some of that digital past is preserved, through web archiving, as well as preservation of heritage collections and the history web. Archiving the web is based on understanding the web as structured record of historical experience and the process of disaggregation of that record into data. Reaggregation involves technology but also human processing by means of interpretive methods. The history web includes curated collections of primary source materials and institutional websites. We could see that new forms of memory are already being integrated in these archives through shared interpretations. Thus, these emerging forms of memory may contribute to the discussion of history and continuous interpretation of human experience participating in a humanities project outlined by Johanna Drucker.

Curating collections for a humanities project and could involve:

- Supporting point-of-view history and capturing viewpoints as the expression of users of digital media;
- Supporting grassroots journalism and providing place for the community to make spontaneous documentation;
- Creating collections for historical and institutional research that allow both for direct and incidental uses of material;
- Constructing the historical evidence to extend the conventional archive beyond the selectivity of paper collections and towards comprehensive cross-institutional archives and invented archives.

Current methods for web archiving and collecting need to be expanded to the grassroots culture of the social web. How can the collecting frenzy of social spaces and the blogosphere be transferred to archival records for the Knowledge Continuum (i.e., using existing knowledge and records to create new knowledge)? Some of these forms are developed on the commercial platforms that present problems of long-term availability and intellectual ownership. How are the vernacular photo-archives in *flickr* [16], social bookmarking (*del.icio.us*) [11], and shared personal library catalogs and booklists on *Library Thing* [23] relevant as cultural record? And how do the universal libraries such as Project Gutenberg, Million Books Project, Open Content Alliance's Open Li-

brary, and Google Books connect to institutionally preserved collections, literary canons, reading? Innovation in the commercial sector produces forms of record that do not have the assurances of quality found in traditional cultural institutions [15]. Researchers' web pages and researcher-developed scholarly archives and the open access repositories are other cultural forms that can be seen as part of the Knowledge Continuum and expression of a particular form of cultural experience.

Clearly, the work of institutional collecting has just begun. The initiatives that I have highlighted in this essay are just some of the attempts to preserve and interpret the emerging digital record. This essay excludes a number of commercial endeavors and Web 2.0. The historical timelines for these efforts are brief in an overall history of electronic media in libraries; it constitutes a fraction of the recent history of the Internet and the web since 1996^{*}; – brief indeed, especially when considered against traditional libraries' institutional histories and collections. Nevertheless, the "Memory in Digits" is changing the nature of scholarly communication and the humanities. That preservation of culture, social memory, and living history is a project for the humanities. This brief survey of the current status emphasized that the future of the preservation and circulation of culture for the Knowledge Continuum can be an opportunity for entrepreneurship for the memory institutions.

REFERENCES

1. *AOLA. Austrian On-Line Archive* [interactive]. [accessed 24 October 2007]. Access through Internet: <<http://www.ifs.tuwien.ac.at/~aola>>.
2. *Benedicte Wrensted: An Idaho Photographer in Focus* [interactive]. 1997 [accessed 24 October 2007]. Access through Internet: <<http://etext.virginia.edu/VAR/bw>>.
3. BLOCH, Marc. *The Historian's Craft*. New York: Vintage Books, [1984]. Originally published in 1953.
4. *Political Communications Web Archiving: An Investigation Funded by the Andrew W. Mellon Foundation* [interactive]. Center for Research Libraries, et al. June 2004. [accessed 24 October 2007]. Access through Internet: <<http://www.crl.edu/PDF/PCWAFinalReport.pdf>>.
5. COHEN, Daniel J.; ROSENZWEIG, Roy. *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web*. Philadelphia: University of Pennsylvania Press, 2006. 316 p.
6. DALBELLO, Marija. *Cultural Dimensions of Digital Library Development, Part I: Theory and Methodological Framework for A Comparative Study of the Cultures of Innovation in Five European National Libraries*. Forthcoming.
7. DALBELLO, Marija. *Cultural Dimensions of Digital Library Development, Part II: The Cultures of Innovation in Five European National Libraries (Narratives of Development)*. Forthcoming.
8. DALBELLO, Marija. Institutional Shaping of Cultural Memory: Digital Library as Environment for Textual Transmission. *Library Quarterly*, 2004, 74 (3), pp. 265–299.

^{*} This reference is to "Wayback machine" search engine of the Internet Archive, which was started in 1996.

9. DALBELLO, Marija. A Phenomenological Study of an Emergent National Digital Library, Part II: The Narratives of Development. *Library Quarterly*, 2005, 75 (4): pp. 28–70.
10. DALBELLO, Marija. A Phenomenological Study of an Emergent National Digital Library, Part I: Theory and Methodological Framework. *Library Quarterly*, 2005, 75 (4), pp. 391–420.
11. *del.icio.us social bookmarking* [interactive]. [accessed 24 October 2007]. Access through Internet: <<http://del.icio.us>>.
12. DERRIDA, Jacques. *Archive Fever: A Freudian Impression*. Chicago: Chicago University Press, 1996. 128 p.
13. The Library of Congress Report to the Digital Library Federation, Spring 2004. Digital Library Federation. In *DLF Newsletter* [interactive]. 2004 [accessed 24 October 2007]. Access through Internet: <http://www.diglib.org/pubs/news04_02/loc.htm>.
14. DRUCKER, Johanna. Humanities Games and the Market in Digital Futures. In *Criticism*, 2005, 47 (2), p. 241–247.
15. DUGUID, Paul. Inheritance and Loss? A Brief Survey of Google Books,. In *First Monday* [interactive]. 2007, 12 (August). [accessed 24 October 2007]. Access through Internet: <http://www.firstmonday.org/issues/issue12_8/duguid/index.html>.
16. *flickr* [interactive]. 2007 [accessed 24 October 2007]. Access through Internet: <<http://www.flickr.com>>.
17. GOULD, Sara; VARLAMOFF, Marie-Thérèse. Digital Disappearances. In *UNESCO Courier*, 2000 (October), p. 46–47.
18. HAKALA, Juha. *Archiving the Web: European experiences* [interactive]. Presented at CONSAL XII, 20–23 October 2003. [accessed 24 October 2007]. Access through Internet: <<http://www.lib.helsinki.fi/tietolinja/0203/webarchive.html>>.
19. *The Institute for Advanced Technology in the Humanities* [interactive]. [accessed 24 October 2007]. Access through Internet: <<http://www.iath.virginia.edu>>.
20. *Internet Archive* [interactive]. 2007 [accessed 24 October 2007]. Access through Internet: <<http://www.archive.org>>.
21. *InterPARES (International Research on Permanent Authentic Records in Electronic Systems)* [interactive]. [accessed 24 October 2007]. Access through Internet: <<http://www.interpares.org>>.
22. KORNBLUM, Janet. Web Archives Will Preserve Attacks, Our Reaction. In *USA Today* [interactive], 10 November 2001. [accessed 24 October 2007]. Access through Internet: <<http://www.usatoday.com/life/2001-10-11-web-archives.htm>>.
23. *Library Thing* [interactive]. [accessed 24 October 2007]. Access through Internet: <<http://www.librarything.com>>.
24. *The Library of Congress Web Capture* [interactive]. [accessed 24 October 2007]. Access through Internet: <<http://www.loc.gov/webcapture/index.html>>.
25. LIU, Alan. *The Laws of Cool: Knowledge Work and the Culture of Information*. Chicago: University of Chicago Press, 2004. 552 p.
26. LIU, Alan. Understanding Knowledge Work. In *Criticism*, 2005, 47 (2), p. 249–260.
27. *MINERVA (The Library of Congress Web Archiving & Preservation Project)* [interactive]. 1 October 2006 [accessed 24 October 2007]. Access through Internet:

- <<http://lcweb2.loc.gov/cocoon/minerva/html/minerva-home.html>>.
28. *Netpreserve.org: International Internet Preservation Consortium* [interactive]. 2007 [accessed 24 October 2007]. Access through Internet: <<http://netpreserve.org/about/index.php>>.
 29. NORA, Pierre. *Realms of Memory: The Construction of the French Past*. Vol. II. Traditions. New York: Columbia University Press, 1997. 591 p.
 30. *The Nordic Web Archive* [interactive]. 31 August 2005 [accessed 24 October 2007]. Access through Internet: <<http://nwa.nb.no/>>.
 31. *PANDORA, Australia's Web Archive* [interactive]. [accessed 24 October 2007]. Access through Internet: <<http://pandora.nla.gov.au>>.
 32. PANOFSKY, Erwin. The History of Art as a Humanistic Discipline. In *Meaning in the Visual Arts*. New York: Doubleday, 1955, p. 1–25.
 33. PHILLIPS, Margaret E. What Should We Preserve? The Question for Heritage Libraries in a Digital World. In *Library Trends*, 2005, 54 (1), p. 57–71.
 34. RAUBER, Andreas; ASCHENBRENNER, Andreas; WITVOET, Oliver. Austrian on-line Archive Processing: Analyzing Archives of the World Wide Web. In *Research and Advances Technology for Digital Technology : 6th European Conference, ECDL 2002, Rome, Italy, September 16-18, 2002*. Proceedings. 2002, p. 16–31.
 35. RAUBER, Andreas; Oliver WITVOET; Andreas ASCHENBRENNER; Robert BRUCKNER; and Max KAISER. Uncovering Information Hidden in Web Archives: A Glimpse at Web Analysis building on Data Warehouses. In *D-Lib Magazine* [interactive]. 2002, 8 (12) [accessed 24 October 2007]. Access through Internet: <<http://www.dlib.org/dlib/december02/rauber/12rauber.html>>.
 36. *The September 11 Digital Archive* [interactive]. 2007 [accessed 24 October 2007]. Access through Internet: <<http://911digitalarchive.org>>.
 37. THELWALL, Mike; VAUGHAN, Liwen. A Fair History of the Web? Examining Country Balance in the Internet Archive. In *Library & Information Science Research*, 2004, 26 (2), p. 162–176.
 38. *Valley of the Shadow: Two Communities in the American Civil War* [interactive]. 2007 [accessed 24 October 2007]. Access through Internet: <<http://valley.vcdh.virginia.edu>>.
 39. VASSALLO, Paul. The Knowledge Continuum – Organizing for Research and Scholarly Communication. In *Internet Research*, 1999, 9 (3), p. 232 – 242.
 40. *WebArchivist.org* [interactive]. [accessed 24 October 2007]. Access through Internet: <<http://webarchivist.org>>.