

Enhanced Exploration of Oral History Archives through Processed Video and Synchronized Text Transcripts*

Michael G. Christel, Scott M. Stevens,
Bryan S. Maher

Entertainment Technology Center
Carnegie Mellon University
Pittsburgh, PA 15213 USA
+1 412 268 {7799, 7796, 8970}

{christel, sms, bsm}@cs.cmu.edu

Julieanna Richardson

The HistoryMakers
1900 South Michigan Ave.
Chicago, IL 60616 USA
+1 312 674 1900

jlr@thehistorymakers.com

ABSTRACT

A digital video library of over 900 hours of video and 18000 stories from The HistoryMakers was used by 266 students, faculty, librarians, and life-long learners interacting with a system providing multiple search and viewing capabilities over a trial period of several months. User demographics and actions were logged with this multimedia collection, providing quantitative and qualitative metrics on system use. These transaction logs were complemented with heuristic evaluation, interviews, and contextual inquiry with representative users. Collectively, these mixed methods informed the development of the next generation web-based interface for the HistoryMakers video oral histories to improve access to and dissemination of this rich cultural resource. In particular, the feature of a synchronized text transcript in the video player for the narratives merited further investigation. Such an interface has not seen widespread use in digital video players available on the web, yet was valued highly by oral history archive viewers. A user study with 27 participants measured the utility of the HistoryMakers web interface incorporating the synchronized transcript video player for stated fact-finding and open-ended tasks. For life oral histories, an aligned text transcript is valued for both tasks, with the video rated significantly more useful for open-ended tasks over fact-finding. These results suggest a task-dependent role of modality in presentation of oral histories, with synchronized transcripts rated highly across tasks.

Categories and Subject Descriptors

H5.1 [Information Interfaces and Presentation]: Multimedia Information Systems.

General Terms

Design, Experimentation, Human Factors.

Keywords

Oral histories, video retrieval, exploratory search.

*© ACM, 2010. This is the author's version of the work. It is posted here by permission of the ACM for your personal use. Not for redistribution.

The definitive version will be published in:

Proc. of the 2010 International Conference on Multimedia MM'10, October 25–29, 2010, Florence, Italy.

1. INTRODUCTION

Oral history has long been recognized as offering a valuable method for recording and studying memories, with extensive life histories representing a particularly valuable dimension for preserving cultural heritage. Oral history has been embraced as a tool “for giving voice to those who have been excluded from the historical record” [5]. Oral histories can “add richness and personal perspective to the historical record and can engage students and scholars in a lively study of history” [6]. Contextual inquiry has found that oral history recordings are considered a central historical artifact [11].

While the value of oral history has been recognized, its access medium has often been purely text, rather than audio or video [11]. In discussing recorded speech, Vemuri and colleagues discuss one reason why: aural speech delivery presents unique challenges [17]. The average speech rate of an English speaker is over twice as slow as the average reading rate. This large disparity suggests that automatically transcribing audio and then accessing it as a written document would be most effective for information retrieval tasks. However, in reading a text transcript, the prosodic cues, which make speech rich in meaning and subtlety, are lost [17].

Furthermore, although video has a level of emotion not available in text transcripts, and the video recordings can help transcript readers clarify text and observe non-verbal cues, “these recordings sit unused after a written transcript is produced” [11]. De Jong, Oard, and colleagues note that significant hurdles in oral history access are brought upon by the linear nature of speech and the extended length of some interviews. They comment that multimedia processing and delivery technologies offer the potential to radically transform the way in which oral histories are made accessible, encouraging and promoting the use of audio and audiovisual records [10]. Automatic transcription, improved alignment of existing transcripts, and additional metadata, particularly at the story segment level, are suggested as means of overcoming the tediousness of navigating through thousands of hours of linear audio and video.

The Informedia research group at Carnegie Mellon University has worked with an oral history archive, The HistoryMakers, to apply such automated techniques and generate time-aligned metadata for use in accessing the video narratives [3]. A lab study in 2007 showed the value of representing the oral histories in video form [4]. This paper reports on recent field studies conducted across a number of institutions and broad set of users, Empirical data of *what happens* through transaction logs is coupled with subjective

reports on *why*, through interviews and contextual inquiry, leading to the development of a new web multimedia interface. One interface element in particular enabled archive users to have greater utility with the oral history videos: the display of a synchronized transcript in the video player, which is investigated further in a study of a web portal interface.

The value of improved navigation into linear media through text transcripts has been acknowledged for webcast lectures [15] and discussed in the context of experimentation with error-laden transcripts from automatic speech recognition (ASR) [17]. The inherent value of a searchable transcript for navigating into linear audio (or the narrative audio track of linear video) can be seen with recent efforts by major Internet corporations such as Google and Microsoft to search *within* video as opposed to only searching *to a* video. Google has their Google audio indexing labs project, Gaudi, with which users can search on words within the spoken narrative [1]. Microsoft Research has their Project Tuva advanced video player in which users can also search within a video based on a synchronized text transcript [14]. Both of these examples leverage from searchable transcripts, but without providing an interface showing the full transcript. The work presented in this paper makes a significant contribution to the field of multimedia processing for cultural heritage by reporting on the use and utility of a video player interface that represents both the full transcript and full video story at the same time in a synchronized player, and conducting a user study looking at interface utility across different task types. The concluding study confirms the utility of the web interface and underscores the relationship of valued multimedia features to user tasks. This work has direct merit for oral history presentation, but also for the continued iterations of web video player interfaces facilitating information exploration of cultural archives where the voice and opinion of the individual carries great significance.

2. THE HISTORYMAKERS DIGITAL VIDEO ARCHIVE

The HistoryMakers, established in 1999, is a non-profit institution whose purpose is to record, preserve and disseminate the content of video oral history interviews highlighting the accomplishments of individual African Americans and African-American-led groups and movements [8]. Its aim is to provide a unique scholarly and educational resource for exploring African American history and culture. It is unique among collections of African American heritage because of its large and varied scope, with interviewees from across the United States, from a variety of fields, and with memories stretching from the 1890s to the present. Rather than focus on one particular part of a person's life or a single subject, such as a career or participation in the civil rights movement, the interviews are life oral histories covering the person's entire span of memories as well as his or her own family's oral history.

Collaboration between The HistoryMakers and the Carnegie Mellon University (CMU) Informedia research group made use of speech alignment, image processing, and language understanding technologies to promote multiple levels of access and fuel the viewing of the actual video recordings in a large oral history corpus [3]. Specifically, The HistoryMakers provided 913 hours of videotaped interviews with 400 prominent African American individuals conducted primarily from 2001-2005. Archivists and

staff hired by The HistoryMakers also provided human-generated transcripts, story boundaries, and reviewed and corrected the key frame per story (a representative image for the story extracted from the video). The resulting HistoryMakers Digital Library, augmented with automated multimedia processing to provide additional metadata such as time-aligned transcripts at the spoken word level of granularity and named entities (people, locations, and organizations), has 18254 interview story segments. The HistoryMakers Digital Library is a sufficiently sized corpus on which to conduct research into user reactions to digital delivery of oral history libraries.

2.1 Preparation for Field Study

A consortium of initial institutional testers was developed, with the corpus and database installed at the History Makers office in Chicago, the University of Illinois (involving the Chicago, Springfield, and Urbana-Champaign campuses), the Schomburg Center for Research in Black Culture of the New York Public Library, the State University of New York at Buffalo (UB), and Drexel University. A planning workshop at UB brought together corpus administrators, staff, and faculty from this testing community. At the workshop, participants designed a range of explorations and experiments to be developed at each site, some by expert users and others, in classroom settings, involving students first encountering the corpus and database. These discussions refined a set of input forms to gather user demographics and satisfaction data, used later in this paper. Hence, the demographics and survey instruments were designed with insights from experts in computer and library science, History, and African-American Studies. In addition, the archive owner decided on a closed form of delivery and testing to best protect the video intellectual property, which had implications for the field study as discussed later in this paper.

Users of the HistoryMakers Digital Video Library Archive (henceforth referred to as DVL) granted permission for online demographics forms, transaction logs, and survey questionnaires to be electronically sent to CMU for processing. Participants from The HistoryMakers Chicago office, the University of Illinois campuses, and Drexel University granted such access, and their work is reported collectively here.

Three pieces of infrastructure were installed at each test site: an Oracle database holding all of the HistoryMakers metadata, file system space for holding the MPEG-1 digital video files and binary search indices, and a Windows XP workstation for running the DVL interface application. DVL was written using the Microsoft .NET Framework accessing the Oracle database to produce video sets and different views into those sets, and using DirectX to play video. At most sites the video files, database, and DVL were dispersed across multiple machines connected through a local area network.

2.2 Producing/Displaying Video Sets in DVL

The opening screen seen by participants after the demographics and consent sign-in forms is shown in Figure 1. Three entry points to interview story segments are provided: text search, a Windows Explorer nested movie view acting as a table of contents arranged alphabetically by interviewee name, and a query preview browser window showing histogram distribution of stories across

time and HistoryMakers “Maker” category ArtMaker, BusinessMaker, CivicMaker, etc. (see [8]).

The video set produced by a search is shown in the main panel in a tabbed layout (so prior result sets can be tabbed to the front). Figure 2 is a screen-shot excerpt showing prior searches of the one-word query “art” and a focused search on art filtered to those stories with 1950s references. Figure 2 also shows the default presentation of a video set, i.e., the story set returned from a user query or browse action. The 54 story segments in the video set are shown in a grid of one thumbnail image per story, with a relevance vertical thermometer color-coded to matching query terms to the left of the thumbnail and tooltips showing details. The Segment Grid allows for a quick visual scan and an appreciation of how stories match the given query terms.

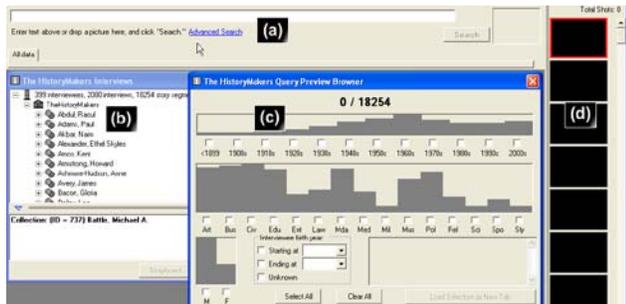


Figure 1. Entry points for oral history interface: (a) issue text query; (b) explore with interviewee table of contents; (c) browse by category such as time; and (d) capture stories to the Collector pane, initially shown as all-empty slots on the right.

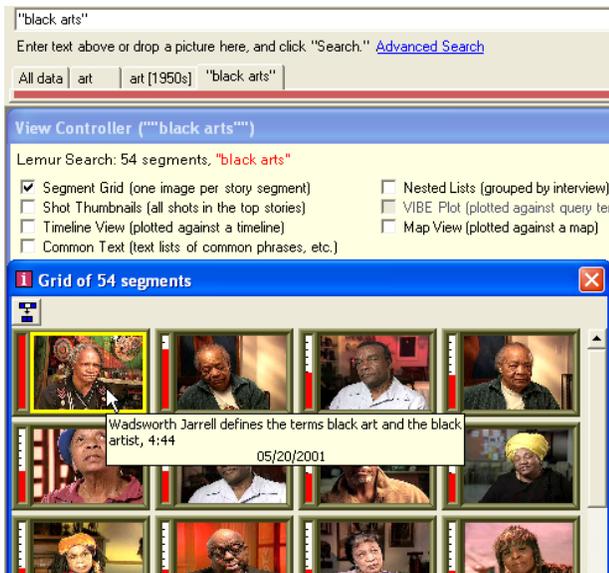


Figure 2. Interface following text query on phrase “black arts” showing the 54 video stories matching the query in the default view: a “Segment Grid” of a thumbnail, relevance vertical thermometer color-coded to matching query terms, and tooltip text with title, duration, and record date details.

The associated Figure 3 shows the video player for the first result of Figure 2 (with tooltip), with synchronized transcript and marked match points (as vertical tick lines in the video scroll line)

where “black art(s)” is mentioned, shown at 9:16 into the video. Showing the matching position is based on early work from TileBars in which the distribution of matches across a document (here, a video document) aids the user in deciding whether the document is worth further inspection [7]. There are also interface widgets to assist in quick navigation just to those match points: arrow buttons that jump to previous/next match points in both pause and play mode, as well as the ability to drag the video play bar along the timeline.



Figure 3. Video player with synchronized transcript, which scrolls as video plays, highlighting the text being spoken; line ticks show match points (here, to “black art” phrase).

Figure 4 shows other views into the video set emphasizing imagery alone, time, text, interviewee names, and geography – in the Shot Thumbnails, Timeline, Common Phrases, Nested Lists, and Map views, respectively. The views were designed to afford easy filtering by these different attributes. For example, the name “Amiri Baraka” occurs in 3 of the segments and is listed in the Common Phrases view. By drilling down on this item (a menu selection or keyboard shortcut), the user can be presented with a view of only those 3 segments, and quickly discover that Amiri Baraka is a poet, playwright, and activist formerly known as Leroi Jones.

3. DVL TRANSACTION LOG ANALYSIS

Hundreds of users of DVL agreed to have their actions with the system logged. They provided data on what was exercised and overlooked, and other insights into DVL use and HistoryMaker expectations.

3.1 Procedure

Representatives from the HistoryMakers testing community communicated periodically through phone conference calls. Based on feedback from these calls, the test systems were known to be deployed either as stand-alone computer workstations or in a computer lab with a set of machines. Participants worked primarily individually with an Intel® Pentium® 4 class machine, a 1280 x 1024 pixel 18-inch or better color monitor, and headphones for private listening. Participants’ keystrokes and mouse actions were logged within the retrieval system during the session. They first filled out an online demographic form designed as an artifact of the UB workshop. They were presented with a final survey form on exit which they could submit or opt out of completing. On subsequent visits to DVL, they could log in with an ID and password to bypass filling out the demographics form.

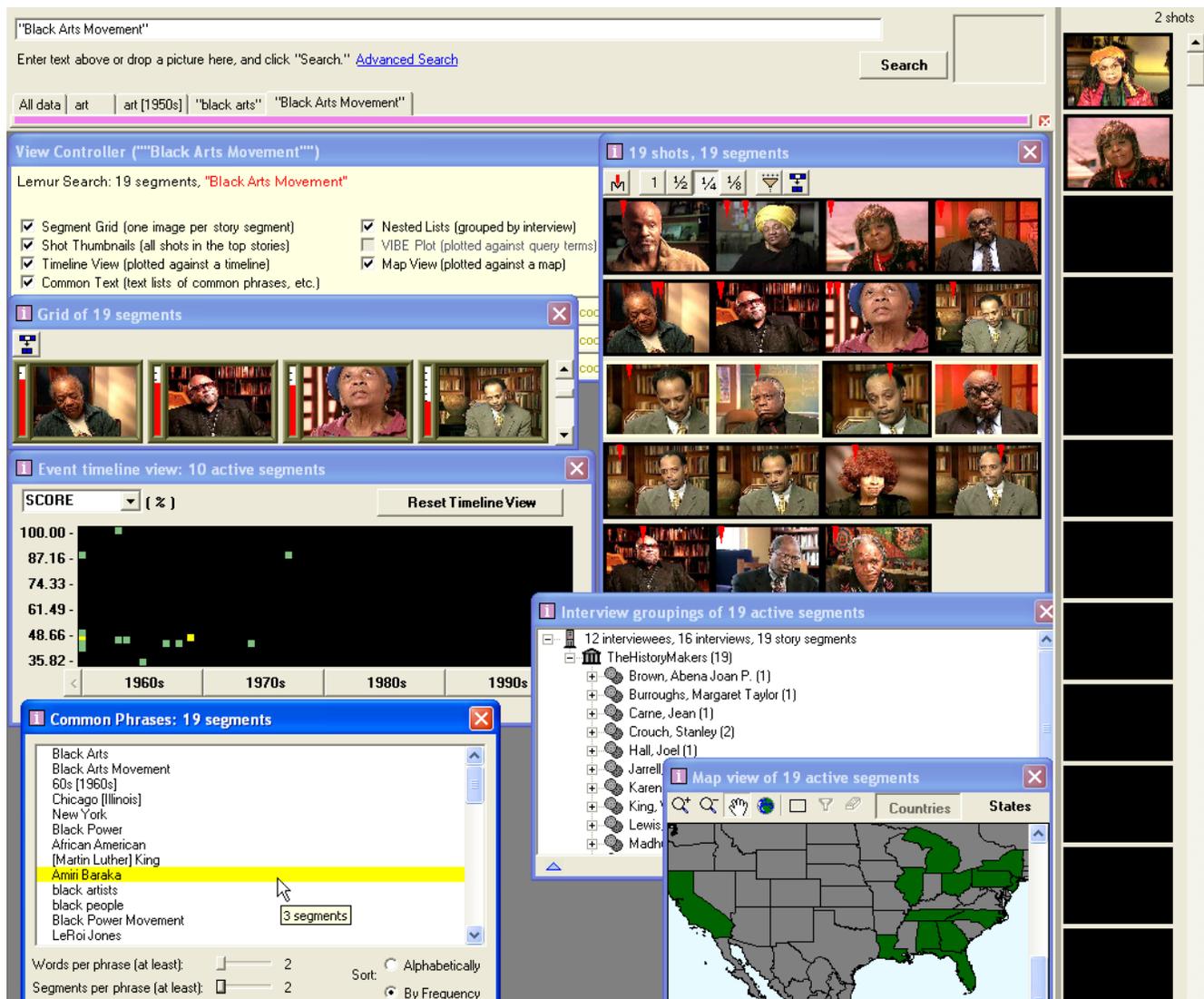


Figure 4. Assorted views into 19 results from "Black Arts Movement" query. Clockwise from top left in the results pane, these are the View Controller, Shot Thumbnails, Nested Lists, Map, Common Phrases, Timeline, and Segment Grid views.

The participants' tasks with the system were varied. They could express the tasks explicitly in an open-ended text input area on the concluding survey form (i.e., "What task were you working on?"). From the phone conversations and transaction logs, it is clear that blocks of students were working on instructor-provided assignments requiring use of the DVL workstation during a class period of an hour. The sites provided 381 transaction session logs and 258 session surveys from 266 participants. Based on conference call feedback, scores of other sessions were conducted using DVL as witnessed visually at the test sites, but with participants opting out of any transaction logging or form submission.

3.2 Participant Demographics

266 participants (158 female, 108 male) used DVL, with 210 people using it for only one logged session, 34 participants used it twice, 18 people used it three to five times, with 4 people using it six or more times. Most users were in the age bracket typical of

traditional undergraduate students, and by far the most frequent occupation typed into the demographics form was "student" with 42 varied occupations repeated at most twice (cashier, law, etc.). All but 2 participants had at least a high school education, with 162 having at least some college education and 37 having some graduate school education. The subjects' reported race/ethnicities using U.S. Census 2000 Form wording were 43% "Black, African American, or Negro," 37% "White," 4% "Asian," 4% "Hispanic or Latino," 3% "Other," 1% for each of "Pacific Islander" and "American Indian or Alaska Native," and 7% multiple selections. Figure 5 summarizes the demographics data.

The participants were spread across the range of familiarity with African American oral histories. On a 5-point scale for "How familiar are you with African American oral histories and/or The HistoryMakers?" (1=Not at all, 5=Very much), the distribution for 1-5 were 51-70-78-40-27 (mean 2.71). The participants were somewhat experienced web searchers, but less experienced searchers of video.

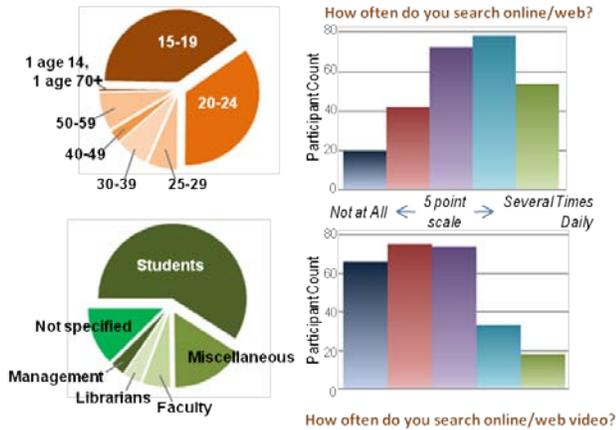


Figure 5. Demographics summary from 266 participants.

3.3 Transaction Logs and Surveys

381 sessions were logged, representing 329 hours of use, based on recorded session start and end times. The logs revealed an interesting pattern when partitioned into the 266 first-time sessions and 115 sessions from a returning DVL user. The averages are presented in Table 1, with the bolded lines in Table 1 show significant differences (session time and text query count at $p < 0.05$; average answer size in terms of number of video stories at $p < 0.01$).

Table 1. Average counts per session and times per session for 266 initial sessions and 115 sessions from returning user (bolded lines significant).

	First Time	Return Sessions
Average Session Time (as mm:ss)	44:13	52:43
Number of Text Queries Issued	4.20	3.39
Avg. Answer Size per Query	413	318
Number of Played Video Stories	10.2	10.7
Time Spent Playing Video (mm:ss)	32:39	35:58
Use of Query Preview Browser	0.18	0.32
Use of Advanced Search Options	0.61	0.97
Use of Movie Table of Contents	0.64	0.12
Use of Timeline View	0.21	0.10
Use of Common Phrases View	0.15	0.18
Use of Storyboard View	0.14	0.19
Use of Nested Lists View	0.14	0.30
Use of Geographic Map View	0.06	0.13

First-time sessions were characterized by more text queries with less precision (larger answer sizes). Users then spent the majority of their time selecting video stories from the answer sets and playing them. The table of contents organizing videos by interviewee also was used as an entry point to video by first-timers (see Figure 1), but little else was used in the interface. The 56 users (36 female) who came back to DVL did not use the table of contents in return sessions, issued fewer, more precise text queries, still spent a majority of their time playing video, but experimented a bit more with advanced search options and the

various views into video sets. For the few users accessing DVL three or more times, the use of the advanced search options and Common Phrases and Nested Lists views, all text-based, increased in later sessions, with use of the map view increasing as well.

One question of interest to the UB workshop participants was what would students do with assignments requiring use of DVL? Perhaps naively or optimistically, it was hoped that students would fall into an immediate exploration of the oral histories through the views provided and different access points, addressing the assignment but also moving beyond it to witness other portions of the corpus. What the transaction logs show is that 81% of self-described students never moved beyond the default segment grid of Figure 2, using no optional views.

A look at repeated queries made by five or more students shows that much use of DVL was prompted by required homework assignments. The following queries (listed alphabetically) were made by 5 or more students in separate sessions: black arts, black arts movement, black aesthetic, Black Panthers, boycotts, busing, Chicago, Fannie Lou Hamer, Hurricane Katrina, Obama, race, Resurrection City, rock and roll, Stokely Carmichael, sundown towns, what is race, women's rights.

Through interaction with the advanced search form, or by entering shortcut syntax of square brackets and angled brackets, the user could issue a more complex query to just a portion of the corpus (e.g., just ArtMakers or those stories discussing the 1950s) or to indexed text from certain fields (with only two fields in the HistoryMakers corpus: story titles or transcripts). The advanced search operations were not utilized frequently, only 8% of the time for queries returning results.

Perhaps the most obvious pattern seen in the query logs is confusion over the default behavior of the DVL text search service, which was set up to "or" terms together to better address small research corpora in various affiliated research efforts at the university. Even for a significant corpus like HistoryMakers, the default "or" behavior was left in place, so that a search on the two words *race riot* would return more than a search on just *riot* for example (because stories containing *race*, *riot*, or both would be returned). Within the advanced search form, there are instructions on how to specify queries matching all terms and matching phrases, specifically to use parentheses for an "and" search so that (*coffee cup*) returns segments mentioning both terms *coffee* and *cup*, and quotes to specify phrases so "*coffee cup*" with quotes in the text input area to launch searches matches stories with the phrase *coffee cup*. As noted from the transaction logs, many users never visited the advanced search tab, and instead were confused by the default "or" behavior as that did not match how Google or other web search engines work by default. In scores of sessions, users added terms to a query with the expectation that the longer word query would return fewer results. With the default "or" of terms, the opposite occurred: the longer query when specified without additional modifying punctuation produced more results.

178 of the participants completed surveys at the end of one or more sessions, producing a total of 258 session surveys regarding your "experience with the system." The survey was constructed at the UB workshop and used in prior studies. One concern was that only positively inclined participants would opt in to provide a survey, skewing interpretation. However, comparison of the results, shown in Table 2, with data from the same system and

survey in a published 2007 experiment that required surveys from all 22 users [4], show no significant differences on any of the measures. Then and now, users rate the system highly, but with room for improvement. Overall, the participants found their experiences with the system satisfying, historically meaningful, inspiring, entertaining, and accurate based on scores near 4 out of 5 on these scales. Open-ended typed comments anecdotally further supported the overall positive rating, with positive comments outnumbering the negative by a three to one ratio. Most negative comments wanted more information in the collection, e.g., more on African American women and the civil rights movement. Positive comments were often generic praise (e.g., “simply amazing,” “great job – this was an excellent experience!”), or statements on the value of oral history to cultural research and appreciation (e.g., “information is more beneficial when coming directly from the primary source,” “monumental archive achievement”, “essential tool for historical research”).

Table 2. Results of post-session questionnaires.

5-Point Scale (left = 1, right = 5)	Average (N=258)
Unsatisfying – Satisfying	3.84
Not Historically Meaningful – Historically Meaningful	3.94
Not Inspiring – Inspiring	3.81
Not Entertaining – Entertaining	3.79
Inaccurate – Accurate	3.95
Too little content – Too much content	3.12

4. CONTEXTUAL INQUIRY AND INTERVIEWS

In parallel with the transaction log collection and analysis, HCI usability techniques of contextual inquiry, heuristic evaluation and interviews were conducted regarding the interface into the HistoryMakers digital archive. The target users were students and professors of history, as the archive was employed in university history departments and used as a resource for students and professors of history. Two contextual inquiries and two interviews are summarized here as representative of this work, which complemented the interviews and group discussions from the UB planning workshop that initiated DVL field testing.

A discount usability engineering method, heuristic evaluation, was employed to get a basic sense of usability issues in DVL. An assessment team’s heuristic evaluation produced 39 violations of Jakob Nielsen’s ten basic usability heuristics [16], while performing lookup and exploratory search tasks with the system. The heuristic violations were distilled into categories of aesthetics and clutter, navigation and history, points of entry, and video set view control.

To gain a sense of understanding of how students are expected to use the system, and how professors of history create assignments, interviews were held with professors in a university History department. These interviews were meant to find out how and to what extent professors involve digital oral histories in assignments, and specifically what they expect of students using DVL. Two separate interviews were conducted with professors

that teach African American Studies courses in which The HistoryMakers digital oral histories, and thus DVL, provide a role in student assignments. Both professors created assignments for African American studies courses without actually using the DVL themselves (both have subsequently used the system and praise the HistoryMakers corpus highly). Due to their subject expertise and familiarity with interviewees in the database, both professors are confident of the existence of oral histories on given topics.

Assignments for students are research papers generally based on literature – a textbook and personal account, and digital oral histories can be used as a supplement to this. However, this by no means reduces the importance of digital oral history: one professor specifically stated that “for some people, like the lower classes, oral histories are the only way their voices can be heard.” The importance of oral histories increases for graduate students, and even more so for doctoral candidates, who may use oral histories as the backbone of a thesis.

Both professors enjoy having their students experience oral histories. Traditionally, they and their History students are much more familiar with leveraging text assets for historical research, as reflected in the lack of digital video experience in the demographics of Figure 5 as well. They see the potential for searchable multimedia oral histories. The ability to meaningfully explore the video corpus is important for students, because their African American studies classes may emphasize a narrow region of time or topic. Students need to track down relevant material from interviewees’ recent memory or recent family histories. Furthermore, when watching video and hearing the narrative, students are able to experience history in an emotionally rich context, rather than through pure text.

The team also conducted contextual inquiry [2] sessions, two of which are reported here: both involved graduate students with research theses, one familiar with the system from previous use, and one new to the system. The purpose of the research in the contextual inquiries was to learn how users work with the system, and gain insight about their mental models of the system, as well as their goals when using the system to support research for an assignment. The contextual inquiries also served the purpose of reaffirming the issues discovered in the heuristic evaluation after examination of breakdowns in sequence models. Contextual inquiries began given the task of finding interviews that supported users’ research theses, and lasted for 30 minutes.

The first user, U1, had previously taken a course with one of the interviewed professors, and had used the DVL to find HistoryMakers videos that could support a thesis. U1 immediately employed text search and was generally able to navigate the system. U1’s contextual inquiry was characterized largely by attempting to find functionality that had been used previously and forgotten, and by trying multiple search queries with very similar terms but a slight change of one term or search syntax. After a search, U1 would generally click through marked search query match time-points on the video timeline to quickly seek to relevant sections of video, and scan the transcript for useful text.

The second user, U2, had no experience with The HistoryMakers corpus of video or the DVL system. U2 had difficulty finding DVL’s primary point of entry: a text search. U2 had many breakdowns and hesitations analogous to categories or specific

issues identified during the heuristic evaluation. Also interesting is U2's desire to explore video related to certain demographics, but inability to do. Searching "women", as U2 finds out, does not bring up a subset of video interviews with women, but rather videos whose transcripts contain the word. U2 demonstrated a desire to explore the corpus based on demographics or perhaps a certain subset of a search query, but was unable to do so.

The pattern of U1's and U2's misuse of DVL text search points to a familiarity with major internet search engine syntax, and repeated similar searches display a desire to explore related content to the current video results, but the system does not easily afford this. The users expressed a desire through words and actions for better exploration through interviewee demographics. They also wanted to know the breakdown of results according to interviewee, not just how many story segments are returned which is made obvious in the interface, but how many interviewees produced those results (which is available but buried within the Nested Lists view illustrated in Figure 4).

5. LESSONS LEARNED FOR WEB DVL DEVELOPMENT

The methods of the prior section produced usability priorities set from categorical analysis of the DVL's heuristic violations, an understanding of the importance and relevance of digital oral histories to students and professors, and concrete ideas of users' tasks and how to better support their work. The transaction logs and surveys provided metrics of use in additional settings. This information was used collectively to brainstorm on developing a new web site portal into the HistoryMakers Digital Library.

5.1 Improved Delivery Architecture

Access to the oral histories was available only through DVL, which was installed on only a limited number of Windows machines on campuses or in The HistoryMakers office. In some cases, the machines were unavailable for extended periods, such as over the summer, explaining why the transaction log counts dropped to near zero in summer. Overall, there was an expressed hope for 24/7 access to the HistoryMakers corpus from users' own computers.

The transaction logs showed another problem with the test site setup: the dependency on three components – a Windows client, Oracle database server, and file system access to indices and video – caused numerous frustrations. Many of the negative comments and failed queries in transaction logs were traced back to one or both of the database server or file system becoming unreachable. On one campus there was a renaming of network servers causing a temporary lookup failure. On another there was a power outage rendering key resources unavailable.

With much encouragement from HistoryMakers test sites, and a realization by the archive owner of the negative implications caused by a closed delivery infrastructure, a web portal has been initiated, where the metadata for the multimedia collection is available through web services and the interfaces to the HistoryMakers search, organize, and view operations are provided by any standard web browser. The application is written in Adobe Flex, with the produced Adobe Flash application being accessible by a vast majority of Internet consumers.

5.2 Clean, Simple Access

The new visual style for web-DVL was implemented with a darker background and lighter text so that HistoryMaker videos would stand out from the interface. The entry screen for DVL (Figure 1) was too busy, based on the collected logs, surveys, and interviews. The new entry page (Figure 6) for the web portal clearly shows text and map search options. The map search link brings up a standard GUI into web map displays, by using freely provided APIs for embedding rich interactive map functionality in a consistent manner (e.g., showing selected locations with pins on the map). The entry point also notes characteristics of the corpus, e.g., that interviews were collected from 1999 through July, 2005. Through such communication, the user knows the bounds of the corpus and will be less likely to ask against it current events like the Obama presidency.



Figure 6. Entry page for web portal to HistoryMakers video.

Multiple views were added, including a packed thumbnail grid view, and a list view with text and various information (blurb and date of interview; Figure 7). The main reason for adding the packed grid view was for the visual appeal and dense visual overview to address users' comments on the utility of a pictorial overview of the HistoryMakers speakers. If the window is expanded, the layout adjusts accordingly to fill the space. Simplicity was emphasized over the complexity of views of Figure 4, with the intent of encouraging more interactions by first-time visitors, getting them to return, and attracting casual browsers and others looking for quick entry points into the oral histories.



Figure 7. Web portal, showing filter controls, video pane with synchronized transcript, results pane, and collector pane.

Text transcripts are always shown with the video player, given the traditional utility and familiarity of the text modality for historical research. These transcripts are automatically aligned with the

spoken stories, and the timing metadata used to support marking where matching terms are spoken and allowing quick click access to navigate to those points in the video.

5.3 Intuitive, “Standard” Use of Text Search

Two fixes to the text search service account for 45% of the queries that frustrated users by returning no results as evidenced in the transaction logs and brought up as well in contextual inquiry: (1) the use of Lemur [12] allowing single letter searches and searching on phrases with stop words, and (2) the adoption of typical web search engine behavior (e.g., Google). By default, all words in a query are required to match rather than any word, so longer queries produce fewer results as expected based on transaction query patterns and web search experience. The complicated syntax and “advanced search” form for specifying field searches and filtered searches are disabled in favor of explicitly showing filters together in a panel under user control.

5.4 Intuitive Use of Filtering

Through continuous user testing, we narrowed down to several main filters that were strongly suggested by the users based on logged queries, surveys, and interviews: Interviewee Gender, Interviewee Birth Year, and Interviewee Occupation (“Maker” category). In the old system, there were an overwhelming number of filters and ways of narrowing down videos. In the transaction logs, we witnessed the lack of use of these advanced search options and specialized viewing options, especially for students and first-time visitors. Through additional user testing, we found that simple, yet easy to understand filters were appealing to History students and faculty. Considering that many had concerns that it would extend the time they take to research, we focused on cutting down the extra time they might have to look up meanings of filters. The History students and professors noted that they wanted to maximize time spent with the video recordings and not the interface into them.

6. UTILITY OF VIDEO AND TEXT IN WEB DVL

To investigate whether the lessons outlined in the previous section resulted in a web interface useful for interacting with the HistoryMakers video collection, a study was formulated in which users were asked to do two tasks, one fact-finding and the other exploratory. Marchionini breaks down three kinds of search activities: lookup, learn, and investigate, noting exploratory search as especially pertinent to the latter two activities [13]. The fact-finding “treasure hunt” task is solely in the “lookup” category, as are the tasks in many traditional information retrieval forums studying fact retrieval or question answering. The open-ended task brought in learning (knowledge acquisition) and investigating (analysis, synthesis) activities. Not surprisingly, the humanities faculty and library science professionals at the UB workshop were quite interested in the use of The HistoryMakers oral histories for more than just traditional fact retrieval and lookup, but also for these learning and investigating activities. Hence, the study incorporated both tasks.

6.1 Participants

Study participants were recruited through an online forum attracting predominantly but not exclusively college students. 28 subjects signed up, with one quitting early due to a schedule

conflict. The 27 subjects who fully participated in this study had no prior experience with the interface or data under study and no connection with the research group. They began the study by completing a demographics questionnaire. The subjects were 13 female and 14 male with a mean age of 24.6. All subjects had at least some college education and 16 had some graduate school education. The subjects’ reported race/ethnicities were 1 “American Indian or Alaska Native”, 1 “Black, African American, or Negro,” 11 “White”, and 14 “Asian.” The participants were generally not familiar with African American oral histories. On a 5-point scale responding to “How familiar are you with African American oral histories and/or The HistoryMakers?” (1=Not at all, 5=Very much), most indicated little familiarity (distribution for 1-5 were 14-11-0-1-1, mean 1.6). The participants were experienced web searchers and modest digital video searchers. For “Do you search any web/online information systems?” (1=Not at all, 5=Very frequently (several times daily), the answer distribution was 1-1-5-9-11 (mean 4.0) while for “Do you use any digital video retrieval system (video stored and searchable on a computer)?” with the same scale, the distribution was 6-8-5-4-4 (mean 2.7). In listing search engines used, all 27 listed Google, with Yahoo a distant second at 5; YouTube was named by just 1.

6.2 Procedure

One task was a treasure hunt task, finding and collecting the video stories that answered one of twelve stated topics, listed in Table 3. These topics were derived to cover a broad span of the Maker categories (e.g., not all sports- or politics-related). The treasure hunt task was used to study the new web interface illustrated in part in Figures 6 and 7. The second task was an open-ended exploration of collecting evidence and interesting stories for a topic of your own choosing. For both tasks the collector pane at lower left of Figure 7 served as the answer space. Participants randomly were assigned to either start with the open-ended exploratory task and then do the treasure-hunt task or vice versa; each task was limited to 20 minutes after which a survey questionnaire was filled out.

Table 3. Topics for treasure hunt (fact-finding lookup) task.

Kerosene ball for baseball
Photo of uniformed military with Secretary of Defense Melvin Laird
Dignity from James Edwards to Sidney Poitier to Denzel Washington
Remembering scents of childhood holidays: turkey, pies, and ocean
Discussing chat with Bishop Tutu about Mandela, held in California
Meeting Bobby Kennedy in a South American country
Jimmy Carter’s family stays in his Colorado home
Citing a “history of the world...is of groups” DuBois quote
Hired onto TV show staff to make show “psychologically believable”
Photo showing three with same middle name Proctor
Naming three areas in the American South that were mostly Catholic
Arabic welcome in album with Louis Armstrong’s “Wonderful World”

6.3 Results and Discussion

On the treasure hunt task, subjects found an average of 5.74 topics out of 12, standard deviation 2.92, minimum of 1 and maximum of 12. To satisfy the fact-finding task, subjects used the system in significantly different ways, and rated the system significantly

differently as well, from their use and ratings for the open-ended task. There were no significant differences brought on by task time ordering (i.e., first task encountered vs. second task).

Participants working on the treasure hunt performed significantly more text queries, $F(1,52) = 57, p < 0.001$; with significantly greater precision (fewer stories per query), $F(1,52) = 8.3, p < 0.01$. Working on the treasure hunt, participants watched significantly more video stories, $F(1,52) = 25, p < 0.001$. The video watching was in brief bursts resulting in significantly less play time than when working on the open-ended task, $F(1,52) = 49, p < 0.001$. These results are shown in Table 4.

Table 4. Summary of interface use, by task.

	Treasure Hunt	Open-Ended
Session Time (as mm:ss)	20:00	20:00
Number of Text Queries Issued	45.2	14.1
Avg. Answer Size per Query	52	139
Number of Played Video Stories	18.4	8.7
Time Spent Playing Video (mm:ss)	7:08	13:41

Participants rated the utility of the text transcript highly for both tasks, as shown in Table 5. The utility of the video player was rated significantly higher for the open-ended exploration than for fact-finding, $F(1, 52) = 14, p < 0.001$.

Table 5. Results of post-task questionnaires.

5-Point Scale (left = 1, right = 5)	Treasure Hunt	Open-Ended
Unsatisfying – Satisfying	4.0	4.0
Not Historically Meaningful – Historically Meaningful	3.9	4.1
Not Inspiring – Inspiring	3.6	4.0
Not Entertaining – Entertaining	3.7	4.0
Inaccurate – Accurate	4.2	4.0
Too little content – Too much content	3.3	3.0
Text Search: Not Useful – Useful	3.8	3.8
Map Search: Not Useful – Useful	2.4	2.8
Search filtering: Not Useful – Useful	3.3	3.7
Video player: Not Useful – Useful	3.4	4.5
Text transcript: Not Useful – Useful	4.3	4.2

When users were addressing the treasure hunt, they were consumed by searching and refining searching to locate stated information needs. Their actions concentrated on confirming a story’s relevance through a quick review of its matching terms, made trivial through the match lines on the video play bar, navigation buttons to seek to previous/next match, and colored match terms in the transcript. The transcript text was rated highly useful. Somewhat surprisingly, the transcript text was still rated as highly useful in the open-ended task. The open-ended task underscored the value of video representation for life oral

histories, as users spent a significant amount of time listening to and appreciating the video accounts and rated the video player most useful.

This study does not conclude our work with HistoryMakers, as more design and development for the web portal is of course planned. As with the prior methods, the study itself contributes more data to inform this future work. The search filtering at the left of Figure 7, for example, was underutilized in the study, and “search filtering” rated low. So, too, was the map search. Folding in others’ faceted search interface developments, and highlighting those facets subdividing the story answer space in interesting ways rather than hiding it all within the filter pane, are future directions we will pursue.

In agreement with published work on information visualization interfaces, more advanced and less familiar features are often underutilized in a user’s first exposure to a system. The users in this study only spent 40 minutes with the system. Future longitudinal studies are planned during which we expect to see greater use of the filtering and map search features. When these features were used in this study, they were used in support of the open-ended task 73% of the time. For fact-finding, participants made use of only text search and text search-in-search, avoiding the use of the additional navigation mechanisms.

7. CONCLUSIONS

The low cost of digital storage now makes feasible hosting thousands of hours of video recordings of life oral histories in a digital video library. Access to the tales, formerly mediated through text transcripts or limited to the analysis and interpretation offered by a synthesizing historian or documentary producer, can be granted to end users directly in multimedia form. This paper reports on interfaces to one such collection.

The lessons reported as subheadings in Section 5 apply not only to oral histories, but to audiovisual interfaces into cultural heritage collections in general. An improved delivery architecture allows broader access to the collection. Clean, simple access encourages first time use drawing viewers into the materials. Intuitive, “standard” use of text search reduces frustration and perceived error. Intuitive use of filtering maximizes time spent with the primary source materials (e.g., video) and not the interface into them. These lessons have been incorporated into the design of the latest web interface into The HistoryMakers Digital Archive [9].

Lessons learned from The HistoryMakers work strongly suggest that a fully searchable, viewable, synchronized text transcript presented alongside the video player will be valued highly and used often for both fact-finding and exploration tasks. Major Internet players are now researching search within video through transcripts. This work offers empirical data on the value of keeping text transcript representation as well. Moreover, the value of the text and video modalities together with indexing (e.g., Lemur) and processing (e.g., automatic transcript alignment to temporally locate word matches) address the hurdles in oral history access brought upon by the linear nature of speech and the extended length of some types of interviews.

8. ACKNOWLEDGMENTS

This work is supported by the National Science Foundation under Grant No. IIS-0705491. We thank all the users who volunteered their input for this study. Special thanks go to CMU students Lee Byron, David Hora, Anna Ly, and Emily Tong who conducted the contextual inquiries and other HCI discount usability techniques informing the development of this work.

9. REFERENCES

- [1] Alberti, C., Bacchiani, M., Bezman, A., Chelba, C., Drofa, A., Liao, H., Moreno, P., Power, T., Sahuguet, A., Shugrina, M., and Siohan, O. An Audio Indexing System for Election Video Material. In *Proc. ICASSP 2009*, IEEE (2009), 4873-4876.
- [2] Beyer, H., Holtzblatt, K. *Contextual Design: Defining Customer-Centered Systems*. Academic Press, London, UK, 1998.
- [3] Christel, M., Richardson, J., and Wactlar, H. Facilitating access to large digital oral history archives through Informedia technologies. In *Proc. JCDL 2006*, ACM Press (2006), 194-195. DOI= <http://doi.acm.org/10.1145/1141753.1141795>.
- [4] Christel, M., and Frisch, M. Evaluating the contributions of video representation for a life oral history collection. In *Proc. JCDL 2008*, ACM Press (2008), 241-250. DOI= <http://doi.acm.org/10.1145/1378889.1378929>.
- [5] Frisch, M. 2003. Working-Class Public History in the Context of Deindustrialization: Dilemmas of Authority and the Possibilities of Dialogue. *Labour/Le Travail* 51 (2003).
- [6] George Mason University. *Oral History Program*. http://sca.gmu.edu/oral_history.htm.
- [7] Hearst, M.A. TileBars: Visualization of Term Distribution Information in Full Text Information Access. In *Proc. CHI 1995*, ACM Press (2005), 59-66. DOI= <http://doi.acm.org/10.1145/223904.223912>.
- [8] HistoryMakers. *The HistoryMakers.com African American History Archive*. <http://www.thehistorymakers.com>.
- [9] HistoryMakers and Informedia Digital Video Library. *The HistoryMakers IDVL Digital Archive*, <http://www.idvl.org/thehistorymakers>.
- [10] Jong, F. D., Oard, D. W., Heeren, W., and Ordelman, R. Access to recorded interviews: A research agenda. *J. Comput. Cult. Herit.* 1, 1 (2008), 1-27. DOI= <http://doi.acm.org/10.1145/1367080.1367083>.
- [11] Klemmer, S. R., Graham, J., Wolff, G. J., and Landay, J. A. Books with voices: paper transcripts as a physical interface to oral histories. In *Proc. CHI 2003*, ACM Press (2003), 89-96. DOI= <http://doi.acm.org/10.1145/642611.642628>.
- [12] *Lemur Toolkit for Language Modeling and Information Retrieval*. <http://www.lemurproject.org>.
- [13] Marchionini, G. Exploratory Search: From Finding to Understanding. *CACM* 49, 4 (2006), 41-46. DOI= <http://doi.acm.org/10.1145/1121949.1121979>.
- [14] Microsoft Research. *Project Tuva*. <http://research.microsoft.com/apps/tools/tuva/index.html>.
- [15] Munteanu, C., Baecker, R., and Penn, G. Collaborative Editing for Improved Usefulness and Usability of Transcript-Enhanced Webcasts. In *Proc. CHI 2008*, ACM Press (2008), 373-382. DOI= <http://doi.acm.org/10.1145/1357054.1357117>.
- [16] Nielsen, J. and Molich, R. Heuristic evaluation of user interfaces. In *Proc. CHI 1990*, ACM Press (1990), 249-256. DOI= <http://doi.acm.org/10.1145/97243.97281>.
- [17] Vemuri, S., DeCamp, P., Bender, W., and Schmandt, C. Improving Speech Playback Using Time-Compression and Speech Recognition. In *Proc. CHI 2004*, ACM Press (2004), 295-302. DOI= <http://doi.acm.org/10.1145/985692.985730>.