CS-TR-3838
UMIACS-TR-97-73
ISR-TR-

Previews and Overviews in Digital Libraries: Designing Surrogates to Support Visual Information Seeking

Stephan Greene, Gary Marchionini[1], Catherine Plaisant, and Ben Shneiderman[2]

Human-Computer Interaction Laboratory,
University of Maryland Institute for Advanced Computer Studies
[1]also College of Library and Information Services
[2]also Department of Computer Science and Institute for Systems Research
College Park, Maryland 20742-3255, +1 301 405 2725
{greene,plaisant,ben}@cs.umd.edu, march@oriole.umd.edu
http://www.cs.umd.edu/projects/hcil

ABSTRACT

To aid designers of digital library interfaces and web sites in creating comprehensible, predictable and controllable environments for their users, we define and discuss the benefits of *previews* and *overviews* as visual information representations. Previews and overviews are graphic or textual representations of information abstracted from primary information objects. They serve as surrogates for those objects. When utilized properly, previews and overviews allow users to rapidly discriminate objects of interest from those not of interest, and to more fully understand the scope and nature of large collections of information resources. We provide a more complete definition of previews and overviews, and discuss system parameters and aspects of primary information objects relevant to designing effective preview and overviews. Finally, we present examples that illustrate the use of previews and overviews and offer suggestions for designers.

Previews and Overviews in Digital Libraries: Designing Surrogates to Support Visual Information Seeking

Stephan Greene, Gary Marchionini[1], Catherine Plaisant, and Ben Shneiderman[2]

Human-Computer Interaction Laboratory,
University of Maryland Institute for Advanced Computer Studies
[1]also College of Library and Information Services
[2]also Department of Computer Science and Institute for Systems Research
College Park, Maryland 20742-3255, +1 301 405 2725
{greene,plaisant,ben}@cs.umd.edu, march@oriole.umd.edu, http://www.cs.umd.edu/projects/hcil

ABSTRACT

To aid designers of digital library interfaces and web sites in creating comprehensible, predictable and controllable environments for their users, we define and discuss the benefits of *previews* and *overviews* as visual information representations. Previews and overviews are graphic or textual representations of information abstracted from primary information objects. They serve as surrogates for those objects. When utilized properly, previews and overviews allow users to rapidly discriminate objects of interest from those not of interest, and to more fully understand the scope and nature of large collections of information resources. We provide a more complete definition of previews and overviews, and discuss system parameters and aspects of primary information objects relevant to designing effective preview and overviews. Finally, we present examples that illustrate the use of previews and overviews and offer suggestions for designers.

Keywords

Browsing, Information Seeking, Surrogates, Previews, Overviews, Visualization.

INTRODUCTION

Movie-goers are enticed and informed by well-designed previews. Book-lovers roam the aisles of bookstores scanning jackets for titles and authors, and flip pages for catchy phrases or images. And now web surfers scan home pages or search result sets to decide where to go next. Unfortunately, in emerging digital information environments, designers often fail to provide appropriate preview materials to give an overall sense of the structure and materials available. This paper seeks to address these problems by recommending designs for *previews* and *overviews* as information representations.

Users prefer comprehensible, predictable and controllable environments in which they can rapidly and safely explore and use information. Design guidelines intended to help realize such systems abound, but recent work in the HCI literature has produced several characterizations of user interface design parameters that have advanced more thorough, concrete and practical concepts for the design space of interactive systems. Furnas [8] proposed useful criteria for designing and evaluating representations of information structures in support of traversal and navigation. Ahlberg and Truve [1] described the design space of query devices. Tweedie [22] outlined several design dimensions for describing and comparing the interactive aspects of information representations. Shneiderman [18] characterized information visualizations in terms of data types and user tasks.

Creation and evaluation of information surrogates are long-standing research issues in library and information science. Heilprin [9] argues that the abstractions and compressions he terms *homomorphic reductions* are fundamental to information science; Borko & Bernier [4] in their work on abstracting list two dozen types of text surrogates. Surrogates are essential for analytical searching because metadata such as Library of Congress MARC records, NASA DIF records, or data codebooks are often composed of standard sets of identifiers that searchers use to make decisions about what objects to examine next. In browsing, surrogates provide an important alternative to primary objects as they take far less time to examine and provide enough semantic cues to allow users to assess the need for further processing of other surrogates and the primary object. In digital libraries and archives, surrogates are crucial for browsing large distributed collections that result from filtering programs or analytical queries of the data space. The need for inventing new types of surrogates underlies much of the research in digital libraries.

Our analysis of previews and overviews builds upon and complements these previous efforts, within the context of a theory of information seeking. Our analysis is based on the design and evaluation of several digital library projects (with NASA, the Library of Congress, and the National Library of Medicine), and most of our examples are drawn from this work. (See Figures 1 and 2).

PREVIEWS AND OVERVIEWS IN CONTEXT

Previews and overviews are graphic or textual representations of objects of interest. As a first approximation:

- A preview is extracted from, and acts as a surrogate for, a single object of interest.

- An overview is constructed from, and represents, a collection of objects of interest.

The design of previews and overviews is guided by the view that information seeking is a dynamic and iterative decision-making process [10]. This view holds that except for known-item searches, information seekers initiate search through a query or by browsing in a subset of the dataset. Users scan objects rapidly to determine whether to examine them more closely, or move on in the dataset. This process continues until the information need is satisfied or satisficed, or the search is abandoned.

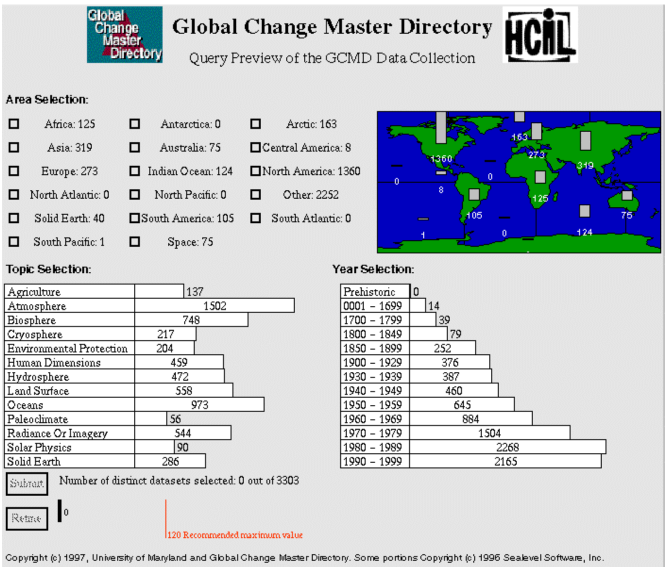

Figure 1a. The interface for NASA earth science data [6] provides a single screen overview of available data (i.e. numbers of data sets).

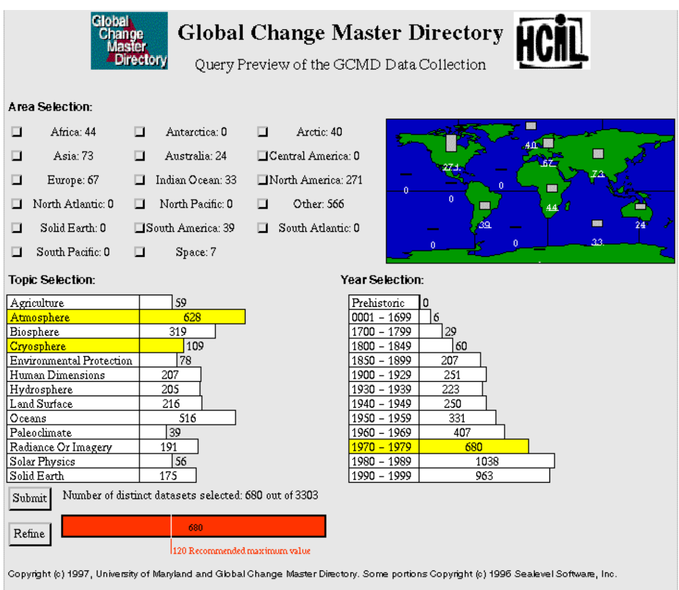

Figure 1b. When users select attribute values, count totals are updated, giving users a preview of the size and distribution of the result set, before the query is submitted. This eliminates the zero-hit or mega-hit query problem.

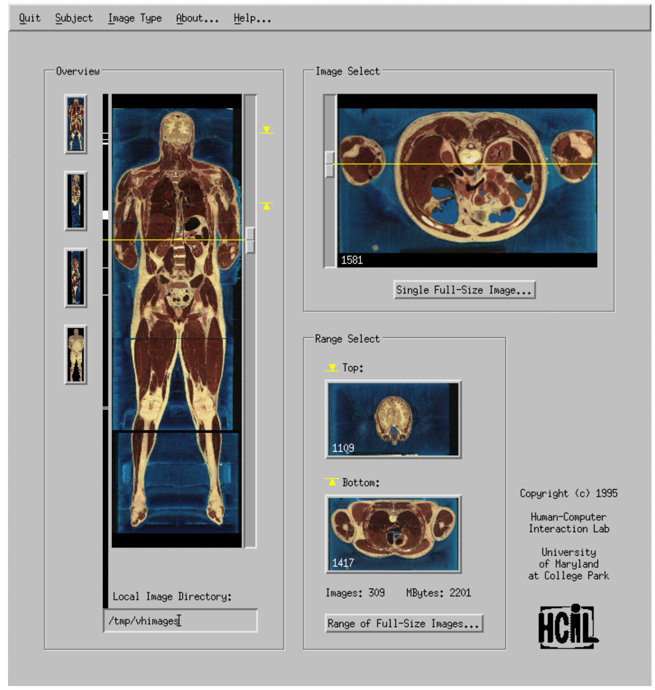

Figure 2. The Visible Human Explorer [12] allows users to rapidly browse thumbnail previews of high resolution images that would take days to download. On the left, a composite full-body thumbnail overview, constructed from the 1800 separate images, allows users to browse individual image previews.

This view of information seeking requires system designs that provide multiple levels of representation for information objects. The levels may range from terse but informative surrogates such as titles, to increasingly detailed surrogates that become successively more indicative of primary objects. To help ensure the success of such processes of progressive refinement, it is important to provide the user with appropriate intermediate surrogates that facilitate the rapid elimination of objects from further examination. Our design strategy provides users with overviews of collections which are used to select successively more detailed overviews or specific previews.

Previews are analogous to bibliographic records and overviews are analogous to catalogs. It is important to note that a catalog may itself have a bibliographic record (preview) and a bibliographic record may serve as a catalog (overview) for a complex object such as chapters of an anthology. Depending on the user's task and the nature of the information represented, some designs fit the definitions of both previews and overviews, or fit one definition and then the other at different points in the interaction. For example, the thumbnail image of the Visible Human Explorer (top right of Figure 2) acts as a preview when the user is selecting images to download, but might be used as an overview to browse the full size image in a classic overview/detail image browser [14]. Any information surrogate will embody, to some degree, preview or overview aspects. Thus, the preview-overview characterization provides a fundamental dimension along which to design and evaluate a wide variety of information representations.

Browsing is a human activity that closely couples search and use as parallel activities. Unlike analytical search strategies, which can be batched and automated, browsing is an inherently interactive process that demands ongoing human attention. Thus browsing is ultimately constrained by the limits of human physiology and cognition. Well-designed surrogates help exploit human perceptual strengths while not exceeding cognitive limitations.

Previews and overviews provide a useful family of information representations that can help to meet these challenges. An effective preview is an information surrogate that communicates to the user, at the appropriate time, sufficient information about the primary object it represents to support users in making a correct judgment about the relevance of that object to the user's information need. An effective overview provides users with an immediate appreciation for the size and extent of the collection of objects the overview represents, how objects in the collection relate to each other, and, importantly, what kinds of objects are not in the collection.

DESIGN DIMENSIONS OF PREVIEWS AND OVERVIEWS

The design of effective previews and overviews requires a thorough analysis of the application domain with respect to three interdependent concepts: (1) what information objects are available to users; (2) how those objects are related and displayed; and (3) how users can manipulate those objects. The analysis guides the design of surrogates by helping to determine what is to be represented, how representations are to be organized and displayed, and how the representations are controlled during interaction. However, this does not imply that surrogate representations are organized and manipulated in ways analogous to the way primary objects are organized and manipulated. Especially for overviews, innovative representations will be abstractions of primary objects and their attributes. Nonetheless, all these designs result from the careful analysis of primary objects and users' tasks.

Information Objects and Object Surrogates

The optimal identification of information objects, and designing appropriate representations for them, requires carefully balancing the demands posed by data characteristics, user needs, and system constraints.

Identifying Primary Objects

Primary information objects should be identified from a user-centric perspective. They are logical entities useful to the user, not necessarily identical to entities as stored in the system. For example, although an image and its caption may be physically retrieved from separate system files, they are likely to be considered a single object.

When primary objects are too diverse or complex, an aggregate form of those objects can be chosen for use in previews and overviews. In the NASA example, the "ultimate" primary object useful to the user is the earth science data itself. But the data comes in such great variety of form (numeric observations, images, series, etc.) that objects cannot be uniformly described. Dataset records were a better choice of primary object for the initial query interface because they have consistent metadata that can be used to create surrogates.

The user task determines the most appropriate level at which to consider an information object. That is, function establishes the optimal granularity of the relevant data type. In the NASA example, scientists need to identify data to conduct a particular inquiry. They are better served by an interface that lets them start by reviewing and selecting a small number of datasets, as opposed to wading through millions of datapoints.

Defining the Role of Object Surrogates

Surrogates for objects are created to serve several purposes. A primary motivation generally dominates the design process, although additional purposes are usually served in tandem. Many surrogates may already be available to the user interface designer because they are inherent to the data (e.g. title or section headings in structured documents) or because they were created to serve other system requirements (e.g. bibliographic catalog entries are created for information management and retrieval purposes and are also useful as basic previews). In all cases, an information surrogate adds value that complements the primary information object.

The main roles for preview and overview surrogates include:

- Aiding retrieval (i.e., used as the basis for indexes)
- Aiding users in quickly making relevance decisions
- Reducing network data transfer (i.e. reducing time and resource needs) to make practical the examination of numerous resource-intensive primary objects
- Supporting the user's need to capture the gist of complex heterogeneous information
- Providing indicators of scope, size and structure of large information spaces
- Informing and enticing with representative samples, while helping users define productive queries

Issues and Options in the Surrogate Design Space

System perspective: Surrogates may be literal extracts which are extracted without change from the primary object, or original constructs, which are alterations or original creations that convey meaning from the primary object. This distinction is similar to the distinction between constructed and converted data as defined by [22]. Examples of literal extracts include thumbnails, tables of contents, pull quotes, video streamers [7], low resolution images, and sampled video frames. Examples of original constructs include text abstracts, controlled vocabulary term lists, collage images or salient stills [21]. Original constructs are generally more difficult to create automatically than literal extracts, and thus are often considered to have "more" added value than literal extracts.

This dimension of surrogate design shows that the original definition of previews as extracted surrogates and overviews as constructed surrogates is relative to the degree of content transformation (opposed to system operation or level of effort). For example a randomly generated collage of sample images from a photo collection has been constructed but did not require any intelligent intervention and remains a literal surrogate.

User perspective: From the user's point of view, previews are very sensitive to time, while overviews are less so. Users preview items before examining their details, and usually need not reference the preview again. In contrast, users can make good use of overviews before, during, or after examining details.

Preview and overview surrogates can give users representations that differ in their degree of abstractness and exhaustiveness. Abstractness refers to the degree to which a surrogate is removed from a literal representation of the primary object. The vivid images of the Visible Human Explorer sit toward one end of the scale and the bars of the NASA example toward the other end. Exhaustiveness is a related concept that refers to the degree to which a surrogate reflects the entirety of the primary object(s) it represents. Overviews by definition should be fairly exhaustive but problems like overlaps or a limited number of represented attributes might limit the coverage of the overview. Previews are much less likely to be exhaustive.

Additionally, previews can influence the attention of users as they decide whether to look at the whole. A preview has the potential to be misleading, eliciting imprecise or incorrect generalizations about the primary object. Overviews, by contrast, can be used to understand an object in context. It is safer to generalize about an object from an overview because it aims to be a general summary or description, and supports comparison among related objects. Thus, previews may be more affective while overviews are more cognitive.

Surrogate Structure and Presentation

Successful browsing depends on the attributes of individual units of information, how different units are related to one another, and how they are displayed to users. Effective previews and overviews need to take advantage of data type-specific structures. Preview displays are mainly determined by what attributes are available and the design decisions revolve around choosing attributes and how best to display the resulting surrogate. Overviews are more challenging as their display depends on commonality of attributes across objects. Design decisions are often complicated by a dearth of common attributes and the need to create abstract attribute classes.

Browsing previews is facilitated by regular object structures such as single or multiple columns, linear sequences of scenes or MTV-like montages, and comma delimited or tabular layouts, which can be used in the design of the surrogate. A good example is TableLens [17] which uses a graphical overview that mirrors the original information object (i.e. the numeric table).

Browsing overviews created from collections with common attributes (e.g., a list of retrieved records in a database or online catalog) allows users to exploit consistent surrogate structure diplayed in lists, tables, and other common displays. However, search engines and filters that work across many information sources often yield heterogeneous results. A result set may include representations of primary objects such as full texts, videos, or datasets employing different levels of object surrogates such as titles, abstracts, samples, or summary statistics. For example, on the Web, users may have to browse vertically across several levels (e.g., across specific pages, sites, or indexes of sites). Another example is a query in the Library of Congress National Digital Library that may return a set of hits that mix collections as a whole (e.g., the name of a photograph collection), finding aids (e.g., a full-text document that describes and gives access to a manuscript collection), and individual items (e.g., a specific photograph or sound recording). In these cases, common attributes may be absent and more abstract attributes must be invented.

In all cases, the key is to create previews and overviews that communicate the level and scope of objects to users so that comprehension is maximized and disorientation is minimized, and the best browsing control mechanisms can be applied. There are still few if any examples of surrogates that address this problem. One attempt is described in our Library of Congress National Digital Library prototype [16].

Surrogate Control Mechanisms

The third major design parameter is the mechanisms for controlling surrogates. Individual surrogates can be compared along interactivity and tailorability dimensions. For example the preview of a collection made of a collage of samples can be merely a static graphic, or can be designed to be more interactive with a slideshow of samples, each selectable to jump to the corresponding primary object. In complex applications tailorability of the surrogates can be very important. The most common examples are found in overviews in which users are likely to want to choose which attributes should be used or modify display attributes such as color, size or labels. In some cases the overview surrogate can become a complete application such as dynamic query interface coupling a zoomable overview with a real-time filtering mechanism using sliders and buttons (Figure 3) [2].

In the dynamic and interactive information seeking process appropriate control mechanisms are also needed to navigate *among* surrogates at the same level or different levels of representation. The two classic control mechanisms are scroll and jump but zooming is becoming a viable alternative mechanism as workstations become more powerful (e.g. [3]). Zooming offers users the continuity of scrolling as well as a natural metaphor for browsing vertically across different levels of surrogates.

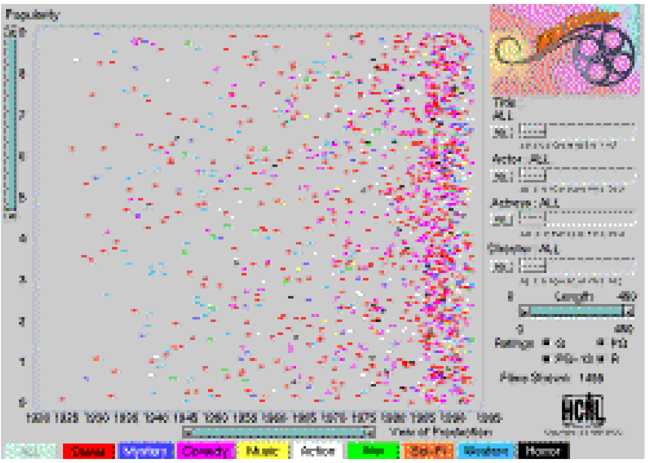

Figure 3. The Film Finder. A very interactive and tailorable, but rather abstract, overview.

Another issue is the use of multiple coordinated windows or views for presenting surrogates. Current work enhances information coordination and organization across multiple windows [13]. The Visible Human Explorer is a good example of an overview coordinated with a preview (Figure 2). Users can select a preview by sliding the cursor on the overview, or change the overview forward and backward by sliding the cursor of the preview.

Specialized manipulation mechanisms are being developed for certain data types. For example video previews [23] can be presented one or several at a time, at a speed adjustable by users [19]. Newly gathered empirical data on optimal frame rates for video previews can guide designers of systems providing access to video resources[5]. Additional empirical data on user performance with alternative representations and control mechanisms for temporally-dependant information is needed.

EXAMPLES FROM APPLICATIONS

The following design examples illustrate uses of previews and overviews.

NASA EOSDIS Query Preview Prototypes

NASA's Earth Observing System Data Information System (EOSDIS) allows users to retrieve earth science data from hundreds of thousands of data resources containing images, measurements, or processed data from data archive centers. Metadata is available and is used to search the archives. Standard EOSDIS metadata includes spatial coverage, time coverage, type of data, sensor type, etc. Traditional form fill-in interfaces for EOSDIS permit searches of the large holdings but zero-hit queries are a problem. It is difficult to estimate how much data is available on a given topic and what to do to increase or reduce result set sizes. Additional problems include data diversity, data complexity, network growth, a varied user base, and slow network access. Our approach to these challenges is to employ overview and preview surrogates of highly abstracted metadata that allows users to rapidly and dynamically eliminate undesired datasets. The reduced volume of the abstracted metadata allows queries to be previewed and refined locally before they are submitted over the network.

Our query preview prototype features a single-screen overview of all earth science data available from a designated information service provider (Figure 1a)[6]. The data are characterized with three high-level attributes: location, years of coverage, and general topic area. These three attributes were generally agreed to be the most salient and universal from among all the attributes typically associated with earth science data resources. For each of the three attributes, the large set of attribute values was aggregated up to around ten high-level attribute designations. The granularity of the attribute values chosen for use in the interface is deliberately crude, in order to be able to represent vast amounts of diverse scientific data in a single overview. The careful choice of appropriate attribute values and attribute value aggregations represents a significant portion of the design effort. For each attribute value in the interface, the number of available data products bearing that attribute value is shown. This gives users an overview indicating the size and extent of the available data. By clicking on combinations of attribute values, users can prune to reflect only those data items that satisfy selected attributes. In this way, users generate a preview of the result set that would be returned if a query matching the selected attributes were sent over the network. Zero hit queries are thus eliminated and users can explore the data without penalty or network delay. When the preview indicates a reasonable result set size, users submit the query, which returns the list of dataset records, or move the query to a refinement phase. In our current prototype. this consists of an additional overview phase that operates on the subset of data items from the first phase. Thus, what was used as a preview of a result set from the entire archive is now treated as an overview of a new subset of the archive. At this point, users are able to discriminate data objects using three additional attributes. This progressive refinement, where overviews lead to previews of subsets which in turn serve as overviews for more fine-grained objects illustrates how interactive overviews and previews can scale to large, complex digital libraries. An empirical study of a query preview interface demonstrated that user performance was improved for non-analytic search tasks [20].

While Furnas [8] discusses methods specifically for achieving effective view navigation, the query preview prototype meets Furnas's criteria within the analogous context of data browsing. Furnas states that high-level semantics play a dominant role in designing navigable structures, and the success of the query preview prototype hinges upon capturing attribute metadata at a carefully chosen, highly abstracted level. In addition, Furnas outlines the need to carve up an information space in a way that supports efficient navigation among distinct objects. Overlap patterns of "to-sets," the set of desired objects that may be reached by navigating across a link, must be able to "whittle down the space of alternatives in a small number of intersections." If one considers the dataset records as the primary objects of interest in the query preview prototype, then the selection of attributes is a way of navigating toward the dataset records that fit the user's need. In the overview state, before selection occurs, all datasets are represented in the count totals for each attribute value. In the preview state, users are shown the degree to which they have moved closer to their desired result by the revised count totals. The attributes as chosen carve up the information, and user selections quickly whittle down the space to a small set of objects.

Furnas also discusses the value of *residue*, referring to the degree to which a node in a navigable structure has the property of indicating what nodes are linked to it. The query preview interface has residue from all the primary objects it represents, in the count totals. In addition, the high-level attribute value bars indicate the nature and distribution of the properties of the represented objects. From this perspective, the design of preview and overview representations is essentially the task of determining what residue is to be captured, and how.

Library of Congress National Digital Library Prototypes

We have worked with a team at the Library of Congress (LC) to develop and test interface designs for LC's National Digital Library Program (NDLP) [16]. Our effort focused on the American Memory collections, which consist of historical multimedia materials: photographs, voice recordings, catalogs, manuscripts, journals, films etc. Five million items are expected to be digitized by 2000. Some will be richly described with metadata such as bibliographical records. But the volume of materials is so great that a large proportion of the individual items will not be described at all (and therefore never "hit" by traditional search). Collections of items are more likely to be described and the browsing of the undescribed items of those collections will be important.

Our designs emphasize providing previews and overviews of materials to users. The American Memory digital collections contain enormous volumes of material in many different formats and levels of descriptive metadata. This makes searching difficult and browsing more important. A preview of a collection can be samples from that collection, while a preview of a photograph can be a thumbnail. An overview of all the collections shows on one screen each of the collections with some of their common attributes. An overview of all the items in a collection shows a representation of all items, possibly with some attributes.

We developed several novel browsers to address the problems of navigating large and diverse LC collections. For the complete family of American Memory digital collections, we developed a Collections Browser applying dynamic query principles to the access of collections. For specific collections within American Memory, we developed a generic collection browser (called WebTOC), especially designed to be usable for both unprocessed and processed collections. These browsers provide overviews of sets of collections or items, and help users learn about topical coverage and the kinds of available materials. Users form concrete and realizable expectations, and can then formulate successful search queries.

The Collection Browser uses dynamic queries [2]. The interface features a visual overview of all the collections (a zoomable timeline with a bar for each collection showing its time coverage), a visual representation of a query using a collection of widgets (three menus for location, topics and format attributes) and tight coupling between all components (timeline, filters, list of collections and short descriptions) (Figure 4). Users browse the database by interacting with widgets (filters, zooming of timeline or simple cursor selection). Each change is in effect a new query, and the results are immediately and continuously shown.

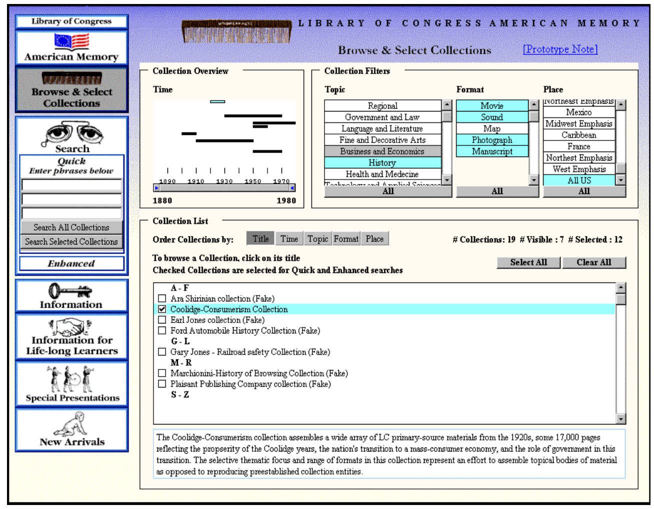

Figure 4: The NDLP Collections Browser. The overview showing the time coverage of the collections has been interactively modified to show only collections including manuscripts on the topic of business and economics.

The collection browser requires consistent metadata for all the collections. Early design work uncovered problems with the metadata. Often metadata was inconsistent across collections, or was not available for materials at the item level, making it impossible to apply this interface to entire collections at the item level. It became clear that it was important to define a set of consistent attributes at the high level of the collection if any effective browsing across the collections was to be provided. A metadata format was defined by the LC team, and some data was generated for the collections used. The timeline, collection list and menus are created automatically from this metadata.

The individual collection browser tool we developed is WebTOC (Figure 5) [11]. WebTOC generates a table of contents using two different strategies: following existing HTML links or using the underlying file system structure. Following links is appropriate for existing web sites, while using the file system is appropriate for newly digitized collections which have not yet been linked, indexed or annotated.

A two frame layout allows users to expand or contract the TOC presented in outline on the left frame, while inspecting any selected item in the right frame. As summarized in the WebTOC legend (top left of Figure 5)

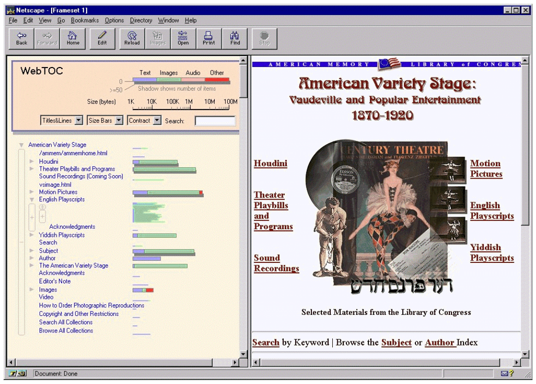

Figure 5. The WebTOC overview of the American Variety Stage Collection. The outline can be expanded to reveal more details.

individual pages are shown with individual lines while bars aggregate the files behind a link (or directory) to represent the total size of the included documents. Color is used to represent file type, length, the overall size, and the shadow below the bar is proportional to the number of documents included. The number of documents and their size is displayed in the browser status area when the user's cursor is over the bar. Figure 5's TOC shows that some parts of the site are very small (nearly empty) relative to others. It would have been impossible to learn this from the home page on the right. In addition, a small shadow suggests that only a few documents exist under a topic.

We designed and refined WebTOC to satisfy our own need to understand the huge directory of the Coolidge collection—what it contained and how it was organized. WebTOC has proven useful working with the unprocessed Coolidge collection and it will be useful to library staff preparing or reviewing new collections. It is likely to be useful for the general public as it provides an alternative paradigm for access to the materials. It is particularly helpful for exploring a collection when searching is inadequate due to the lack of metadata, finding aids, or searchable text; or because users are not able to specify their search appropriately. It allows users to discriminate at a glance large from small collections, or collections made primarily of links from sites including digital materials. The graphical overview gives an appreciation of the type of materials. Finally, it provides a way for expert users to quickly jump to objects deep in a hierarchy without necessarily following a series of explicit links on pages that each take time to load, and can be used to display results in the context of the hierarchy. WebTOC is an overview tool that can help users to browse, and librarian staff to develop and organize, the collections.

In addition we proposed a series of preview tools so that sets of homogenous materials could be reviewed at once: an image browser lets users preview thumbnails of search results or collections (Figures 6 and 7); a video browser allows users to examine key frames of motion. A page-oriented browser allows users to peruse on a single screen: surrogates, thumbnails or full views of journals, manuscripts, and other linear sets of page images.

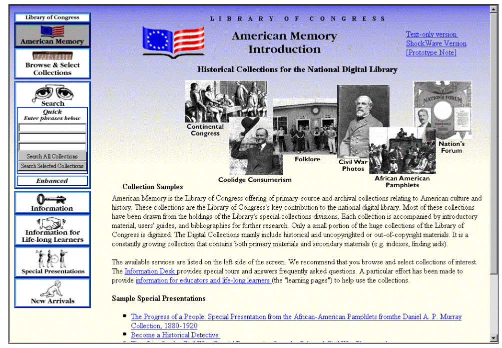

Figure 6: Opening screen for the American Memory Collections, showing animated samples of the collections.

Figure 7: With a preview showing 50 thumbnails users can browse a 200-photo collection in 4 screens.

Lifelines

LifeLines was designed to provide a rapid overview of a personal history record (Figure 8) [15]. It uses the simple metaphor of timelines, making good use of the common attribute (time) of all information objects. This is an exhaustive overview. Rich interactivity allows users to reveal more details, or jump to documents. The display can be tailored to present useful attributes via color, size or label coding.

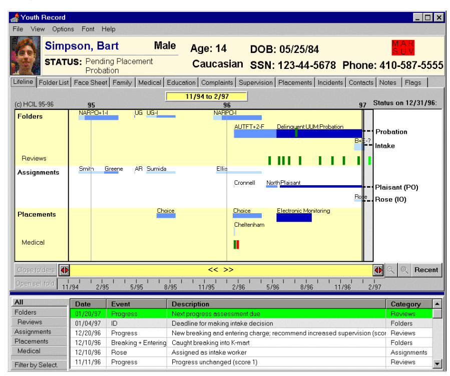

Figure 8: LifeLines provides a single screen overview of an entire case history of a Juvenile Justice record.

CONCLUSION

Interfaces for digital resources regularly fail to provide honest representation of what they include, wasting users' time and increasing their frustration with online systems. Some museum web sites give the impression of having large amounts of digitized art but in fact only hours of operation and links to other sites are available. Likewise, readers of corporate web sites often are led to believe that extensive product information and user support is available online, but following links may only lead to advertisements and a few frequently asked questions. Text summaries for Web search results remain too cryptic, indicating little of significance about the item previewed, nor its relation to the query as expressed.

Designers of digital libraries and web sites should make the nature of their resources more apparent to users. The use of well-designed previews and overviews can help to achieve this goal. Many home pages offer site maps, but these can be elaborated with more attribute information, including the type, size, and media of available information. Overviews should communicate the full range of several or all data attributes common across the site.

Previews should be available at a high level within a site so users get a taste of what is to come early in their visit. Exemplary samples should be selected and brought to the surface. The use of preview and overview concepts in visual information representations is clearly beneficial to the design of interactive information systems. This paper has outlined an initial set of design parameters and discussed some specific, successful designs. The good design of previews and overviews results from a careful analysis of the information to be represented, and the ability to determine the high-level semantics of the information most relevant to capturing its essence, and making it quickly understandable and navigable for the user. The design and testing of previews and overviews for additional types of systems and information resources will refine the theory and improve practice.

ACKNOWLEDGEMENTS

The preparation of this report was supported in part by the Library of Congress and NASA. In addition, this material is based in part upon work supported by the National Science Foundation under Grant No. IRI9615534.

REFERENCES

[1]Ahlberg, C. and S. Truve. 1995. Exploring terra incognita in the design space of query devices. In *Proc. of Engineering for Human Computer Interaction, EHCI '95*. North Holland.

[2]Ahlberg, C. and Shneiderman, B.. 1994. Visual Information Seeking: Tight coupling of dynamic query filters with starfield displays , In *Proc. of ACM CHI94*, 313-317.

[3]Bederson, B. & Hollan, J. 1994. Pad++: A zooming graphical interface for exploring alternative interface physics. In *Proc. of ACM UIST '94*. 17-26.

[4]Borko, H. & Bernier, C. 1975. Abstracting concepts and methods. NY: Academic Press.

[5]Ding, W. and Marchionini, G. 1997. A Study on Video Browsing Strategies. Tech. Report CS-TR-3790, UMIACS-TR-97-40, CLIS-TR-97-06, University of Maryland.

[6]Doan, K., Plaisant, C., Shneiderman, B. 1995. Query previews in networked information systems. In *Proc. of the Third Forum on Research and Technology Advances in Digital Libraries, ADL '96*, IEEE CS Press, 120-129.

[7]Elliott, E. 1993. Watch, grab, arrange, see: Thinking with motion images via streams and collages. MSVS Thesis Document. MIT Media Lab: Cambridge, MA.

[8]Furnas, G. W. 1997. Effective view naviation. In *Proc of ACM CHI97*, 367-374. ACM : New York.

[9]Heilprin, L. 1985. Paramorphism versus homomorphism in information science. In L. Heilprin, (Ed.), Toward foundations of information science. White Plains, NY: Knowledge Inducsty Pub. pp 115-136.

[10]Marchionini, G. 1995. *Information Seeking in Electronic Environments*. Cambridge University Press.

[11]Nation, D.A., Plaisant, C., Marchionini, G., and Komlodi, A. 1997. Visualizing web sites using a hierarchical table of contents browser: WebTOC. In *Proc. of the 3rd Conference on Human Factors and the Web*.

[12]North, C., Shneiderman, B. and Plaisant, C. 1996. User Controlled Overviews of an Image Library: A Case Study of the Visible Human. In *Proc. ACM Digital Libraries '96*, ACM: New York.

[13]North, C. and Shneiderman, B. 1997. A Taxonomy of Multiple Window Cooridation. In preparation.

[14]Plaisant, C., Carr, D., Shneiderman, B. 1994. Image browsers: Taxonomy, guidelines, and informal specifications, IEEE Software, March 1995, 21-32.

[15]Plaisant, C., Milash, B., Rose, A., Widoff, S., Shneiderman, B 1995. Life Lines: Visualizing personal histories, *In Proc. ACM CHI96*. ACM: New York.

[16]Plaisant, C., Marchionini, G., Bruns, T., Komlodi, A., Campbell, L. (1997) Bringing Treasures to the Surface: Iterative design for the Library of Congress National Digital Library Program. *Proceedings of CHI '97,* March 1997, ACM New York. 518-525.

[17]Rao, Ramano, and S.K. Card. 1994. The Table Lens: Merging graphical and symbolic representations in an interactive focus+context visualization for tabular information. In *Proc. of ACM CHI94*, 318-322.

[18]Shneiderman, B. 1996. The eyes have it: A task by data type taxonomy for information visualizations. In *Proc. of 1996 IEEE Symposium on Visual Languages*, 336-343.

[19]Slaughter, L., Shneiderman, B., & Marchionini, G. 1997. Comprehension and object recognition capabilities for presentations of simultaneous video key frame surrogates. Research and Advanced Technology for Digital Libraries, Proceedings of the First European Conference, ECDL '97. Berlin: Springer, 39-54.

[20]Tanin, E., Lotem, A., & Haddadin, I. Evaluation of Query Previews: User Preference and Performance. Manuscript: www.otal.umd.edu/SHORE

[21]Teodosio, L. & Bender,. W. 1993. Salient video stills: Content and context preserved. In *Proc. of ACM Multimedia 93*, ACM: New York, 39-46.

[22]Tweedie, L. 1997. Characterizing interactive externalizations. In *Proc. of CHI97*, 375-382. ACM : New York.

[23]Wactlar, H., Kanade, T., Smith, M., Stevens, S. 1996. Intelligent access to digital video: Infomedia Project. IEEE Computer, May, 46-52.