

# Guided Paths through Web-Based Collections: Design, Experiences and Adaptations

Frank M. Shipman III, Richard Furuta, Donald Brenner, Chung-Chi Chung, Hao-wei Hsieh

Department of Computer Science & Center for the Study of Digital Libraries  
Texas A&M University  
College Station, TX 77843-3112  
(409) 862 - 3216  
E-mail: {shipman, furuta}@cs.tamu.edu

## ABSTRACT

Digital libraries need to facilitate the use of digital information in a variety of settings. One approach to making information useful is to enable its application to situations unanticipated by the original author. Walden's Paths is designed to enable authors to collect, organize, and annotate information from on-line collections for presentation to their readers. Experiences with the use of Walden's Paths in high-school classrooms have identified four needs/issues: (1) better support for the gradual authoring of paths by teachers, (2) support for student authoring of paths including the ability for students to collaborate on paths, (3) more obvious distinction between content of the original source materials and that added by the path author, and (4) support for maintaining paths over an evolving set of source documents. These observed needs have driven the development of new versions of Walden's Paths. Additionally, the experiences with path authoring have led to a conceptualization of meta-documents, documents whose components include complete documents, as a general domain where issues of collaboration, intellectual property, and maintenance are decidedly different from traditional document publication.

**KEYWORDS:** Computers and education, Meta-documents, Guided tours, Digital libraries, Directed paths, World-Wide Web, Walden's Paths, Presentation authoring, Spatial hypertext, VIKI.

## 1. INTRODUCTION

Digital libraries provide access to a wide variety of information resources. Indeed, due to the volume of materials and variety of media and content types, much of the work around digital libraries has emphasized the storage and retrieval of this information. But traditional libraries do more than provide access to information, they provide a setting in which information may be put to use — it is this less-considered role of a digital library, as a facilitator for the *use* of digital information, that is the focus of this paper.

Use of information, paper-based or digital, occurs in a variety of contexts. Sometimes users have specific information needs, such as looking for the answer to a specific question. Other times, users want to browse information in order to get a “feel” for the content area or collection. Both of these examples (as described)

are solitary endeavors — one person working with the information source towards their own goal. Another class of digital library use is as an artifact around which to build communication and collaboration. This includes using the library contents as the basis for shared annotation (Marshall, 1997), argumentation (Reeves, 1993), and discussion (Robertson, Jitan, & Reese, 1997), as well as selecting elements of the collection as being of relevance to others (Resnick & Varian, 1997). A social use of information that combines selection and annotation is the creation of presentations.

Creating presentations is a common activity. Teachers, scholars, and many business people spend a significant part of their time producing materials they will use to disseminate information, ideas, and views. Different types of presentations rely on background materials taken from different sources. Teachers traditionally gather materials from textbooks and curriculum guides. Scholars reflect on published results as well as on the outcomes of their own research. Business people may use market analyses, trade publications, and internal reports as the basis for their presentations.

The common element of all these characterizations is that much of the information underlying the presentations is derived from paper-based sources. But this is changing. More and more newspapers, magazines, and academic and trade publications are becoming available on-line. Additionally, digital libraries provide access to new information types, like geographic information system data (Kacmar et al., 1995), and collections of physical artifacts or specimens (Schneider et al., 1998). These on-line resources frequently include new facilities that make them better suited for creating information-rich presentations: for example, the external information is readily accessible through full-text search and alternative indexing schemes, and document components are easy to reuse, given their manipulable electronic representations.

Information services like Yahoo and OpenText are examples of the growing number of businesses providing better on-line access to information. Professionally generated and edited information is becoming available as the World-Wide Web becomes more commercialized. Many newspapers, magazines, and edited academic journals now have on-line subscription services. Of course, the Web is also a source of unedited opinions, conjecture, and gossip. How does this ready availability of external sources affect the needs of and possibilities for presentation authors?

Most presentation software, such as Microsoft PowerPoint, Aldus Persuasion, or Adobe Premiere, exists as stand-alone applications. This type of software assumes that an author creates a self-contained presentation in which the content is produced strictly in service of the presentation itself. But the growth in on-line information resources suggests the development of new presentation genres, e.g., forms that draw on materials from external sources and reuse content in new, unanticipated contexts.

One alternative model for presentations that we have been investigating generalizes Zellweger's "directed paths" (Zellweger, 1987; Zellweger, 1988; Zellweger, 1989) and Trigg's "guided tours" (Trigg, 1988; Halasz, Moran, & Trigg, 1987). As originally defined, a directed path provided the means for guiding a reader's traversal along a path of components extracted from a set of documents or an existing hypertext network. The ordering of components on the path is not constrained by the structure of the source documents — in other words, the components do not have to follow the temporal orderings of the source. The directed path defines a meta-structure that is layered on top of the underlying documents' existing structures.

This work is also reminiscent of Vannevar Bush's vision of a scholar blazing trails through reference materials in a memex (Bush, 1945). The trails that Bush describes are coherent linear paths through collected materials, with possible digressive side trips. Rather than emphasizing the expressiveness of the path mechanism as is the focus of Zellweger's and Trigg's subsequent work, Bush describes a trail as analogous to gathering together physical items from widely separated sources; having simple mechanisms for flipping through the materials and commenting on them; and being able to transmit this annotated linear trail to a colleague.

Walden's Paths, our variation of a directed path facility, supports the presentation of an annotated linear path of Web-based information by teachers for their students (Shipman et al., 1996; Furuta et al., 1997). In these presentations, the teacher creates a list of Web pages and associates annotations with them. By adding some text or other annotations to the content of a Web page, the teacher may provide a rhetorical structure to the path as a whole, create transitions to fill in gaps between pages, and emphasize particular aspects of the materials.

The goal of Walden's Paths is to allow teachers to incorporate Web-based information into their curriculum. Creating a Web-derived presentation for students is different than creating a presentation for a conference or other professional presentation. Most notably, educational paths typically take information authored by others and present it within a new context. Authoring such a path is then a combination of locating relevant Web sites, selecting which materials to include, and creating the context in which the materials will be presented to the students.

In this paper we will first discuss Walden's Paths and some of our experiences with its use by high-school teachers and students. We use these experiences to motivate new development and educational pedagogy. The next two sections provide a description of the Walden's Paths prototype that was placed in schools and the selection and training of the teachers and classrooms, respectively. Following this is a discussion of four high level issues that have become more apparent through our experiences and how these experiences

have resulted in new prototype development. Finally, we discuss what we think of as the emerging area of meta-documents and our conclusions.

## 2. WALDEN'S PATHS

To take advantage of the variety of information and software that makes up the Web, a guided path mechanism must work with the common Web browsers and Web servers. We adopted a strategy of developing a “Path Server” that mediates between a student’s browser and the source material’s server, and a “Path Authoring Tool” that supports teachers and media specialists during the creation of paths. The Path Authoring Tool stores paths in the “Path Database”, which “publishes” completed paths for access through the Path Server. The Path Database stores information specific to the path — the order of pages in the path, annotations, and the URL for the original information. Figure 1 is a diagram of this architecture.

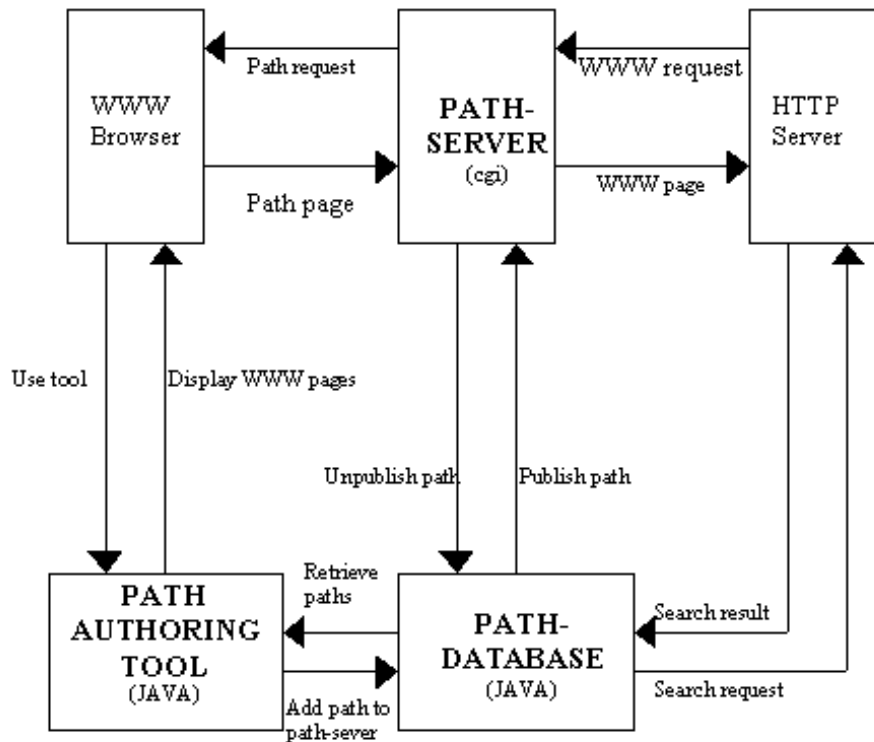


Figure 1: Communication between Web browsers, HTTP servers, and Walden's Paths Path Server, Path Database and Path Authoring Tool.

Early work on directed paths situations where the source document authors and the path authors are part of the same community. Walden’s Paths emphasizes the distributed authorship found in the context of the Web. Other research on guided path mechanisms for the Web have used a similar strategy of creating an intermediary between original source documents and the reader (Nicol, Smeaton, & Slater, 1995; Greer & Philip, 1997).

## 2.1 Path Server

The Path Server acts as an intermediary between the student's browser and the original source information on the Web. By leaving the content of the pages at the source Web server and retrieving it when a Web browser requests a path page, the Path Server can take advantage of the evolution of the preexisting Web structure and content. Once a page is retrieved, the Path Server adds path controls and the path-specific annotations, then returns the resulting page to the Web browser.

Figure 2 shows the view of a high-school art student browsing a path about Duane Hanson, a modern sculptor. At the top of the Web browser are the control flow buttons for moving forward and backward through the path. The numbers below the title of the path show the length of the path and indicate which page is currently being displayed. The introductory text or annotation provided by the teacher is shown below the page numbers and above the three thick horizontal bars separating the original source material from the information added by Walden's Paths.

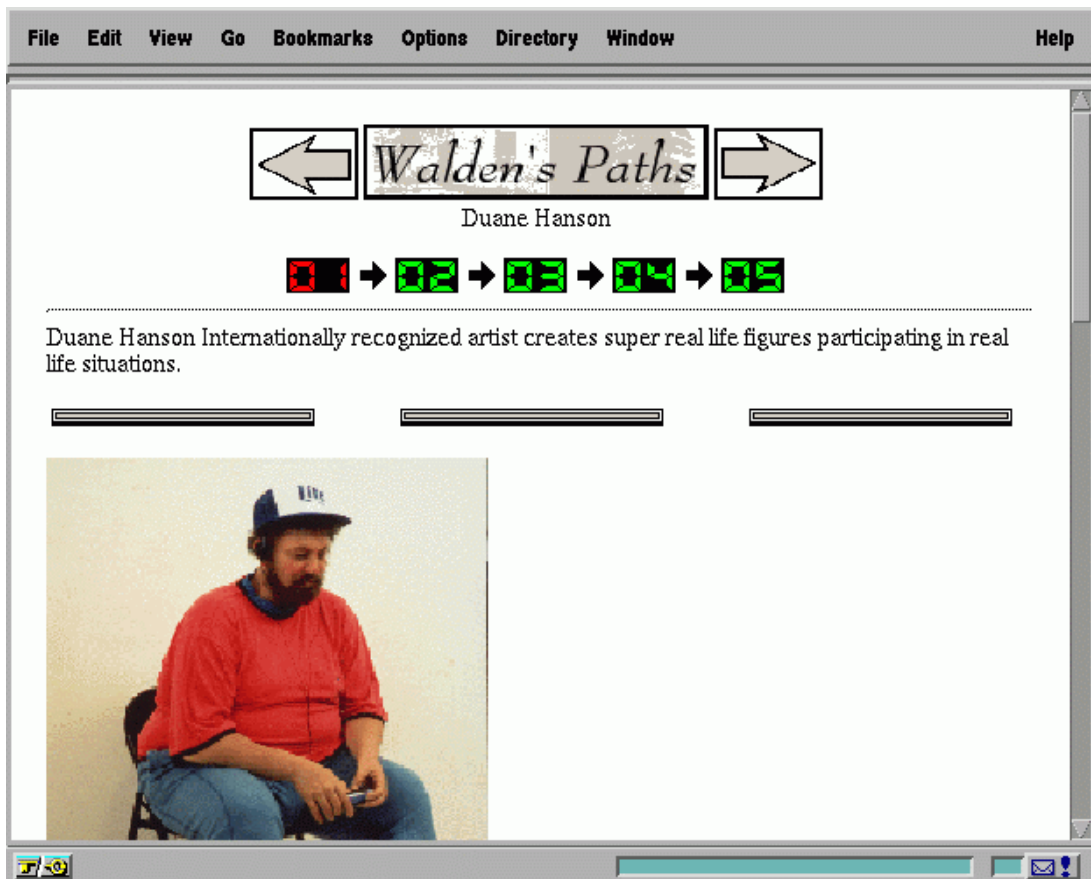


Figure 2: The path page contains control flow buttons, annotation, and the original source materials.

While reading a path, students may take any of the links embedded in the original material to explore based on their own interests. No matter how many links are taken, there is a “return to path” button available, as shown in Figure 3, that returns the student to the point on the path where the side-trip began.

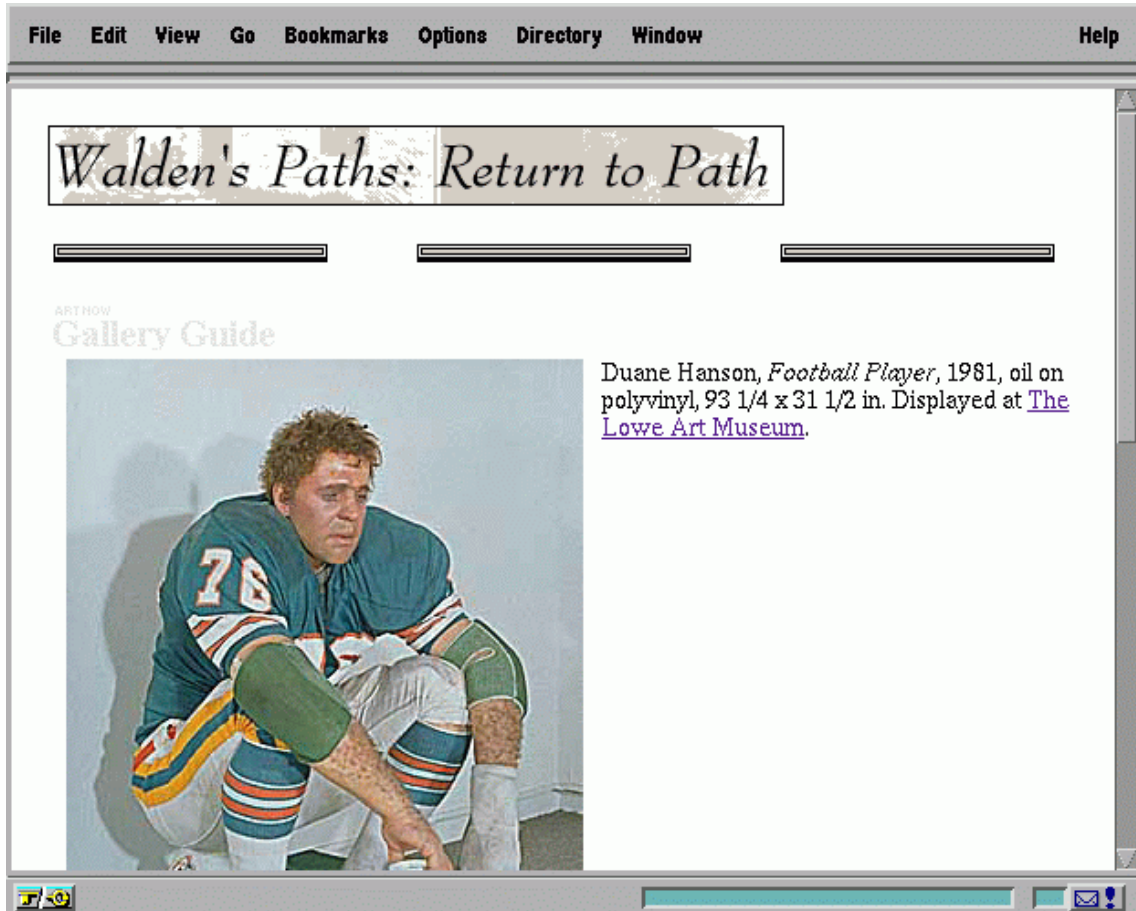


Figure 3: When browsing off the path there is always a quick way to return to the path.

In addition to piecing together the original source material with the annotations and other path-specific content, the Path Server provides a variety of caching functions. The combination of slow network connections at schools and unpredictable access to sites on the Internet leads to the desire to store pages in anticipation of their later access. Because each path identifies a series of pages that are likely to be accessed in order, there are a variety of potential caching strategies.

*Traditional caching*, like that done by Netscape and some other browsers, simply stores Web documents as they are retrieved from the Internet. The advantage of having this occur at the Path Server over at the Web browser is that with the Path Server cache each page is only retrieved once (and stored once) for everyone using the Path Server. *Pre-caching* is a caching strategy that caches pages before they are ever accessed.

The Path Server can use its record of the URLs that are likely to be accessed during a session — in the case of guided paths, it is natural to assume that once a student has begun traversing a path that all the pages on the path are apt to be viewed later.

As with Web browser caches any choice of caching strategy inevitably leads to issues concerning versioning, cache size, and cache time-out policies. In some cases there is no clear solution — different use environments favor incompatible caching policies. Not all information in a guided path should have the same time-out policy — static information may sometimes be an asset. A teacher may want to be sure that the material on this page will not change before the students access it. On the other hand, if the page is a real-time display, such as the current local weather map, presenting an old version may not be acceptable. This implies that each page should have its own policy regarding caching. There are at least four possibilities: (1) cache now and don't remove (sets the exact information for the path); (2) cache during execution and don't remove (ensures all students see the same up-to-date version); (3) cache during execution and remove as needed (implements the traditional caching scheme); and (4) don't ever cache (retrieves dynamic information each time).

## **2.2 Path Authoring Tool**

In the course of our early work on the Walden's Paths mechanism we have gained experience with authoring Web-based presentations. This experience shows that path authoring is a complex task which includes:

- locating promising Web sites,
- browsing and evaluating materials at these sites,
- selecting information elements for use in the path,
- developing an outline for the presentation,
- placing pages within the sections of the path, and
- writing the introductory text and annotating the pages.

Not all of these activities occur during every authoring task; nor is the list above intended to imply a strict sequence of actions. Rather, an authoring process may require many iterations through these commonly identifiable subtasks. The Path Authoring Tool (PAT) provides a single, integrated, straightforward interface to do just this.

The prototype Path Authoring Tool (shown in Figure 4) consists of four main areas. The top portion displays three lists: (1) the pages returned from the most recent Web search using an external search engine such as Alta Vista or Opentext (top left list), (2) a "holding area" for pages set aside for later use but not

included in the path (lower left list), and (3) the pages in the path (right list). The pages of the path can be reordered via simple control buttons and the teacher adds annotations to the path pages through a dialog window. The bottom portion of the authoring tool displays basic information about any page selected from the above lists, such as its title, URL, and annotation. Double-clicking on an element of any list will cause the page to be displayed in a browser.

When the teacher is finished working on a path and wants to install it at the Path Server, the Path Authoring Tool provides an interface to the Path Database for saving, loading, and publishing their paths. This interface includes a login screen for the teacher to ensure that the server can verify the teacher's authorization to access, save, or publish a path.

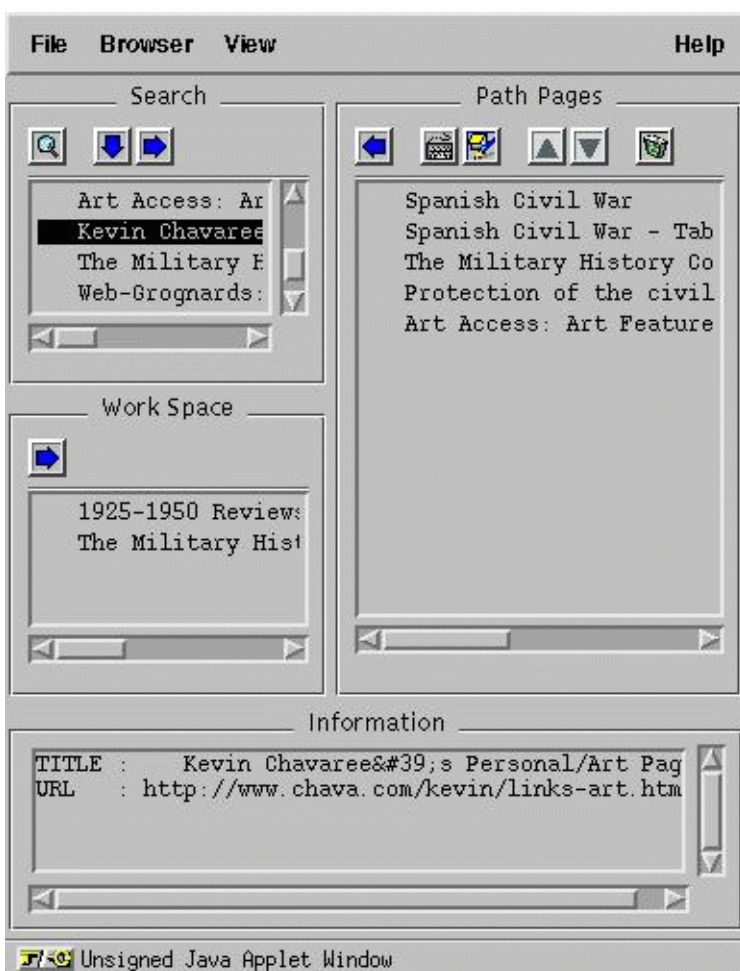


Figure 4: The Path Authoring Tool.

### 3. TEACHER SELECTION AND TRAINING

Walden's Paths has been used by eight high school art and music teachers. These teachers were initially introduced to the concept of Walden's Paths, along with a variety of other educational technologies, at a

“technology fair” held at their schools.

Teachers, based on class topics and their interests, were matched with educational technologies and were provided in-depth demonstrations along with question and answer sessions during a second visit to the schools. At this point teachers began to provide feedback on the technology to help focus new development and suggest changes. It was also at this stage that one art teacher began to author paths, collecting pages from on-line museums for use in her classroom.

Once Walden’s Paths was ready for distribution, a training workshop was held that seven of the eight teachers were able to attend. This workshop provided the teachers with a chance to use the software with developer supervision and some opportunities to learn basic skills necessary for path authoring. At this workshop each teacher authored at least one path, with many of the teachers authoring two or more.

After the workshop, Walden’s Paths was available for use in these eight classrooms. Due to network problems and timing the system ended up in use in only four of the eight classrooms. Students in these classes did follow paths that their teachers had authored but there were other modes of use that developed as well. The following section will discuss our experiences with use and how these experiences have driven several new lines of development.

#### **4. EXPERIENCES AND REDESIGN**

Our experiences with the use of Walden’s Paths in high school classrooms have raised technological, social, legal and practical issues. Many of these issues are not specific to the particular instantiation of Walden’s Paths but pose more general concerns for the use of Web-based materials in classrooms.

The next four sections report on our experiences in detail and describe new development and practices that have developed as a result. The first section describes experiences with teachers authoring paths containing Web materials. The second section discusses student interactions with Walden’s Paths. The third and fourth sections explore issues of intellectual property and path maintenance, respectively.

##### **4.1 Teachers Authoring Paths**

In the original conception of Walden’s Paths, teachers would use a Path Authoring Tool (shown in Figure 4) to search for Web pages, collect and organize them into paths, and annotate the paths to provide context for the student readers. Our experiences point to two serious difficulties for this process: first, many teachers have very limited experience with the Web and more generally with information location strategies; second, elementary and secondary school teachers frequently have five hours of classes a day and only one hour at school to prepare for classes.

The first of these problems became apparent in the workshop to train the teachers on the use of Walden's Paths. We expected to spend some time working with the teachers on the use of search engines and indices, such as AltaVista and Yahoo, to help them develop information location skills. Due to many teachers not having used a Web browser before, much of the workshop time ended up devoted to explaining the general hypertext concepts, such as links and navigation history, and the specifics of how to use Netscape. Once basic browsing skills were acquired, teachers began to learn strategies for information location using search engines including issues of terminology and broadening or narrowing queries.

A positive result of these training sessions was that once teachers became accustomed with Netscape and the search engines, the idea of guided paths and the use of Walden's Paths was very natural and questions primarily focussed on the materials being collected or the Web sites being visited.

The second problem, that teachers have very limited time to prepare materials for class, is an overarching concern for any type of system that asks teachers to take on new class preparation tasks. Because path authoring could only be done at school, where the teachers had network connections with which to search for materials on the Web, only a few teachers found time during the course of the semester to author new paths. The paths which teachers authored tended either to present materials from one or two Web sites or to provide students with a few top-level pages of interesting sites for the students own exploration. One music teacher described Walden's Paths as the "Mother of all Bookmarks", adding that her assistants had found hundreds of useful pages that could be used to write paths for many types of computer music projects and commenting positively that "As Web sites and Walden's Paths are developed by teachers everywhere, I can see that this could easily become my full time job" (CAETI, 1997). But teachers already have a full time job.

A specific problem with the Path Authoring Tool was its expectation of a short-term authoring process. Since teachers had such small chunks of time to look for materials, a single path authoring process could take the form of several different uses of the tool. Because the tool was designed to support a streamlined search, select, organize, and annotate process, not much context of prior searches and effort was kept so teachers had to reacquaint themselves with Web sites on the current topic each time they started the tool.

To support a longer-term, more gradual path authoring process we have been investigating the use of the VIKI spatial hypertext system as a path authoring environment (Marshall & Shipman, 1995; Shipman, Furuta, & Marshall, 1997). VIKI provides a hierarchy of two-dimensional workspaces where a teacher can collect materials from the Web by issuing queries, and organize and collect the materials through color coding and spatially arranging Web documents.

The path authoring process with VIKI follows a process of searching for pages, evaluating the search results, organizing useful pages into the presentation order, and annotating the pages. Figure 5a shows the VIKI interface for searching the Web, in this case a science teacher is putting together a path on antimatter. The results are placed in a VIKI collection, as shown in Figure 5b. Each query's results initially appear in a newly created collection.

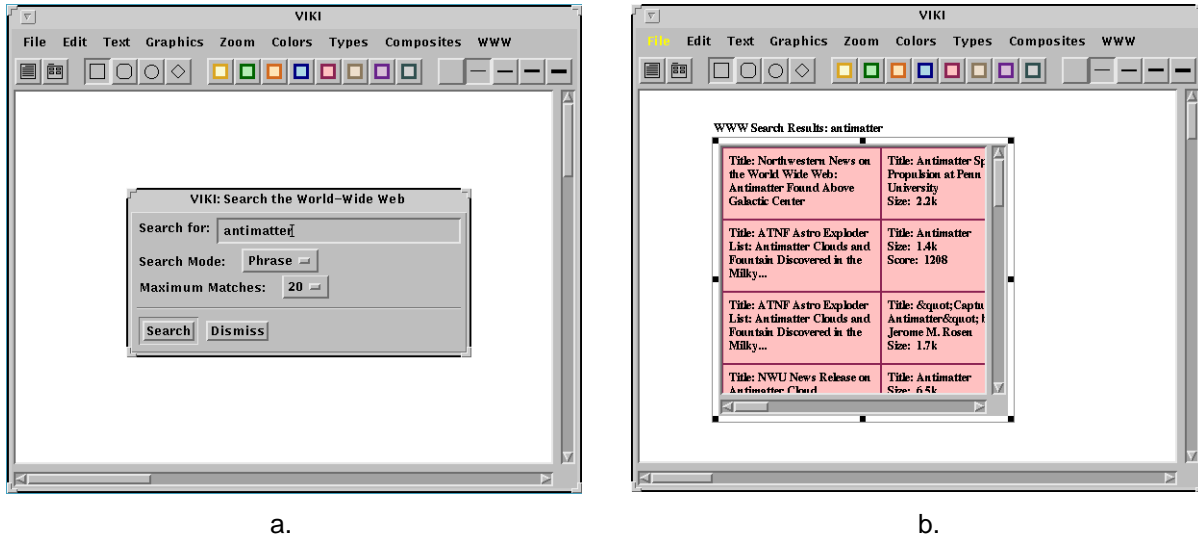


Figure 5: (a.) Entering a Web query in VIKI. (b.) Results of Web query appear in new VIKI collection.

The teacher browses the Web pages returned by double-clicking on the visual symbol for each of the pages. By coloring and moving these symbols, the teacher can move from early categorization, in this case identifying which pages provide basic information about antimatter, which discuss antimatter detection, and which are about antimatter production and use.

After organizing and annotating a set of pages into the structure shown in the right of Figure 4, the teacher uses the “Create Presentation” dialog (shown on the left in Figure 6) to publish the path for the students. The publication dialog asks for a path title and whether only Web pages should be included in the path. Symbols like the headers in the structure could be turned pages with only annotation in the final path, if the teacher wants.

The advantage of teachers using VIKI is that each time they go to author they are provided their personal visual workspace of Web documents. The visual and spatial cues they have developed within their workspace allows teachers to more quickly remember where they were in the authoring process and what information they had already located.

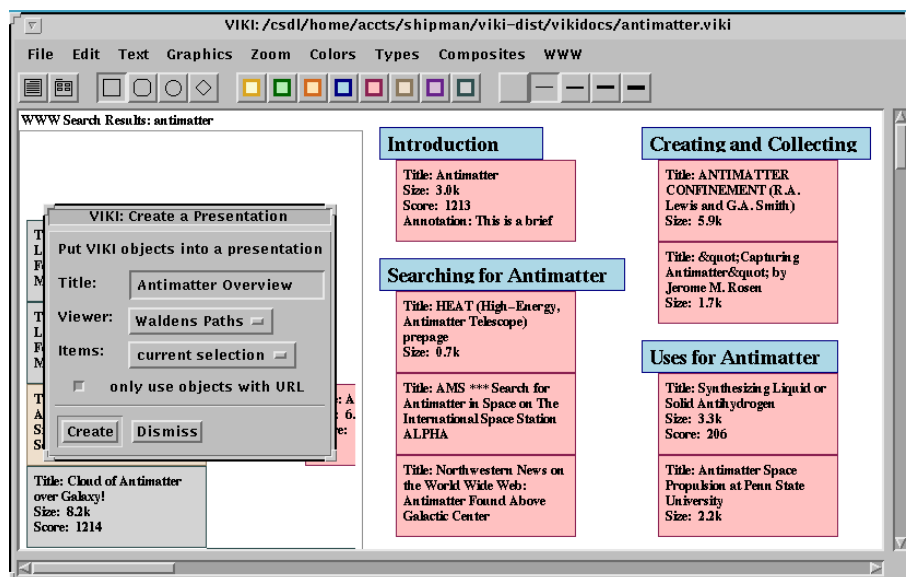


Figure 6: Once organized and annotated, the teacher creates a path with the collected materials.

#### 4.2 Students and Walden's Paths

The paths authored by teachers during the workshops and during the normal course of the semester were read by students as part of their in-class efforts. Depending on the particular emphasis of the class and the types of materials collected by the teacher, students made use of paths in a variety of ways.

In a computer music class emphasizing composition skills a path of Web pages on MIDI (a computer music protocol) were used as reference materials. One art teacher used Walden's Paths to collect images supplementing the old, and too expensive to be replaced, textbook with images from modern artists and images not included in the textbook. In this case, paths were augmenting the existing curriculum and classroom activity by providing access to new materials.

Paths were not just used as information resources. During the art history phase of one art class, paths with images of paintings and sculptures were used in the place of traditional slide show that had previously been part of the curriculum—the students filling out a form describing each art work that they would have to turn in at the end of class. In this case, the art teacher commented on how useful it was that each student could allocate their time answering the questions about the images independently—something not possible in prior years. This is an example of how paths could be used to modify existing classroom practices.

The above uses of Walden's Paths are close to those that were anticipated—the teacher authoring paths for the students to read and explore later. During the course of the semester, a number of classrooms began alternative uses of Walden's Paths by assigning authoring roles to students. In one music class, the students

divided up into groups that each collected information on a particular topic for a path. In another music classroom, the students each were asked to submit one or two pages to a path authored by the whole class. In this case, the students were asked to provide rationale why their particular suggestion should be included. These examples show how teachers used assignments involving path authoring in place of essay writing with the goal of students learning information location and evaluation strategies.

Student authoring of paths was not limited to projects to be turned in to the teacher. In one art classroom, a student who was ahead of schedule in class was given the extra-credit task of authoring paths for the other students. The teacher in this case was allowing the advanced student to take on extra responsibilities that could also help the other students. Students, perhaps being more comfortable with computers and the Web, seemed to pick up path authoring quite quickly. When one student was asked about how he had learned the software, he said the teacher had just handed him the Walden's Paths user's manual.

The variety of student activities surrounding Walden's Paths is largely due to paths being a general purpose medium for collecting and presenting material, and not being tied to any particular educational pedagogy. For a more in-depth discussion of the uses of Walden's Paths see (Shipman et al., 1999). These experiences have led us to develop new functionality within the Path Server to promote more student interaction. Some of these extensions promote specific educational practices that appeared naturally in the classroom.

The observations reported here, coupled with those reported on previously (Furuta et al., 1997), emphasize the central nature of student-student collaboration in the educational environments in which Walden's Paths is being studied. During exploration of information, students draw their neighbors' attention to interesting resources. More formally, teachers assign group research projects, with report and results to be created collaboratively.

We experimented with issues of supporting collaboration in Walden's Paths by extending the Path Server to retain information about the identity of and location of students traversing paths. By maintaining state information for path reading sessions, such as what pages were previously visited, the Path Server can provide context-dependent information to the reader. In addition, students were asked to "login" before browsing paths, thereby allowing implementation of status-aware displays such as one that allows students in a class to jump directly to a page "tagged" by one of their classmates.

Figure 7, in the area labeled "Tag Menu", shows the preliminary interface used by students to tag a page

for their fellow students and to see the list of pages tagged by other students.



Figure 7: Tagging interface appended to path page.

The primary application of these architecture enhancements was in implementation of a new environment in which students, either individually or in groups, could make modifications and extensions to an existing path, creating a new version of the path. Figure 7 also shows how students browsing a path can enter their team name in order to edit the path in the Path Collaboration Tool.

The Collaboration Tool allows multiple students to modify the path simultaneously, and a new collaboration server (working in conjunction with the traditional Walden's Paths Path Server) was added to the architecture to manage independent versions (i.e., separate groups working on modifications to the same original path) and to linearize editing operations. Figure 8 shows the Collaboration Tool, which allows the student to see the current list of pages and annotations in the group's path.



Figure 8: The collaboration tool lets groups of students work together to author a path.

Modification operations, accessed through a form added to the original pages, include the ability to reorder path entries, to add annotations (which augment any annotations provided by the teacher in the original path), and to add new pages to the path. Once completed the new path could then be submitted to the teacher for evaluation.

These enhancements also involved creation of management tools for teachers. Teachers are, for example, allowed to specify whether paths submitted for evaluation are to be “locked”—further changes are prohibited unless the teacher permits. Teachers also can attach general comments and grades to paths, return a path to a student group for revision and resubmission, or can make the path available for general use by the other students.

Informal evaluation of the prototyped enhancements by a group of thirteen teachers and teacher trainees was generally positive but, as would be expected, pointed out a number of specific ways in which the environment could be improved. We are currently studying the results of this evaluation to determine the ways in which the prototyped enhancements should be incorporated into the generally available implementation.

### 4.3 Intellectual Property

Over the course of the Walden's Paths project issues regarding intellectual property have been raised by both the authors of the paths and by the authors of the materials included in paths. Path authors have asked for more control than complete inclusion of a document into their paths—they want the ability to select parts of documents for inclusion and the ability to turn on and off specific links within the documents. Additionally, there have been requests for changing the visual appearance of the included material, such as changing fonts, or text and background coloring schemes. However these requests conflict with the interests and rights of a document's authors who often desire to control not only the content of the document and its reuse but also its presentation. Indeed, as we discuss in (Furuta et al., 1997), control of the presentation *context* of use is often of importance to the document's author.

Undoubtedly, the first, and primary concern of document authors is one of correct attribution of authorship. Our existing design has gone through several iterations as we have tried to emphasize that the original material included in the path is a complete and separable document. In the current version, we have provided graphically-distinctive thick bars delimiting the page of included materials and included a printed reference to the source materials' URL. Figure 9 shows the separation of the control buttons and annotation from the original Web page material in the original version.

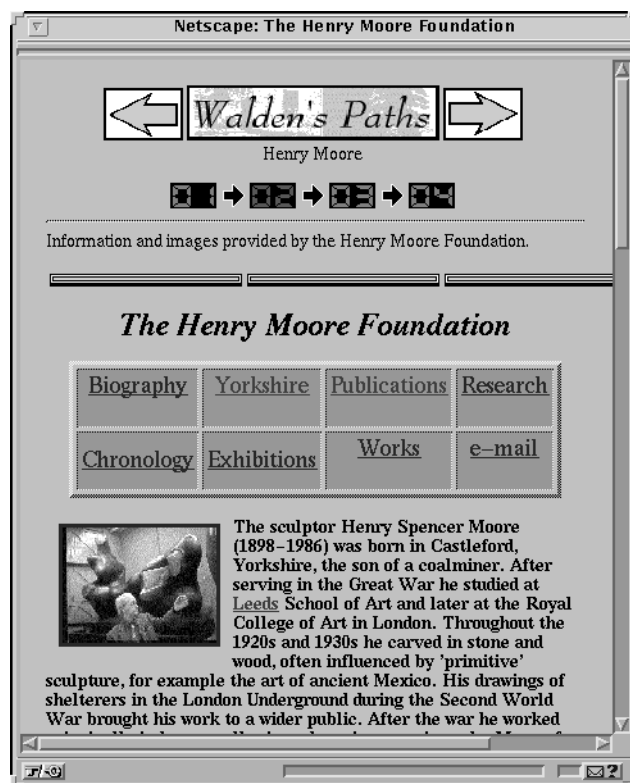


Figure 9: Three thick bars separate the Walden's Paths header from the original Web page.

Even with this level of attribution to the original source we have received comments from authors who have found their Web pages being used in Walden's Paths bringing up the issue of intellectual property and providing permission or suggestions. For example, the following message was sent regarding a path authored by a group of students:

“While searching the Internet for links to my Duke Ellington page, I discovered today that parts of Rude Interlude have been included in pages published by your project. I am honored, but also perturbed by some questions that must arise:

1. My pages are copyrighted. You are republishing them without my permission. I hereby grant you permission to do so, but I retain the right to withdraw that permission at any time. ...”

In essence the students' work was the electronic analog of a term paper. There would be no confusion or difficulty about authorship of material if the term report had been presented in paper form; however in the electronic presentation, the issues of ownership clearly are blurred in the reader's mind. As this message implies, there is still not a strong enough distinction between material authored by the path authors and the original material.

We are investigating two approaches towards making the separation between original and added material more obvious. In the first case, we plan to make use of HTML frames to put all the material from Walden's Paths in one frame and the original source document in another. This is an approach taken by many Web sites which want to include information provided by other sites while still providing context for that information. For example, when browsing from NASDAQ's Web page for a particular company's stock to information provided by the company, the NASDAQ banner and access to information about the company from other sources are provided in surrounding NASDAQ-provided frames.

This design still has the potential for confusing the reader about the proper attribution of materials—the increasing commonality of frame-based designs tends to reinforce a reader's perception that *all* material contained within the browser's window falls under the same administrative control. Consequently, we are also working on a design, to be discussed in the next section, that further separates Walden's Path's material into a separate window.

#### **4.3.1 Implementing paths in multiple windows.**

We have developed a multi-window version of Walden's Paths to try out techniques to address two issues. The first is the one just discussed; the need to increase separation between path controls and path content to counteract the tendency of readers to become confused about the source and ownership of materials contained on a path. The second is to consider issues surrounding path stops that logically reference

multiple content elements. The prototype, implemented in Java and JavaScript, provides the path control buttons and annotations in one window, separated from the original materials.



Figure 10: Three steps in a path reading session using the multi-window path server.

Figure 10, left, shows the initial stop in a path. Note that the controls and annotation are contained in a separate window (back of the two) and are retained while off-path browsing takes place, as in Figure 10, center. The controls in the dark bar on each window permit the reader to bring another window to the front and, in the case of off-path browsing, to return immediately to the path's page. As in the original version of Walden's Paths, stepping along to the next path stop updates controls, annotation, and content; see Figure 10, right.

Often a stop in a path logically involves multiple Web pages—for example, a description of parallel research activities carried out by a group. In the current implementation of Walden's Paths, such stops would be linearized, creating several actual stops corresponding to the single logical item. Ladd, et al. (1995), have discussed the usefulness of applying multi-headed multi-tailed links in the context of the World-Wide Web. We also incorporated these links into our experimental path structure, permitting stops with multiple content elements, such as that shown in Figure 11.



Figure 11: A path stop involving multiple content elements.

Use of multiple windows introduces practical design issues in the Web environment. In particular are issues of initial window placement, which are resolved in different ways from window system to window system, effective use of screen real estate, and, in the Java/JavaScript environment, interaction with implementation security restrictions. Consequently we are continuing to develop, evaluate, and compare the two separate approaches of within-window (e.g., frame-based) and separated-window path presentations.

Our efforts to promote easier attribution of authorship of source material and path material emphasize the goal that students distinguish material authored by their teacher or classmates from material collected from the Web. The mechanisms described use frames or multiple windows to minimize any confusion between the author of the path and the authors of the source material. The growth of public information available in topical digital libraries, from government and university agencies, and by individual Web authors, means there is plenty of information that can be used by teachers and students to author paths.

#### **4.4 Paths and Maintenance**

A growing concern for the continued use of Walden's Paths is the issue of maintaining paths authored over a continually evolving set of source documents. Some of the paths authored within Walden's Paths are now almost two years old—a long time in the time-frame of Internet materials—and many of the

documents originally included have moved, changed, or are no longer available. Such fluidity in the materials included in Internet-based paths indicates a need for supporting maintenance and archival. Agents which identify Web pages that have changed, and thus indicates the need for potential path maintenance, is a topic that others have started looking at (Starr, Ackerman, & Pazzani, 1996). How is it different for paths? Archiving paths again brings up intellectual property concerns. Also, many of the documents on the Internet, such as the daily weather report, change over time and archiving a path at a specific time limits the timeliness of the path's information.

An additional issue that makes paths difficult to maintain is that of path complexity. The paths authored by our teachers tend to be relatively short—perhaps 15 nodes at the maximum. In large part this is related to the application environment. The shorter length of the path seems to correspond to a classroom “chunk”—a cohesive unit that can be used in a day's curriculum, especially when assuming some exploratory browsing off-the-path. A second factor is the teacher's busy schedule resulting in limited time to create more elaborate presentations.

Consequently over time the teacher accumulates a number of discrete path fragments, each of which may be useful in a variety of different environments (e.g., in both a basic and also an advanced art class). Supporting reuse of these fragments (and hence reducing the time needed to create new presentations) involves a number of factors including those of supporting collection of fragments into higher-level structures and enabling contextualization of paths.

To explore issues of coalescing fragments, we have prototyped a subpath mechanism in the Path Server. As might be expected a subpath represents a branch, itself a linear structure, off the main path. A subpath, once selected, is traversed in the same manner as a path. Once readers complete the traversal of the subpath, they continue from the point of departure in the original path. Our implementation allows the specification of subpaths in a path entry's annotation, using a unique tag to mark the subpath's anchor. The Path Server then modifies this tag (we recognize the WALDEN tag in our implementation) to create a normal HTML link, which when selected invokes the subpath. Consequently, identification and contextualization of the subpath is achieved as part of the annotation's narrative.

In addition, we have explored contextualization of annotation through implementation of multiple, conditionally-selected, annotations in a node. The annotation selected for display by the Path Server will depend on an “annotation level” assigned to each student by the teacher.

Examining enhancements to the basic linear path structure raises questions of the tradeoffs between expressiveness and usability. Clearly expressiveness is enhanced as the path structures become richer.

Eventually, the path evolves into a more general “meta-structure”. We discuss some aspects of meta-structures in the next section.

## 5. DISCUSSION AND CONCLUSIONS

The path is an instance of a *meta-document*; a more general information organizing structure. The definition of “meta-document” parallels that of the “document.” In short, a *document* author, for example the author of a Web page or the author of a book, ties information together implicitly through organization of material and explicitly through links between information elements. In essence, the author provides *structure* relating the information elements to each other. A *meta-document* provides *meta-structure*—adding an additional layer of structure that relates together elements from within disparate documents but without altering the organization of the elements within those sites. The term “meta-document” emphasizes that the meta-structures built are themselves cohesive, not randomly selected.

In future work, we plan to examine general meta-documents, particularly in support of the processes of organizing, contextualizing, and communicating information discovered on the World Wide Web. Information discovery—gaining knowledge necessary to reach a goal—is a critical use of today’s libraries, both traditional and digital. So too is communication of discovered information. Communication can be from person to person, for example from teacher to student, or from one time to a later time, for example records to be consulted at a later time when the information is needed again. Information communicated directly between two peers can be telegraphic in form, but as the differences in experience level or in elapsed time increases, so too does the need to contextualize the found information. The World-Wide Web in some sense presents a worst-case scenario—information discovery and communication is a critical need for its effective application in education but its distributed administrative structure makes cohesive enabling policy difficult to implement.

The World-Wide Web is an enormous, dynamically-changing, decentralized, unregulated information repository. Unlike more-traditional media such as books, television, and newspapers, the Web’s distribution channels are not strongly controlled and much of the information is not professionally edited or organized. Additionally there is no desire or possibility for establishment of a centralized authority to create policy and procedures on coordinated acquisition, deacquisition, and archiving of Web information. Equally importantly, the lack of centralized authority eliminates the ability to tightly synchronize the state of catalogs and other finding aids with the state of source information.

The existing information location resources on the Web tend to be either search-based or index-based. Both present Web information items as sets of independent resources with little contextualization; in other words as lists of potential sites of interest rather than as resources embedded in a broader narrative.

Consequently, information location and winnowing must be carried out repeatedly each time information on a particular topic is sought.

The experiences reported in this paper suggest to us that it will be valuable to investigate the generalization of path-based architectures to provide more general meta-document-based architectures. Although many issues remain open for further study, our experience with Walden's Paths suggest that meta-documents can be a natural and effective means for organizing, communicating, and contextualizing information in decentralized, distributed, digital environments such as the Web.

### **ACKNOWLEDGEMENTS**

This material is based in part on work supported by the Texas Advanced Research Program under Grant Number 99903-230 and the National Science Foundation under Grant Numbers IIS-9812040 and IIS-9734167.

### **REFERENCES**

- Bush, V. (1945). As We May Think. *The Atlantic Monthly*, 6 (1), 101-108.
- The CAETI Newsletter. (Number 18, May 2, 1997). Available on-line at <http://isx.com/CAETI/newsletters/newsletter.18.html#composers>.
- Furuta, R., Shipman, F., Marshall, C., Brenner, D., & Hsieh, H. (1997). Hypertext Paths and the World-Wide Web: Experiences with Walden's Paths. In *Hypertext '97: the Eighth ACM Conference on Hypertext* (pp. 167-176). New York: ACM.
- Greer J., & Philip, T. (1997). Guided Navigation Through Hyperspace. In *Proceedings of the Workshop on Intelligent Educational Systems on the World-Wide Web, 8th World Conference of the AIED Society*, [http://www.contrib.andrew.cmu.edu/~plb/AIED97\\_workshop/Greer.html](http://www.contrib.andrew.cmu.edu/~plb/AIED97_workshop/Greer.html).
- Halasz, F., Moran, T., & Trigg, R. (1987). NoteCards in a Nutshell. In *Proceedings of the ACM CHI+GI Conference* (pp. 45-52). New York: ACM.
- Kacmar, C., Jue, D., Stage, D., & Koontz, D. (1995). Automatic Creation and Maintenance of an Organizational Spatial Metadata and Document Digital Library. In *Proceedings of Digital Libraries '95* (pp. 97-105).
- Ladd, B., Capps, M., Stotts, D., & Furuta, R. (1995). Multi-head/Multi-tail Mosaic: Adding parallel automata semantics to the Web. In *Proceedings of the 4th WWW Conference* (WWW Journal, vol. 1) (pp. 433-440). O'Reilly and Associates Inc.
- Marshall, C.C. (1997). Annotation: from paper books to the digital library. In *Proceedings of ACM Digital Libraries '97* (pp. 131-140). New York: ACM.
- Marshall, C.C., & Shipman, F.M. (1995). Spatial hypertext: Designing for change. *Communications of the ACM*, 38 (8), 88-97.
- Nicol, D., Smeaton, C., & Slater, A.F. (1995). Footsteps: Trail-blazing the web. In *Proceedings of the Third International World Wide Web Conference*.

- Reeves, B. (1993). *Supporting Collaborative Design by Embedding Communication and History in Design Artifacts*. Ph.D. Dissertation, Department of Computer Science, University of Colorado, Boulder, Colorado, 1993.
- Resnick, P., & Varian, H.R. (1997). Special Issue on Recommender Systems, *Communications of the ACM*, 40, 3.
- Robertson, S., Jitan, S., & Reese, K. (1997). Web-based Collaborative Library Research. In *Proceedings of ACM Digital Libraries '97* (pp. 152-160). New York: ACM.
- Schneider, E.R., Leggett, J.J., Furuta, R.K., Wilson, H.D., & Hatch, S.L. (1998). Herbarium Specimen Browser: A Tool for Accessing Botanical Specimen Collections, In *Proceedings of ACM 1998 Conference on Digital Libraries* (pp. 227-234). New York: ACM.
- Shipman, F., Marshall, C.C., Furuta, R., Brenner, D., Hsieh, H., & Kumar, V. (1996). Creating educational guided paths over the World-Wide Web. In P. Carlson and F. Makedon (editors). *Educational Telecommunications, 1996: Proceedings of ED-TELECOM 96* (pp. 326-331). Association for the Advancement of Computing in Education.
- Shipman, F., Furuta, R., & Marshall, C.C. (1997). Generating Web-based presentations in spatial hypertext. In *1997 ACM International Conference on Intelligent User Interfaces* (pp. 71-78). New York: ACM.
- Shipman, F., Furuta, R., Brenner, D., Chung, C.C., & Hsieh, H. (1998). Using Paths in the Classroom: Experiences and Adaptations. In *Proceedings of ACM Hypertext '98* (pp. 267-276). New York: ACM.
- Shipman, F., Furuta, R., Hsieh, H., Francisco-Revilla, L., & Karadkar, U. (1999). Using the Internet in the Classroom: Variety in the Use of Walden's Paths. In *Proceedings of Ed-Media '99*. Association for the Advancement of Computing in Education.
- Starr, B., Ackerman, M., & Pazzani, M. (1996). Do-I-Care: A collaborative web agent. In *CHI 96 Conference Companion* (pp. 273-274). New York: ACM.
- Trigg, R. (1988). Guided tours and tabletops: Tools for communicating in a hypertext environment. *ACM Transactions on Office Information Systems*, 6 (4), 398-414.
- Zellweger, P. (1987). Directed paths through collections of multi-media documents. In *Hypertext '87*. Position paper.
- Zellweger, P. (1988). Active paths through multimedia documents. In J.C. van Vliet (editor). *Proceedings of the International Conference on Electronic Publishing, Document Manipulation, and Typography* (pp. 19-34). Cambridge University Press.
- Zellweger, P. (1989). Scripted documents: A hypertext path mechanism. In *Hypertext '89 Proceedings* (pp. 1-26). New York: ACM.