CHAPTER 2

Object-Oriented Web Historiography

Kirsten Foot & Steven Schneider

INTRODUCTION

In the fall of 2001, we worked with the U.S. Library of Congress and the Internet Archive identifying and collecting tens of thousands of websites, organizing these materials into the September 11 Web Archive (viewable at both http://september11.archive.org and http://lcweb2.loc.gov/diglib/lcwa/html/sept11/sept11-overview.html), and completing several scholarly articles and chapters based, in part, on analyses of archived materials. One of us recently attended a meeting with some of the staff of the National September 11 Memorial & Museum—the organization that is building a memorial and a museum at the site of the World Trade Center in New York. The purpose of the meeting was to explore ways of explaining the role of the web in the aftermath of the 2001 terrorist attacks in the United States. The September 11 Museum staff were interested in exploring how they might tell this part of their story. Although budget constraints prevented us from moving the project forward, the discussion illustrated for us some of the central issues of web historiography that we hope to highlight in this chapter. To start with, of course, the Museum staff wanted to understand what, how much, and how many: What had been archived? How much material had been collected? How many times had we collected each site? This led to a discussion of why: Why had we collected web materials? Why had we made the choices we did? Why did we start? Why did we stop? And finally, we turned to the notion of: How did the archiving software work? How were archived pages displayed? And how did our technologies and techniques influence what was available to be shared with others? The answers to these very practical questions begin to provide a helpful point of departure for this chapter.

As the chapters in this book illustrate, there are a variety of approaches to conducting web history, from the building of archives and developing tools for historical analyses, to retrospective studies of web production, to analyses of web-

mediated practices such as the use of webcams. We hope the ideas presented in this chapter have some value for practitioners of each of these approaches. We have three aims in this chapter: (1) to propose an "object-oriented" approach to researching and writing web history; (2) to suggest ways that such an object-oriented approach makes developmental analyses of the web more rigorous, and (3) to encourage social researchers to be proactive in building web archives, alongside archivists and librarians—as well as interfaces and tools for analysis to facilitate robust web historiography (and in this we concur with several other contributors to this volume). We come to this subject of web historiography as social researchers—we have novices' appreciation of and enthusiasm for history but do not share the kinds of expertise cultivated by historians and archivists. Whether it is an advantage to be free from the constraints of practice imposed by those professions or a disadvantage to be unaware of the epistemological and methodological controversies of these fields is a judgment yet to be rendered. We write as practitioners with experience initiating and completing studies of the web based largely on analyses of objects contained in archives of web objects we collected on our own and with other institutions, contemporaneously or prospectively during our studies. Our patterns of practice as web archivists, and our motivations to archive, have evolved since we began archiving in 1999, and we welcome the opportunity this chapter has afforded to reflect and critique, in a serious way, on what we have done and what our choices have meant for the types of web histories we have been able to write.

Our initial foray into web studies was born of an opportunity—we were both hired in 1999, independently, to work on a large, well-funded research project seeking to assess the impact of the web on the U.S. federal elections in 2000. We quickly concluded that a reasonable starting point would be to analyze the development of election-oriented websites over the course of several months leading up to the election. We recognized the impossibility of personally observing, on a daily basis, all or even a substantial percentage of election-related websites and determined that some sort of archive or collection of sites would be necessary to complete a post-hoc developmental analysis. The principal investigator we worked with was fond of doing experiments, so we pitched our idea for an archive as the mother of all stimuli for experiments on citizens' reactions to online election phenomena. He bought our pitch, so we jumped into the challenge of pro-

spective archiving to enable the retrospective analyses we wanted to be able to conduct.

Our experiences in that initial study of the 2000 election connected us to a small and eclectic group of scholars, librarians, and archivists who had similarly recognized the challenge—and desirability—of archiving web materials for future analysis and assessment. In the summer of 2001, we contacted staff at the U.S. Library of Congress (LC), who were at the time starting to create collections of born-digital materials thought likely to have historical significance. We discovered that the LC, not surprisingly, approached web archiving from the perspective of librarians: expert selection of materials with no attention to (or even recognition of) sampling strategies; cataloging basic meta-data about an entire collection, with no consideration (or even recognition) of the need for site-level cataloging, or tools to facilitate analysis. We sketched a plan for our newly emergent research group, WebArchivist.org, to work with the LC, and later the Internet Archive, to create web archives for social researchers as well as the U.S. Congress and the public at large. When terrorist attacks occurred in the United States a few weeks later, the three organizations, joined by the Pew Internet & American Life Project, quickly agreed to immediately implement our sketchy plan for collaboration. The result was the September 11th Web Archive, consisting of daily captures between September 11 and December 1, 2001, of all the URLs that the librarians at LC, the researchers in our group, and web users around the world added to the Internet Archive's seed list for this collection. We conducted a preliminary categorization of 2,500 sites, including creation of meta-data describing the type of producer and some assessment of content—this was a big advance for LC as it moved beyond the collection-level cataloging it had done for previous collections. We also conducted detailed analysis of 250 sites by examining the kinds of activities in which site visitors could engage. We developed two distinct interfaces to the archive to facilitate interaction with collected materials. Eventually, other researchers joined us in producing a compilation of reports on a range of post-9/11 web phenomena captured in the archive, and at least three journal articles were published based on materials from the archive. We have gone into our origins as web archivists and web historians in some detail because the observations that follow have been deeply colored and shaped by our experiences.

We refer to the September 11 Web Archive project throughout this chapter, because it was a formative experience for us, and because it has been the basis for several web histories to date—thus, it is a useful case study in the emergent field of web historiography. In contrast to our earlier study of the 2000 election, we had no particular research questions in mind on September 11, 2001, as we launched into helping build an archive of whatever was happening on the web related to the terrorist attacks. But our election research convinced us that a contemporaneously collected archive would document and preserve some of the ways that the web served as a significant surface for social, political, and cultural activity in the wake of the attacks. Over time, we and other researchers formulated a variety of questions, and retrospective analyses were conducted on a range of topics, including the rise of do-it-yourself journalism, governmental web responses, visual imagery, religious organizations' web actions, and shifts in online personal expression regarding the attacks during the autumn of 2001 (see Rainie, Schneider, & Foot, 2002). Throughout this chapter we refer to the experience of building the September 11 Web Archive, conducting our own studies based on materials drawn from it and other archives, and interacting with other scholars about researching and writing histories of the post-9/11 web.

THE COMPLEXITY OF OBJECT

In proposing an object-orientation to research and writing web history, we first offer a two-pronged definition of object summed up as object as motive and object as artifact. We then turn to some implications of this dual notion of object for the practice of web historiography. We identify potential motivating aims for this kind of scholarship, suggest levels and units of analysis that we consider especially important to think through, and consider the relationship between these motives, units of analysis, and the practices of doing web historiography. Developmental analyses of any aspect of the web, whether engaged in contemporaneously or retrospectively, entail dynamics within and between the (co)producers of web artifacts, production practices and techniques, and web artifacts themselves—as well as between the researcher(s) and the phenomena under investigation. These dynamics make it difficult but very important for scholars to identify and situate their object(s) of analysis historically, theoretically, and as methodologically constructed.

"Object" is used within fields as diverse as object-oriented programming and museology to reference discrete units and distinguish them from each other, as well as to define and demarcate those things that are being described or analyzed. Similarly, we suggest that the practice of researching and writing web history is enhanced when it is artifact-aware—that is, grounded in and shaped by a researcher's interactions and experiences with web artifacts as well as the artifacts themselves. This notion of object is consistent with the German term *Objekt*. A second, more abstract definition of object is "the end toward which effort or action is directed." This definition is resonant with the German philosophical term *Gegenstand*, which entails the concept of the embeddedness-in-activity of objects—both material and immaterial—that serve as motivating but largely unattainable horizons. This is the concept of object employed by Lev Vygotsky (1978) and the activity theorists who succeeded him, and it is the one with which we begin our discussion. In the next sections of this chapter, we explore these two aspects of "object" in greater depth.

AN ACTIVITY THEORY PERSPECTIVE ON OBJECT AS MOTIVE

The activity theory notion of object is richly complex, but English-language elucidations of this essential concept are scanty (see Foot, 2002; Kaptelinen, 2005; Leont'ev, 1978; Miettinen, 1998, 2005; Tuunainen, 2001). The activity theory concept of object can be difficult to grasp, in part because the German and Russian terms in which it developed are not easily translatable into English. Stated briefly, an object (*Gegenstand*) may be understood in the framework of activity theory as a collectively constructed entity, in material and/or ideal form, through which the meeting of a particular human need is pursued. To elaborate, activity theorists (Engeström, 1990, 1999; Engeström & Escalante, 1996; Lektorsky, 1984; Leont'ev, 1978) argue that the process of object formation arises from a state of need on the part of one or more actors. The need state, which is usually unconscious and thus not clearly definable, precipitates a set of "search actions" (Engeström, 1999, p. 381), during which any number of potential artifacts may be encountered. In most cases, it is only when search actions result in an encounter between the need and an artifact that the need begins to be experienced consciously. A social researcher as subject(s) orients toward one of these artifacts,

such as a website, through actions mediated by both personal experiences of the researcher and reifications of cultural-historical experience, and a "motive" arises out of the encounter of the need state and the artifact. Thus a social researcher's motive for doing web historiography can be understood as the interpolation of any one of a number of potential needs, such as the need to understand, the need to create knowledge, or the need to achieve fame and fortune with any kind of web artifact or virtual phenomenon.

Since the relationship between object (*Gegenstand*) and motive is dialectical in the activity theory framework, in that motive energizes object-oriented activity, and the conjoining of object and need state evokes motive, it is essential to maintain a clear analytical distinction between the two concepts (Foot, 2002). For our purposes in this chapter, it is sufficient to note that this notion of object-orientation is useful for understanding and advancing web historiography as a particular kind of activity that always involves researchers' own need states, which give rise to particular motives that may differ among researchers studying same pool of web artifacts or may vary over time for a single researcher. The activity-theoretic distinction between web artifacts and the motivating/orienting horizon of the object (*Gegenstand*) of the activity of web historiography is also very useful. In this perspective, the object-as-motive (*Gegenstand*) embedded in the activity of researching and writing web history can be understood as activity-context dependent and socio-culturally formed and thus historically evolving.

Drawing together these insights from activity theory, we view one prong of object-orientation in web historiography as the researcher's motive for researching and writing web history, that is, the horizon that is being reached for in the activity of web historiography. Others have written on the motives that give rise to the creation of archives. In an essay on representations of the 9/11 attacks in internet news stories, Brown et al. (2003) suggest that there is an "archival impulse" that prompts attempts to capture and preserve instances of media performance. Similarly, social theorist Jacques Derrida (1995) argues that archiving is an ancient practice, reflecting a deeply embedded human drive, and that archives are shaped by individual psyches and sociocultural formations that are remarkably persistent over time. Derrida argues that in spite of this, and simultaneously, humans in general (and elites in particular) are intrinsically motivated to manage social memory in ways that are often distortive and sometimes destructive of memory, and ultimately of society. He terms this malevolent mo-

tive "archive fever." Whether due to this notion of archive fever or not, in our experience web archives are few and far between. While we are not suggesting that our fellow scholars are afflicted by a strain of archive fever (although Derrida might), the use, much less the creation, of web archives by social researchers is relatively rare.

Certainly there are instances when people (ourselves included), whether expert archivists or not, build web archives to preserve web phenomena that they (or their institutions) find meaningful. Other motives for web archiving are more interactional in nature, such as preserving what others would rather erase or expunge. For example, the Internet Archive collection (accessible at http://www.archive.org) includes pages produced by U.S. government agencies and industry groups which were removed from the web in the weeks after the attacks because they were deemed to be interesting to potential terrorists (Soraghan, 2001; Toner, 2001). Similarly, the archive preserves a site that may have been produced by the Taleban (http://www.taleban.com) in Afghanistan that was subsequently defaced (Smetannikov, 2001) and later removed from the web altogether (Di Justo, 2001). But in general, at least in North America, far more of the web is over-written, erased, or deleted than captured, due to the dominant ideology of perpetual technological innovation and the widespread cultural impulses to revise or forget (web) history. Perhaps ironically, the more macro cultural-historical context of technological determinist ideology and historical amnesia in which we work as scholars is part of what motivates us to archive and engage in web historiography. It is our way of being countercultural as Americans, and as social researchers working in fields where ahistorical, point-in-time studies are the norm.

Going beyond our own motives for researching and writing web history, we suggest several other plausible motives for doing web historiography, for example, the need to make sense of socio-cultural-political relations and events or to try to understand development and evolution on the web at different levels and/or over time. Some social researchers may be motivated to retrospectively trace the emergence of a web phenomenon in order to get a read on its trajectory.

Illustration: Our motives for retrospective analyses of web memorializing.

Although neither of us lost anyone we knew in the 9/11 attacks, we, along with many, experienced a persistent sense of collective loss and grief along with empathy for those who had lost loved ones. Somewhat consciously, and perhaps more so subconsciously, we desired to validate the losses through our research expertise in the realm of technology and society. As academics who get paid to do research as well as to teach, we write to live, and we want to invest our research and writing in subjects that hold significance for us and others. We also happen to both be drawn to the emergent, to web phenomena that are nascent and unfolding. And so, nearly two years later, when the topic of memorializing on the web came up in a conversation with another colleague, Barbara Warnick, one of us (we can no longer recall which one) said "someone should study this," and the other said "we should." In that conversation the need states we have just described encountered the artifacts of memorial websites, and gave rise to the object-as-motivating-horizon of understanding and rendering web-based memorializing more visible by researching it in its historical context on the post-9/11 web (see Foot, Warnick, & Schneider, 2005, to view the outcome of our collective research activity).

OBJECT AS ARTIFACT

The second perspective on which we focus conceptualizes object as artifact. The sense of object as artifact is derived from the conceptualizations of object offered by a diverse set of professionals, including object-oriented programmers, museologists, art curators and social theorists. When thinking of object as artifact, we suggest that consideration be given to the demarcation of objects, the properties of objects, and the process by which objects become artifacts. We examine each in turn.

The first consideration with respect to object as artifact is to demarcate the object, a process in which boundaries around objects are determined. Web objects can be considered along a continuum ranging from "bits" to "experiences" (Arms, Adkins, Ammen, & Hayes, 2001), and each step along this continuum can be considered as an object. Few historians are likely to engage the web at the bit level—though those interested in the use of the web to distribute hidden messages would certainly be the exception (Provos & Honeyman, 2001). Some historians studying the web may consider page elements as artifacts—for example, examining patterns of images found within a set of web pages (Dougherty, 2003;

Sillaman, 2000). More commonly, either web pages—groups of elements assembled by a producer and displayed upon request to a server—or websites—groups of pages sharing a common portion of their URL—are treated as objects for the purposes of analysis. Alternatively, some historians might be interested in examining links between pages as objects. At another level altogether, some historians may wish to examine the experience of the web, without particular attention to sites or pages or elements, and define a browsing session as an object. In many respects, the challenges facing historians and archivists as they grapple with this aspect of object as artifact are similar to those actively engaged with curating objects for museum exhibits or art exhibitions. Within museology, "object" refers both to the "specimens" of artistic and cultural activity and the "re-presented" or "staged" associations between specimens crafted by historians and museologists (Preziosi, 2006). In short, we need to bring to the surface the underlying assumptions made when examining web objects as artifacts and to recognize the role of the researcher in constructing the object as artifact.

At the same time, it is important to establish some common framework of web objects in order to give the artifact aspect of the concept some meaning. To that end, we turn to a consideration of the term "object" as it is conceptualized in the field of object-oriented programming—the inspiration for the title of our chapter. In this field, "object" is considered as a discrete unit that has "the same power as the whole" (Kay, 1993). Objects, within this perspective, are sets of programming instructions that stand on their own, can be re-used in multiple applications, and carry with them a set of properties or inheritances that affect the way they "behave" or are interpreted by other objects within different contexts. Objects, wrote Alan Kay, a pioneer in object-oriented programming "are a kind of mapping whose values are its behaviors" (Kay, 1993). The "values" of objects are expressed in the properties assumed to be inherent in them. Encapsulation highlights the independent standing of an object and allows objects to be used or referenced by other objects under terms established and enforced by the object itself, even if not visible to other objects. Polymorphism suggests that some characteristics of objects are dependent on the context in which they are encountered. Inheritance assures that an object defined in relationship to another object possesses the traits and characteristics associated with those other objects. In

short, from the perspective of object-oriented programming, objects as artifacts are stable, dependable, and predictable.

Finally, we should reflect on the process by which an object becomes an artifact. From the perspective of the historian, considering an object as an artifact requires that we consider the tension between the object as "performed" and the object as archived (Schimmel & Ferguson, 1998). Jacques Derrida (1995) is particularly attuned to this distinction and draws our attention to the impact of the archival act on the impression of the object being viewed. While archived web objects may not obviously suffer the same degradation in meaning and expressive capacity as objects representing performance art do upon being re-presented in artifactual form (Greenstein, 1998), the challenges associated with re-presenting web artifacts remain significant and should be acknowledged by historians. For example, consider the re-presentation of a page from a web archive: the page is likely to consist of several elements (html code, images, etc.) assembled together and presented as a single artifact; however, each of the elements, archived individually over a period of time, may have changed during the archiving process, thus potentially rendering the archived object differently than it would have been performed (Arms et al., 2001). Furthermore, the viewing of web objects as artifacts may take place using browsers and displays significantly different than the technology available at the time the object was archived. In short, objects as artifacts are less fixed and more fluid (Levy, 2001) than some might have initially perceived.

Illustration: The website as artifact encompassing potential action.

As we began our analysis of the post-September 11 web through an intensive observation of pages archived in the immediate aftermath of the terrorist attacks, we began to notice that visitors to sites had been invited to participate in a variety of specific actions, some online and some offline. We further noticed that some of these potential actions were found within sites produced by types of actors not often associated with the actions observed, especially in 2001. For example, some national governments created pages within their sites, providing the opportunity for unrestricted text to be entered on memorial pages. We decided to expand our notion of online structure potentiating action (see Schneider & Foot, 2002) by focusing on the artifacts representing potential actions.

Our subsequent analysis (Schneider & Foot, 2003) focused on the types of online and offline actions in which visitors to websites in the September 11 Web Archive could have potentially engaged, and examined the distribution of actions potentiated across different types of site producers. This analysis illustrates the use of a website as an object, as we ascribed the potential action observed to the site encompassing the action. Our analysis further abstracts the observed objects by tallying actions across sites produced by types of actors (i.e., government agencies, educational institutions, etc.) as an estimate of how different types of site producers developed the web during the observed time period.

THE PRACTICE OF OBJECT-ORIENTED HISTORIOGRAPHY

Having proposed a two-pronged notion of object as motive and artifact, we now turn to the implications of this notion for the practice of object-oriented web historiography. We begin our discussion of object-oriented practice by considering different approaches to the web as a focus of study. If some aspect of web is to be the focus of an historical analysis, it is helpful to examine the underlying practices associated with a variety of approaches that are taken by social researchers and that can be useful in web historiography. We have elsewhere (Schneider & Foot, 2004, 2005) identified three sets of approaches that have been employed in web studies and examine each to highlight their relevance to distinctly historical research. We conclude with a focus on the techniques associated with scholarly web archiving.

Approaches that employ discursive or rhetorical analyses of websites—treating artifacts (sites, pages, elements) essentially as texts and drawing conclusions based on an analysis of content—allow historians to focus on the emergence of specific communicative phenomena that occur on the web. However, this approach is likely to downplay or ignore the impact of structuring elements and fail to provide an opportunity to assess the role of links among pages and between sites. Structural/feature analysis methods—which tend to use individual websites as their unit of analysis and focus on the structure of sites, such as the number of pages, hierarchical ordering of pages, or on the features found on the pages within the site—draw historians attention to aspects of web production and development but tend to understate the contribution that content may make to the over-

all experience of web users. We term a third set of approaches that analyzes multi-actor, cross-site action on the web as sociocultural analyses of the web (see several examples in Beaulieu & Park, 2003). Lindlof and Shatzer (1998) pointed in this direction in their article calling for new strategies of media ethnography in "virtual space." Hine (2000) presented one of the first good examples of sociocultural analysis of cross-site action on the web. By appropriating the term "sociocultural" to describe this set of approaches, we seek to highlight the attention paid in this genre of web studies to the hyperlinked context(s) and situatedness of websites—and to the aims, strategies, and identity-construction processes of website producers—as sites are produced, maintained, and/or mediated through links. This approach affords the historian an opportunity to take a broader view of the development of websites in historically evolving contexts.

Through a consideration of these approaches, whether employed concurrently with or retrospectively to evolving web phenomena, the differentiated attention and focus on object become clear. We have found that a multi-method approach that balances the analysis of evolving content with an analysis of changing structure and includes an assessment of artifacts within the context of the web provides the most illuminating and comprehensive outcomes in retrospective developmental analyses.

As part of the practice of object-oriented web historiography, we find it necessary to involve ourselves as researchers in the collection of artifacts. This may be temporary: as the practice of web archiving becomes standardized over the next decade or so, the robustness of archiving may become sufficient to support scholarly work. The first generation of web archiving—illustrated by the Internet Archive—has not provided us with archives that are robust enough to write some kinds of scholarly web histories. We suggest that historians become familiar with, and seek to influence, web archiving at several levels. To that end, we turn to a discussion of the techniques associated with web archiving. Our view of web archiving encompasses a wide range of activities, from the identification of objects to archive, through the stages associated with "getting bits on disk" as well as associating meta-data about the objects with the objects, and finally to representing archived objects in a web browser. Web archiving began in earnest in the mid-1990s, as sufficient technologies were developed and the recognition of rapidly disappearing content was acknowledged. The impetus behind web archiving activity was clear: The web was doubling in size every three to six months

from 1993 to 1996, and it appeared that it had the potential to become a significant platform on which a wide variety of social, political, scientific, and cultural phenomena would play out. Individuals at different types of institutions, such as libraries and archives, whose mission included the preservation of cultural and historical artifacts and materials, recognized the challenge that digital materials presented.

The use of web archives by scholars, including historians, obligates us to identify the specific techniques used by both archivists and researchers when examining web artifacts. Archivists engage in a number of discrete processes on the path from conceptualizing an archive to making archived objects and associated meta-data available to others (Brügger, 2005). The historian employing web archives to support historical claims must account for the potential impact of these processes on their conclusions. This is especially important when claims are being made about the web itself (as opposed to claims being made about individuals or organizations whose materials happened to be on the web). Together with Paul Wouters, we have explicated our view of these practices elsewhere (Schneider, Foot, & Wouters, 2009) and here will only briefly touch on those relevant to historians and other research-oriented users of archives. We intend to draw attention to the techniques associated with making web objects accessible as artifacts for future analysis.

One set of processes involves management tasks associated with the selection and representation of artifacts in archives. *Identification* includes the steps necessary to make known to an archiving system those web objects to be considered for inclusion. Obviously, any archiving process involves the selection of some web artifacts and the exclusion of others. There are innumerable techniques associated with this process. Experts can identify websites of interest, often from a published directory. Query results from a search engine can be processed using a fixed set of rules to select artifacts of interest. In any case, it should be expected that historians using archives specify the underlying assumptions by which objects were identified. A separate but related process, *curation*, involves creating a set of rules and procedures necessary to collect the desired objects. These rules might specify, for example, the instructions to be given concerning whether to follow or to ignore links to other artifacts. This process, often a highly technical and specialized procedure, is frequently opaque to researchers and other users of

archives. However, these procedures determine the specific artifacts that are included in archives, and, more importantly, those that are excluded. Making curatorial processes visible allows historians and others who develop evidence from archives to document, as necessary, those factors which may influence the observed results. Finally, *representation* is the process of retrieving archived artifacts from a collection and presenting them in a web browser. Considering this step as a process associated with web archiving will draw attention to the fact that rendering archived artifacts involves affirmative choices that affect the ways in which the rendering is performed. Historians and other researchers need to be aware of how the practice of representation is shaping their perspective on archived artifacts.

We have also suggested that there are three distinct processes of associating and presenting meta-data, or data about the artifacts, with the objects themselves. We distinguish these meta-data processes from each other on the basis of the techniques utilized. *Indexing* refers to the process of generating meta-data about collected artifacts or groups of objects algorithmically, while *categorization* is the process of generating meta-data through human observation and analysis. Indexing can involve developing meta-data from one of the available sources of information about archived artifacts, including the artifact itself, log files from crawling programs, and data developed externally to the archiving project. Categorization of meta-data can be applied to different types of artifacts (such as page elements, pages, or sites). The process of *interpretation* provides meta-data about collected artifacts, derived through the processes of categorization and indexing, to support sense-making activities such as discovery and search and to facilitate selected representation of collected artifacts. This process may include the design and implementation of an interface to a web archive, allowing users to select archived artifacts for examination or analysis. Providing full-text search of both meta-data and archived artifacts is an interpretation technique especially well suited to presenting unstructured data generated through annotation of artifacts, as well as providing access to archived artifacts containing text matching submitted queries.

Our aim in elucidating the techniques associated with web archiving is to increase their transparency and enhance the ability of social researchers to assess archives and interfaces that they confront while writing web histories, as it will often be the case that historians will address archives through an interpretative

interface that masks categorization, indexing, and representation processes. In addition, we hope to provide a framework that will encourage social researchers to actively participate in constructing archives and designing interfaces that support their ongoing efforts to write web histories.

Illustration: The web archive interface as an interpretative frame.

Our work creating two distinct interfaces to the September 11 Web Archive illustrates the importance of interpretative frames to archived collections. The interface visible at http://september11.archive.org focused on the type of actions in which visitors to websites in the collection could engage. This interface incorporated specific meta-data fields for 250 analyzed sites; these meta-data were collected as part of our research process. The interface links to web pages, archived on a specific date, that were identified as potentiating the actions identified.

A second interface, visible at http://www.loc.gov/minerva/collect/sept11, allows visitors to browse 2,313 websites for which other fields of meta-data were collected. Visitors to the archive can browse sites by selecting the first letter in the producer name, producer type, producer country, language, and the presence/absence of content associated with bioterrorism, September 11, and the Afghan War. Successively selecting meta-data fields narrows the search parameters to matching sites; clicking on the producer name provides access to a web archive record hosted by the U.S. Library of Congress. The web archive record links to a listing of dates on which the website was archived for this collection.

These interfaces provide dramatically different access to the same underlying archive. The first interface described would be valuable to those visitors particularly interested in the specific actions about which meta-data was collected, and would direct visitors to specific objects that potentiated these actions. The second interface described would serve visitors interested in framing their own questions about objects in the archive and may enable them to develop samples of sites for further analysis.

CONCLUSION

In summary, we have proposed an object-oriented approach to web historiography based on a two-prong definition of object, involving object as motive and artifact. We have sought to demonstrate that such an object orientation serves to make web historiography more transparent and more rigorous. We hope this

object orientation will evoke greater reflexivity among social researchers regarding their motives for doing web historiography, the ways in which web elements are made into and selected as artifact-objects of research, and the full range of techniques employed in web historiography. This proposal of object-orientation in web historiography will also, we hope, cause social researchers to become more proactive in (co)constructing the archives and interfaces in/through which they seek to conduct their scholarship.

So what does this kind of object orientation imply for historians of the web in contrast to historians of other media? Scholars who desire to research and write web histories face significant challenges. The affordances of the web render object orientation very important. Some characteristics of the web that are generally agreed upon by new media scholars include its scale and scope, its pervasiveness and ubiquity, the immeasurable and evolving volume of its content and its ephemerality, and the widespread access to the means of production that an unprecedented number of non-professional users have. In view of these affordances and a general lack of awareness of the connection of current web phenomena with the past, web phenomena must be understood as (already) history that is relevant to human activities in the present and the future.

William Thomas (2004) cites literary scholar Espen Aarseth's (1997, p. 62) insightful characterization of cybertexts as non-linear, dynamic, explorative, configurative narratives and argues that the "ergotic" nature of web artifacts, combined with the selective and iterative nature of web archiving, holds significant implications for historians:

> For historians the first stage in such textual developments for narrative have [sic] already been expressed in a wide range of digital archives. While these archives might appear for many users as undifferentiated collections of evidence, they represent something much more interpreted. Digital archives are often themselves an interpretative model open for reading and inquiry, and the objects within them, whether marked-up texts or hypermedia maps, derive from a complex series of authored stages.

A commitment to object-oriented historiography can shape our work as researchers in how we archive, how we identify/articulate our motives and the aims of our research, and in how we conduct our research practice in view of the research questions we pursue.

REFERENCES

Aarseth, E. (1997). *Cybertext: Perspectives on ergotic literature*. Baltimore and London: Johns Hopkins University Press.

Arms, W., Adkins, R., Ammen, C., & Hayes, A. (April 15, 2001). Collecting and preserving the web: The MINERVA prototype. *RLG DigiNews, 5*.

Beaulieu, A., & Park, H. W. (Eds.). (2003). *Journal of Computer-Mediated Communication, 8*(4) (special issue Internet networks: The form and the feel), retrieved September 1, 2009, from http://jcmc.indiana.edu/vol8/issue4/.

Brown, M., Fuzesi, L., Kitch, K., & Spivey, C. (2003). Internet news representations of September 11: Archival impulse in the age of information. In S. Chermak, F. Y. Bailey & M. Brown (Eds.), *Media representations of September 11* (pp. 103–116). Westport, CT: Praeger.

Brügger, N. (2005). *Archiving websites: General considerations and strategies*. from http://www.cfi.au.dk/publikationer/archiving/.pdf.

Derrida, J. (1995). Archive fever: A Freudian impression. *Diacritics, 25*(2), 9–63.

Di Justo, P. (2001). Does official Taliban site exist? [Electronic Version]. *Wired*. Retrieved January 25, 2009 from http://www.wired.com/politics/law/news/2001/10/47956.

Dougherty, M. (2003). Images of September 11th on the web. In L. Rainie, S. M. Schneider & K. Foot (Eds.), *One year later: September 11 and the internet* (pp. 62–65). Washington, DC: Pew Internet and American Life Project.

Engeström, Y. (1990). *Learning, working and imagining: Twelve studies in activity theory*. Helsinki: Orienta-Konsultit.

Engeström, Y. (1999). Innovative learning in work teams: Analyzing cycles of knowledge creation in practice. In Y. Engeström, R. Miettinen & R.-L. Punamaki (Eds.), *Perspectives on activity theory* (pp. 377–406). New York: Cambridge University Press.

Engeström, Y., & Escalante, V. (1996). Mundane tool or object of affection? The rise and fall of the postal buddy. In B. Nardi (Ed.), *Context and consciousness: Activity theory and human-computer interaction* (pp. 325–374). Cambridge, MA: MIT Press.

Foot, K. A. (2002). Pursuing an evolving object: Object formation and identification in a conflict monitoring network. *Mind, Culture and Activity, 9*(2), 132–149.

Foot, K. A., Warnick, B., & Schneider, S. M. (2005). Web-based memorializing after September 11: Toward a conceptual framework. *Journal of Computer Mediated Communication, 11*(1), retrieved September 1, 2009, from http://jcmc.indiana.edu/vol11/issue1/foot.html.

Greenstein, M. A. (1998). The act and the object. *Art Journal, 57*(3), 85-88.

Hine, C. (2000). *Virtual ethnography*. Thousand Oaks, CA: Sage.

Kaptelinen, V. (2005). The object of activity: Making sense of the sensemaker. *Mind, Culture, and Activity, 12*(1), 4-18.

Kay, A. (1993). The early history of smalltalk: 1960–66—Early OOP and other formative ideas of the sixties. Retrieved September 4, 2008, from http://www.smalltalk.org/smalltalk/TheEarlyHistoryOfSmalltalk_I.html

Lektorsky, V. A. (1984). *Subject, object, cognition*. Moscow: Progress.

Leont'ev, A. N. (1978). *Activity, consciousness, and personality*. Englewood Cliffs: Prentice-Hall.

Levy, D. (2001). *Scrolling forward*. New York: Arcade Publishing.

Lindlof, T. R., & Shatzer, M. J. (1998). Media ethnography in virtual space: Strategies, limits, and possibilities. *Journal of Broadcasting and Electronic Media, 42*(2), 170–189.

Miettinen, R. (1998). Object construction and networks in research work: The case of research on cellulose degrading enzymes. *Social Studies of Science, 38*, 423–463.

Miettinen, R. (2005). Object of activity and individual motivation. *Mind, Culture, and Activity, 12*(1), 52-69

Preziosi, D. (2006). Art, art history, and museology. *Museum Anthropology, 20*(2), 5–6.

Provos, N., & Honeyman, P. (2001). *CITI technical report 01-11: Detecting steganographic content on the internet* (PDF). Ann Arbor: Center for Information Technology Integration, University of Michigan.

Rainie, L., Schneider, S. M., & Foot, K. A. (Eds.). (2002). *One year later: September 11 and the internet*. Washington, DC: Pew Internet and American Life Project.

Schimmel, P., & Ferguson, R. (1998). *Out of actions: Between performances and the object, 1949–1979*. New York: Thames and Hudson.

Schneider, S. M., & Foot, K. A. (2002). Online structure for political action: Exploring presidential Web sites from the 2000 American election. *Javnost (The Public), 9*(2), 43–60.

Schneider, S. M., & Foot, K. A. (2003). Crisis communication & new media: The web after September 11. In P. N. Howard & S. Jones (Eds.), *Society online: The internet in context* (pp. 137–154). London: Sage.

Schneider, S. M., & Foot, K. A. (2004). The web as an object of study. *New media & society, 6*(1), 114–122.

Schneider, S. M., & Foot, K. A. (2005). Web sphere analysis: An approach to studying online action. In C. Hine (Ed.), *Virtual methods: Issues in social research on the internet* (pp. 157–170). Oxford: Berg.

Schneider, S. M., Foot, K., & Wouters, P. (2009). Web archiving as e-research. In N. Jankowski (Ed.), *E-research: Transformation in scholarly practice* (pp. 205–221). London: Routledge.

Sillaman, L. (2000). *The digital campaign trail: Candidate images on campaign websites*. Unpublished Master's thesis, University of Pennsylvania, Philadelphia, PA.

Smetannikov, M. (2001). Taleban site defaced after terrorist attacks [Electronic Version]. *Extreme Tech*. Retrieved February 20, 2009 from http://www.extremetech.com/article2/0,2845,120263,00.asp.

Soraghan, M. (2001). In wake of terrorist attacks, industry Web sites remove information [Electronic Version]. *Denver Post*. Retrieved January 25, 2009 from http://www.accessmylibrary.com/coms2/summary_0286-7842438_ITM.

Thomas, W. G. (2004). Computing and the historical imagination. In S. Schreibman, R. Siemens & J. Unsworth (Eds.), *A companion to digital humanities*. Oxford: Blackwell.

Toner, R. (2001). Reconsidering security, U.S. clamps down on agency Web sites [Electronic Version]. *New York Times on the web*. Retrieved January 25, 2009 from http://query.nytimes.com/gst/fullpage.html?res=9C06E5D81F31F93BA15753C1A9679C8B63.

Tuunainen, J. (2001). Constructing objects and transforming experimental systems. *Perspectives on Science, 9*(1), 78–105.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.