The Open Video Digital Library: A Möbius Strip of Research and Practice

Gary Marchionini and Barbara M. Wildemuth
School of Information and Library Science, University of North Carolina, Chapel Hill, NC 27599. E-mail: marchionini@unc.edu, wildem@ils.unc.edu

Gary Geisler
School of Information, University of Texas, Austin, TX 78712. E-mail: geisler@ischool.utexas.edu

The Open Video Digital Library (OVDL) provides digital video files to the education and research community and is distinguished by an innovative user interface that offers multiple kinds of visual surrogates to people searching for video content. The OVDL is used by several thousand people around the world each month and part of this success is due to its user interface. This article examines the interplay between research and practice in the development of this particular digital library with an eye toward lessons for all digital libraries. We argue that theoretical and research goals blur into practical goals and practical goals raise new research questions as research and development progress—this process is akin to walking along a Möbius strip in which a locally two-sided surface is actually part of a globally one-sided world. We consider the gulf between the theories that guide current digital library research and current practice in operational digital libraries, provide a developmental history of the OVDL and the research frameworks that drove its development, illustrate how user studies informed its implementation and revision, and conclude with reflections and recommendations on the interplay between research and practice.

Introduction

There are inherent tensions between academic researchers and practitioners, who each have distinct objectives and reward systems. Researchers are motivated by publications, high-risk innovative thinking, and formal models, whereas practitioners are motivated by implementations, high return on investment thinking, and the lessons they have learned from their professional experiences. These tensions are especially manifested in professional schools in medicine, engineering, business, and information in which practice dwarfs research in terms of public perception, number of participants, resource allocation, and direct human impact. It is informative to examine how these tensions develop in a new field such as digital libraries with an eye toward discovering ways to minimize or even capitalize on these tensions. This article presents the case of the Open Video Digital Library (OVDL) as an example of how research and practice have gone hand in hand to advance both the development of new knowledge through research and the application of that new knowledge to improved information service. The article first considers the gulf between the theories that guide current digital library research and current practice in operational digital libraries, then provides a developmental history of the OVDL and the research frameworks that drove it, illustrates how user studies have informed implementation, and concludes with reflections and recommendations on the interplay between research and practice.

Digital libraries (DLs) are a derivative of advancing information technologies and the needs of individuals and organizations to manage huge volumes of digital data better in all manner of formats and knowledge domains. As such, they have dependencies on engineering, psychology, and organizational behavior, among other disciplines. The initial motivation for digital library research and development was substantial funding by governments (e.g., the National Science Foundation [NSF] and other agencies in the United States, and Esprit, the information technologies program in the European Union), and this funding spurred computer and information scientists to develop basic infrastructures for working digital libraries (e.g., storage and retrieval algorithms, multimedia representation techniques, metadata schemes). Few of the early research projects directly created working digital libraries but rather prototypes that demons trated underlying research solutions. It remained up to libraries, businesses, and government agencies to transfer and adapt the fruits of

Received July 18, 2005; revised October 10, 2005; accepted October 10, 2005

 • Published online 17 August 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asi.20336

research to actual practical digital libraries. This two-phase pattern (i.e., theoretical and empirical research conducted in the first phase, followed by business and industry identification of useful findings and application of them to practice) is the traditional modus operandi of research and practice. This rather *ad hoc* strategy for reaping the fruits of research is being increasingly challenged by globalization and advancing technology. Universities are increasingly forming partnerships with industry (e.g., in genomics, in information technology) to do problem-based research, and government funding has begun to require linkages between research and impact. For example, NSF now requires all proposals to discuss impact explicitly; in fact, the NSF Digital Government Program explicitly links government partners with university researchers to ensure that research results are put in place in work-a-day agencies; and the National Technology Transfer Center (http://www.nttc.edu/) is specifically chartered to support such linkages. Marchionini and Levi (2003) provide a case study of how government practice and academic research have worked hand in hand to improve government information services. The tensions between research and practice have long been at play in high-profile professional schools such as those for medicine and business and are becoming increasingly apparent in information schools in research universities. Digital libraries are one realm in which this hybridized state of affairs is clearly explicated because testbeds must be realistic to demonstrate efficacy. One purely practical point of view is that if we build a DL well, users will come. A more theoretical point of view is that we must have a complete understanding of user needs before building a DL. Clearly, the prudent approach lies somewhere between these two perspectives, and it is this balanced approach to DL research and development that has motivated the Open Video DL efforts over the past half-decade.

Digital library research sits at the intersection of information science, computer science, and social science. Volumes from Lesk (1997) and Arms (2000) emphasize the economic and technical aspects of DLs, and Borgman (2000) emphasizes the social and library aspects of DLs. Fox and his colleagues (Goncalves, Fox, Watson, & Kipp, 2004) have developed a technical framework, S5, for DL architectures; Reddy and his colleagues have mobilized an international effort called the Million Book Project (http://www.rr.cs.cmu.edu/mbdl.htm) to "democratize" knowledge; Google Print has initiated an ambitious project to digitize books at five major research libraries (http://print.google.com/); and a 2003 workshop drew many DL leaders together to consider DLs and information resources as key components to the emerging cyberinfrastructure (http://www.sis.pitt.edu/~dlwkshop/paper_smith.html). The Open Video Project has evolved on the basis of the theory proposed by Douglas Engelbart in 1963 that computational tools can augment the intellect. Our view is that DLs are beginning to augment individual as well as collective memories and ultimately will augment other human cognitive processes as well. This general belief about DLs is manifested in two theoretical frameworks that continue to evolve and have provided a foundation for the Open Video work for more than a decade.

The first framework considers DLs to be such powerful extensions of traditional libraries that qualitative shifts in form and function arise and new properties emerge. This notion is manifested in the concept of *sharium*, a term meant to suggest that a DL is a forum for mutual sharing of intellectual resources. A sharium goes beyond providing information in a curated collection, to inviting active participation in the form of collaboration and contributions from all users and to providing flexible means for reusing information resources (Marchionini, 1999). In the sharium, there are four key extensions to traditional libraries that allow personal and collective augmentations of intelligence:

1. Rather that only visiting the library to find information, people also contribute ideas, time and expertise, and information products to the common infrastructure.
2. Tools for finding and using textual as well as multimedia resources and interacting with different online communities are available, as are tools. These tools are easy to use and provide alternative interfaces and representations to suit the capabilities and preferences of a global population of users.
3. Tools and services are available for small group (e.g., collaborative writing and annotation) or massive collaboration (e.g., SETI [Search for Extraterrestrial Intelligence]-like efforts, http://setiathome.ssl.berkeley.edu/).
4. Policies and tools for easily reusing and reintegrating information are in place (e.g., open source and Creative Commons licenses).

Clearly, no sharium exists today; however, selected elements are merging in projects such as ibiblio (http://ibiblio.org/), in which more than 1,500 separate collections are maintained by different communities and indexed across the repository; SourceForge (http://www.sourceforge.org), in which thousands of open source computer programs are available for download; the Alexandria Digital Library (http://www.alexandria.ucsb.edu/), where powerful tools for finding spatial information and incorporating the results into classroom are available; the Perseus Digital Library (http://www.perseus.tufts.edu/), in which text and images are accessible through powerful multilingual search and analysis tools; the Open Directory Project (http://www.opendirectory.org), which supports community-based cataloging of Web sites, and wikipedia (http://www.wikipedia.org), which supports community-contributed encyclopedia entries. Beginning with the Baltimore Learning Community Project and now with the OVDL, we aim to instantiate some of these elements of the sharium framework.

The second framework addresses how people interact with electronic information, and we have been developing and testing what we call the *AgileViews* approach to interface design (Geisler, 2003; Marchionini, Geisler, & Brunk, 2000). The idea is that people should be able effortlessly (i.e., with agility) to shift among different representations for information spaces and objects as they seek and use information. In an AgileViews environment, it is easy to shift among overviews of collections, previews of objects, historical views

DOI: 10.1002/asi

(reviews) previously experienced, views shared with others, and peripheral views of context. The framework is rooted in three human-computer interaction principles. First, people should be in control of systems and information. The design objective is to provide direct manipulation widgets to control representations. We have leveraged simple mouse activity in most of our instantiated work. Second, people should have alternative representations for information. The design objective is to provide a manageable set of alternative displays or inputs. We have focused on designing and testing visual surrogates for video in the OVDL. Third, design is advanced through systematic empirical user testing. We have focused our user evaluations on very specific design decisions in the overall OVDL interface. This article focuses mainly on the AgileViews framework and our evaluation methods and how they evolved as we built and revised a working DL.

The Evolution of the Open Video Digital Library

The Open Video DL Project grew out of a project known as the Baltimore Learning Community (BLC) that aimed to make digital video available to teachers in the mid- to late 1990s (Marchionini et al., 1997). The BLC project digitized U.S. government–produced documentaries as well as commercial video from the Discovery Channel and incorporated them into an educational digital library that included video, Web pages, text, and images from the National Archives and Records Administration. The BLC created novel dynamic query user interfaces and conducted several user studies to assess visual surrogates (Ding, Soergel, & Marchionini, 1999; Komlodi & Marchionini, 1998; Tse, Marchionini, Ding, Slaughter, & Komlodi, 1998). Although the project led to many new approaches to managing and delivering multimedia content, the commercial content limited its application to four schools in Baltimore. Additionally, the Java interfaces and connectivity limitations of that time limited its application even in those schools. The BLC experience stimulated a number of interests in digital video research and development and, in order to build upon what was learned in the BLC, we decided to focus on digital video research questions and build a fully open digital library testbed for this research. In addition to the BLC experience, we benefited from the Informedia Digital Library Project (Wactlar, Christel, Gong, & Hauptmann, 1999; Christel, Smith, Taylor, & Winkler, 1998) in several ways, including their willingness to share approximately 50 hours of U.S. government–produced video that they had digitized and indexed with their extensive metadata scheme.

To inform the research issues and technical challenges associated with digital video and user-centered interfaces to support video retrieval and use, we held two symposiums: one to examine issues of browsing in 1998 and one focused on video retrieval in 1999. One of our early goals was to develop a video testbed that the research and development community could use to compare results (Slaughter, Marchionini, & Geisler, 2000). This effort was inspired by the Text Retrieval Evaluation Conference (TREC) (http://trec.nist.gov/testbeds) for text retrieval, and TREC initiated a video retrieval track in 2000 that used some of the Open Video content. The National Institute of Standards and Technology (NIST) used some OVDL content for the 2005 TREC video retrieval evaluation conference. In 2000 we received a NSF Interactive Systems Program grant to develop and evaluate visual surrogates and incorporate them into highly interactive user interfaces based on our AgileViews user interface design framework (Geisler, Marchionini, Nelson, Spinks, & Yang, 2001; Marchionini et al., 2000). Our main efforts were (1) to create visual surrogates and systematically test their efficacy with user studies and (2) to develop the testbed by acquiring additional video for the library, developing tools and techniques to manage these materials, determining metadata schemes, and building interface prototypes. By 1999 a preliminary design that made approximately 100 Mb of video available in MPEG-1 format was implemented. We worked with the Prelinger Archive to digitize some of their BetaSP content and with the University of Maryland's Human-Computer Interaction Laboratory to add the annual videos they made over a 15-year period. We expanded partnerships to the Internet-2 Distributed Storage Initiative, which made the OVDL a channel in its network and provided distributed storage support for MPEG-1 files. An early summary of the OVDL was provided by Marchionini and Geisler (2002).

In addition to acquiring content, building tools for processing and managing digital video, and developing procedures for managing digital libraries, we focused on our primary research goal of creating and evaluating visual surrogates for video. Between 2001 and 2004 we conducted eight user studies for a variety of visual surrogates. The early user studies and first year of experience led to a major redesign for the OVDL in 2003. The redesign proved highly successful as demonstrated by the usage and media attention received. Since 2003, major new collections from the National Aeronautics and Space Administration (NASA) and Association for Computing Machinery (ACM) conferences were added, as well as smaller collections on a variety of topics. In the spring of 2005 the DL was serving about 15,000 unique visitors each month (more than 1 million hits per month) and included about 2,500 video segments. New studies are under way to continue development in the years ahead, and a new effort to develop a curatorial decision template for digital video preservation has received NSF/Library of Congress funding. These efforts and the theoretical and practical extensions they spawned are discussed in the following sections, but first a high-level summary of the practical and theoretical goals of the project and how they evolved over the years is presented.

The original motivation for developing the Open Video DL was both practical and theoretical. On the practical side, we aimed

- To create a digital video testbed of downloadable files for video retrieval researchers
- To leverage this testbed to serve our own research interests in developing and evaluating highly interactive user interfaces for digital services
- To create useful surrogates for video content

DOI: 10.1002/asi

On the theoretical side, we aimed

- To use the testbed to continue to develop and test a framework for the design of user interfaces (the AgileViews framework)
- To continue to develop a theory of information seeking and information interaction

As the project matured over the years, we extended the practical goals

- To provide a useful general-purpose digital video library appropriate for all users
- To develop a set of tools for backend processing of video assets
- To develop tools for end users to use video
- To develop methods and metrics for assessing user interactions with video surrogates in particular and digital libraries in general

Likewise, the theoretical goals matured and expanded:

- To understand people's interactions with video, especially how understanding emerges as people interact with various metadata and surrogates
- To promote a model of digital video reuse that is open and fine-grained rather than proprietary and based on entire programs
- To develop a generalized cost-benefit model of human interaction with digital video

Most importantly, although we can provide lists of practical and theoretical goals such as these, it is the interplay between these goals that led to the most interesting results. For example, although we had to develop and adapt methods for conducting user studies, a research framework that is an important component of both the design framework and our theoretical understanding of human video interaction emerged. Likewise, our original goal of providing a testbed was greatly expanded over time to creating and maintaining a general purpose video DL as currently instantiated. In effect, theoretical and research goals blur into practical goals and practical goals raise new research questions as research and development progress—this is akin to walking along a Möbius strip in which a locally two-sided surface is actually part of a globally one-sided world.

A Foundation for Both Research and Practice

The technical architecture of the Open Video DL has been key to enabling us to pursue both research and practical goals easily and to use what we have learned in one arena to further work in the other. The original goal of creating a digital video testbed, for example, was a practical one intended to serve real needs of digital video researchers. At the same time, because no publicly available video repositories had yet been developed, we were also interested in furthering understanding of the issues in creating an open source digital video library (Slaughter et al., 2002). As we pursued these goals, the open, nonproprietary infrastructure of the Open Video DL enabled us to add new tools and processes easily while also making the tools, processes, and video content very accessible to those developing ideas for new applications, user studies, and theoretical ideas.

Figure 1 illustrates the basic architecture of the Open Video DL. Users of the public DL connect through Web browser clients to the Open Video server, which runs on the Linux operating system with an Apache Web server. The DL consists of a MySQL database storing video metadata and a PHP-driven Web site that provides a range of services for finding, viewing, and downloading video. The video files themselves require a significant amount of disk space and are thus stored on other, distributed servers, although this technique is predominantly transparent to the end user.

The digital video files offered by OVDL, along with the associated video surrogates and metadata, are generated through a multistage production system. Note that our principle of open acess is reflected in our policy of providing downloadable files that people can reuse as they wish rather

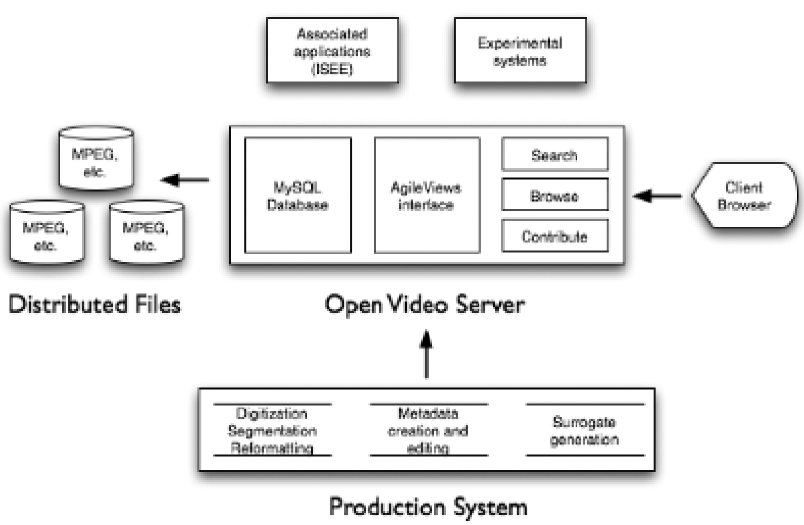

FIG. 1. Open Video Digital Library system architecture.

DOI: 10.1002/asi

than providing video streams that only support playback functionality. Furthermore, keeping with the open source nature of our project and our goal of supporting both research and practice, the production system relies almost exclusively on open source software and homegrown tools. Video has been contributed in a wide range of formats: Beta SP, VHS, and PAL videotapes; previously digitized formats sent to us on CD-ROM and DVD; and existing WWW-based digital files we link to or download. To maximize the potential usefulness of the video content to our broad audience there are two primary decisions we make regarding incoming video: whether the video should be segmented and in which formats it should be offered for download. The segmentation decision is important because digital video files are relatively large (1 hour of medium-quality MPEG-1 video, for example, is approximately 600 Mb) and thus place an additional burden on users when deciding whether they can afford the time, bandwidth, and storage space to download a given video. Our goal has been to try to balance the need for reasonable file sizes with the convenience of a structure of fewer segments for a single video title. As we segment longer videos into smaller segments, we base our segmentation decisions on logical breaks in content rather than segmenting at fixed intervals. This method enables us to produce segments that are meaningful and potentially useful on their own, independently of the other segments that belong to the same video title/program. Incoming video is digitized and segmented primarily with Apple Final Cut Pro.

Because the audience of the Open Video DL is very broad we also try to make videos available in a range of compression formats. For instance, we recently added nearly 500 video titles from National Aeronautic and Space Administration's (NASA's) K-16 Science Education Programs and made each one available in MPEG-1, MPEG-2, MPEG-4, QuickTime, and RealPlayer formats. A user interested in video for technical or research purposes might prefer the significantly larger but higher-quality MPEG-2 format, but a student accessing the site from an international location with limited bandwidth might be satisfied with the lower-quality but much quicker-to-download QuickTime version of the video. For technical reasons not all the video in the Open Video DL is available in each of these formats, but such availability is a general goal of the DL.

While digitizing, segmenting, and reformatting incoming video content are obviously a critical stage of our production process, equally important stages are generating metadata and producing surrogates for each video. Good, complete metadata are crucial in order (1) to give users multiple access points (e.g., genre, date, topic, keywords) they can use to find video in the DL and (2) to describe each video title fully. Our experience has been that we rarely receive contributed video with complete metadata, and, thus, a time-consuming part of the production process is enhancing the metadata for incoming video. Thus far, we have relied primarily on spreadsheet templates for enhancing and correcting video metadata, although we have experimented with dedicated cataloging interfaces (Yang, Mu, & Marchionini, 2003a), and the creation of a video cataloging interface is one of the primary goals of a related project, the Open Video Digital Library Toolkit (http://www.open-video-toolkit.org).

The visual aspect of digital video is clearly a central component to its use, and thus much of our research and practical work has been in the area of visual surrogates. As described later, the visual surrogates we provide in the Open Video DL have been influenced by our user studies; we currently provide three types of visual surrogates for each video segment, in addition to a single-frame representative posterframe: a storyboard, a fast-forward, and a 7-second excerpt. To generate the storyboard surrogates we start with a donated (University of Maryland at College Park UMIACS) program called MERIT, which we ported to Linux, to extract candidate keyframes from each video automatically. A simple Web-based interface displays the candidate keyframes for a given video and enables us to prune them to a reasonably small but representative set. Selected keyframes are stored in the database and are displayed in a storyboard via PHP code when a DL user views a given video's Details page.

To create the fast-forward surrogate we first run an application we created called the *Video Annotation and Summarization Tool* (VAST; Mu & Marchionini, 2003) to extract keyframes at regular intervals from the full video. A script then assembles all the extracted keyframes into a QuickTime file that in effect shows the video as if it were played at 64 times normal speed. We also developed a script to automate the extraction of a 7-second excerpt from a video to create the third form of visual surrogate. Not part of the public DL at the time of this writing are scripts we have also developed, and we have worked with other open source software to explore the potential of other forms of video surrogates. For example, we have experimented with speech to text by using the Sphinx software produced by Carnegie Mellon University (http://www.speech.cs.cmu.edu/) and developed scripts to create various text- and audio-based surrogates.

While the production system is primarily intended to increase the size of the DL to offer more content to our end users, this architecture also benefits the research goals of our project in multiple ways. First, the video files, surrogates, and metadata serve as source content for the research studies of our and other groups. For example, in the first 2 years of the annual TREC Video Retrieval Evaluation workshop (TREC VID), a program sponsored by NIST to promote digital video retrieval research (Smeaton, Over, & Kraaij, 2004), a portion of the test collection of video used by participants was drawn from the OVDL. In each of our own user studies, all content used to conduct the studies was drawn from the Open Video DL and all customization of video and surrogates necessary for a given study was done with the existing production system.

The existence of our production system and the OVDL has also facilitated the rapid development of a wide range of experimental applications and tools and has enabled students to incorporate the resources in their coursework and independent studies. The Interactive Shared Educational Environment (ISEE; Mu, Marchionini, & Pattee, 2003) application, for

DOI: 10.1002/asi

example, began as an experiment to create an interface that would enable users at different locations to view, discuss, and annotate video content from the Open Video DL collaboratively and synchronously in a shared environment (Mu, 2004). A University of North Carolina (UNC) student used the DL for a master's thesis that compared the results of manual indexing of video by either viewing the full video and by viewing surrogates only (Stachowicz, 2002). In addition, the OVDL was instrumental in our development of the AgileViews interface design framework.

The AgileViews Design Framework at the Intersection of Research and Practice

The development of the AgileViews framework (Marchionini et al., 2000; Geisler, 2003; Geisler, Marchionini, Nelson, Spinks, & Yang, 2001; Geisler et al., 2002) is a prime example of the interplay between research goals and practical goals in the development of the Open Video DL. The building of the OVDL guided the development of the AgileViews design framework. Simultaneously, the design framework motivated the development of new tools and designs for application in the OVDL.

The AgileViews framework began with the recognition that there exists today an unprecedented amount of digital information. As more people gain access to this information and the information itself continues to expand in both quantity and richness, it becomes increasingly critical that system developers and interface designers find ways to enable people to access this information more effectively. Information seeking systems must not only help people find the information they need to accomplish their goals but also make the process of finding it more efficient, transparent, and satisfying.

Current information seeking systems often rely on navigation as a metaphor, which, although sometimes appropriate, often contributes to interfaces that fail to provide helpful information structures and visual cues, impose high demands on short-term memory, and force the user to divide attention constantly between primary and task management processes. The AgileViews framework suggests that an alternative metaphor—one based on views rather than navigation—might enable us to improve the information seeking experience. Views—coherent units of information representing subsets of an information environment—are quite common in information systems, but interfaces that utilize a formal, coordinated set of views such as that presented in the AgileViews framework are much less common.

The AgileViews framework includes six fundamental types of views: *Overviews* and *previews* are two types of views commonly found in information systems today. The distinctions between them are sometimes blurred, but in general overviews are representations of an entire collection of objects or a subset of the collection and previews are representations of a single object (Greene, Marchionini, Plaisant, & Shneiderman, 2000). *History views* are essentially representations of the past activities of a user or, in aggregation, a group of users. Information that reveals the activities, opinions, and recommendations of other people in the same context makes up a *shared view*. The *primary view* is the view that the user is primarily focused on at a given time. Finally, contextual information—including other views when they are out of focus—creates *peripheral views*.

Although the development of the AgileViews framework was guided by other projects as well (e.g., the Relation Browser interface; http://idl.ils.unc.edu/rave/), the practical needs of the OVDL were particularly influential. As more collections of digital video such as Open Video are accessed by more people, it will become increasingly important for these systems to provide interfaces with features and mechanisms that are optimized to enable users to browse and retrieve video more easily. We also believe that in a general purpose collection such as Open Video there are a variety of sources of evidence that can help a user find the most relevant resources for a given information need and that for distinct types of users or information needs one type of evidence might be more useful than another. Formally categorizing these distinct types of evidence into a set of views that are based on the different ways people can change their views in the physical world led us to the six fundamental types defined. At the same time, our ongoing work with the OVDL gave us the opportunity to see how well the framework applied to practical systems by implementing aspects of the framework in our operational and experimental systems.

Overviews and Previews in the OVDL

Overviews and previews are two components of the AgileViews framework that have been implemented in the public Open Video DL. Because of the fundamental *visual* nature of video, each of these views is instantiated in the Open Video DL as a way for the user quickly to shift from one view of the collection (in an overview) or a particular item in the collection (in a preview) to another in which the view options in each case differ in their visual characteristics. For example, Figure 2 shows the default search results page in the Open Video DL. As is common on a search results page in many systems, the page summarizes the resources in the collection that match the search criteria with a thumbnail image, title, date, short description, genre, keywords, duration, and popularity and enables the user to sort the results by different attributes (relevance, date, duration, popularity) and select how the number of results to display on a single page.

As shown in Figure 3, however, our overview also enables the user to change the visual display of the results to any one of four layouts, from a text-oriented view in which the size of the thumbnail image is minimized to enable a large number of results to be displayed, to a thumbnail-only view in which the display emphasizes the visual aspect of the resources. Thus, users who have different individual preferences or information needs can easily select the form of overview most appropriate to their situation. It seems likely, for example, that a student looking for a long video about the *Apollo* space mission might benefit more from the text-oriented overview that enables him quickly to scan title,

DOI: 10.1002/asi

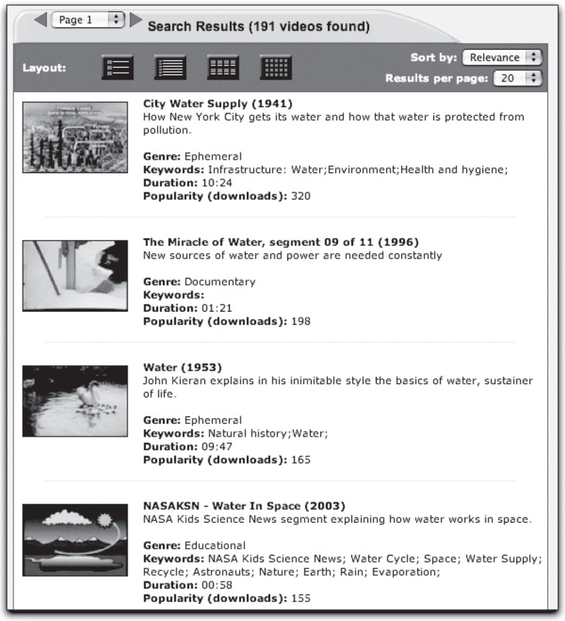

FIG. 2. Open Video Digital Library default search results display.

descriptions, and durations, and a multimedia artist searching for a short clip with a certain style might find that the large thumbnail view is much more efficient for his or her goal of quickly scanning the resources to find those with the right visual feel.

Whereas overviews represent collections of objects, previews are representations of a single object (Greene et al., 2000). The purpose of a preview is to provide the user with enough information about a primary object to determine whether the object is of interest (i.e., whether it moves the user closer to resolving the information need). Publishers and librarians have long created surrogates (for example, abstracts and synopses) as alternative representations of primary objects to aid retrieval. Traditional text-oriented surrogates, however, are of limited use in video collections because they cannot efficiently convey information about a video object's visual style, layout, or changes over time. In collections of digital video, visual surrogates are especially critical because, although they require much less time to access than the full video, they still provide enough cues to enable the user to extract gist and assess relevance.

The previews currently available in the Open Video DL are surrogates for the individual video objects in the video collection. As with the overviews, we offer the user the capability of quickly and easily switching among several types of preview, each of which might be the most appropriate in a given situation: a storyboard, a 7-second excerpt of the video, and a fast-forward version of the entire video. The storyboard is a small collection of keyframes extracted from the entire video and might be considered analogous to keywords for a document. The 7-second segment provides the user with the sound and action of the video, played in real time, and might be considered analogous to a brief excerpt from a document. The fast-forward surrogate provides a very quick overview of the entire video (but without sound) and might be considered analogous to an abstract of a document. As shown in Figure 4, switching among these different previews is quick and intuitive; our transaction logs show that our users view the fast-forward and 7-second excerpt surrogates with equal frequency and the storyboard twice as often as the other two preview surrogates.

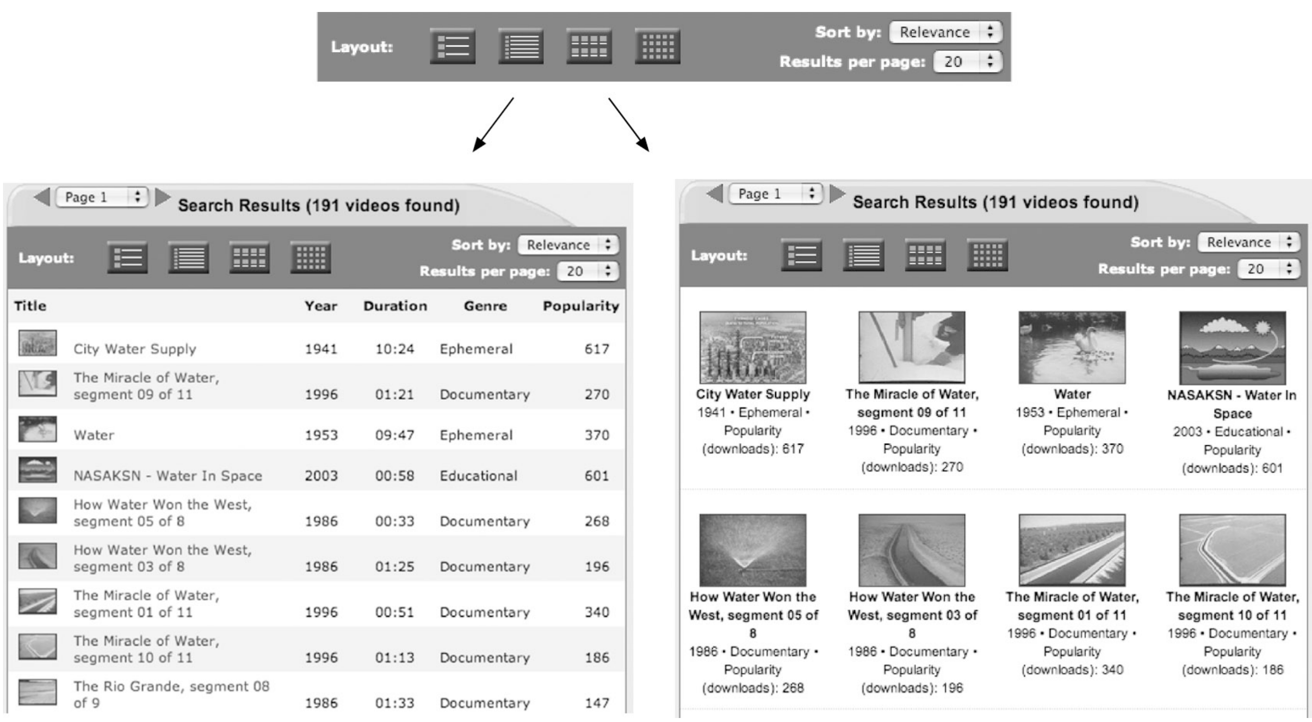

FIG. 3. Open Video Digital Library overview alternative displays.

DOI: 10.1002/asi

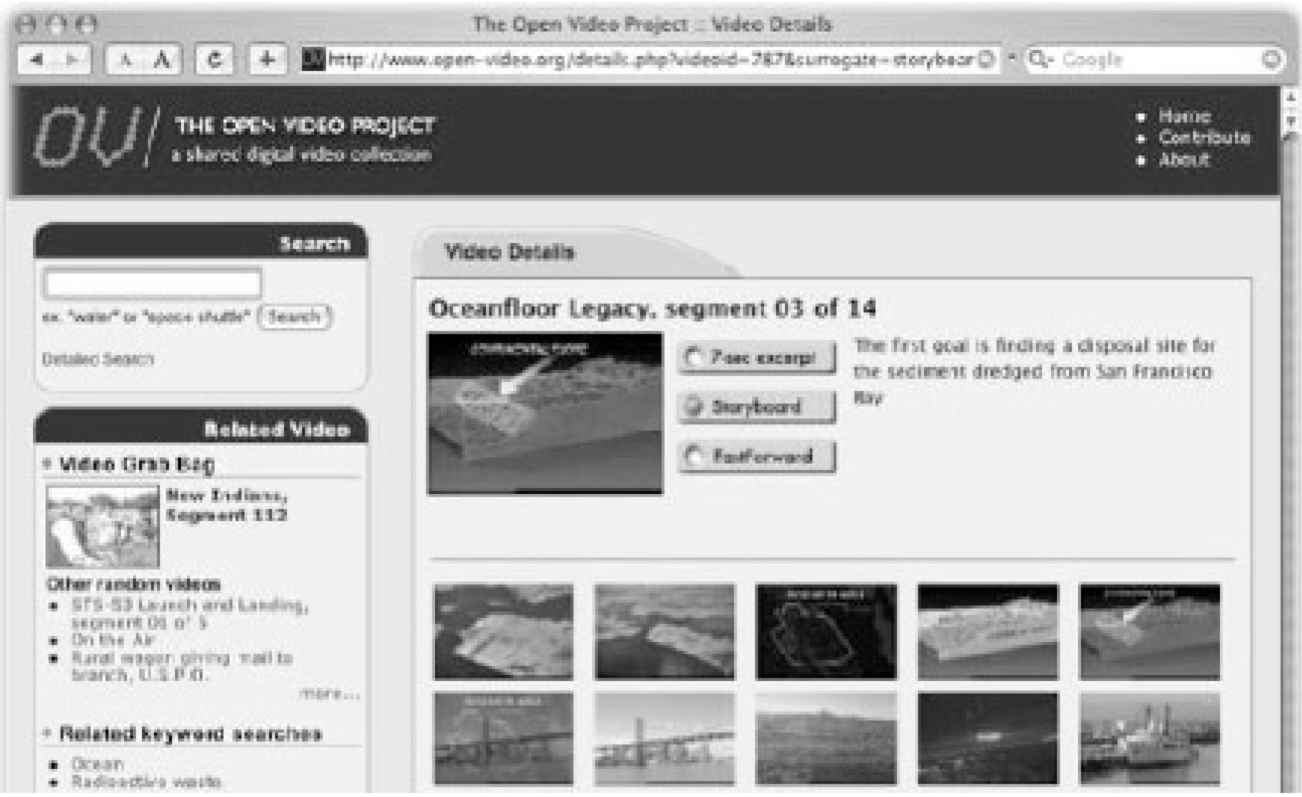

FIG. 4. Open Video Digital Library alternative preview options.

Shared and History Views in the OVDL

Not part of the public OVDL at the time of this writing are other components of the AgileViews framework we have implemented in experimental versions of the DL, and we expect that features such as these will be part of future versions of the DL. For example, in addition to overviews and previews, one experimental version of the DL includes both history and shared views. Shared views are based on the way we sometimes rely on other people's opinions and current activities when making decisions in the physical world. Similarly, in the digital environment there is often a variety of information that reflects other people's opinions, suggestions, and activities related to an information resource. These opinions, suggestions, and activities are from different people but they are focused on the same resource; collectively they form a shared view of the resource. Information systems that enable their users to rate, review, or otherwise comment on resources have become increasingly common. A prominent example is Amazon.com, which provides what we would consider both literal shared views (customer reviews and ratings) and implied shared views (system-generated recommendations derived through copurchase data) of its resources.

Our experimental version of the Open Video DL included both forms of shared views. Each shared view displayed a list of video titles that were related to the currently selected video title in some way. The first type displayed video titles that were related by virtue of having been downloaded by other users who had downloaded the currently selected title. This is a form of implicit recommendation as used by Amazon and many other systems. The second type of shared view (shown in Figure 5) provided a similar list of titles related to the currently selected one, but this list consisted of more explicit recommendations made by other users for the current title.

History views are based on the fact that in the everyday world we are surrounded by evidence of our own and of other people's previous activities, and, in a very natural way, we use this evidence to help us make decisions about things to buy or use, restaurants or museums to visit, even which direction to travel in an unknown area. Considerable research exists on the value of histories in the digital environment, from visualizing the interaction history of a specific resource (Hill, Hollan, Wroblewski, & McCandless, 1992) to using aggregated history information to make recommendations within a collection of resources, such as Usenet news articles (Sarwar et al., 1998) or Web pages (Pirolli, Pitkow, & Rao, 1996). By making available histories of use and patterns of activity, either implicitly or explicitly, history views can benefit the information seeker by providing a perspective that is not based on the content of the resource but instead on how it has been used by other people, presenting an alternative to surrogates based solely on topical relevance and reflecting how people often make decisions in the physical world.

There were two types of history views in our experimental version of the Open Video DL. The two types were similar, but one showed data reflecting how often the selected video title has been downloaded, while the other showed

DOI: 10.1002/asi

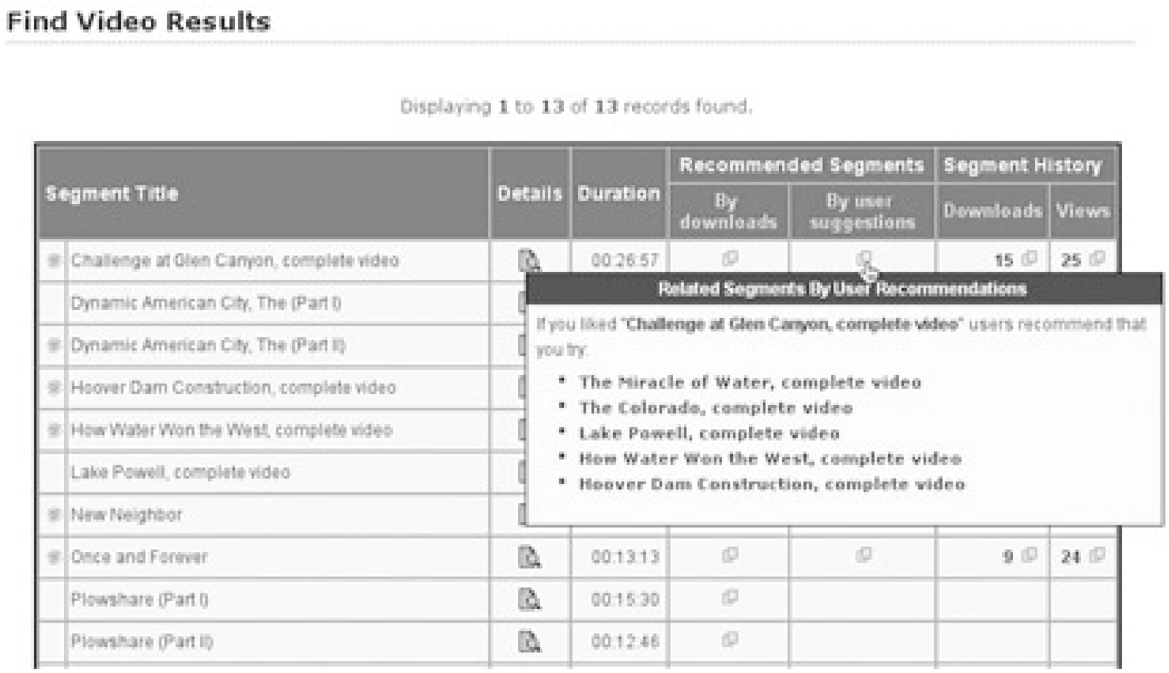

FIG. 5. Open Video Digital Library shared view display.

data reflecting how often the metadata for the selected video title had been viewed. Figure 6 shows the details of the history view based on downloads. The ranking of the currently selected video title in terms of how often it has been downloaded compared to other titles in the collection was provided in both textual and graphical form, and the number of times the title has been downloaded was broken down by user type, as determined by the profile information of registered users.

This experimental version of the Open Video DL was used to evaluate the effectiveness of the AgileViews framework in a formal user study (Geisler, 2003). The study demonstrated that participants both used and liked the shared and history views when completing the assigned tasks. Specifically, participants most often used the shared views to narrow down the number of videos they would examine in detail, to confirm that a candidate video would be a good one to select, or to help them decide between two candidate videos. The history views were generally used during tasks in which participants felt knowing how often a video had been previously downloaded and what type of users had downloaded it helped them to decide whether they should download it. We have not yet implemented fully formed shared and history views in the public Open Video DL, but this experimental version not only provided us with a practical basis for refining the AgileViews theoretical framework but also gave us, along with other user studies, a basis for a complete redesign of the Open Video DL interface that was made public in August 2003.

An Evaluation Framework Serving Both Research Goals and Design Goals

Our primary objective over the past several years has been to create and evaluate surrogates for video that take advantage of human visual capabilities. This objective serves

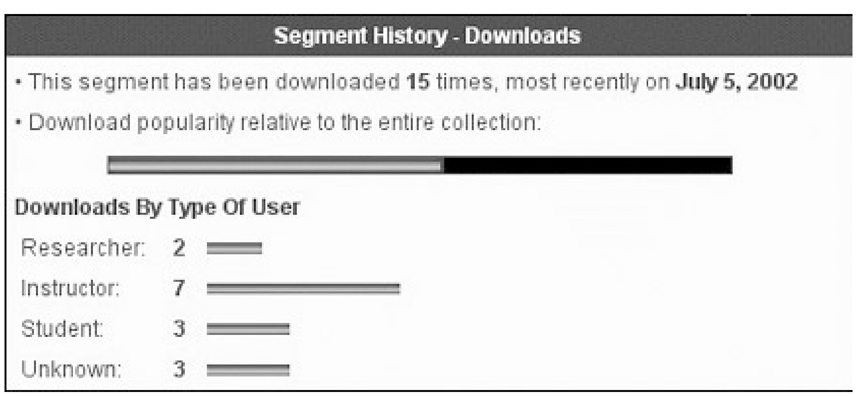

FIG. 6. Open Video Digital Library history view display.

DOI: 10.1002/asi

both our research goals and our design goals. In addition, it raises two challenges and we claim that our solutions on both are significant contributions to video retrieval research progress. The first challenge was the creation of visual surrogates that would help people quickly make judgments about whether it is worth looking for additional details or downloading the video. This is the classical relevance judgment assessment based on some surrogate. We began with three types of visual surrogate: slide show (a sequential display of selected frames), storyboard (a static array of selected frames), and fast-forwards (a rapid display of every *n*th frame, in order). We looked at these three types of surrogate under a variety of conditions, including adding speech or text keywords to slide shows, different fast-forward rates, and selections and displays of different poster frames (best examples of selected frames). One key issue for these surrogates is how to select the frames, which are termed *keyframes* in the literature (first suggested by O'Connor, 1985). Because it is impractical to select the most representative keyframes for video manually, there are a variety of techniques for automatically identifying them. Most are based on identification of scene changes, which are easy to detect with intraframe pixel differences (e.g., Zhang, Low, & Smoliar, 1995; Yeung & Yeo, 1997; Wang, Liu, & Huang, 2000). Our solution was a hybrid that started with an automatic technique for keyframe extraction from MPEG-1 streams. We first used the MERIT system from the University of Maryland (Kobla, Doermann, & Faloutsos, 1998) to extract candidate keyframes on the basis of large intraframe differences. We then created scripts to display these candidate keyframes and manually selected up to 36 of the best candidates to use as keyframes for the video. The resulting keyframes are the basis for posterframe, slide show, and storyboard surrogates in OVDL. For fast-forward surrogates, we automatically extracted every *n*th frame. For the 7-second excerpts, we manually selected the first 7 seconds of the video after the initial credits. Thus we have a hybrid set of procedures for creating the OVDL surrogates. Many of these design decisions were heuristic, based on literature reviews, experience, and extensive discussion. Our intention was to test design parameters empirically to assess best default settings for the system. This desire led to the second primary challenge of determining how to evaluate the effectiveness and efficiency of the visual surrogates.

We had substantial previous experience conducting user studies and, for the OVDL user studies, used the facilities of the Interaction Design Laboratory at UNC-Chapel Hill, which include a usability workstation with multiple cameras and video mixer for different inputs and an eye-tracking workstation. Our preliminary evaluation model aimed to assess the surrogates while controlling other factors such as individual user characteristics, video characteristics, and tasks. Figure 7 illustrates the overall evaluation framework with performance and preference as the primary dependent variable constructs.

User studies are complex, in many ways resembling the kinds of decisions people make when designing systems. Designing a user study requires a host of decisions about dependent and independent variables. Of course, the human characteristics are complex and are usually controlled by using multiple subjects so that individual differences can be removed with statistical power. Creating realistic tasks for people who participate in studies in a usability laboratory is another challenge, and over the years we created several

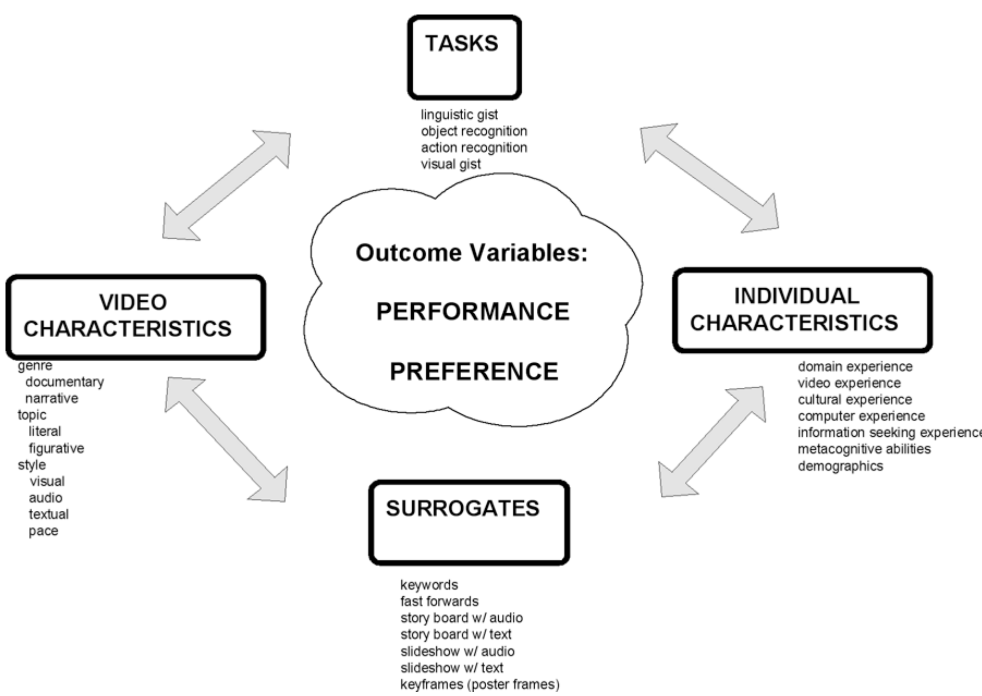

FIG. 7. Preliminary evaluation framework.

DOI: 10.1002/asi

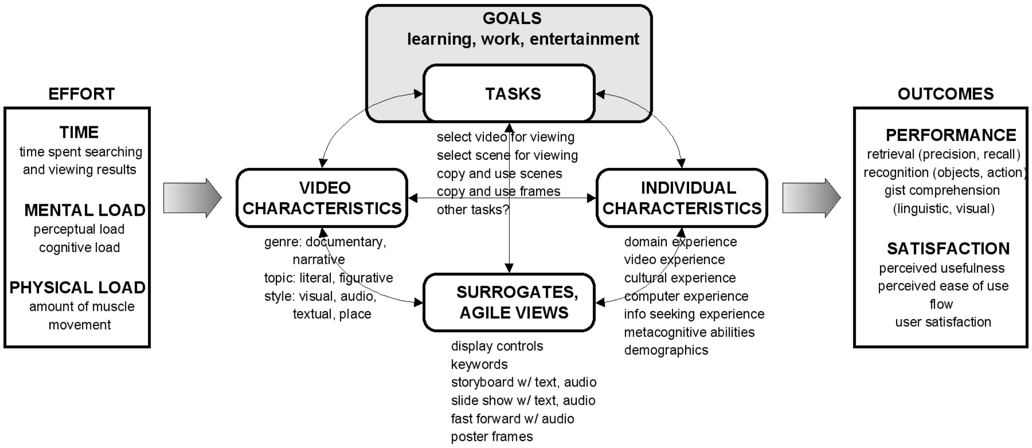

FIG. 8. Revised evaluation framework.

kinds of video retrieval tasks. In our early studies we learned that the video characteristics also determine human performance so we endeavored to control the genre and visual style properties of the videos used in our studies. Because our primary independent variable of interest was the surrogate, we aimed to limit the number of surrogates we created, not a simple process because surrogate efficacy interacts with video genre and task. Finally, because there has been an enormous amount of work on video retrieval techniques and systems but very little in the way of user studies, we had to invent a series of dependent measures for performance and preference. We claim that these metrics are one of our primary contributions to video retrieval, and other groups have begun to adopt them in their user studies. The specific studies are summarized later, but before presenting them, we present a revised evaluation framework that reflects our current thinking about what constructs influence human video interaction. Figure 8 reflects a framework that embeds tasks in larger life goals and considers the overall interaction of goals, tasks, individual characteristics, video characteristics, and video surrogates as part of an activity flow that requires human effort to achieve outcomes. Outcome measures have been refined into two classes, performance and satisfaction with particular metrics used in each class. We are concerned not only with search efficiency outcomes but with the ways people understand and make sense from video surrogates. Most importantly, we have begun to consider various classes of effort, including time, mental load, and physical load, that influence the interaction of the overall video retrieval and understanding activities.

Summary of User Studies

Over a 3-year period, a number of studies were conducted with an eye to informing our OVDL designs and developing a theory of how people interact with video surrogates. The studies helped us to evolve the overall evaluation framework as well as make specific design decisions for the system. The studies have been reported in various publications and are briefly summarized here to illustrate how our ideas about system design, evaluation, and human information behavior evolved hand in hand with the investigations.

A preliminary study (Wildemuth et al., 2002) examined the efficacy of slide shows with and without audio keywords, storyboards with and without audio keywords, and fast-forwards. Ten participants, all of whom watch videos at least monthly and search for videos at least occasionally, viewed the surrogates for seven video segments. During and after viewing the surrogates, participants commented directly on the surrogates' strengths and weaknesses for particular tasks (e.g., selecting frames or short clips for reuse). They then completed a series of performance measures by using three surrogates of their choice. The results of this study indicated that no surrogate was universally judged "best." However, the fast-forward surrogate garnered the most support, particularly from experienced video users, and we were able to reject the slide show with textual keywords from further consideration.

Because the fast-forward surrogate was relatively novel, it was investigated further. Specifically, we were interested in how fast a fast-forward surrogate could be viewed and still be meaningful to the viewer. The fast-forward surrogate is created by selecting every *n*th frame from the full video; therefore, a "faster" surrogate samples fewer frames. Our second study (Wildemuth et al., 2003a) compared four fast-forward surrogate speeds, with sampling rates of 1:32, 1:64, 1:128, and 1:256. Forty-five participants interacted with all four surrogates and completed six performance measures. It was concluded that the tradeoff between efficiency and performance was balanced at the 1:64 speed, and we therefore used 64× fast-forwards in the Open Video site redesign.

Each of these studies evaluated the effectiveness of the surrogates under study by using six new performance measures we developed (Yang et al., 2003b, 2003c): two object recognition measures, one action recognition measure, two gist comprehension measures, and one visual gist (vist) comprehension measure. The object recognition measures provided the study participants with a set of stimuli (nouns/textual for one measure and frames/graphical for the other), and the participants were asked to mark those that

DOI: 10.1002/asi

had been seen in the video surrogate. The action recognition measure was analogous, but the stimuli were short (2- to 3-second) clips from the surrogate. The free-text linguistic gist comprehension measure asked participants to write a summary of the video's gist on the basis of the surrogate they had viewed. The multiple-choice linguistic gist comprehension measure asked the participants to select the best summary from a stimulus set of five. *Visual gist comprehension* was defined as the viewer's overall understanding of the video, including both its content and its style. The visual gist comprehension measure provided stimulus frames not yet seen by the participant and asked the participant to identify those that "belong" in the video represented by the surrogate. As part of our investigations, the reliability and validity of these measures have been investigated, and we continue to refine them in our ongoing studies.

Our early studies tended to focus on isolated variables (e.g., speed of a fast-forward) in highly constrained task environments so as to establish basic design parameters. As we integrated the surrogates into the OVDL redesign, we began to consider more *in situ* studies that looked at larger portions of the whole system (e.g., the full browsing component, which included several surrogate options, and the results displays after a search has been executed). We began to apply satisfaction measures in our participation in Text REtrieval Conference VIDeo track (TRECVID) 2003 (Wildemuth et al., 2003b), and included measures of perceived ease of use, perceived usefulness, and two measures of flow (enjoyment and concentration) (Davis, 1989; Ghani, Supnick, & Rooney, 1991). Using a video repository (made up entirely of television news stories) and feature identification data provided by NIST and other TRECVID participants, 36 study participants interacted with three different search systems: one providing access through the transcripts of the news stories, one through the features automatically identified in the news stories, and one through both access mechanisms. The transcript-only and the combination systems outperformed and were preferred to the features-only system. Another study aimed to see whether users could use slider settings to weight different feature parameters such as amount of luminosity, presence of objects (e.g., people, vehicles), and setting (e.g., landscape, cityscape). Although people were able to set these parameters, no strong effects or preferences were found (Gruss, 2004). These results parallel other findings in the literature with respect to the importance of including textual information for full-function video retrieval (Christel & Conescu, 2005).

During our first study, the participants often commented on the narrative structure (or lack of it) in the videos and the way that structure affected their interactions with the surrogates. To explore this idea further, we investigated people's perceptions of narrativity in short video segments by inviting ACM SIGCHI 2002 participants to rate narrativity for videos (Wilkens, Hughes, Wildemuth, & Marchionini, 2003). We found that narrativity has two components (organization by cause/effect relationships and the presence of agency), both of which are necessary to people's perceptions of narrativity.

Another issue that arose from our early studies was the relationship between graphical representations and textual representations of the videos. More specifically, as we worked with the design of the Open Video Web site and the way it provided access to the collection, we were unsure of how people might use poster frames in combination with brief textual descriptions of the videos. We explored this issue further with the aid of eye tracking, comparing people's use of two alternative overviews: one with the poster frame on the left and the text on the right, and one with the text on the left and the poster frame on the right (Hughes, Wilkens, Wildemuth, & Marchionini, 2003). Twelve study participants selected items from a displayed results list; their eye movements were tracked to see where, when, and how long they viewed the text and image representations. Most participants used the text as an anchor from which to make their relevance judgments and the images as confirmatory evidence of their selections. A follow-up study (Hughes, 2003) investigated the role that task plays in users' preferences for textual or visual surrogates. Twelve subjects thought out loud as they worked on five different types of retrieval tasks (ranging from simple noun objects to topical searches to personally interesting videos). Participants were shown results pages for the various search tasks and could use either textual descriptions or storyboards to make decisions about relevance to the task. The number of surrogates selected was counted and subjects were asked to say which surrogate mainly led to their decision. As expected, participants used more visual surrogates for the concrete tasks and more textual surrogates for the topical tasks. However, most subjects used both types of surrogate in combination—hypothesizing relevance with one and using the other to confirm or refute their hypothesis. These studies parallel results from other researchers and demonstrate the need for a variety of surrogates for full function video retrieval systems.

In addition to being directly applicable to concrete design decisions, our previous work has informed our theoretical framework in several ways. First, we have gained a clearer understanding of the role of the users' goals in shaping the tasks they perform when interacting with a video repository. For example, a teacher searching for materials that will aid his or her students' learning will approach a retrieval task much differently than an amateur film producer looking for just the right clip to juxtapose over new footage in a video production. This understanding was strongly demonstrated in Yang's (2005) dissertation, which focused on how instructors, video librarians, and video editors and producers make relevance judgments about video. Second, we realize the importance of the video's structure (narrative vs. categorical) in people's comprehension of surrogates. This particular characteristic of the video must be accounted for in developing alternative representations and may be important for archival decision making. Third, our research framework now makes explicit the role of user effort in interacting with a digital video repository. We could see the importance of examining user effort in Mu's (2004) dissertation, as he examined perceptual, cognitive, and collaborative loads on learners as they used the Interactive Shared Educational Envrionment (ISEE) tool to view and discuss videos remotely in teams. Finally, these studies have led us to recognize that the many components of video (e.g., sound track, frames, action sequences, text superimpositions) can themselves be useful adjuncts to communication and learning. We think of these small units of video or extracts from video as "chunks." For example, a 10-second clip showing an animated exponential curve for cellular growth may be extracted and applied for other scientific processes, or a short explanation of the concept of lift given by a test pilot might become an advance organizer for a discussion of dynamics. This chunking and reuse capability has important implications for preservation and context specification, problems that we are addressing in our ongoing work.

Lessons Learned and Recommendations

The experience of conducting user studies to inform the design of a system has taught us several lessons:

- First, our experience demonstrates that user studies inform good design. The studies of how people use different visual surrogates helped us select which surrogates to include and which variants to use.
- Second, the studies also helped determine which overviews, previews, and reviews to include in the redesigned interface. More generally, the user studies in total not only helped us make these practical design decisions but also reinforced and extended the AgileViews design framework. The studies demonstrated the efficacy of giving people multiple views and easy control mechanisms for shifting among those views.
- Third, the range and number of studies conducted also illustrate that there are no "silver bullets" to solve the DL user interface design challenge. There are many factors that affect people's performance with and preferences for DL interfaces and features. The user study evaluation framework that evolved over the course of the studies illustrates these factors and highlights the challenges of design. Clearly, the design of DLs must take into account the user community, the nature of the content, and the current state of technology. However, it is equally important that designers and digital librarians recognize that there is no such thing as a perfect design or a completed design—their designs must continue to evolve over time as the user community evolves and the global cyberinfrastructure develops.
- Fourth, the studies demonstrate the importance of using a multifaceted approach to global evaluation in DLs. The fact that so many factors determine performance with and preferences for aspects of a DL interface must also apply to the larger questions of overall DL effectiveness.

The experience of building OVDL also offers several lessons related to both DL research and the practice of building DLs. First, we have learned the importance of establishing a sustainable model for a DL. Our efforts were supported with research grants; however, we have leveraged our institutional resources at UNC-CH to maintain the DL. Although it is feasible to maintain servers and bandwidth for the foreseeable future without special funding, the higher costs of curating new content present problems. We have no immediate solution other than continuing to involve our students in the day-to-day operations of digitizing, creating surrogates and metadata, and maintaining the database. We have initiated discussions with various campus units, such as the campus library, to assume responsibility for operating the OVDL. However, there is a big difference between maintaining a DL in status quo state and actively developing and revising the DL over time. Working out how these responsibilities are managed will take many years as hundreds of DLs cope with this basic sustainability challenge of technology transfer from research-oriented operations to full-scale production systems that dependably serve people's information needs. Ultimately, we see DLs as part of the global and public cyberinfrastruture. As such, many DLs will be supported just as the Internet is now supported—through a combination of public and private investments using a mix of economic models similar to the current arrangements for libraries, which range from national libraries to localized public or corporate libraries, and for for-profit organizations, which range from independent book shops to Amazon-like services.

Second, we have learned that multiple partnerships and collaborations are necessary to build and sustain a DL. These collaborations are institutionally diverse and represent different kinds of resources, ranging, in the case of OVDL, from universities (UNC-Chapel Hill, University of Maryland, Carnegie Mellon University) to government agencies (NASA), to individuals (e.g., the Prelinger collection). Some partnerships are built around sharing or contributing content, and some are based on technical infrastructure (sharing hardware and software and expertise). Of course, creating and sustaining these collaborations require time and effort to build trust and establish *quid pro quo* values. Moreover, the collaborations require basic levels of standardization or agreement. For example, if one DL points to files held elsewhere, naming schemes must be fixed, and, if they do change, all interested parties must be notified. Clearly, DLs are not independent entities that stand as islands in cyberspace, but rather are rooted in communities of interest that have varying levels of commitment and connections to other communities.

Third, we have learned that although the notion of the sharium may be appealing from idealistic and global resource allocation points of view, there are several practical issues that constrain participation. Contributions can overwhelm a DL's curatorial routines. In the case of video, content must be examined in real time, and even when metadata are provided, they too must be verified. Because there are so many parameters for video contributions (e.g., formats, rights, metadata), a fully automatic contribution process is impractical as there are often questions and answers that consume the time of all participants. Our practical solution in OVDL was to focus on contributions of collections from well-known sources (e.g., university groups, government agencies) and deal with contributions from individuals only when there was time to do so. Although DLs such as ibiblio demonstrate that contributor-run DLs can work, a small staff

DOI: 10.1002/asi

is required to lend support to those who wish to contribute Web sites to the DL. There is a corollary lesson about information reuse in general that arose from the BLC and OVDL experiences. Although it makes sense in principle to say that a short segment of video should be able to be "plugged in" to classroom lessons, there are challenges of indexing at fine grains (e.g., to state and local instructional objectives) as well as of incorporating these chunks without context into a derivative work. It is our observation, based on e-mail and serendipitous conversations with users, that teachers tend to use the original video rather than small chunks they edit themselves (video editing is still very awkward with today's software in schools, and this practice could change over time). Artists currently seem to be the users who reuse video chunks as parts of derivative works. It will be interesting to see whether others reuse video chunks over time as better tools for creating, editing, and transmitting video are available for the school and home. Cell phones with video and the host of video search services available from the popular WWW search engine companies suggest that these tools will soon be available.

Fourth, we have learned that digital video does offer new potential for learning and communication. Our experience with the ISEE demonstrates how remote learners can study video together and how it can be used in hybrid classroom–remote access settings (Mu, 2004). Our studies of relevance judgments have begun to illustrate how people make sense of video by using visual characteristics as well as topical ones, building an understanding of visual gist (vist) that helps them evaluate possible relevance (Yang, 2005).

In conclusion, we argue that research and practice are both necessary aspects of scholarship, complementing each other in ways that benefit the development of both. This article lays out the beauties and warts of a decade of building and testing digital video libraries, a dynamic process of research and practice that resembles a walk along a Möbius strip: Research leads to practice that leads back to more research.

Acknowledgments

This work was supported by a grant from the National Science Foundation (IIS 0099638), by a contract from NASA, and by the School of Information and Library Science's Interaction Design Laboratory. Thanks are due to the many organizations and individuals who provided content and software, the dozens of students who have worked on the project at various times, and the scores of people who participated in the various user studies.

References

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

Borgman, C.L. (2000). From Gutenberg to the global information infrastructure: Access to information in the networked world. Cambridge, MA: MIT Press.

Christel, M., & Conescu R.M.. (2005). Addressing the challenges of visual information access from digital image and video libraries. In Proceedings of the ACM/IEEE Joint Conference on Digital Libraries (pp. 69–78). New York: ACM Press.

Christel, M., Smith, M., Taylor, C.R., & Winkler, D. (1998). Evolving video skims into useful multimedia abstractions. In Proceedings of CHI '98: Human Factors in Computing Systems (pp. 171–178). New York: ACM Press.

Davis, F.D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340.

Ding, W., Soergel, D., & Marchionini, G. (1999). Performance of visual, verbal, and combined video surrogates. In Proceedings of ASIS 99 (pp. 651–664).

Engelbart, D. (1963). A conceptual framework for the augmentation of man's intellect. In P. Howerton & D. Weeks (Eds.), Vistas in information handling (Vol. 1, pp. 1–29). Washington, DC: Spartan Books.

Geisler, G. (2003). AgileViews: A framework for creating more effective information seeking interfaces. Unpublished doctoral dissertation, University of North Carolina at Chapel Hill.

Geisler, G., Marchionini, G., Nelson, M., Spinks, R., & Yang, M. (2001). Interface concepts for the Open Video Project. Proceedings of the 2001 ASIST Annual Meeting, 38, 58–75.

Geisler, G., Marchionini, G., Wildemuth, B.M., Hughes, A., Yang, M., Wilkens, T., et al. (2002). Video browsing interfaces for the Open Video Project. ACM SIGCHI Conference on Human Factors in Computing Systems, CHI 2002: Extended Abstracts (pp. 514–515). New York: ACM Press.

Ghani, J.A., Supnick, R., & Rooney, P. (1991). The experience of flow in computer-mediated and in face-to-face groups. In Proceedings of the 12th International Conference on Information Systems (pp. 229–237). Atlanta, GA: Association for Information Systems.

Goncalves, M., Fox, E., Watson, L., & Kipp, N. (2004). Streams, structures, spaces, scenarios, societies (5s): A formal model for digital libraries. ACM Transactions on Information Systems, 22(2), 270–312.

Greene, S., Marchionini, G., Plaisant, C., & Shneiderman, B. (2000). Previews and overviews in digital libraries: Designing surrogates to support visual information seeking. Journal of the American Society for Information Science, 51(4), 380–393.

Gruss, R. (2004). The VisOR system: Testing the utility of user interface components for feature-based searching in video retrieval software. Unpublished master's thesis, University of North Carolina at Chapel Hill, School of Information and Library Science. Retrieved June 20, 2005, from http://ils.unc.edu/MSpapers/2936.pdf

Hill, W.C., Hollan, J.D., Wroblewski, D., & McCandless, T. (1992). Edit wear and read wear. In Conference Proceedings on Human Factors in Computing Systems (CHI '92) (pp. 3–9). New York: ACM Press.

Hughes, A. (2003). When do you look and why do you read? An exploratory study regarding user interaction with visual and textual surrogates when searching for video information. Unpublished master's thesis, University of North Carolina at Chapel Hill, School of Information and Library Science.

Hughes, A., Wilkens, T., Wildemuth, B., & Marchionini, G. (2003). Text or pictures? An eyetracking study of how people view digital video surrogates. In Proceedings of the International Conference on Image and Video Retrieval (CIVR 2003) (pp. 271–280). Springer Lecture Notes in Computer Science 2728. Berlin.

Kobla, V., Doermann, D., & Faloutsos, C. (1998). Developing high-level representations of video clips using VideoTrails. In Proceedings of the SPIE Conference on Storage and Retrieval for Image and Video Databases VI (pp. 81–92). Bellingham, WA: International Society for Optical Engineering.

Komlodi, A., & Marchionini, G. (1998). Key frame preview techniques for video browsing. DL'98, Proceedings of the Third ACM Conference on Digital Libraries (pp. 67–75). New York: ACM Press.

Lesk, M.E. (1997). Practical digital libraries: Books, bytes, and bucks. San Francisco: Morgan Kaufmann.

Marchionini, G. (1999). Augmenting library services: Toward the sharium. In Proceedings of International Symposium on Digital Libraries 1999 (pp. 40–47). Tsukuba, Japan: University of Library and Information Science. Retrieved from http://www.ils.unc.edu/~march/sharium/ISDL.pdf

Marchionini, G., & Geisler, G. (2002). The Open Video Digital Library. D-Lib Magazine 8(12). Retrieved from http://www.dlib.org/dlib/december02/marchionini/12marchionini.html

Marchionini, G., Geisler, G., & Brunk, B. (2000). AgileViews: A human-centered framework for interfaces to information spaces. In Proceedings of the Annual Meeting of the American Society for Information Science (pp. 271–280). Silver Spring, MD: American Society for Information Science and Technology.

Marchionini, G., & Levi, M. (2003). Digital government services: The Bureau of Labor Statistics case. Interactions: New Visions of Human-Computer Interaction, 10(4), 18–27.

Marchionini, G., Nolet, V., Williams, H., Ding, W., Beale, J., Rose, A., et al. (1997). Content + connectivity = community: Digital resources for a learning community. In Proceedings of ACM DL '97 (pp. 212–220).

Mu, X. (2004). Smartlinks in a video-based collaborative distance learning system: A cognitive model and evaluation study. Unpublished doctoral dissertation, University of North Carolina at Chapel Hill.

Mu, X., & Marchionini, G. (2003). Enriched video semantic metadata: Authorization, integration, and presentation. Proceedings of the Annual Meeting of the American Society for Information Science and Technology, 40, 316–322.

Mu, X., Marchionini, G., & Pattee, A. (2003). The interactive shared educational environment: User interface, system architecture and field study. In Proceedings of the ACM/IEEE Joint Conference on Research on Digital Libraries (pp. 291–300). New York: ACM Press.

O'Connor, B. (1985). Access to moving image documents: Background concepts and proposals for surrogates for film and video works. Journal of Documentation, 41(4), 209–220.

Pirolli, P., Pitkow, J., & Rao, R. (1996). Silk from a sow's ear: Extracting usable structures from the Web. In Conference Proceedings on Human Factors in Computing Systems (CHI '96) (pp. 118–125). New York: ACM Press.

Sarwar, B.M., Konstan, J.A., Borchers, A., Herlocker, J., Miller, B., & Riedl, J. (1998). Using filtering agents to improve prediction quality in the GroupLens Research Collaborative Filtering System. In Proceedings of CSCW '98 (pp. 345–354). New York: ACM Press.

Slaughter, L., Marchionini, G., & Geisler, G. (2000). Open Video: A framework for a test collection. Journal of Network and Computer Applications, 23(3), 219–245.

Smeaton, A., Over, P., & Kraaij, W. (2004). TRECVID: Evaluating the effectiveness of information retrieval tasks on digital video. Proceedings of the ACM Multimedia '04 (pp. 652–655). New York: ACM Press.

Stachowicz, C. (2002). The effectiveness of storyboard surrogates in the subject indexing of digital video. Unpublished master's thesis, University of North Carolina at Chapel Hill, School of Information and Library Science. Retrieved June 20, 2005, from http://ils.unc.edu/MSpapers/2765.pdf

Tse T., Marchionini, G., Ding, W., Slaughter, L., & Komlodi, A. (1998). Dynamic key frame presentation techniques for augmenting video browsing. Proceedings of Advanced Visual Interfaces (AVI'98) (pp. 185–194). New York: ACM Press.

Wactlar, H., Christel, M., Gong, Y., & Hauptmann, A. (1999). Lessons learned from building a terabyte digital video library. Computer, 32(2), 66–73.

Wang, Y., Liu, Z., & Huang, J. (2000, November). Multimedia content analysis using both audio and visual cues. IEEE Signal Processing Magazine, 12–36.

Wildemuth, B.M., Marchionini, G., Wilkens, T., Yang, M., Geisler, G., Fowler, B., et al. (2002, September). Alternative surrogates for video objects in a digital library: Users' perspectives on their relative usability. Presented at the European Conference on Digital Libraries (ECDL), Milan, Italy. Retrieved June 20, 2005, from http://www.open-video.org/papers/ECDL2002.020620.pdf

Wildemuth, B.M., Marchionini, G., Yang, M., Geisler, G., Wilkens, T., Hughes, A., et al. (2003a). How fast is too fast? Evaluating fast forward surrogates for digital video. Paper presented at the ACM/IEEE Joint Conference on Digital Libraries, Houston, May 2003. Winner of the Vannevar Bush Award for Best Paper. Retrieved June 20, 2005, from http://www.open-video.org/papers/p221-wildemuth.pdf

Wildemuth, B.M., Yang, M., Hughes, A., Gruss, R., Geisler, G., & Marchionini, G. (2003b). Access via features versus access via transcripts: User performance and satisfaction. SILS Technical Report 2003-05. University of North Carolina at Chapel Hill, School of Information and Library Science. Retrieved June 20, 2005, from http://ils.unc.edu/ils/research/reports/TR-2003-05.pdf

Wilkens, T., Hughes, A., Wildemuth, B.M., & Marchionini, G. (2003). The role of narrative in understanding digital video: An exploratory analysis. Proceedings of the Annual Meeting of the American Society for Information Science, 40, 323–329. Retrieved June 20, 2005, from http://www.open-video.org/papers/Wilkens_Asist_2003.pdf

Yang, M. (2005). An exploration of users' video relevance criteria. Unpublished doctoral dissertation, University of North Carolina at Chapel Hill.

Yang, M., Mu, X., & Marchionini, G. (2003a). Vivo—a video indexing and visualization organizer. Proceedings of the Third ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2003) (p. 403). Retrieved June 20, 2005, from http://www.open-video.org/papers/yang_jcdl_2003.pdf

Yang, M., Wildemuth, B.M., Marchionini, G., Wilkens, T., Geisler, G., Hughes, A., et al. (2003b). Measures of user performance in video retrieval research. SILS Technical Report 2003-02. University of North Carolina at Chapel Hill, School of Information and Library Science. Retrieved June 20, 2005, from http://www.ils.unc.edu/ils/research/TR-2003-02.pdf

Yang, M., Wildemuth, B.M., Marchionini, G., Wilkens, T., Geisler, G., Hughes, A., et al. (2003c). Measuring user performance during interactions with digital video collections. Proceedings of the 66th Annual Meeting of the American Society for Information Science and Technology (ASIST 2003), 40, 3–11.

Yeung, M., & Yeo, B.L. (1997). Video visualization for compact presentation and fast browsing of pictorial content. IEEE Transactions Circuits and Systems for Video Technology, 7(5), 771–785.

Zhang, H., Low, C., & Smoliar, S. (1995). Video parsing and browsing using compressed data. Multimedia Tools and Applications, 1, 89–111.

DOI: 10.1002/asi