A Digital Library for Recovery, Research, and Learning from April 16, 2007 at Virginia Tech

Edward A. Fox, Christopher Andrews, Weiguo Fan, Jian Jiao, Ananya Kassahun, Szu-Chia Lu, Yifei Ma, Chris North, Naren Ramakrishnan, Angela Scarpa, Bruce Scarpa-Friedman, Steven D. Sheetz, Donald Shoemaker, Venkat Srinivasan, and Seungwon Yang

Virginia Tech, Blacksburg, VA 24061

Abstract

We are developing a digital library to help during the long-term recovery from the mass shooting on April 16, 2007 at Virginia Tech. Content comes from uploaded texts, images, videos, and other files, as well as pages crawled from the Web, and information collected, with permission, from those working with Web 2.0 sites like Facebook and Flickr. We are applying data/text mining, social network analysis, and information visualization methods to facilitate systems science, providing key added-value services, especially to social and behavioral scientists seeking faster and easier ways to analyze, model, understand, and test hypotheses. We consider how technologies influence communications. We aim to support the university, local community, region, nation, and the world, as it seeks to learn from a tragic event.

A Digital Library for Recovery, Research, and Learning from April 16, 2007 at Virginia Tech

This paper is about our vision, early work, and plans to help with the recovery from the mass shooting on April 16, 2007 at Virginia Tech – to support research and learning using a digital-library-based approach. We also hope this will lead to faster and more effective methods for applying the social and behavioral sciences to address tragic events and their aftermath, much as advanced techniques have been applied to the natural sciences in e-science initiatives (Atkins et al., 2003).

The context of our work is a world in which radical changes have occurred as a result of the convergence of many innovations and "flattening" trends, including globalization, advanced information technologies, openness (Raymond, 1999; Van de Sompel & Lagoze, 2000), collaboration/competition/workflows, new economies, and socio-political realignments (Friedman, 2006). One response to these changes is the move, in projects like ours, to develop cyberinfrastructure (NSF, 2007). Another is to help prepare professionals, as well as the broad base of students in countries like the United States of America, for Living in the Knowledge Society (LIKES) (Fox, 2007a, 2007b). In particular, for April 16, we are building a digital library (Arms, 2000; Chen & Fox, 1996; Fox, Suleman, Madalli, & Cassel, 2004; Fox, 1993; Fox, Akscyn, Furuta, & Leggett, 1995; Fox & Marchionini, 1998; Fox & Sornil, 1999, 2000; Fox & Urs, 2002; Lesk, 1997, 2005), which combines content and flexible services to support the needs of a targeted user community. Thus, we advance beyond the usual highly sequenced and pre-planned approach to research and education (see Figure 1).

<Have Figure 1 around here.>

We believe this is the right type of support for research and learning related to the events and aftermath of April 16. Traditionally, as can be seen in Figure 1a, research deals with hypotheses and related models or theories, that then lead to data collection, analysis, hypothesis testing, and subsequent publishing. The cycle can be extended, such as when data suggest revisions to the original hypotheses, or improvements to models or theories. However, when an interdisciplinary community seeks to study different aspects of an important phenomenon, such as the tragedy on April 16 and its aftermath, the approach shown in Figure 1b could be more efficient and effective. Based on some definition of scope, with regard to data and other types of content, as well as to studies of interest, a broad collection effort can proceed. Much as a university or town supports a local library, or a nation addresses archiving of its culture and heritage, in the digital era, a community can be supported by a comprehensive digital library. Furthermore, while "brick and mortar" libraries typically provide only a small set of services (e.g., searching, browsing, reference-desk support), digital libraries can provide an open-ended set of services, much like diverse companies operating on the Web help the millions who are interested in web pages. As can be seen in Figure 1b, digital library services can provide add-on content, an extensible set of services, and direct aids for analysis, synthesis, collaboration, hypothesis testing, etc.

One added benefit from the approach illustrated in Figure 1b is that it is feasible to add to the digital library all results of studies, including curricular materials, educational resources, temporary data sets, partial analyses, visualizations of findings, and varied publications (reports,

theses, conference papers, journal articles, etc.). Thus, building rapidly upon the work of others is facilitated, and collaboration can lead to more comprehensive insights. Also, once a large collection is available that can support teams of interdisciplinary scholars, new studies can be launched without having to wait for additional (and expensive) data collection activities. Requests for Institutional Review Board review can be "batched" to cover a range of studies, further reducing related delays in follow-on research schedules.

These improvements, however, must be carefully arranged, else they may be constrained because of legal and social concerns, which often lead to limited access, even for those with helpful intentions, to sensitive or personal information. Privacy must be respected, and anonymity must be preserved when desired. In our litigious society, complaints and suits may arise in the aftermath of tragedy, any related data may be subpoenad, and those providing data may be called as witnesses. Thus, even if a research design has separated identifying information from survey responses, a subpoena may lead to identification of individuals. Groups providing data may be viewed as classes, as for class-action suits. Individuals in special communities, e.g., those injured during the tragedy, or administrators directly involved in the event, may be off limits with regard to data collection, for long periods of time, typically three years or more.

Nevertheless, to the extent possible, this project aims to provide a range of advanced technologies to serve the community of those interested in April 16. In addition, we seek to study technology as a factor in the activities and processes related to April 16, as is discussed in the next section.

Approach

In this section, we consider how information visualization, data and text mining, and social network analysis, supported by and integrated through digital libraries, can be applied to help regarding April 16, 2007. But, as in any research, we first must understand (at least partially) the needs and requirements for such application.

Broadly, we aim to support the global community, with regard to their diverse interests regarding April 16. We will provide access to millions of pieces of news. We will help connect with the hundreds of thousands of individuals who have used blogs or other mechanisms to express their concern. We will support access to tens of thousands of photos, and thousands of hours of video. We aim to support students writing reports, at all levels. Ultimately there will be scholarly theses and dissertations related to the event. In spring 2008, several of the authors will run an Honors class for some of the university's brightest students, leading to class reports that will be online accessible by May 2008.

We also will support the community, as it remembers, analyzes, mourns, grieves, learns, and provides assistance for those involved. People will view the tragedy from a variety of personal perspectives, as well as from viewpoints related to disciplines like psychology and sociology. Since there are so many angles to consider, in the following, to illustrate our plans, we focus on what may be an unusual aspect of the Virginia Tech scene, namely the widespread deployment of advanced computing, communication, and information technologies.

Research Agenda for the Social and Behavioral Sciences Related to Technology

We plan to develop, through discussions and focus groups, a research agenda regarding the application of such technologies during and after a crisis. Indeed, we must consider more broadly the relationship between technology and crisis, not only as an aid, but also as a cause, correlate, and part of the context of crisis. One aim of the research agenda will be to address how using technology during and after a crisis can serve to reduce anxiety and manage distress. Accordingly, we are involved in survey activities to collect data about the use of technology (during and immediately after the crisis), feelings of distress, and coping strategies.

Technology has become embedded into the lives of people across the global society. It is now commonplace for communication to occur through multiple technologies, e.g., cell phones, Personal Digital Assistants (PDAs), and computers, resulting in multi-media traces of the communications being saved as voice, text, video, and instant messages. Similarly, the internet and the emergence of virtual communities, e.g., Facebook (Facebook, 2007) and MySpace (MySpace.com, 2007), have enabled communication on a scale previously unimaginable. However, we do not know how these technologies were used during and after crises to reduce or mitigate distress. Indeed, we do not know whether the use of technology reduces distress; it is possible (perhaps likely) that problems encountered with normally dependable technologies that fail during crises exacerbate rather than mitigate stress for those involved. Thus, a series of questions results from the current nexus of technology and tragedy: How do the established theories of dealing with distress apply in the context of the current state of technology? What capabilities should technology have, to provide support during crises? How can technology be enhanced to aid in reducing stress immediately after crises? How are social networks used

through technology to mitigate stress during and after crises? Is the use of the social network through technology different for individuals experiencing high stress versus those with lower stress? Does the emergence of unique virtual communities necessitate evolution of the theories related to distress? Do the communication patterns of distressed individuals (such as the gunman on April 16) provide insights that suggest ways to identify those in need of counseling or other interventions?

An important sociological issue to be addressed is the social network used to relay and receive information concerning a traumatic event, such as what occurred on the Virginia Tech campus on April 16, 2007. Specifically, where do students, faculty, staff, and others go for information about what happened? Whom do they contact first for such information, and to whom do they first relay the information they have received? Whose opinion do they trust? Do people turn to relatives, friends, and authority figures known to them, such as professors, or to the media, or some other source of information? In turn, to whom do people relay the information they have received? In the aftermath of a traumatic occurrence, do these contacts change? Also, are these social networks contextualized by social characteristics, such as age, gender, or race and ethnicity? Of particular interest would be how technology is used to receive and convey information and reactions to traumatic events, and the sociological contextualization of the use of technology. For example, one hypothesis would be that college students use technology more often to receive and convey information, than do professors or adults in general. The purpose of our discussions and focus groups is to generate such hypotheses and effective methodologies for testing them. In particular, we are seeking to encourage interaction

between social and behavioral scientists and experts in technological fields to produce new and innovative ways of developing and testing hypotheses concerning how people react and respond to traumatic violent episodes, both in the short-term and over the long-term.

It has been estimated that 10-40% of individuals who experience trauma will subsequently develop post-traumatic stress disorder (PTSD) (Michaels et al., 1999). In light of these figures, PTSD is considered one of the most common anxiety disorders experienced among young adults (Breslau, Davis, Andreski, & Peterson, 1991). The Diagnostic and Statistical Manual of Mental Disorders (Association, 2000) defines a traumatic event as "an extreme traumatic stressor involving direct personal experience of an event that involves actual or threatened death or serious injury, or other threat to one's physical integrity" that elicits "intense fear, helplessness, or horror" (pg. 463). Specific traumatic experiences can include natural or technological disasters, violent crimes, abusive relationships, and sexual assaults. Undoubtedly, the Virginia Tech shootings fall centrally within the realm of trauma. Nonetheless, not all who experience a trauma will develop PTSD; many individuals show resilience in the face of tragedy. Conceptual models of PTSD explain its emergence in terms of interactions among the trauma experience (e.g., proximity to the event, degree of life threat, loss/disruption), survivors' characteristics (e.g., ethnicity, pre-existing psychopathology, prior trauma experiences), cognitive processing (e.g., positive or negative coping), and the recovery environment (e.g., perceived social support) (Green, Wilson, & Lindy, 1985; LaGreca, Silverman, Vernberg, & Prinstein, 1996). The digital library, and specifically the surveys of students, faculty, staff, and community, will allow us to test aspects of this model in the immediate aftermath of this tragedy

as well as in a longitudinal manner over the following year. Three hypotheses (of many) are of specific interest in this regard. First, does the psychosocial model predict which individuals will or will not develop PTSD as well as other stress-related reactions (such as depression, substance abuse, and physical health problems)? Second, which aspects of coping (e.g., spiritual, cognitive, active/passive, emotion-focused) and social support (e.g., friends, family, classmates, teachers) are most influential in buffering survivors from these negative outcomes? Third, to better understand protective factors, how do coping strategies change and how are social support networks formed and utilized?

Another potential area of exploration is the impact on physical health of acute and chronic psychological stress from traumatic events such as April 16. Both short- and long-term negative health effects have been previously documented in response to the stress of traumatic community events (Lovallo, 2005). For example, increased incidence of sudden cardiovascular events such as myocardial infarction and stroke have been noted in acute response to earthquakes (Leor, Poole, & Kloner, 1996), missile attacks (Kark, Goldman, & Epstein, 1995), and even non-threatening events such as an upset loss by a major sports team (Carroll, Ebrahim, Tilling, Macleod, & Smith, 2002). It also is possible that the lingering effects of trauma such as the April 16 shootings may negatively impact health by way of chronic increases in negative emotional states such as anxiety, anger, and depression, which are known to have negative relationships to physical health (Everson-Rose & Lewis, 2005). Two common psychological factors across stressors are well-known to have deleterious physical effects: control and predictability. Stress and trauma that is unpredictable, and which individuals exert little or no control over, will

engage autonomic nervous and endocrine system responses to threat that can have negative effects on the immune and cardiovascular systems (Porges, 1995; Seligman, Maier, & Solomon, 1971; Thayer & Lane, 2007; Weiss, 1991) . The April 16 shootings were both unpredictable and uncontrollable, and as such may confer similar negative effects on the health of the community in Southwest Virginia. Possible directions of exploration include how specific coping styles predict health outcomes, as well as the roles of factors such as social support and religiosity in buffering the negative effects of stress. Findings might not only help document the broad effects that this tragic event has had on our community, but also might be generally applied toward optimal preparation for and response to unpredictable stressors, traumas, and crises.

Special Analysis

Our project aims to address the needs of the public, as well as the research and education goals of various interested communities, by providing important add-on services that will be closely coupled to the digital library of related content. In particular, we focus on three analysis aids: information visualization, data mining, and social network analysis.

Information visualization (North, 2005; Spence, 2007) helps us understand the data and information under consideration, and facilitates discussion and presentation of results. Advanced techniques have been developed to help us deal with urgent problems faced by the intelligence community, such as terrorism (Thomas & Cook, 2005), and so can be applied to April 16. One key benefit is gaining additional insight (Saraiya, North, Lam, & Duca, 2006), a very important goal for those trying to understand tragedies.

Visualization can be applied to texts (Wong, Whitney, & Thomas, 1999), geospatial information (Jung, 1999), retrieval results (Hearst, 1995; Nowell, 1997), biological pathways (Saraiya, North, & Duca, 2005), and cross-jurisdictional information (Kaza et al., 2005). They can lead to 3D presentations (Amavizca, Sánchez, & Abascal, 1999), handling of graphs (Tollis, 1996), or work with internet collections (Viegas, Wattenberg, Ham, Kriss, & McKeon, 2007).

An important challenge in our work is the integration of information visualization with digital libraries (Nowell & Fox, 1995), such as through special protocols and transformations (Shen, Wang, & Fox, 2002; Wang, 2002). Even more broadly, information visualization in digital libraries can be integrated with related searching and browsing services, under the broader rubric of exploration (Shen, Vemuri, Fan, da S Torres, & Fox, 2006).

Of particular value, as we engage in discussions with social and behavioral scientists, is to demonstrate the benefits of information visualization. Fortunately, Virginia Tech has a laboratory with a flexible GigaPixel-scale display arrangement (Ball & North, 2007), so different approaches simultaneously can be applied to related data, and then compared and discussed in a collaborative setting. Later, we can select methods that are shown to be of interest, and migrate them to smaller display configurations, making the services available remotely over the Web.

Data and text mining are key analysis techniques, which in turn can feed into information visualization, as above. Data mining can work with large data collections or databases (Fayyad, Piatetsky-Shapiro, & Smyth, 1996). It can lead to discovery of patterns, shifting from data analysis to development of scientific models (Ramakrishnan & Grama, 1999). Data mining can be helpful when working with crimes (Chen et al., 2004). Text mining (Fan, Wallace, Rich, &

Zhang, 2006), a type of data mining, focuses on the large amounts of text available on the internet, a key resource for what is being collected in connection with April 16.

Social network analysis, that is, analyzing contact patterns among people, has proven to be a valuable research tool in various domains. This approach has been applied most notably to study the spread of epidemics (Eubank et al., 2004) and to devise viral marketing strategies (Domingos & Richardson, 2001). Such analyses have revealed that the study of the underlying properties of a network can yield valuable clues regarding the dynamics of the processes in the network.

<Have Table 1 around here.>

In the context of this project, we propose to study the "technology network", during and shortly after the campus shootings. We will work to construct networks of interactions between people who communicated, especially using the means shown in Table 1. Some of the most important questions to address include:

1. What were the most common modes of communication among different groups of people (faculty, students, staff, etc.)? Which were most effective for each group?
2. Who are the most influential, central, or well connected individuals/groups (dean's office, department heads, focus groups, Greek societies, etc.), if any? Under circumstances like those on April 16, who would be able to communicate information most effectively?
3. How long does it take for information to spread among different groups of people? Are there any bottlenecks?

4. What were the patterns of technology usage in the aftermath? Which ones were used most, and in what ways (blog entries expressing condolences, memorial videos, etc.)?

Digital Libraries

A context and environment to support information visualization, data and text mining, and social network analysis must integrate these and other services with related content. Thus, our project builds atop work in the field of digital libraries.

<Have Figure 2 around here.>

The traditional library (see Figure 2a) assumes that authors publish works, which are collected by librarians, and then can be discovered and accessed by readers. This works both with "bricks and mortar" libraries and with digital libraries, but many new interaction patterns are only feasible using digital technologies. For example, as shown in Figure 2b, authors or creators can put their works onto web sites. With the help of search engines, as from Google, Microsoft, or Yahoo, those works are "crawled" and placed into large collections, to be indexed and made searchable for web surfers.

More recently, another paradigm has arisen, whereby millions of people can publish on the web, with simpler schemes like blogs and wikis. Further, these creative web users can work in contexts of specialized types of content (e.g., photos in Flicker, or videos in YouTube) or community collaboration (e.g., Yahoo groups, Facebook, MySpace). Figure 2c summarizes this "Web 2.0" (O'Reilly, 2005) approach, that also is discussed in the next sub-section. Here we focus on how to organize such diverse and distributed content, a key goal of libraries.

Table 2 illustrates a conventional library organization, giving the key layers or perspectives. Users or "patrons" work through interfaces, either in the physical world where one walks around bookshelves, or in the virtual world where one has computer screens (or PDAs or cell phones …). These lead to services, such as searching and browsing, which make use of catalogs (of descriptive information, i.e., metadata (Weibel, 1995)) or indexes, and eventually lead to desired content.

<Have Table 2 around here.>

In our work with digital libraries we have abstracted and generalized the handling of information, developing the "5S" framework, a theory-based formal model and environment for describing, understanding, developing, and innovating (Fox, 1999; Goncalves, 2004; Goncalves & Fox, 2002; Goncalves, Fox, Watson, & Kipp, 2004; Gorton, 2007; Zhu, 2002). Table 3 explains each of the five "Ss", which can be formally defined, but at the same time give an intuitive way to characterize the underlying aspects behind any advanced information system. Thus, digital libraries manage content (streams), supporting a variety of organizational aids (structures), across time and space and presented typically in a 2D form (spaces), for various types of uses (scenarios, implemented as services), for diverse societies – from the personal to the global.

<Have Table 3 around here.>

If we focus on the content portion of a digital library, we refer broadly to "digital objects" that can be of any media type. Sometimes a group of related digital objects, like the parts of a

dissertation, can be viewed as a composite (Fox & France, 1987) or compound document (Lagoze & Van de Sompel, 2007) or package, a larger type of digital object.

Combining metadata and digital objects, along with broad ranges of services, and addressing the distributed nature of modern information systems, we can summarize the layers of a modern digital library as in Table 4. This gives an architectural view of the type of digital library necessary for aiding work related to April 16, where various types of distinct repositories must work together in a federated fashion (Candela, Manghi, & Pagano, 2007). In particular, we are building a system as shown in Figure 3, to provide a range of services (right side) over a range of data sources (left side), to suit the needs of a broad user community (with needs specified through the aid of focus groups – see top center), and ultimately sustained by Virginia Tech University Libraries archival efforts regarding April 16 (see bottom center). The arrows represent the functions and processing to support making all this work together. Of particular interest is the handling of Web 2.0 information, as is discussed in the next sub-section.

<Have Table 4 around here.>

<Have Figure 3 around here.>

Web 2.0

Web 2.0 (O'Reilly, 2005) is a new model of the Web, which collectively refers to the collaborative and value added services, in addition to the usual content hosting services, that websites pervading the Web provide. The collaborative aspect of Web 2.0 refers to the interaction among users visiting the websites. Prominent examples are websites such as

Facebook and MySpace, which provide a platform for users to not just browse the content, but also actively participate in discussions, form focus groups, upload and share content, etc. Value addition is achieved, for example, by providing a platform for building applications and services on top of the already existing services that the websites provide. Google Maps (Google, 2007), for example, provides Application Programming Interfaces (APIs) which can be programmatically used to build services on top of the ones already offered by the website. This provides considerable value-addition to an audience interested in building applications based on Google Maps.

Such value-addition, especially by means of providing APIs, has added a new dimension to digital library research. A new paradigm called Service Oriented Architecture (SOA) has been proposed (Petinot et al., 2004), and successfully used, wherein different digital libraries (possibly heterogeneous, distributed, and geographically separated) can present a single uniform interface to the users and can call upon the services of each other (see Figure 2c). The interaction between the digital libraries is achieved by calling the Web Services hosted at each others' ends.

<Have Table 5 around here.>

We direct our efforts towards analyzing resources related to April 16, from prominent Web 2.0 sites, like YouTube, Facebook, MySpace, Flickr, etc. These probably represent the most popular social networking and content sharing points on the Web today. Table 5 summarizes our user-centered rationale to make use of these sites. The resources hosted at these sites can be accessed either directly from the respective content providers (depending on the copyrights and license restrictions associated with the resource), in which case we will provide the URL to the

resource, or locally at our end, in which case these resources will be cached at our site using the APIs provided by the content providers, or using copies provided by content creators. We will use the DSpace platform (MIT, 2003) as our digital library system. The resources (or surrogates thereof), as well as the metadata associated with them, will be stored in the digital library, and be made available for analysis.

Supporting Systems-Level Science

A key goal in our integrating special services, digital libraries, and the Web 2.0 infrastructure, is to support "systems-level" science on the social dynamics associated with crisis events. Inspired by current research trends in biology and the life sciences, systems-level science seeks to understand the functioning of very large and complex systems, and all the interactions therein, in a wholistic fashion. This is in contrast to more traditional reductionist science, which narrows the problem down to focus on specific individual variables. Systems-level science is more exploratory in nature, looking for broader trends and patterns of higher complexity; it encourages the development of new hypotheses. In the life sciences, the "system" of systems-level science refers to the functioning of complex biological organisms. In Virginia Tech's case, the "system" under study is the complex communities of people and how they respond to crises.

Systems-level science is more challenging, but can offer deeper insights into underlying phenomenon. Systems-level science requires:

1. Rapid and continuous collection of massive data: The recent growth of systems-level science in biology was supported by the invention of micro-array and similar instrumentation that

enables simultaneous collection of data about thousands of genes and proteins at the cellular level, thereby offering a complete picture that is both detailed and broad. Similarly our digital library will be organized to ingest large collections, such as student-submitted collections of email-related data, which are collected and curated over time.

2. Integration of diverse, heterogenous data: Our digital library must be able to bridge diverse data sources, including vast Internet and Web collections, as well as rich personal sources such Facebook pages, surveys, and interviews. We also will investigate new kinds of data that could be captured, such as by volunteers using our two SenseCam devices (Hodges, Williams, & Wood, 2007).
3. Realtime exploratory analysis: To gain deep insight, our digital library must link with analysis tools to support rich exploratory analysis of complex patterns and interrelationships for theory development. Realtime analysis must maintain synchronization with incoming information to promote awareness of breaking hypotheses and quick response. Access, analysis, dissemination, and utilization will be supported by visualizations and data mining tools.

Community, Technology, and Content

Our community of users includes those interested in April 16, e.g., those who through Web 2.0 or other services have contributed to the body of content related to the event and its aftermath. Many of those contributing content made use of mechanisms summarized in Table 1, which summarizes technology support.

Model

From Table 1 we see that a variety of types of devices are employed for communication and for information access. These allow a number of mechanisms for communication, with diverse goals, communicating with individuals or broader groups. Often, selections among these are based on quality metrics, relating to utility, ease of use, and contextual factors.

In a crisis, technology support may have an impact on stress and adaptation to trauma. Table 6 gives a set of propositions we seek to explore in this regard. Figure 4 provides a model of the causality we assume, that needs to be tested, including use of technology. (Recall the earlier section on *Research Agenda for the Social and Behavioral Sciences Related to Technology.)*

<Have Table 6 around here.>

<Have Figure 4 around here.>

Content Collection

Back in 1996, Virginia Tech led the world by encouraging (graduate) students to upload content, in particular electronic theses and dissertations (Fox et al., 1996); as of January 1, 1997 this was a requirement for all masters theses and doctoral dissertations. In 2003, the Networked Digital Library of Theses and Dissertations formalized the role it had been playing since 1996, incorporating as an international non-profit (Fox, 2003), to further spread the idea around the globe, so by August 2007 there were over 333,000 ETDs available in many languages. Oddly, in

spite of the familiarity of uploading to many young people today, many graduate students, faculty, and universities, in parts of the USA as well as in nations with smaller economies, still have not understood how easy and beneficial such uploading is, and loose out with regard to the wider recognition of their research contributions that comes through services that build atop the NDLTD Union Catalog (OCLC, 2004).

Uploading content related to April 16 thus is a natural approach for Virginia Tech, as shown in Table 7. The campus developed a web site, with various pages, around the theme of remembering. University Libraries has committed to a permanent archive, and has been addressing the tens of thousands of physical artifacts contributed. The Center for Digital Discourse and Culture has launched the April 16 Archive to collect content by crawling the Web (in collaboration with the Internet Archive), and by supporting the uploading of information (CDDC, 2007). Our effort is supplementing those activities, handling other types of content as well, such as videos, audio recordings, and email information.

<Have Table 7 around here.>

Content in Web 2.0

Though to many, the actions described in the last section are novel and advanced, requiring a "reach" in technology application, to others the people involved are but "digital immigrants", not "digital natives" who are comfortable and who grew up with the Web (Prensky, 2001). Those who are comfortable with current web services can of course upload content, but often prefer communicating directly by working inside applications that support virtual

communities, like Facebook. In Table 7, the counts are low for all rows except for the Guest Book, which can be viewed as a Web 2.0 application. Thus, as in Table 5, since the effort is less, posting to a Web 2.0 service is much easier, hence to be preferred, relative to uploading (which is called for in applications and sites described in the previous sub-section).

Such users "vote with their mouse", as can be seen in Table 8. The Web 2.0 applications listed are popular among those interested in April 16, leading to large numbers of entries. By far the largest number is the count of users who have joined groups related to April 16, as is detailed in Table 9. This table lists only large groups, i.e., with more than 500 in the group, and only with first letter in range A-I. Nevertheless, the table entries account for roughly the estimated count of 400,000 shown in the first row of Table 8.

<Have Table 8 around here.>

<Have Table 9 around here.>

Clearly, to understand April 16 in the large, it is important to capture information about those who expressed their concern and interest electronically. Fortunately, many of the Web 2.0 services have APIs to help automate interaction, as is shown in Table 10. Our group will be working with these APIs as we develop our digital library. This will facilitate identification and in some cases collection of information, when access permissions allow.

<Have Table 10 around here.>

Preliminary Analysis

We are just beginning to analyze some of the data available to us, and just starting to explore hypotheses and models that have been proposed with regard to April 16. Below we discuss some preliminary observations, to illustrate the potential of the approach.

Example Pattern

One observation, made by a co-author who is a close friend of a person severely injured on April 16, is that frequent patterns of behavior emerged that day, like that illustrated in Figure 5. Those who were in Blacksburg on April 16 were contacted by family and friends concerned about our safety, using a wide variety of communication schemes. Depending upon responses received, those making inquiries engaged in appropriate follow-on behaviors.

<Have Figure 5 around here.>

We seek to validate this pattern, understand what types of communication schemes were used, and to connect actions with stress and trauma and other reactions. It will be hard to collect the data needed, but we are involved in survey efforts to help, and another research project is seeking voluntary submission of phone logs (using both mobile and land lines) to help with related analyses.

Tags

Fortunately, some other types of analyses are easier to undertake, with less labor-intensive data collection requirements. One line of research has to do with categories employed in connection with Web sites for collecting content, in particular the April 16 Archive (CDDC,

2007). That site allows people to tag files being uploaded. We considered about 800 items uploaded, extracted all of the tags (i.e., words or short phrases) used with those, and summarized some key results in Table 11. This shows pairs of tags that co-occurred often in the collection of items. Thus, in 55 of the items in the collection, Blog and Commentary both appeared, showing the popularity of blogs for the purpose of making comments. As a second example, we note that the co-occurrence of Professor Librescu's name and the name of a leading Romanian magazine can help guide us to the news coverage in the country of origin regarding one of the heroes killed on April 16.

<Have Table 11 around here.>

We can learn more about the "cloud" of tags connected with the collection by examining Figure 6, an information visualization that complements the tabular presentation of Table 11. Here we miss out on co-occurrence information, but see the relative frequency of individual tags, the most popular of which is "memorial". We see also that people were involved not only in events like the campus vigil, but also in electronic communication using blogs and other mechanisms to prepare commentaries (a trend mentioned above that also was spotted using a different analysis).

<Have Figure 6 around here.>

Clusters of Text Words

To illustrate how texts in a digital library can be analyzed and presented using information visualization methods, we made use of the IN-SPIRE software suite (McQuerry,

2004) (Wise et al., 1995). We used all the text from all of the contributions (about 800) to the April 16 Archive (CDDC, 2007). The contributions are clustered according to a textual analysis of words and phrases in the text, and the clusters are labeled with frequently used words within the cluster.

Figure 7 provides a Themescapes view, where clusters of contributions are shown as mountains that form a landscape of themes. Figure 8 provides a Galaxy view, where each dot represents one contribution. Other than the differences in visual representation, these two views present essentially the same analysis and clustering layout. In particular, the placement of contributions and their groupings reflect a multi-dimensional scaling analysis, whereby closely related works are placed nearby, and sections of the visualization reflect aspects or issues popular in the collection.

<Have Figure 7 around here.>

<Have Figure 8 around here.>

Thus, for example, on the right side is a large cluster about the student community, with some sub-clusters for the vigil and memorials. In the top-right is a cluster about societal issues. The middle areas are mainly about various events and memorials, like the convocation, drillfield, etc. The left side seems to deal with external issues and foreign languages. Interest in Professor Librescu led to a sizeable cluster here. Issues related to Korea are concentrated toward the top, especially on the left.

Plans and Prospects

This project is just beginning, with an August 15, 2007 starting date for our NSF funded initiative, IIS-0736055, entitled "SGER: DL-VT416: A Digital Library Testbed for Research Related to 4/16/2007 at Virginia Tech". While we have worked between May and August to launch this effort, only now can we purchase equipment to record terabytes of videos, audio files, high quality photos, and other content collected through the many schemes outlined above.

We plan to explore, demonstrate, and help others in their use of a growing suite of services attached to the digital library. Of particular benefit as we work with the local community to demonstrate the potential benefits is the very large display laboratory at Virginia Tech, including the GigaPixel display (Ball & North, 2007). In the area of information visualization we will make use of further tools in the IN-SPIRE suite, such as ThemeRiver (Havre, Hetzler, Whitney, & Nowell, 2002), to show themes and changes over time. In addition, we will use other tools, like Spotfire (Ahlberg & Shneiderman, 1994). In connection with the work on social network analysis, we will use information visualization tools scaled to work with internet collections (Viegas et al., 2007).

Ultimately, we hope that the "systems" approach that has been so successful in science and engineering, and which has evolved to provide effective support in biology as part of the "systems science" trend (Saraiya et al., 2005), can lead to "systems behavioral and social science". We hope that powerful digital libraries can help during the ongoing recovery from April 16, and that when other tragedies occur in the future, tailored digital libraries will be easier to devise and even easier to deploy, for even greater benefit.

Table 1

Technology Support: Issues and Mechanisms

Issue	**Details**
Device	Land line telephone Cell phone Pager PDA Laptop Desktop computer In-person meeting
Mechanism	Talk Leave voicemail Email Listserv Online instant messaging/chat Send text message Website Web 2.0 service (Facebook, YouTube, …) Audioconference Videoconference
Goal	Obtaining information Providing information Understanding another person Being understood
Communicating with	1 person a group a community the public
Quality concern	Ease of use Utility/benefit Compatibility with personal activities/tasks/lifestyle

Table 2

Layers of a (Digital) Library

Interfaces
Services
Catalogs
Contents

Table 3

5S

S	**Roles (examples)**
Societies	Users, groups, organizations, communities, nations, cultures, world; agents, actors, service managers
Scenarios	Services, workflows, procedures, tasks, activities, programs, routines
Spaces	1D, 2D, 3D, multi-dimensional, vector, measure, probabilistic, topological, …
Structures	Data structure, database, ontology, taxonomy, graph, org-chart, schema, pattern, grammar
Streams	Video, audio, voice, character, bit, sensor

Table 4

Layers of a Comprehensive Digital Library

Layer	Description / Constituents
1	Interface: Browser + Java Applets + Plugins
2	End-User Services: Search, Browse, Download, Upload, Collaborate, Analyze, Visualize
3	Content Aggregations: Repository, Catalog, Collection, Ontology, Taxonomy
4	Content Types: Metadata, Digital Object, Package/Composite Content Forms: Video, Audio (Speech, Music), Image (Photo, Graphic), Structured (XML, HTML, SGML, Spreadsheet, Database), Text, …
5	Logical/Accessible Content: Local Content, Mapped Versions of Remote Content, Surrogates for Remote Content, Pointers to Remote Content
6	Federation and Collection Services: Crawling, Harvesting, Federated Search, Retrieving, Mapping, Fusing, Mediating, API-based Accessing
7	Remote Sites: Web Sites, Web 2.0 Sites, Federation Partners

Table 5

Comparison of Collecting with Using Web 2.0

Scenario	**User Effort / Benefit**	**Volume**	**Researcher Effort / Effort with API**
User upload	High / Low	Low	Low / n/a
User post to Web 2.0	Low / High	High	High / Low

Table 6

Propositions / Assumptions Regarding Factors Related to Stress

Proposition	**Description**
P1	Event exposure is positively associated with event related stress.
P2	Pre-event stress is positively associated with event related stress.
P3	Coping is inversely associated with event related stress.
P4a	Social Support is inversely associated with pre-event stress.
P4b	Social Support is inversely associated with event related stress.
P5a	Community Support is inversely associated with pre-event stress.
P5b	Community Support is inversely associated with event related stress.
P6	Using technology associated with event related stress

Table 7

Local Sites and Counts of Entries

Source	**URL**	**Counts**
Main site	http://www.vt.edu/remember/	10 pointers
Memorial site	http://www.vt.edu/remember/memorial/ (pages of condolences)	720
Higher ed support	http://www.vt.edu/remember/higher_ed_support.php (links to pages stating support)	395
Biographies	http://www.vt.edu/remember/biographies/	32
Univ. Relations	http://www.vt.edu/remember/archive/ (releases to VT community, only for 4/16 - 4/22)	30
Photo Gallery	http://www.vt.edu/remember/photos/	24
Office of Recovery & Support	http://www.recovery.vt.edu/	9 pointers
Guest book	http://www.legacy.com/GB/GuestbookView.aspx?PersonId=87298850	4282 pages
The April 16 Archive	http://www.april16archive.org/	150 photos, ~850 items
VT Magazine	http://www.vtmagazine.vt.edu/memorial07/today.html (plus 35 condolences …)	21 photos, etc.
Planet Blacksburg	http://blog.planetblacksburg.com/	12 blog entries, book

Table 8

Public Sites with Entries for "Virginia Tech April 16"

Site	**URL**	**Estimated Entries**
Facebook	http://www.facebook.com	400000 users
Flickr	http://www.flickr.com	400 photos
MySpace	http://www.myspace.com/	700 results
YouTube	http://www.youtube.com/	4000 videos
Technorati	http://technorati.com	9000 blog posts
Digg	http://Digg.com	500 Stories (link elsewhere)

Table 9

Some Facebook Groups Larger than 500, for "Virginia Tech April 16"

Size	**Title**
1531	A Canadian Tribute To All Those Involved In The Virginia Tech Shooting
1070	A DEDICATION TO THOSE HURT BY THE INCIDENTS AT V-TECH
752	A Heartfelt Thank You to the Heroes of 4.16.07
8129	All will be missed - VA Tech: April 16, 2007
3384	Always remember the Virginia Tech Massacre
873	Always Remember 4/16/07-Virginia Tech School Shootings
3896	America Stands in Support of Virginia Tech
1409	AMERICA SUPPORTS VIRGINIA TECH
1201	April 16, 2007 - A Moment of Silence " TODAY WE ARE ALL HOKIES "
40322	April 16, 2007 - A Moment of Silence
898	April 16th 2007 Will Be Remembered Forever
324653	A tribute to those who passed at the Virginia Tech Shooting
561	College Students United for Virginia Tech
859	Condolences For Virginia Tech...
688	Deepest condolences from England to the dead of Virginia Tech
730	Georgians (and others) Praying for VA TECH
506	04*16*07 God Bless the Hokies
633	God bless the Virginia Tech 33, We will always remember.
2195	God Bless the Virginia Tech students
573	God Bless VA Tech - Current Events
3673	God Bless VA Tech - Beliefs & Causes
1176	God Bless Virginia Tech - Beliefs & Causes
1921	God Bless Virginia Tech - History
795	God Bless Virginia Tech - Current Events
848	High Schooler's Taking A Stand. We're Here For You Virginia Tech!
621	Hoosiers Pray for Virginia Tech Tragedy
1379	I'll be praying for Virginia Tech.
765	In Honor of the VA Tech Victims
1882	In loving memory of all those who lost their lives at Virginia Tech 1560
1560	In Memory of the Victims of the Virginia Tech Shootings

Table 10

Web 2.0 Sites and the URL for their API

Site+Description	**URL for API**
Flickr Photo sharing service	http://www.flickr.com/services/
Google Search etc.	http://www.google.com/apis/
YouTube videos	http://www.youtube.com/dev
del.icio.us bookmarks	http://del.icio.us/doc/api
Facebook social networking	http://developers.facebook.com/
MySpace social networking	http://code.google.com/p/myspace-api/
AOL Video Search	http://developer.searchvideo.com/index.php
Yahoo Video Search	http://developer.yahoo.com/search/video/V1/videoSearch.html
Technorati Blogs search	http://technorati.com/developers/
Microsoft Virtual Earth	http://dev.live.com/virtualearth/
Creative Commons Licensing	http://api.creativecommons.org/docs/
Digg social bookmarks	http://www.digg.com/tools/
NewsGator news, websites, and blogs	http://www.newsgator.com/ngs/api/

Table 11

Some high frequency tag pairs from the April 16 archive

Tag 1	**Tag 2**	**Count**
Blog	Commentary	55
Candlelight	Vigil	43
University	Vigil	35
Hokie rock	Hokie stone	25
Drill field	Memorial	21
Blog	Cho	17
Second life	Memorial	16
Librescu	Omagiu	13
Mourning	Vigil	12
Omagiu	Alexandru Ioan Cuza	11
Memorial	Balloon	9
Art	Drawing	9
Blog	Gun control	9

Figure Captions

Figure 1. Standard research sequence of activities (a) and DL-supported version (b)

Figure 2. Information models for traditional libraries (a) and using the Web (b) or Web 2.0 (c)

Figure 3. Architectural view of our digital library

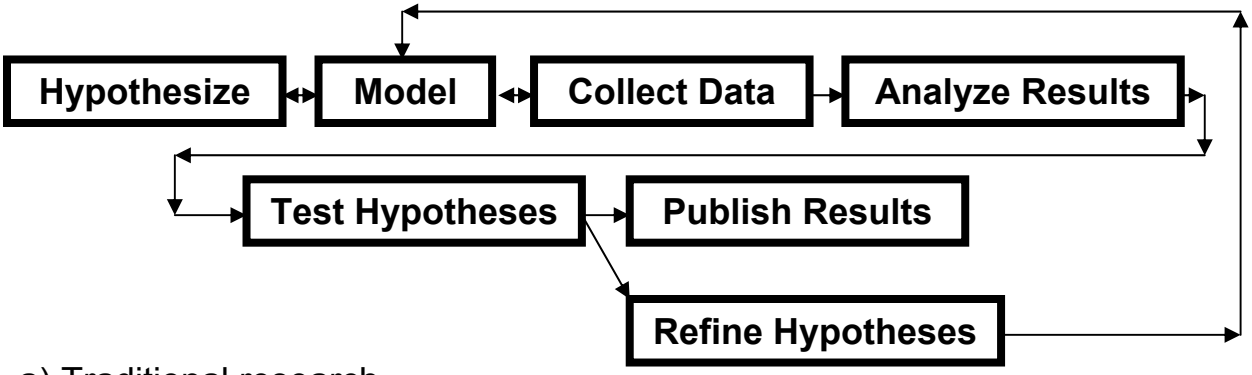
Figure 4. Model of stress from a tragedy and the factors related

Figure 5. Patterns of Facebook blog postings on April 16 and afterwards

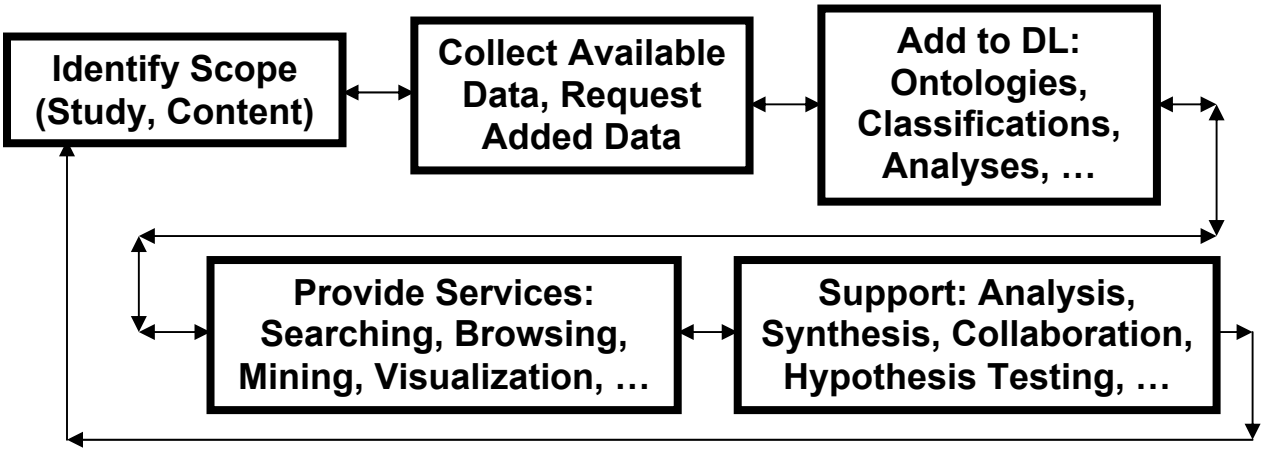
Figure 6. How often tags appear in the submitted works of the April 16 Archive

Figure 7. Themescapes view from IN-SPIRE for all contributions to the April 16 Archive

Figure 8. Galaxy view from IN-SPIRE for all contributions to the April 16 Archive



a) Traditional research

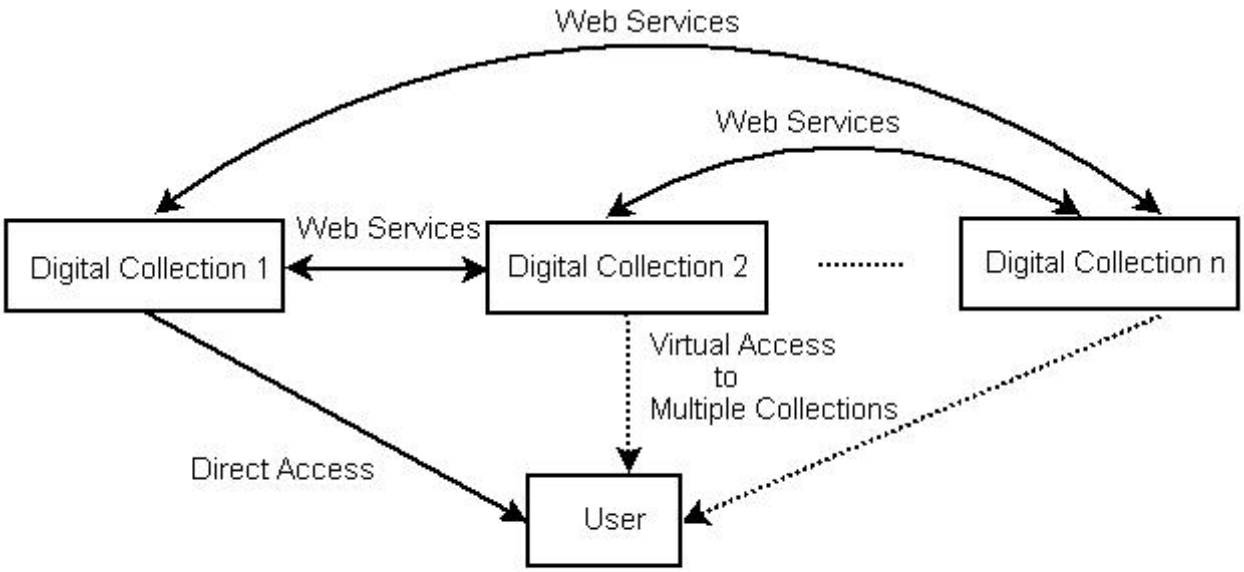


b) DL-supported research

Fig. 1

a) Traditional Library Model

b) Web Information Model



c) Web 2.0 Service Oriented Architecture Model

Fig. 2

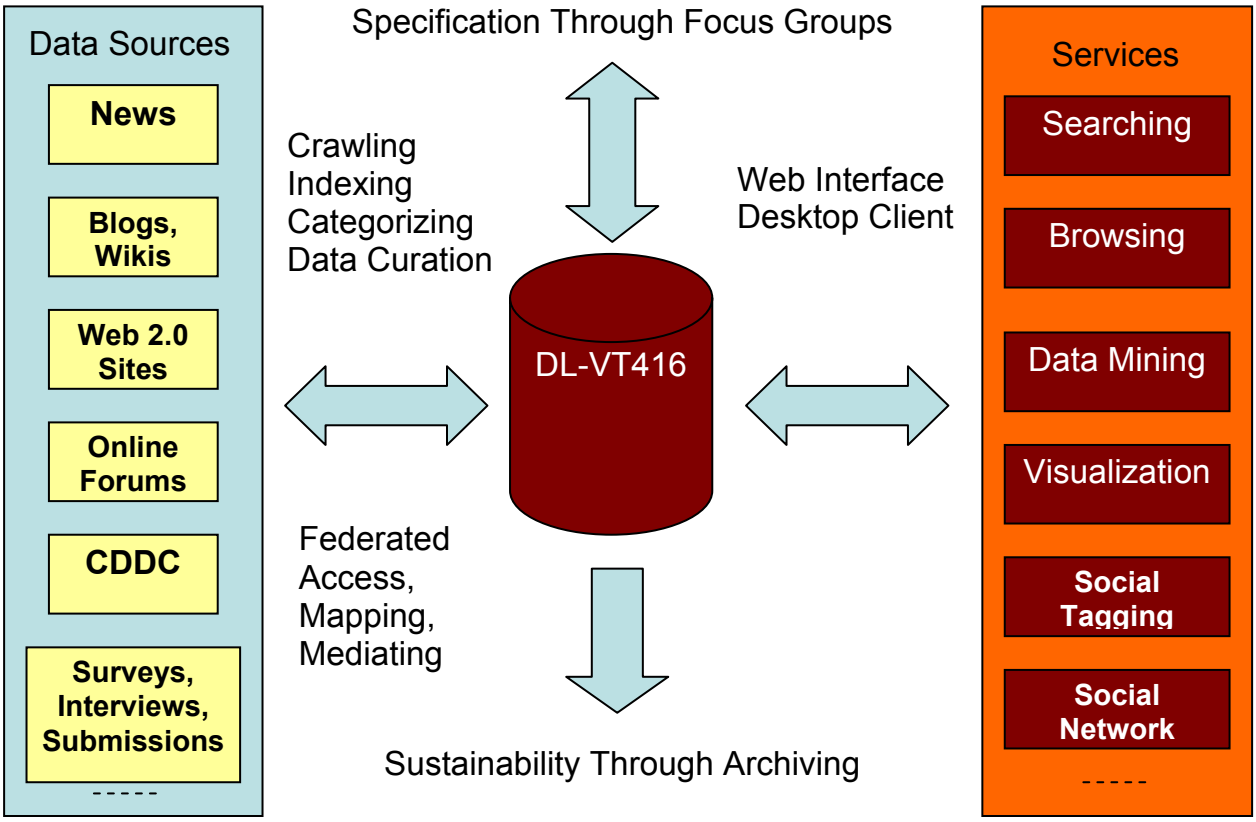

Data Sources
News
Blogs, Wikis
Web 2.0 Sites
Online Forums
CDDC
Surveys, Interviews, Submissions
- - - - -
Specification Through Focus Groups
Crawling
Indexing
Categorizing
Data Curation
Web Interface
Desktop Client
DL-VT416
Federated
Access,
Mapping,
Mediating
Sustainability Through Archiving
Services
Searching
Browsing
Data Mining
Visualization
Social Tagging
Social Network
- - - - -

Fig. 3

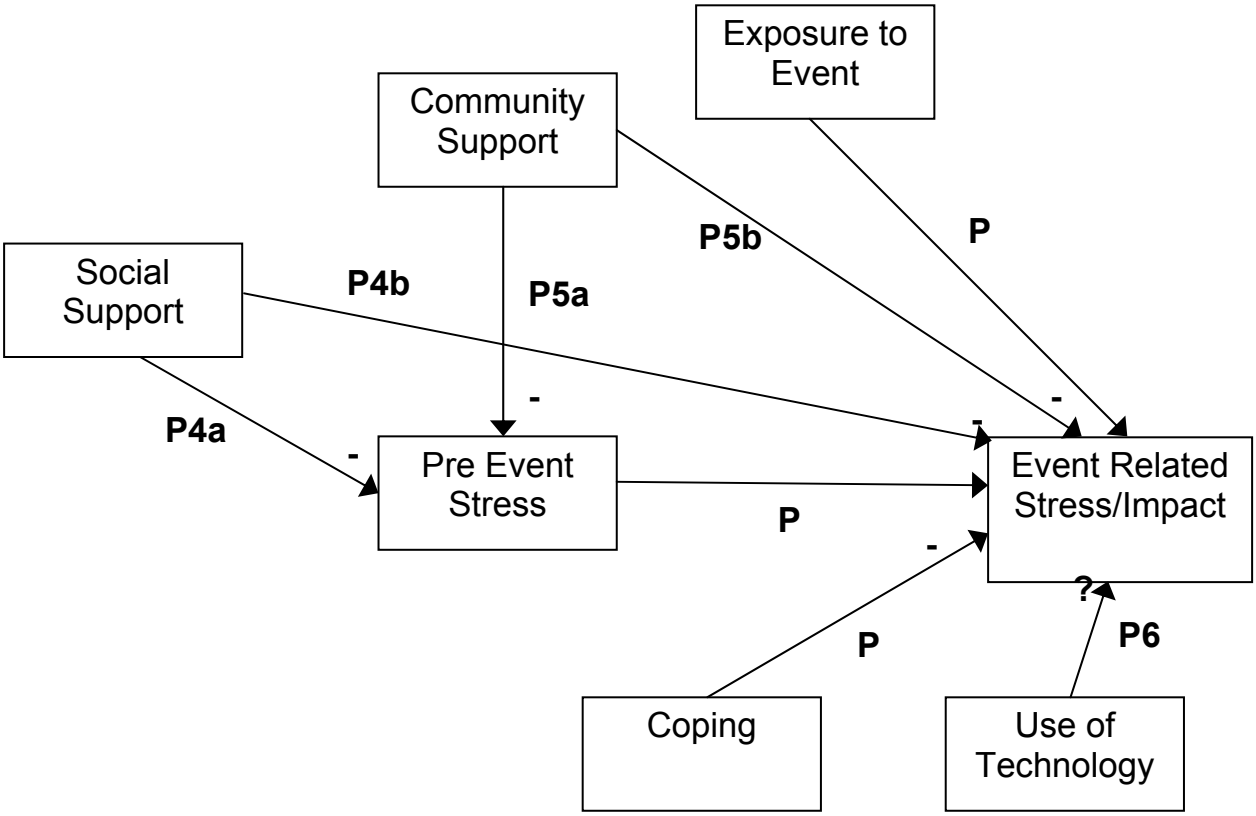

Exposure to Event
Community Support
Social Support
Pre Event Stress
Event Related Stress/Impact
Coping
Use of Technology
P5b
P
P4b
P5a
P4a
-
-
-
-
P
-
?
P
P6

Fig. 4

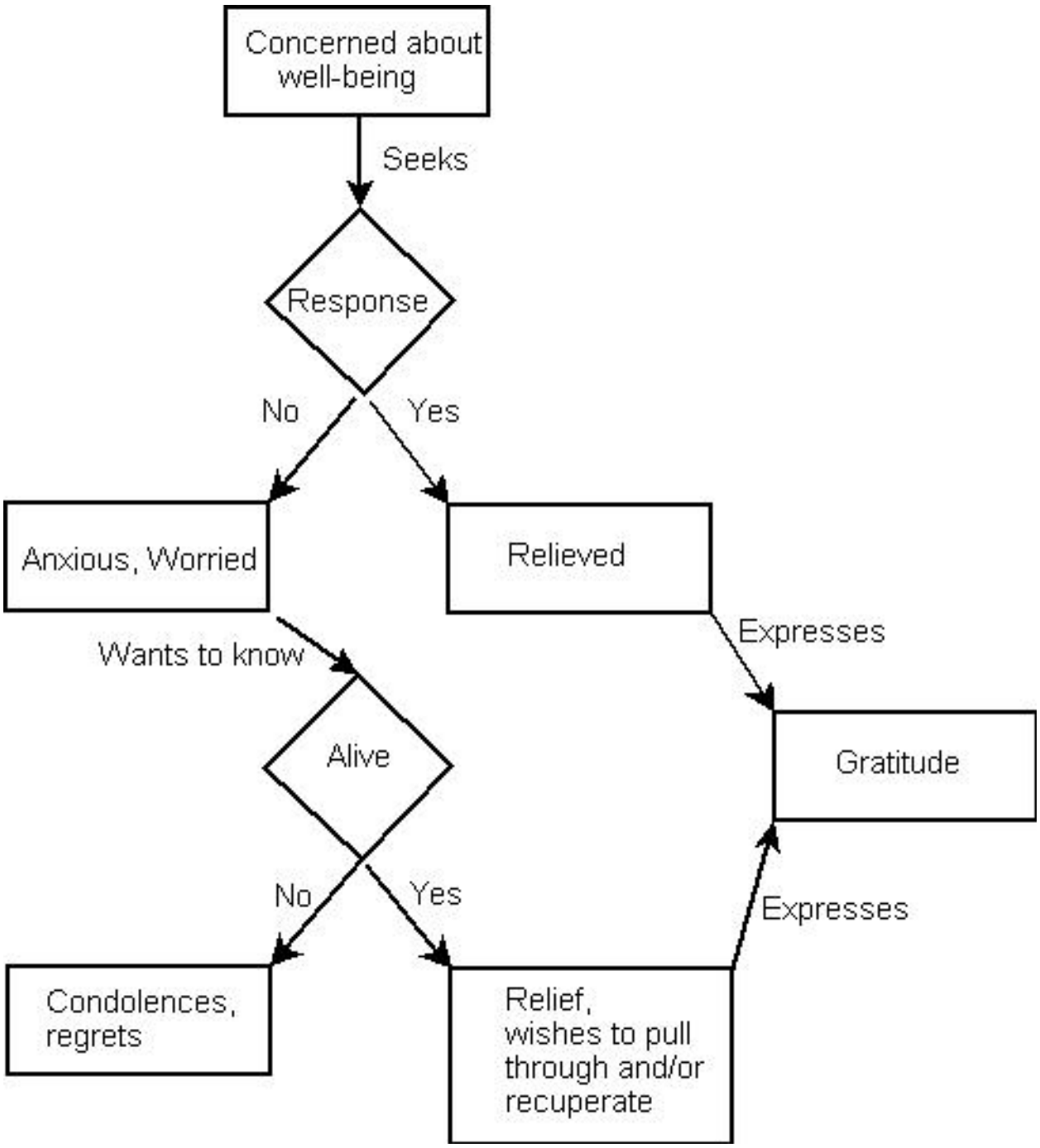
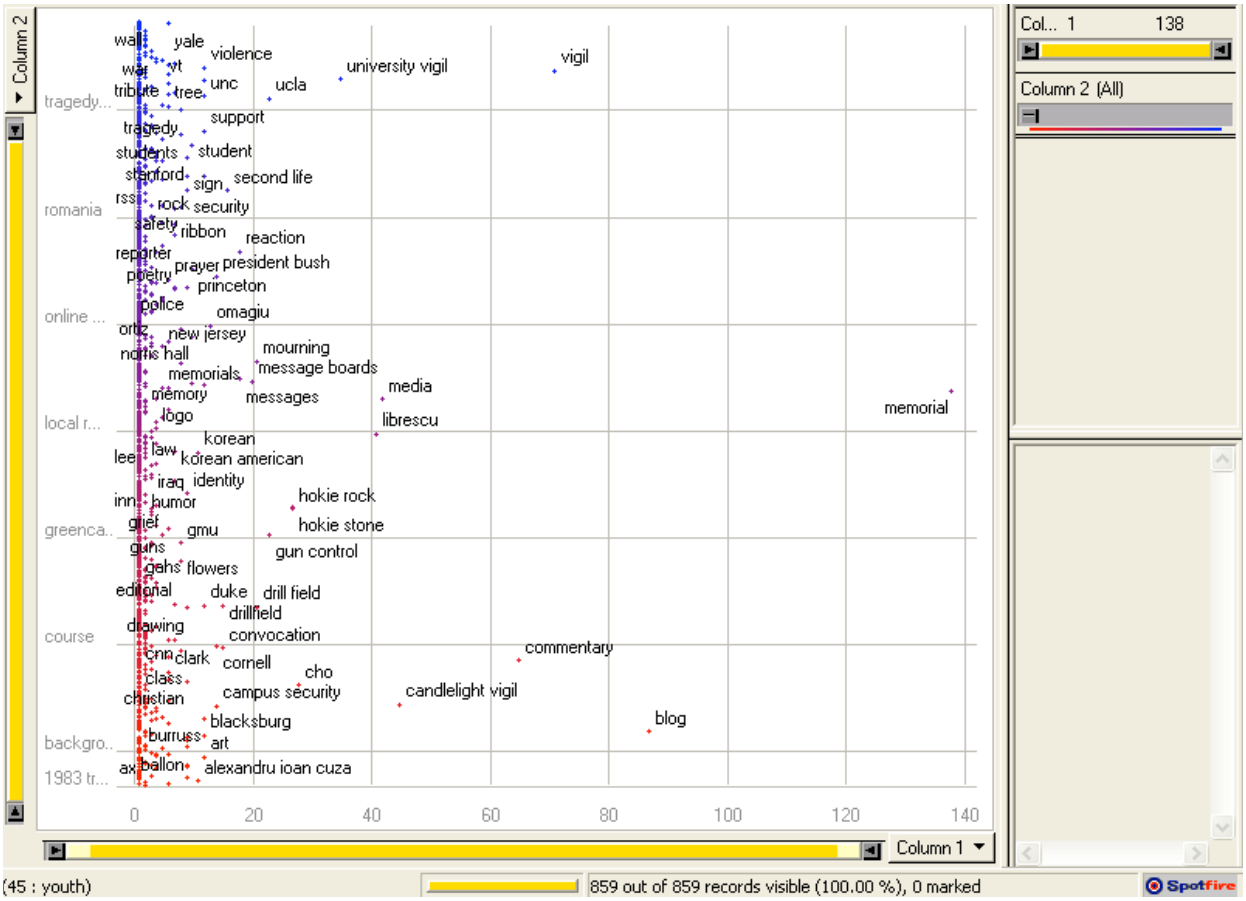

Concerned about well-being
Seeks
Response
No
Yes
Anxious, Worried
Relieved
Wants to know
Expresses
Alive
Gratitude
No
Yes
Expresses
Condolences, regrets
Relief, wishes to pull through and/or recuperate

Fig. 5


Column 2
wall
yale
violence
university vigil
vigil
war
vt
tribute
tree
unc
ucla
tragedy...
support
tragedy
students
student
stanford
sign
second life
rss
rock
security
romania
safety
ribbon
reaction
reporter
poetry
prayer
president bush
princeton
police
online ...
omagiu
ortiz
new jersey
north hall
mourning
message boards
memorials
media
memory
messages
memorial
local r...
librescu
korean
law
korean american
identity
iraq
hokie rock
humor
hokie stone
greenca...
grief
gmu
gun control
flowers
editorial
duke
drill field
drillfield
drawing
convocation
course
commentary
clark
cornell
class
cho
campus security
candlelight vigil
christian
blacksburg
blog
burruss
art
backgro...
ballon
alexandru ioan cuza
1983 tr...
0
20
40
60
80
100
120
140
Col... 1
138
Column 2 (All)
Column 1
(45 : youth)
859 out of 859 records visible (100.00 %), 0 marked
Spotfire

Fig, 6

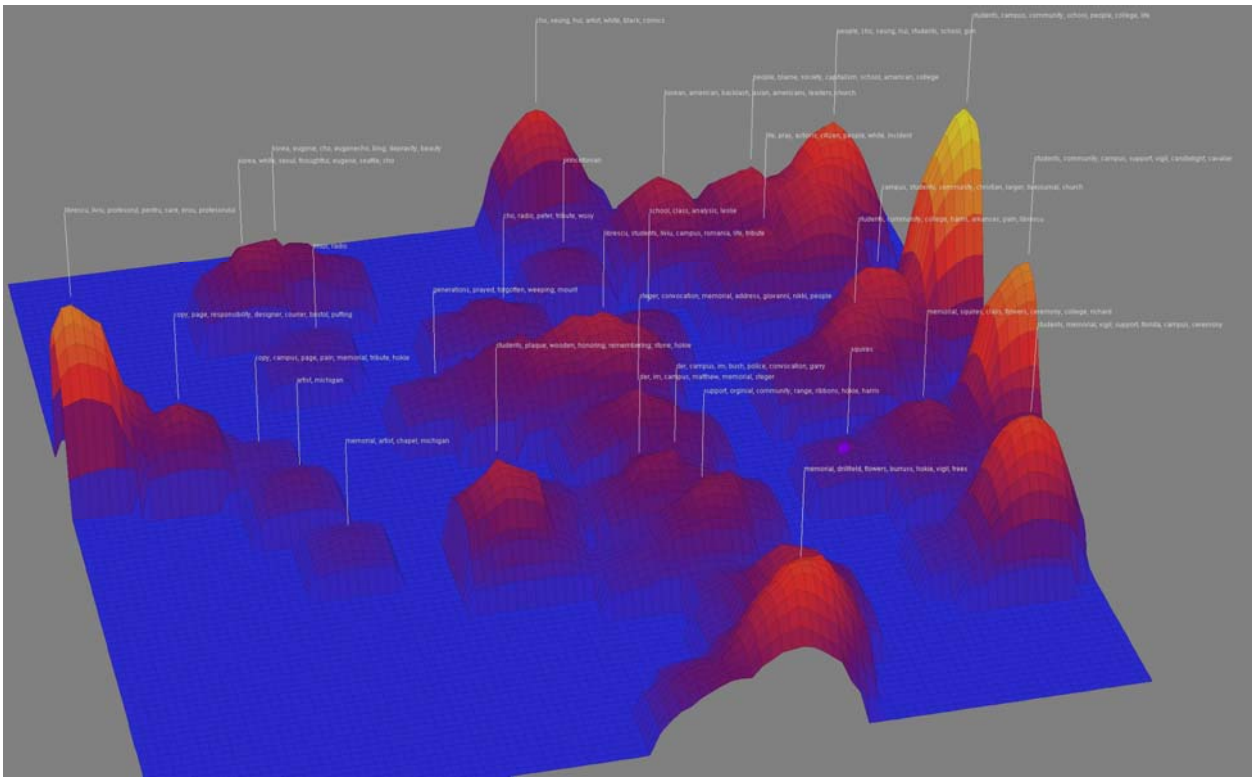


Fig. 7

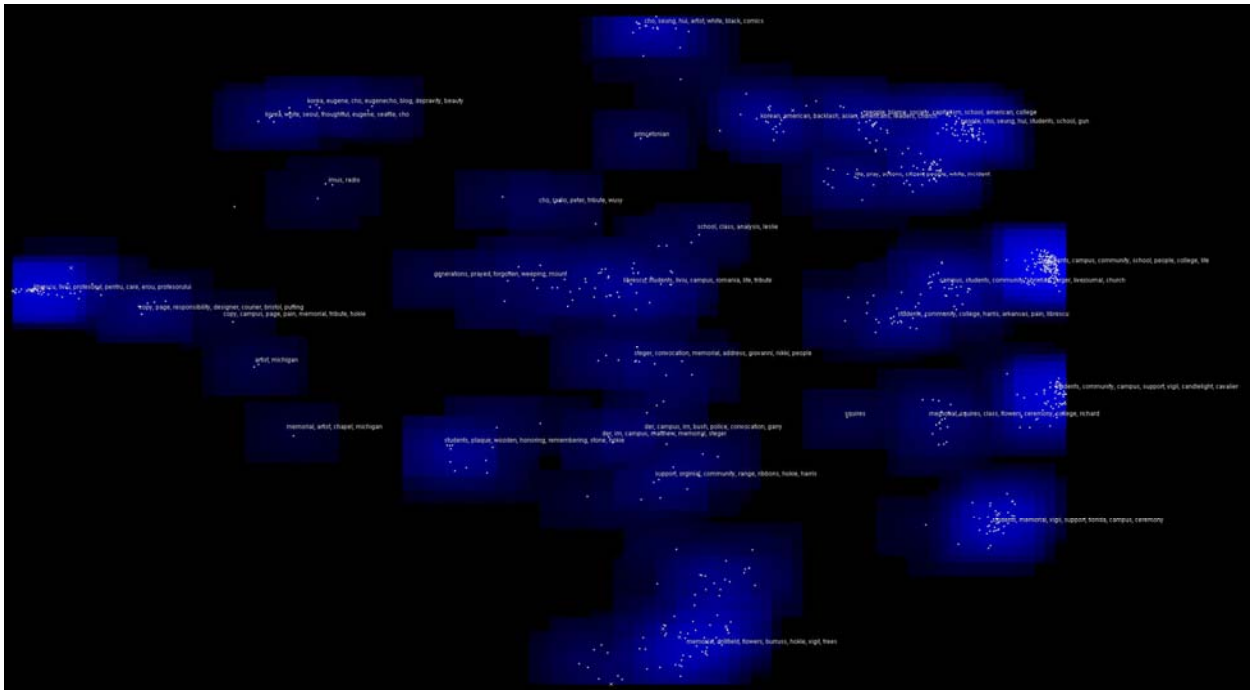


Fig. 8

References

Ahlberg, C., & Shneiderman, B. (1994). Visual information seeking: tight coupling of dynamic query filters with starfield displays. In S. Dumais & J. Olson (Eds.), Proceedings of CHI '94, the SIGCHI Conference on Human Factors in Computing Systems: Celebrating interdependence (Boston, Massachusetts, United States, April 24 - 28, 1994) (pp. 313-317). New York, NY: ACM Press.

Amavizca, M., Sánchez, J. A., & Abascal, R. (1999). 3DTree: Visualization of large and complex information spaces in the Floristic Digital Library, Proceedings of Segundo Encuentro de Computación (ENC'99). Pachuca, Hidalgo, México.

Arms, W. Y. (2000). Digital Libraries. Cambridge, MA: MIT Press.

Association, A. P. (2000). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, D.C.: American Psychiatric Association.

Atkins, D. E., Droegemeier, K. K., Feldman, S. I., Garcia-Molina, H., Klein, M. L., Messerschmitt, D. G., Messina, P., Ostriker, J. P., & Wright, M. H. (2003). Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure. Arlington, VA: NSF.

Ball, R., & North, C. (2007). Realizing Embodied Interaction for Visual Analytics through Large Displays. Computers & Graphics (C&G), 31(3), 380-400.

Breslau, N., Davis, G., Andreski, P., & Peterson, E. (1991). Traumatic events and posttraumatic stress disorder in an urban population of young adults. Archives of General Psychiatry, 48(3), 216-222.

Candela, L., Manghi, P., & Pagano, P. (2007). An Architecture for Type-based Repository Systems. In D. Castelli & Y. Ioannidis (Eds.), Proc. Second Workshop on Foundations of Digital Libraries, in conjunction with 11th European Conference on Research and Advanced Technologies on Digital Libraries (ECDL 2007), Budapest, Hungary, September 20, 2007: DELOS.

Carroll, D., Ebrahim, S., Tilling, K., Macleod, J., & Smith, G. D. (2002). Admissions for myocardial infarction and World Cup football: database survey. British Medical Journal, 325, 1439-1442.

CDDC. (2007). The April 16 Archive. Blacksburg, Virginia: Virginia Tech Center for Digital Discourse and Culture, http://www.cddc.vt.edu/.

Chen, H., Chung, W., Xu, J. J., Wang, G., Qin, Y., & Chau, M. (2004). Crime data mining: a general framework and some examples. IEEE Computer, 37(4), 50-56.

Chen, S., & Fox, E. A. (1996). Guest Editors' Introduction to Special Issue on Digital Libraries. Journal of Visual Communication and Image Representation, 7(1).

Domingos, P., & Richardson, M. (2001). Mining the Network Value of Customers, KDD '01: Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, San Francisco, California (pp. 57-66). New York: ACM Press.

Eubank, S., Guclu, H., Kumar, V. S. A., Marathe, M. V., Srinivasan, A., Toroczkai, Z., Wang, N., & Team, T. E. (2004). Modelling disease outbreaks in realistic urban social networks. Nature, 429, 180-184.

Everson-Rose, S. A., & Lewis, T. T. (2005). Psychosocial factors and cardiovascular diseases. Annual Review of Public Health, 26, 469-500.

Facebook. (2007). Facebook. Retrieved, 2007, from the World Wide Web: http://www.facebook.com

Fan, W., Wallace, L., Rich, S., & Zhang, Z. (2006). Tapping the Power of Text Mining. Communications of the ACM, 49(9), 76-82.

Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD Process for Extracting Useful Knowledge from Volumes of Data. Commun. ACM, 39(11), 27-34.

Fox, E., & France, R. (1987). Architecture of an Expert System for Composite Document Analysis, Representation and Retrieval. International Journal of Approximate Reasoning, 1(2), 151-175.

Fox, E., Suleman, H., Madalli, D., & Cassel, L. (2004). Digital Libraries. In M. Singh (Ed.), Practical Handbook of Internet Computing (pp. chapter 4).

Fox, E. A. (1993). Digital Libraries ("hot topics"). IEEE Computer, 26(11), 79-81.

Fox, E. A. (1999). The 5S Framework for Digital Libraries and Two Case Studies: NDLTD and CSTC, Proceedings NIT'99, Taipei, Taiwan, Aug. 18-20.

Fox, E. A. (2007a). LIKES: Living in the Knowledge Society, invited talk 20 July 2007, University of Arizona, Tucson, AZ. Tucson, AZ.

Fox, E. A. (2007b). Living In the KnowlEdge Society (LIKES) Community Building Project. Virginia Tech. Retrieved, 2007, from the World Wide Web: http://likes.dlib.vt.edu/

Fox, E. A., Akscyn, R. M., Furuta, R. K., & Leggett, J. J. (1995). Guest Editors' Introduction to Digital Libraries, Special Issue. Communications of the ACM, 38(4), 22-28.

Fox, E. A., Eaton, J., McMillan, G., Kipp, N., Weiss, L., Arce, E., & Guyer, S. (1996). National Digital Library of Theses and Dissertations: A Scalable and Sustainable Approach to Unlock University Resources. D-Lib Magazine, 2(8).

Fox, E. A., & Marchionini, G. (1998). Toward a Worldwide Digital Library: Guest Editors' Introduction to Special Section on Digital Libraries: Global Scope, Unlimited Access. Comm. ACM, 41(4), 29-32.

Fox, E. A., & Sornil, O. (1999). Digital Libraries. In R. Baeza-Yates & B. Ribeiro-Neto (Eds.), Modern Information Retrieval (pp. 415-432). Harlow, England: ACM Press / Addison-Wesley-Longman.

Fox, E. A., & Sornil, O. (2000). Digital Libraries. In A. Ralston & E. D. Reilly & D. Hemmendinger (Eds.), Encyclopedia of Computer Science, 4th edition (pp. 576-581). London: Nature Publishing Group.

Fox, E. A., & Urs, S. (2002). Digital Libraries. In B. Cronin (Ed.), Annual Review of Information Science and Technology (ARIST), Ch. 12 (Vol. 36, pp. 503-589): American Society for Information Science.

Fox, E. E. D. (2003). NDLTD: Networked Digital Library of Theses and Dissertations. Retrieved, from the World Wide Web: http://www.ndltd.org

Friedman, T. L. (2006). The World Is Flat [Updated and Expanded]: A Brief History of the Twenty-first Century. New York: Farrar, Straus and Giroux.

Goncalves, M. A. (2004). Streams, Structures, Spaces, Scenarios, and Societies (5S): A Formal Digital Library Framework and Its Applications. Unpublished Doctoral dissertation, Virginia Tech, Blacksburg, VA.

Goncalves, M. A., & Fox, E. A. (2002). 5SL - A Language for Declarative Specification and Generation of Digital Libraries. In G. Marchionini (Ed.), Proc. JCDL'2002, Second ACM / IEEE-CS Joint Conference on Digital Libraries, July 14-18, Portland, Oregon (pp. 263-272): ACM.

Goncalves, M. A., Fox, E. A., Watson, L. T., & Kipp, N. A. (2004). Streams, Structures, Spaces, Scenarios, Societies (5S): A Formal Model for Digital Libraries. ACM Transactions on Information Systems, 22(2), 270-312.

Google. (2007). Google Maps. Google. Retrieved, 2007, from the World Wide Web: http://maps.google.com/

Gorton, D. C. (2007). Practical Digital Library Generation into DSpace with the 5S Framework. Unpublished Master's, Virginia Tech, Blacksburg.

Green, B. L., Wilson, J. P., & Lindy, J. D. (1985). Conceptualizing post-traumatic stress disorder: A psychosocial framework. In C. R. Figley (Ed.), Trauma and its wake (pp. 53-69). New York: Brunner/Mazel, Inc.

Havre, S., Hetzler, E., Whitney, P., & Nowell, L. (2002). ThemeRiver: Visualizing Thematic Changes in Large Document Collections. IEEE Transactions on Visualization and Computer Graphics, 8(1), 9-20.

Hearst, M. A. (1995). Tilebars: Visualization of Term Distribution Information in Full Text Information Access, Conference Proceedings on Human Factors in Computing Systems, CHI'95 (pp. 59-66). Denver: ACM.

Hodges, S., Williams, L., & Wood, K. (2007). Sensors and Devices: SenseCam. Microsoft Research. Retrieved, from the World Wide Web: http://research.microsoft.com/sendev/projects/sensecam/

Jung, V. (1999). MetaViz: Visual Interaction with Geospatial Digital Libraries (Technical report TR-99-017): ICSI, U. Ca. Berkeley.

Kark, J. D., Goldman, S., & Epstein, L. (1995). Iraqi missile attacks on Israel: The association of mortality with a life-threatening stressor. Journal of the American Medical Association, 273, 1208-1210.

Kaza, S., Marshall, B., Xu, J., Wang, A. G., Gowda, H., Atabakhsh, H., Petersen, T., Violette, C., & Chen, H. C. (2005). BorderSafe: Cross-jurisdictional information sharing, analysis, and visualization. Paper presented at the Intelligence and Security Informatics Conference.

Lagoze, C., & Van de Sompel, H. (2007). Compound Information Objects: The OAI-ORE Perspective. Open Archives Initiative Retrieved, from the World Wide Web: http://www.openarchives.org/ore/documents/CompoundObjects-200705.html

LaGreca, A. M., Silverman, W. K., Vernberg, E. M., & Prinstein, M. J. (1996). Symptoms of posttraumatic stress in children after Hurricane Andrew: A prospective study. Journal of Consulting and Clinical Psychology, 64, 712-723.

Leor, J., Poole, W. K., & Kloner, R. A. (1996). Sudden cardiac death triggered by an earthquake. New England Journal of Medicine, 334, 413-419.

Lesk, M. (1997). Practical Digital Libraries: Books, Bytes and Bucks. San Francisco: Morgan Kaufmann Publishers.

Lesk, M. (2005). Understanding Digital Libraries, 2nd Edition. San Francisco: Morgan Kaufmann, Elsevier.

Lovallo, W. R. (2005). Stress and health: Biological and psychological interactions (2nd ed.). Thousand Oaks, CA: Sage.

McQuerry, D. (2004). IN-SPIRE™ Visual Document Analysis. Richland, WA: Pacific Northwest National Laboratory (PNNL), http://www.pnl.gov.

Michaels, A. J., Michaels, C. E., Moon, C. H., Smith, J. S., Zimmerman, M. A., Taheri, P. A., & Peterson, C. (1999). Posttraumatic stress disorder after injury: Impact on general health

outcome and early risk assessment. Journal of Trauma: Injury, Infection, and Critical Care, 47(3), 460-467.

MIT. (2003). DSpace: Durable Digital Depository [Web site]. MIT. Retrieved April, 2004, from the World Wide Web: http://dspace.org

MySpace.com. (2007). MySpace.com: a place for friends. Retrieved, 2007, from the World Wide Web: http://www.myspace.com

North, C. (2005). Information Visualization. In G. Salvendy (Ed.), Handbook of Human Factors and Ergonomics, 3rd Edition (pp. 1222-1246). New York: John Wiley & Sons.

Nowell, L. (1997). Graphical Encoding for Information Visualization: Using Icon Color, Shape and Size to Convey Nominal and Quantitative Data. Unpublished Ph.D. Dissertation, Virginia Tech, Blacksburg, VA.

Nowell, L. T., & Fox, E. A. (1995). Envision: Information Visualization in a Digital Library: Proc. ACM SIGIR'95, Seattle, WA, July 10.

NSF, C. C. o. (2007). Cyberinfrastructure Vision for 21st Century Discovery, NSF 07-28. Arlington, VA: National Science Foundation Cyberinfrastructure Council.

O'Reilly, T. (2005). What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. O'Reilly Media. Retrieved, from the World Wide Web: http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html

OCLC. (2004). XTCat NDLTD Union Catalog. OCLC. Retrieved, from the World Wide Web: http://alcme.oclc.org/ndltd/index.html

Petinot, Y., Giles, C. L., Bhatnagar, V., Teregowda, P. B., Han, H., & Councill, I. G. (2004). A service-oriented architecture for digital libraries. Paper presented at the Proceedings of the Second International Conference on Service Oriented Computing.

Porges, S. W. (1995). Cardiac vagal tone: A physiological index of stress. Neuroscience and Biobehavioral Reviews, 19, 225-233.

Prensky, M. (2001). Digital natives, digital immigrants. On the Horizon, 9(5), 1-6.

Ramakrishnan, N., & Grama, A. Y. (1999). Data Mining: From Serendipity to Science. IEEE Computer, 32(8), 34-37.

Raymond, E. S. (1999). The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary. Sebastopol, CA: O'Reilly and Associates, Inc.

Saraiya, P., North, C., & Duca, K. (2005). Visualization of Biological Pathways: Requirements Analysis, Systems Evaluation, and Research Agenda. Information Visualization, 4(3), 191-205.

Saraiya, P., North, C., Lam, V., & Duca, K. (2006). An Insight-based Longitudinal Study of Visual Analytics. IEEE Transactions on Visualization and Computer Graphics, 12(6), 1511-1522.

Seligman, M. E. P., Maier, S., & Solomon, R. L. (1971). Unpredictable and uncontrollable aversive events. In F. R. Brush (Ed.), Aversive conditioning and learning (pp. 347-400). New York: Academic Press.

Shen, R., Vemuri, N. S., Fan, W. G., da S Torres, R., & Fox, E. A. (2006). Exploring digital libraries: Integrating browsing, searching, and visualization, Opening Information Horizons (pp. 1-10).

Shen, R., Wang, J., & Fox, E. A. (2002). A Lightweight Protocol between Digital Libraries and Visualization Systems, JCDL Workshop on Visual Interfaces to Digital Libraries, Proceedings of the Second ACM/IEEE-CS Joint Conference on Digital Libraries (pp. 425). Portland, USA: ACM Press.

Spence, R. (2007). Information Visualization: Design for Interaction (2nd Edition): Prentice-Hall, Inc.

Thayer, J. F., & Lane, R. D. (2007). The role of vagal function in the risk for cardiovascular disease and mortality. Biological Psychology, 74, 224-242.

Thomas, J. J., & Cook, K. A. (Eds.). (2005). Illuminating the Path: Research and Development Agenda for Visual Analytics: IEEE.

Tollis, I. (1996). Graph Drawing and Information Visualization. ACM Computer Surveys, 28(4es), 19.

Van de Sompel, H., & Lagoze, C. (2000). Open Archives Initiative [WWW site]. Cornell University. Retrieved, 2007, from the World Wide Web: http://www.openarchives.org

Viegas, F., Wattenberg, M., Ham, F. v., Kriss, J., & McKeon, M. (2007). Many Eyes: A Site for Visualization at Internet Scale, Proc. IEEE Information Visualization (InfoVis) 2007, Oct. 2007.

Wang, J. (2002). VIDI: A lightweight protocol between visualization systems and digital libraries. Blacksburg, VA: Virginia Tech, Department of Computer Science Masters thesis.

Weibel, S. (1995). Metadata: The Foundations of Resource Description. D-Lib Magazine, 1(7).

Weiss, J. M. (1991). Stress-induced depression: Critical neurochemical and electrophysiological changes. In I. J. Madden (Ed.), Neurobiology of learning, emotion, and affect (pp. 123-154). New York: Raven.

Wise, J. A., Thomas, J. J., Pennock, K., Lantrip, D., Pottier, M., Schur, A., & Crow, V. (1995). Visualizing the non-visual: spatial analysis and interaction with information from text documents, Infovis, Proceedings on Information Visualization (pp. 51).

Wong, P. C., Whitney, P., & Thomas, J. (1999). Visualizing Association Rules for Text Mining, IEEE Information Visualization '99 (pp. 120-123). Los Alamitos, CA: IEEE CS Press.

Zhu, Q. (2002). 5SGraph: A Modeling Tool for Digital Libraries. Unpublished Master's Thesis, Virginia Tech – Department of Computer Science.