Adaptive Information Visualization for Personalized Access to Educational Digital Libraries

PETER BRUSILOVSKY[1,2], ROSTA FARZAN[2], AND GIRISH CHAVAN[1]

[1]School of Information Sciences and [2]Intelligent Systems Program University of Pittsburgh, Pittsburgh PA 15260, USA

peterb@pitt.edu, gchavan@mail.sis.pitt.edu, rosta@cs.pitt.edu
Abstract. Personalization is one of the emerging ways to increase the power of modern Digital Libraries. The Knowledge Sea II system presented in this paper explores *social navigation support,* an approach for providing personalized guidance within the open corpus of educational resources. Following the concepts of *social navigation* we have attempted to organize a personalized navigation support that is based on past learners' interaction with the system. The study indicates that Knowledge Sea II became the students' primary tool for accessing the open corpus documents used in a programming course. The social navigation support implemented in this system was considered useful by students participating in the study of Knowledge Sea II. At the same time, some user comments indicated the need to provide more powerful navigational support, such as the ability to rank the usefulness of a page.

Keywords: Social navigation, navigation support, user model, group model, knowledge map, SOM

Introduction

Personalization is one of the recognized priority research strands in the field of Digital Libraries (DL). As pointed out by the Joint

NSF/DELOS Working Group on Personalization and Recommender Systems in Digital Libraries: "*Current digital libraries tend to be passive, like Web search engines; they do little to adapt themselves to their patrons. However, digital libraries continue to grow in size and importance: They are now necessary tools for many ordinary people and for many common tasks. In order to improve its effectiveness, a digital library must be proactive in offering and tailoring information for individual users; if it is not personalized for individual users, then a library is defaulting on its obligation to offer the best service possible.*" One of the key problems of DL personalization is to help the users locate resources that are relevant to their goals, knowledge, and interests through *personalized access* to resources. Personalized access is especially important for educational digital libraries (EDL). Not only is it hard for an inexperienced user to formulate the proper query [35], but they also find it difficult to choose the most relevant resources from the list returned by the engine. The results of several research projects have shown that the ability to choose appropriate links requires a relatively high level of background or subject knowledge [4; 37]. We can hardly expect an average student to be able to locate relevant resources, even in a relatively small EDL where a query typically returns tens of hits. If digital libraries continue to grow geometrically, now that they are equipped with large automatic harvesters of resources [2], queries will result in at least hundreds or thousands of relevant hits, making search-only access practically unusable, even for domain experts.

So far, most research on personalized access within DE focused on two simple human-driven approaches:

1) The **user-driven approach** known as “my library” allows users to personalize their views of the library, promoting their favorite resources or categories [11; 13; 22].
2) The **teacher-driven approach,** based on the classic mechanism of guided tours, allows a teacher to provide a narrated path through digital library content, adapting it to what she believes to be the needs of her class [34].

While both kinds of personalization (both typically known as “static personalization”) are good starting points, they can’t guide the user to resources that are beyond their own or their teacher’s current knowledge.

Beyond this, these approaches provide a one-dimensional list-based form of access to resources. Some pioneer work on personalization in EDL attempted to explore dynamic personalization approaches in the educational context, but continued to use simple list-based formats to recommend their information resources [33].

The focus of the research presented in this paper has been to explore possible approaches for providing dynamic personalized information access to educational digital libraries through *adaptive information visualization*. While information visualization is the newest paradigm of information access, it has several unique features that distinguish it from other paradigms, such as information retrieval, filtering, and browsing [8]. Information visualization (IV) allows users to see a relatively large set of information resources as a whole, while still being able to discern individual resources. These resources are usually presented in two or three dimensions using various visual cues to show document properties and the relative positioning of documents, in order to express several different relationships between the documents. As a result, a fully two- or three-dimensional IV has a much higher expressive power for organizing information resources than an IR/IF system, since the latter is limited to a one-dimensional expressive power (a list of links) or hypertext. The higher expressive power of information visualization is usually complemented by a higher level of interactivity (Figure 1): most information visualization systems allow the user to manipulate the presented documents, then observe the changes in visualization. In the context of information access, such pioneer systems as VIBE [32], Envision [21], and MovieFinder [1] have demonstrated the benefits of interactive two-dimensional visualization.

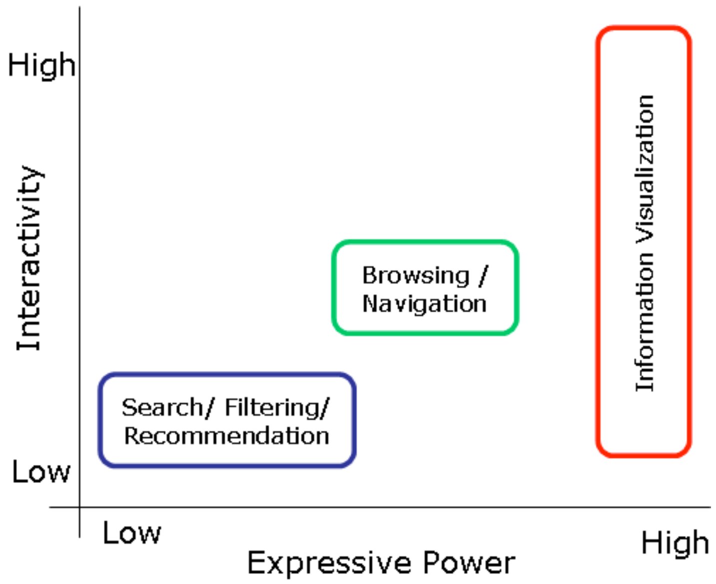


Figure 1 - Interactivity and expressive power of the major information access paradigms

Both expressive power and interactivity are important for personalized access. Expressive power allows the system to present a variety of personalized details about the documents while interactivity supports better user modeling [19]. Despite that, major research on adaptive information visualization has not yet begun. While dozens of projects were devoted to developing techniques for personalized information access within the three older paradigms [3; 28], we have found that only one project – the Lighthouse system – pioneered the use of adaptive information visualization for exploring information resources. Our approach to adaptive information visualization, as presented in this paper, was motivated by our earlier work on adaptive navigation support in educational hypermedia. Adaptive navigation support, an outgrowth of the field of adaptive hypermedia [3], is a group of technologies [3]created to help users find relevant information resources within hyperspace. Within this group of technologies, we have chosen *adaptive annotation*, which provides navigation support by attaching personalized visual cues to hyperlinks. These cues express various attributes of the documents behind the links and help users select the most relevant links to follow.

Our system Knowledge Sea II, presented in this paper, explores the value of adaptive annotation in the context of information visualization. Knowledge Sea II is an extension of our earlier project Knowledge Sea [6; 7], which was focused on information access to educational resources through information visualization. This paper begins by presenting the approach to information visualization used in the original Knowledge Sea system. It then introduces *social navigation support (SNS)* technology and the application of this within our newer system Knowledge Sea II, a personalized version of the original Knowledge Sea. After that, we present the results of three classroom studies of Knowledge Sea II (KSII) and discuss our plans to expand the system.

2. Accessing Multiple Educational Resources with the Knowledge Sea System

Our original Knowledge Sea system was motivated by practical needs – to create an interface that allows students to access relatively large volumes of educational resources in a discerning way. While creating a repository of open corpus educational material for a course on Programming and Data Structures, which was based on the C programming language, we gathered about 10 good C-language tutorials on the Web. Different tutorials used different presentation styles and championed different aspects of the language. It was quite clear to us that these tutorials provided an excellent complement to the printed C textbook, often presenting some features of C language better than in the book or in a way that would match up better with certain categories of students. First, we tried a "simple" way of providing access to these tutorials by listing links to their start pages on the course home page – only to find that nobody was using these tutorials. The resources were simply too many clicks away from the main pathways, hidden somewhere inside the tutorial navigation hierarchy. The Knowledge Sea was an attempt to apply an information visualization approach to place relevant resources virtually "one click away" from students.

The idea of the Knowledge Sea system was to decompose all tutorials into a set of information pages and provide structured goal-oriented access to these pages through a course *knowledge map* (see Figure 3). To build the knowledge map we used the Self-Organizing Map (SOM) approach. SOM is an artificial neural network that builds a two-dimensional representation of the inputs. It is a very attractive technology for developing compact maps for a large hyperspace since it builds a map representing only the neighborhood relationships between the objects [23]. SOM has been used in the past to organize and visualize large collections of documents [9; 20; 24; 36; 39].

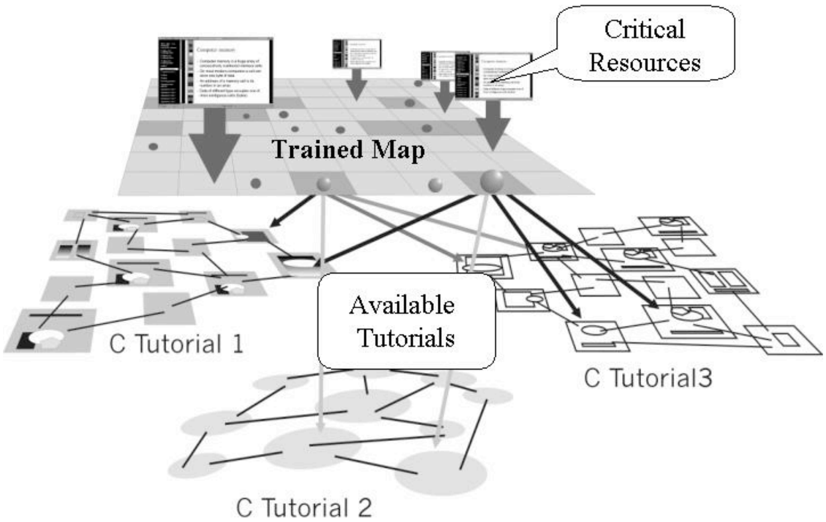


Figure 2 - Organization of the knowledge map.

Knowledge Sea uses the standard TF*IDF page representation [25], the Euclidean similarity measure, and SOM technology to allocate pages of several hierarchically structured tutorials (open corpus) and lecture handouts (closed corpus) on an 8-by-8 map (Figure 2). In building this knowledge map, we relied on the remarkable ability of SOM to group similar pages in the same cells as well as to place reasonably similar groups of pages in neighboring cells. SOM allowed us to provide what we called *map-based* navigation for multiple educational resources [7]. Keywords and critical resources such as lecture landmarks were placed on the map, to help students find cells that were closest to their current goal. When "opened," a cell of the map showed a list of links to similar pages in multiple tutorials. It allowed the students to navigate "horizontally" between tutorials in addition to the vertical hierarchical

navigation supported by each tutorial. More information about the knowledge map construction, the interface, the mechanism, and the motivation behind the Knowledge Sea can be found in [7].
Knowledge Sea was evaluated in several classroom studies. The students praised highly the ability of Knowledge Sea to group similar resources together and to help them find relevant pages [7]. At the same time, a number of students noted that Knowledge Sea provided no help in locating relevant resources within a cell. Since pages were added to a cell on the basis of keyword level similarity, excellent pages were sometimes located next to confusing pages or pages that contained no useful information. These problems are addressed in our new Knowledge Sea II system, which attempts to extend the power of visualization-based access to educational resources by including the power of personalized navigation support.

3. Social Navigation Support: The Interface of Knowledge Sea II

The idea of the Knowledge Sea II project is to augment the power of information visualization with additional navigation support, aimed at helping the students identify the most relevant resources on the Knowledge Sea Map. We were motivated by our experience with adaptive hypermedia systems [3], which use adaptive navigation support to help students select the most relevant links on each page. The problem, however, is that navigation support in existing adaptive hypermedia systems is based on manually-provided content knowledge about each page. As a result, traditional adaptive hypermedia systems can provide guidance within only a relatively limited set of resources (the so-called *closed corpus*). In our project we specifically wanted to deal with a very large and changeable set of resources (*open corpus*). While in some application areas it is feasible to have a team of experts to encode the content knowledge about thousands and thousands of available resources, educational systems cannot afford it.

To cope with the open corpus challenge, Knowledge Sea II explores an alternative approach for providing personalized guidance to students. Following the concepts of *social navigation* [15], we have attempted to organize a personalized navigation support that is based on the learners' past interaction with the system. We call this *social navigation support*. Unlike traditional adaptive navigation support, which relies on expert-provided knowledge about each resource, *social* navigation support relies on the *collective knowledge* of a large community of learners, gathered through different forms of feedback.

Social navigation is a new stream of research that explores methods of organizing users' explicit and implicit feedback to support information navigation. It began with a few now-classic projects [14; 38], which attempted to support a known social phenomenon: when navigating, people tend to follow the "footprints" of other people.

An important feature of all social navigation systems is *self-organization*. Social navigation systems are able to work with little or no involvement of human administrators or experts. They are powered by a community of users. Properly organized community-powered systems such as Web auctions (www.ebay.com) or Weblogs (www.lifejournal.com) are known to be among the most successful

Web applications. Self-organization is critical to the goals of our project. In the educational context, self-organization is one of the most promising approaches for guiding students to the most useful resources, outside of the continuous involvement of human experts to index these resources. The feasibility of social navigation in E-Learning has been explored in a few pioneering projects [16; 26; 30]. These projects were an inspiration for our work. We have attempted to extend these projects by developing a social navigation support technology that can be combined with IV and operate within larger volumes of educational content.

The idea of our social navigation mechanism is to guide the user to resources that have been "appreciated" by similar users. While the most reliable sign of "appreciation" is an explicit rating for a page, this approach is too intrusive and usually does not work in practice. Following other modern work on Web recommender systems we decided to use implicit indicators such as page visits (traffic) and page annotation. We assumed that for a homogeneous group of users, the more visits and annotations the page has, the higher the probability it is relevant for new users. Thus, the goal of the system was to provide social navigation by expressing visually the page traffic and presence of notes made by similar users. This approach extended the original "footprint" approach and combined it with group user modeling.

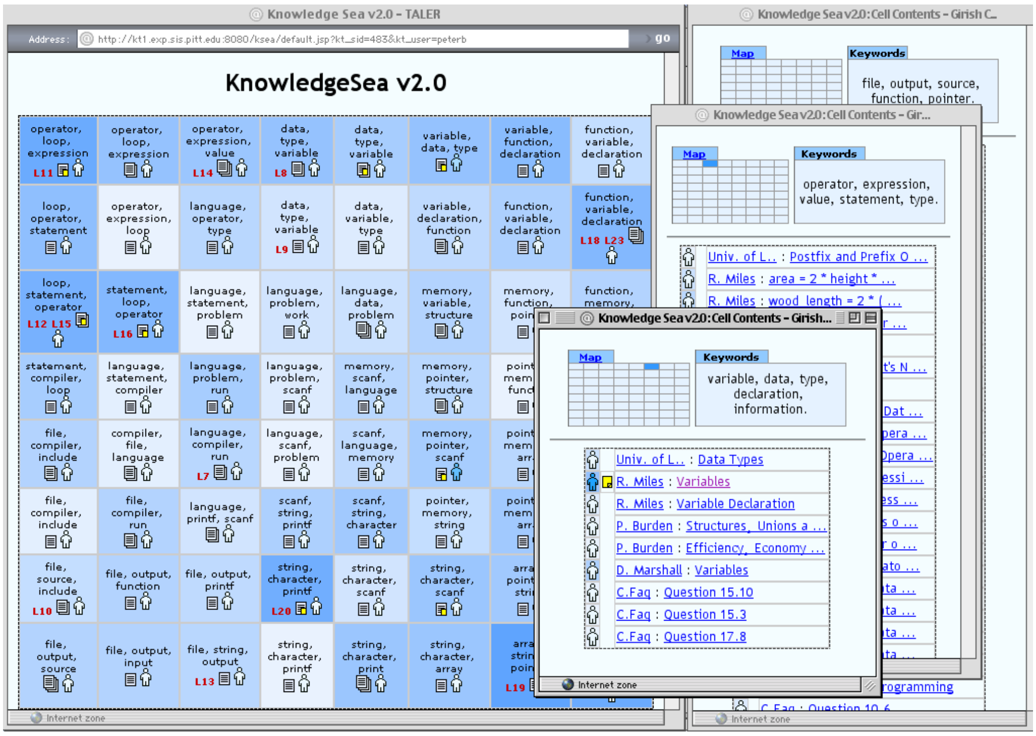
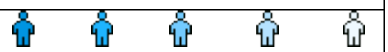


Figure 3 - The map view and three cell views in Knowledge Sea II

Group traffic is the main social factor visualized by Knowledge Sea II. It counts how many times users of the same group access each tutorial page. The system calculates group traffic for each map cell by summing up group traffic for the individual pages belonging to that cell. There are several possible ways to visualize group traffic for tutorial pages and map cells. For the first version of SNS we decided to explore visual cues based on background color. This approach was attractive since it made group traffic clearly visible without cluttering the interface. In Knowledge Sea II, the cell traffic is now visualized by changing the intensity (saturation) of the blue cell background. The more intense the color, the more times the cell's pages have been accessed. As a group of users navigates though the tutorial pages represented on the map, the knowledge map becomes gradually darker and darker. During the classroom study of Knowledge Sea II we observed that the cells with the most useful content become darker much faster. At the end of the course, the web page pathways taken by the course's students are reflected in the color of the cells, thus capturing the usefulness of each map cell remarkably well. Dark cell colors indicate that a good number of pages present the most complicated course topics. Very light colored cells usually focus on topics not covered by the course.

Table 1: Visual cues on the knowledge map

	More Less
Individual cell traffic	
Number of pages in the cell	
Presence of annotated pages in the cell	

In addition to expressing the amount of group traffic through the background color, each cell on the map shows four other visual cues:

1) The list of *top three keywords* associated with documents of that cell help the students to find cells similar in content.
2) *Lecture numbers* act as landmarks, helping users to locate open corpus documents that are similar to well-known closed corpus documents (weekly course lectures).
3) The *icon shaped as a stack of papers* indicates the relative number of documents in the cell and serves as an anchor for accessing the cell. A small *yellow note* on this icon indicates the presence of user notes on some of the cell's pages (Table 1).

4) *Human-shaped icon* indicates the number of the user's own visits to the cell's documents, using the same color intensity approach – a more intensive color indicates a larger number of visits (Table 1). This technique provides an extended form of history-based navigation support, which was explored in some early educational adaptive hypermedia [12]. The user can see which pages she has explored well or not at all.

The intensity of the background and the icon are coordinated. As a result, if the user has visited the cell more than the average group user, the icon is darker than the background; if she has visited it less than average, the icon is lighter than the background. It allows the users to instantly discern potentially useful cells that they have under-explored (as well as displaying possibly over-explored cells).

The same traffic-based approach is used to annotate links inside each of the map cells (Figure 4, right). Each link is annotated with a human-shaped blue icon on a blue background. The color of the icon shows the student's own navigation history. The color of the background shows the cumulative navigational history of the whole class. The intensity of the two colors are also coordinated (i.e., the contrast between the figure and the background indicates the difference between the average and individual traffic for that page).

A note-shaped icon indicates the presence and density of notes on a page. For consistency, we use the same color intensity approach to show the density of notes (but using a yellow color instead of blue). The color of the background indicates the density of public notes made by the group as a whole; the color of the icon indicates the number of notes made by the student herself. Similar traffic-based adaptive annotation was added to all tutorial pages, thus creating standard annotation-based adaptive navigation support within the hypertext. This aspect of the system is beyond the scope of this paper but can be found in [18].

Thus, Knowledge Sea II uses a combination of simple group user models and individual user models to provide a personalized knowledge map for every student and to help the user make navigational choices. In the spirit of good adaptive navigation support, our visual cues do not prescribe where the user must go next. Instead, they provide an additional level of personalized information for decision-making that is not visible in a typical non-adaptive system. The decision of where to go next is made by the user and may also depend on the goals of a specific educational session. For example, if the goal is to explore new material related to a specific lecture, the user may want to target cells and pages that have not yet been visited, but which are appreciated by others. If the goal is to prepare for a test, the user may prefer to find cells that she has visited (and commented on) in the past while also watching for popular neighboring cells and pages that were originally skipped. Our hope is that by providing a layer of adaptive guidance to augment the knowledge map, our adaptive visualization system will serve its users better than previous approaches did.

4. Adaptation Mechanisms

One benefit of socially-based adaptation mechanisms is that the mechanisms and the user models behind them are relatively straightforward. The individual user model stores the number of the user's visits and the number of annotations made for each resource page represented in the Knowledge Sea II. This technology is implemented as a service component of our architecture, KnowledgeTree [5]. Every visit to a map cell or a page and every comment made to a page send an event to the central user modeling server CUMULATE. As a result, at any moment, for any page P_k the number of visits made to this page by the user i (individual traffic) can be directly obtained from the user model as v_{iPk} and the number of annotations made as α_{iPk}. Since it knows which pages are contained in each knowledge map cell, the system can calculate the total of each individual's traffic for cell C_j as

$$v_{iC_j} = \sum_k \{v_{iP_k} \mid P_k \subset C_j\} \qquad (1)$$

Also, CUMULATE supports additive group modeling [29]. Knowing the list of users for each user group G, it calculates and stores the total number of visits made by members of this group to each page P_k as

$$vg_{P_k} = \sum_i \{v_{iP_k} \mid i \subset G\} \quad (2)$$

Similarly, the total number of annotations made by all group members is calculated as

$$\alpha g_{P_k} = \sum_i \{\alpha_{iP_k} \mid i \subset G\} \quad (3)$$

Using page-level data, Knowledge Sea II can calculate the total group traffic for each knowledge map cell as

$$vg_{C_j} = \sum_i \{v_{iC_j} \mid i \subset G\} \quad (4)$$

This data is used by Knowledge Sea II to assign visual cues on the map and cell displays. The page and cell user traffic determines the choice of color for the human-shaped icon in each cell in the map view and next to each link on the cell view. Knowledge Sea II distinguishes 5 different levels of user traffic, which are displayed using the same blue icon with different color intensities (Table 1). The page and cell group traffic is used to calculate the intensity of the cell background color on the map and icon background in the cell view. Knowledge Sea II uses 43 different levels of color intensity to indicate group traffic. This ad-hoc number was selected after several trials. It provides good precision in displaying the traffic while keeping the colors relatively distinct from each other. The number of individual and group annotations for a page is used to determine the foreground and background colors of the "note" icon shown next to the human-shaped icon in the cell view. Only two color choices (different intensities of yellow) are used for the icon and only three for the background since distinguishing the intensity of yellow is much harder for humans.

The basic idea of traffic-based color calculation is straightforward: zero-level traffic is represented by the lightest color, maximal traffic is represented by the darkest color, and anything else is represented by a corresponding color in between these poles. Using the HLS (Hue, Lightness, Saturation) model representations for colors, the numerically expressed color intensity can be directly converted to a specific color. However, choosing the proper mapping function is important. For doubled traffic the target mapping function should generate a color that is perceived as "twice darker." It is known that the relationship between color intensity used by computers and humanly-perceived intensity is not linear [31, p.92]. It is also important to choose the proper traffic level for maximal traffic.

After exploring several alternatives, we decided to use *maximal estimated traffic* for maximum and logarithmic scale for color mapping. Maximal estimated traffic t_{max} for a page was set to 10 hits. Maximal estimated traffic for a cell is calculated as the total sum of maximal page traffic for cell pages. Similarly, maximal group traffic is the total sum of maximal traffic for each user. This approach keeps color annotations for cells and pages relatively stable and independent from each other. The use of logarithmic mapping for color and sound intensity is common in engineering. While Norman advocates the use of the cubic scale, we found logarithmic mapping quite intuitive. When the traffic levels are relatively low, small changes to the traffic are immediately reflected with a change in color. With the growth of traffic, a larger increase is required to produce a visible color change. We defined the mapping function N(x) as follows :

$$N(x) = \left\{ \begin{matrix} \ln(x) & x > 1 \\ 0.2 & x = 1 \\ 0 & otherwise \end{matrix} \right\} \qquad (5)$$

We chose *N(1) = 0.2* because *0.2 < ln(2),* which is the next possible traffic level after 1. To obtain the color level *cl* from the traffic level *t* we use simple normalization:

$$cl = \left(\frac{N(t)}{N(t_{\max})} \right) * N_{cl} \quad (6)$$

N_{cl}

The obtained color is thus logarithmically distributed between 0 and N_{cl}. This approach is used to calculate color for all visual cues. Note that for different cues, the number of color levels N_{cl} is different (i.e., 43 for cell background, 5 for human icon, etc.). To display the proper icon color, the system simply chooses an icon that corresponds to the calculated *cl*. To display the proper background color, *cl* is converted to RGB and used to specify table or cell background color in generated HTML code.

5. Classroom Study of Social Navigation Support

We performed three classroom studies of Knowledge Sea II in the introductory programming class at the University of Pittsburgh. Knowledge Sea II was available to the students for the whole duration of the course. The teacher briefly demonstrated the system during several lectures and welcomed the students to use it in addition to the printed course textbook. The access to Knowledge Sea map was provided through our course portal, KnowledgeTree [5]. Note that Knowledge Sea II was not designed to improve learning directly –

rather, it was designed to provide better access to open corpus learning resources. If the accessed content is of high quality and the students use it properly, it can increase the learning outcome, but the Knowledge Sea II system itself can't ensure high quality learning materials. Therefore, to evaluate the system we checked whether the system helps users to access open corpus content more often and whether the users appreciate its help. We also checked to see if accessibility to the Knowledge Sea II system from home would have any effect on the students' performance on homework.

For objective evaluation of the system we used access log data. Due to the presence of the user model server, all results of the user work with Knowledge Sea II were automatically logged. For subjective evaluation of the system we asked the students who worked with Knowledge Sea II to fill in a non-mandatory questionnaire at the end of the class.

5.1 Log Analysis and System Usage Evaluation

The log analysis showed that out of 73 students in three classes with access to Knowledge Sea II, 52 used the system at least once. Table 2 shows general information about usage of the system over three semesters of its evaluation.

When processing the log data, we were interested in examining two issues: (1) to what extent does the system help students to access open corpus tutorial material and (2) which category of students uses the system most.

Table 2. Knowledge Sea II access data over 3 semesters

Semester	# of Students	# of Students using KSII	Max Accessed Pages	Max Visits	Max Annotations
Fall 2003	30	14	129	201	77
Spring 2004	28	26	81	91	7
Fall 2004	15	12	134	263	13

5.1.1 Knowledge Map as a Tool for Information Access

From the first classroom study of Knowledge Sea II, in the Fall 2003 semester, it was immediately clear from system usage data that SNS helped the students to reach tutorial pages. The amount of system use was impressive given the non-mandatory nature of the system. On average, each class generated about 1000 hits and made about 50 annotations. This was a dramatic comparison to our original attempt to provide access (by offering links to several tutorial roots) in which less than 10 hits were generated by the whole class (and most did not extend beyond the roots themselves). However, we were most interested in determining whether we could demonstrate comparative the value of Knowledge Sea II visualization-based access over traditional ways to access open corpus content (provision of links to tutorial roots and direct page recommendation). Beginning with the Spring 2004 semester, we included two additional ways of accessing the same educational content. First, on the Knowledge Sea II home page, right under the map, we added links to the roots of all tutorials that we identified as useful – including six tutorials represented on the map and two tutorials not represented there. If the students started their browsing from these root pages, they would be able to access every single tutorial page, including pages not accessible through the map. Standard SNS was provided for all links in these tutorials.

Secondly, we added direct links to 15 useful tutorial pages, as recommended readings for several lectures on the course portal KnowledgeTree (KT). The course portal is frequently visited by students who wish to browse program examples and work with quizzes, so we expected that these pages, directly connected to the portal, would be accessed most frequently. Most of these pages were also present on the SNS map, though 3 of 15 direct links were taken from the two tutorials not represented on the map.

The analysis of 20 top-visited tutorial pages for two semesters (Figure 5) demonstrated that the knowledge map became the most important tool for accessing external resources, far surpassing even the traditionally popular access via directly recommended readings for a lecture. For both semesters shown on Figure 5, only four of the 20 most popular pages were those directly linked to the portal (and clearly, even for these pages, some of the accesses were through the map, not through the portal[1]). Similarly, very few non-map pages made it to the top 20 – and these were not the three pages with direct linkage from the KT course portal but were pages on the top of tutorial trees accessed through tutorial root links listed directly under the map.

---

[1] During the Fall 2004 semester, only 56% of the access to these documents was done through the KT portal. We do not have the associated data for Spring 2004.

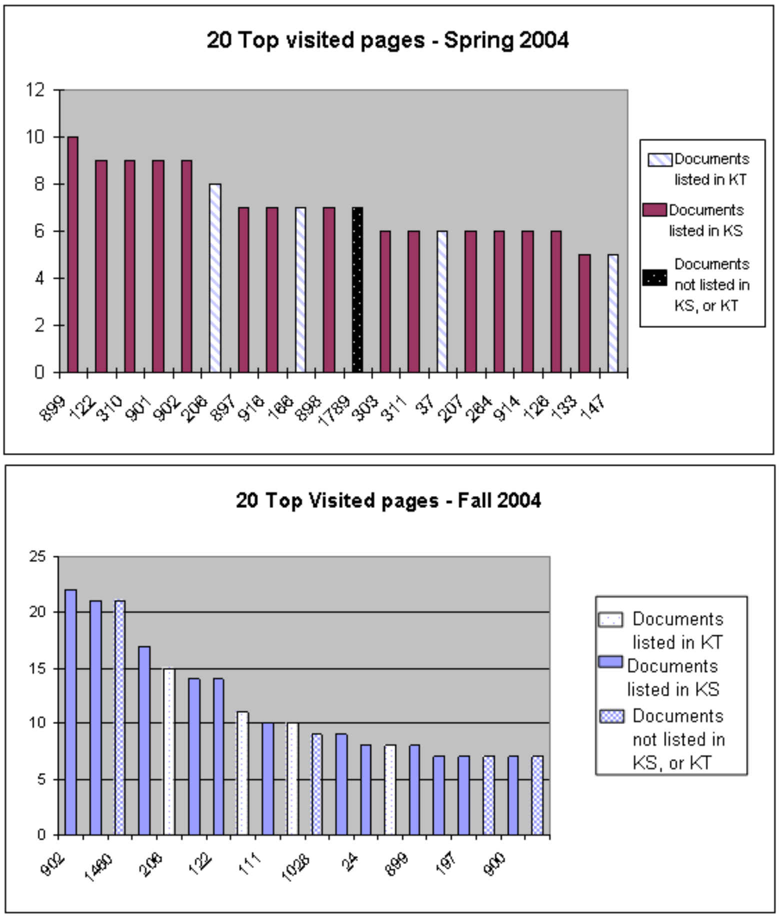


Figure 4 - Most popular tutorial pages (by page IDs) and the number of visits to the page, over two semesters

To better assess the effect of the map and social navigation on the students' navigational behavior, we analyzed the normalized average access rate for pages that can and can't be access from the map; with and without group traffic. If a page had at least two prior visits before, it was considered a page with "Group Traffic" (since two visits makes darker background color clearly visible); otherwise, it is considered a page with "No Group Traffic." To compute the average normalized access, we divided the number of access to pages in each category by the total number of available pages in each category. Table 3 shows the normalized average access rate and how it is computed. As the data shows, the chance of visiting an arbitrary resource in Knowledge Sea II is very low (close to 0) since there are a large number of resources in the system. Both the presence of a document on the map and the presence of social navigation support dramatically affects users' navigation behavior, increasing the chances that a student will access a resource that is useful them. The data also shows that the effect of the map and social navigation support can be combined.

Table 3. Computing normalized average access rate for pages with and without group traffic (Fall 2004)

	All		Pages available through Map	
	Clicks/Pages	Access Rate	Clicks/Pages	Access Rate
No Group Traffic	665 / 24670	0.003	269 / 908	0.296
Group Traffic	299 / 148	2.020	202 / 76	2.668

5.1.2 Effect of adaptation on different categories of users of the system

To determine the category of students who would use the system most frequently, we used a pre-course questionnaire to split the average number of page visits by:

- knowledge of the subject (as shown by the starting level and final grade)

- Web experience
- gender

While the small number of students within each category does not allow us to make a reliable conclusion, it is interesting to observe that students with weaker knowledge of the subject use the system, on average, at least as much as stronger students. It is in sharp contrast to our study of the original Knowledge Sea [7] where we found that the system was used almost exclusively by students with stronger knowledge of the subject. This could be evidence that the adaptive version of the system is friendlier for students who have weaker knowledge, exactly the ones who need more support.

5.1.3 Effect of usage of the system on students' performance

The main goal of KS is helping students in the process of information seeking. One typical case of usage of the system can be when students are working on their C programming homework and they need more information to solve the problems.

We evaluated the general effect of using Knowledge Sea on homework performance by looking at homework grade for different usage levels of Knowledge Sea. For evaluation purposes we categorized usages of the system into four categories based on number of clicks. For each category we computed the average weighted homework grade for students of that category. We weighted the homework grade based on the difficulty of the homework and the score can be more than 100%. As it can be seen in the figure 5 the homework performance is improving as usage of the system increases. The improvement at Fall 2004 is significant at $\alpha=0.1$ but we did not observe significant difference over Spring 2004.

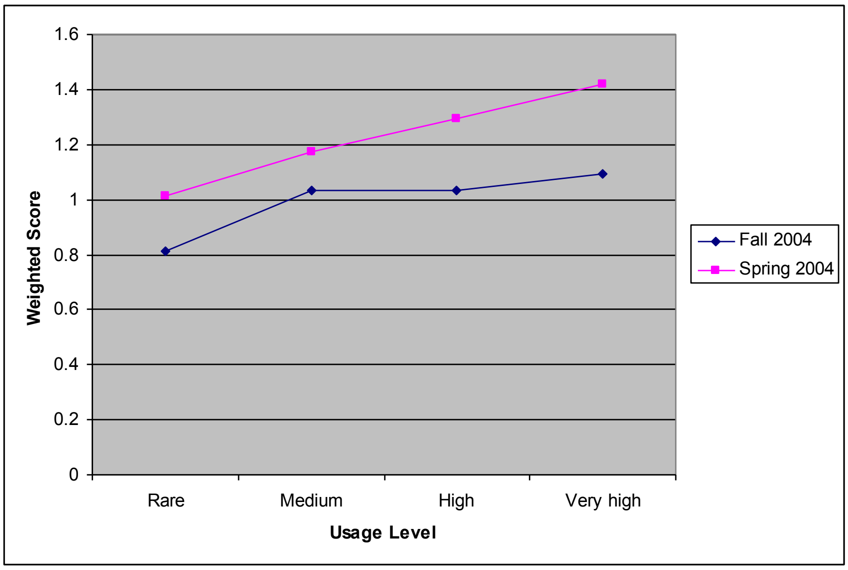


Figure 5 - Homework performance of students with different usage pattern of the system

5.2 Student Feedback Analysis

The students were asked to fill in an optional end-of-course questionnaire that was meant to collect user feedback about various aspects of the system. It included between 19 and 27 questions (depending upon the semester), with all students receiving the same main section but with the addition of a few new questions each semester, in order to evaluate new features of the system. Table 4 presents student participation in answering the questionnaire over three semesters.

Table 4. Student participation in subjective evaluation over 3 semesters

Semester	Total # of students	# of students answering
Fall 2003	30	14
Spring 2004	28	22
Fall 2004	15	10

Table 5. The structure of the questionnaire

Category	Questions	Theme of the questions
Overall impression	3	How well does the system achieve its goal? How good are the interface, and the overall design?
Map-based navigation	5	Evaluate the concept of map-based navigation, which associates the content of different tutorial pages within the same cell, neighboring cell, or cells related to lecture notes.
Interface Innovations	7	Evaluate different interface features of the system, such as the use of different visual cues to provide social navigation support.
Social Navigation Support	4-5	Evaluate the help provided through social navigation support.
Other	2-7	Evaluate different, changing features of the system, over three semesters.

The students' response to the questionnaire was used to evaluate user attitude toward various features of the system. Table 5 describes the questionnaire in general. As presented in the table, the questionnaire includes five categories of questions. The most relevant categories to the discussion of this paper (Interface and Social Navigation Support) are analyzed below (Figures 6 and 7). We have provided some information about their attitude toward the overall value of the system as well. The analysis of the student's attitude to map-based navigation is presented elsewhere.

The group of questions focused on interface features of the system attempted to discover successful or unsuccessful interface innovations. We used a 4-point asymmetric scale where the available answers were:

- strong positive (i.e., very good interface feature),
- positive,
- neutral, or
- negative (i.e., completely wrong idea).

The use of the asymmetric scale is a part of our research on the evaluation of adaptive systems. We have performed a range of studies for adaptive educational systems with both symmetric and asymmetric scales and have noticed that strong negative categories are almost never used. We think that this category's presence makes the data look better than it really is.

The analysis of students' answers in this group demonstrated that adaptive visual cues symbolizing traffic were highly appreciated interface features. With the exclusion of individual map traffic (in Spring & Fall 2004) and map group traffic (Fall 2003), more than 80% of users considered these features either as good or very good (Figure 6). The students' attitude toward adaptive visual cues was typically higher than their attitude toward the system interface in general. The results are consistent over all three semesters.

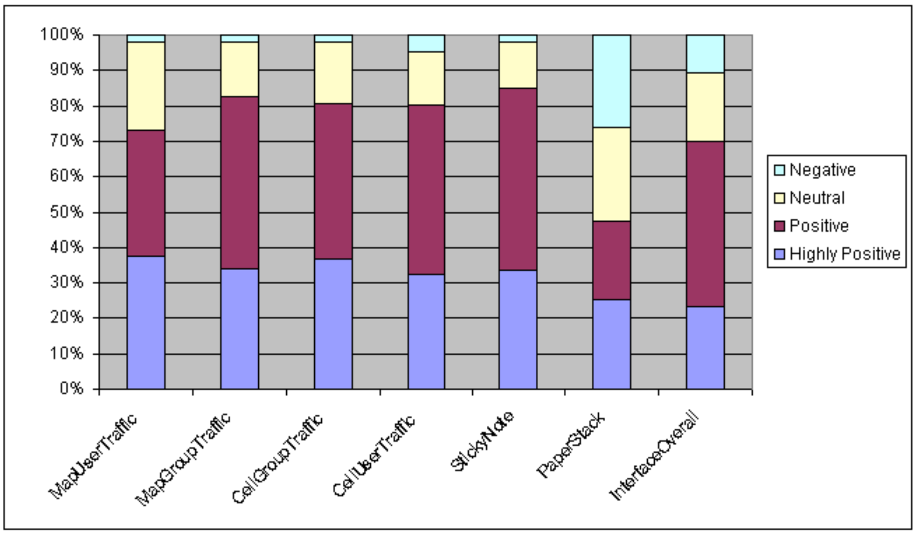


Figure 6 - Students' attitude toward different navigation support cues over 3 semesters (52 students)

Another group of questions was designed to assess the student's overall interest in seeing individual/group visits. In each of two contexts (map navigation and cell navigation), we asked the users what kind of traffic they would like to see visualized. The answer options were:

- only group traffic,
- only user traffic,
- both, or
- none.

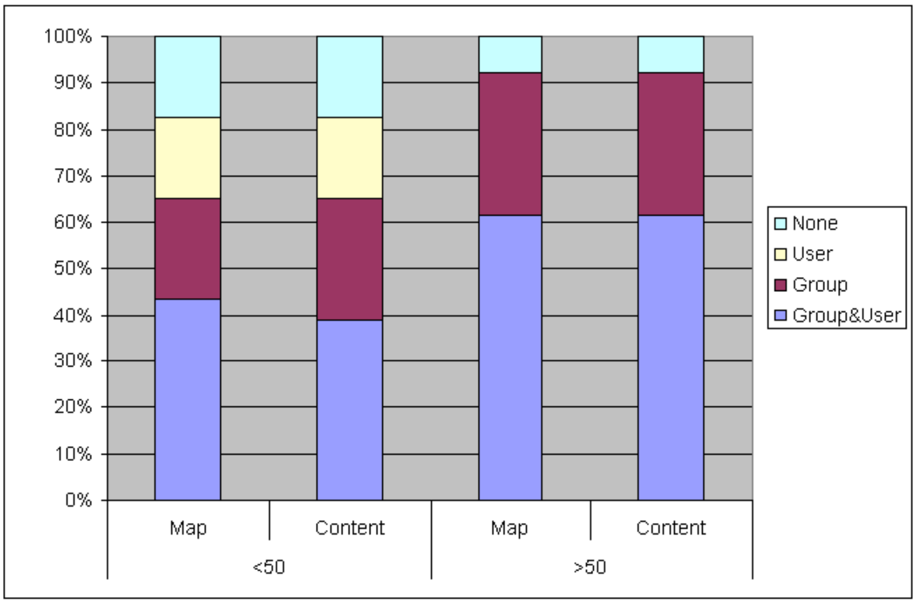


Figure 7 - . Traffic visualization preferences, split by user experience over Fall 2003 and Spring 2004 (42 students)

The analysis (Figure 7) prompts two observations. First, students considered visualization of group traffic more useful than visualization of user traffic. Most students wanted to see either group-only traffic or both kinds of traffic. Only a relatively small percentage of students who have less experience with the system (less than 50 hits) wanted to see user-only traffic. Secondly, students who had reasonable experience working with the system (more than 50 hits) appreciated traffic visualization in general and group traffic visualization specifically, more than less-experienced students.

6. Conclusions and Future Work

In this paper, we have discussed some of the problems involved in providing access to educational digital libraries through adaptive visualization. We have presented the Knowledge Sea II system, which provides map-based access to resources enhanced by social navigation support (SNS) – an open corpus adaptive navigation support based on the concepts of social navigation. While the traffic-based SNS implemented in our system is relatively simple, it was considered quite useful by students participating in our classroom studies of that system. At the same time, some user comments indicated the need to provide more powerful navigation support, such as an indication of the most useful pages, not simply the most visited pages.

Our current work is focused on improving the quality of SNS. Starting from a simple history-based approach, which was advocated in early works on social navigation and is used as the base level in our work, we are expanding SNS technology in two main directions. The first direction of our work is to extract more meaning from the page-visiting trace by measuring time spent reading a page. The preliminary results from this research, found in [17], demonstrate that time-based SNS could provide better results than classic click-based SNS.

The second direction is to provide *additional* knowledge sources for SNS "beyond page visits." So far, we have explored page annotations as evidence of page relevance and quality. Over a number of semesters, we have been exploring and enhancing the annotation interface and annotation-based SNS. Some preliminary reports on this work can be found in [18]. In the future, we intend to improve the quality of social navigation support by extending the use of implicit indicators of user interest to more measures, such as mouse movement and scrolling.

References

[1] Ahlberg C and Shneiderman B (1994) Visual information seeking: Tight coupling of dynamic query filters with starfield displays. In Human Factors in Computing Systems. CHI '94. ACM Press, New York, pp. 313-317

[2] Bergmark D (2002) Models and tools for generating digital libraries: Collection synthesis. In Marchionini G and Hersh W (eds) ACM/IEEE-CS Joint Conference on Digital Libraries, Portland, OR, July 14-18, 2002, pp. 253-262

[3] Brusilovsky P (2001) Adaptive hypermedia. User Mod & User-Adapt Interact 11: 87-110, also available at http://www.wkap.nl/oasis.htm/270983

[4] Brusilovsky P (2003) Adaptive navigation support in educational hypermedia: the role of student knowledge level and the case for meta-adaptation. Brit J Educ Technol 34: 487-497

[5] Brusilovsky P (2004) KnowledgeTree: A distributed architecture for adaptive e-learning. In The Thirteenth International World Wide Web Conference, WWW 2004 (Alternate track papers and posters), New York, NY, 17-22 May, 2004, pp. 104-113

[6] Brusilovsky P and Rizzo R (2002) Map-based horizontal navigation in educational hypertext. J Digital Information 3, also available at http://jodi.ecs.soton.ac.uk/Articles/v03/i01/Brusilovsky/

[7] Brusilovsky P and Rizzo R (2002) Using maps and landmarks for navigation between closed and open corpus hyperspace in Web-based education. New Rev Hypermedia & Multimedia 9: 59-82

[8] Brusilovsky P and Tasso C (2004) Preface to special issue on user modeling for Web information retrieval. User Mod & User-Adapt Interact 14: 147-157

[9] Chen H, Houston Andrea L, Sewell Robin R, and Schatz Bruce R (1998) Internet browsing and searching: User evaluations of category map and concept space techniques. J Amer Soc Inf Sci 49: 582-608

[10] Claypool M, Le P, Wased M, and Brown D (2001) Implicit interest indicators. In 6th International Conference on Intelligent User Interfaces, Santa Fe, NM, January 14-17, 2002, pp. 33-40

[11] Cohen S, Fereira J, Horne A, Kibbee B, Mistlebauer H, and Smith A (2000) MyLibrary: Personalized Electronic Services in the Cornell University Library. D-Lib Magazine 6

[12] de La Passardiere B and Dufresne A (1992) Adaptive navigational tools for educational hypermedia. In Tomek I (ed) Computer Assisted Learning. ICCAL'92, 4-th International Conference on Computers and Learning, Wolfville, Canada, June 17-20, 1992. LNCS 602, Springer-Verlag, Berlin, pp. 555-567

[13] Di Giacomo M, Mahoney D, Bollen J, Monroy-Hernandez A, and Ruiz Meraz CM (2001) Mylibrary, a personalization service for digital library environments. In Second DELOS Network of Excellence Workshop on Personalisation and Recommender Systems in Digital Libraries, Dublin, Ireland, June 2001, also available at http://www.ercim.org/publication/ws-proceedings/DelNoe02/Giacomo.pdf

[14] Dieberger A (1997) Supporting social navigation on the World Wide Web. Int J Human-Comp Interact 46: 805-825

[15] Dieberger A, Dourish P, Höök K, Resnick P, and Wexelblat A (2000) Social navigation: Techniques for building more usable systems. interactions 7: 36-45

[16] Dron J, Boyne C, and Mitchell R (2001) Footpaths in the stuff swamp. In Fowler W and Hasebrook J (eds) WebNet'2001, World Conference of the WWW and Internet, Orlando, FL, October 23-27, 2001, pp. 323-328

[17] Farzan R and Brusilovsky P (2005) Social navigation support in E-Learning: What are real footprints. In Anand SS and Mobasher B (eds) IJCAI'05 Workshop on Intelligent Techniques for Web Personalization, Edinburgh, U.K., August 1, 2005, pp. 49-56, also available at http://maya.cs.depaul.edu/~mobasher/itwp05/final/Paper7Farzan.pdf

[18] Farzan R and Brusilovsky P (2005) Social navigation support through annotation-based group modeling. In Ardissono L, Brna P and Mitrovic A (eds) User Modeling 2005. 10th International User Modeling Conference, Edinburgh, UK, July 24-29, 2005. Lecture Notes in Artificial Intelligence, Springer Verlag, Berlin, pp. 463-472, also available at http://www2.sis.pitt.edu/~peterb/papers/FarzanBrusilovskyUM05.pdf

[19] Fischer G (2001) User modeling in human-computer interaction. User Mod & User-Adapt Interact 11: 65-86

[20] Flexer A (1999) On the use of self-organizing maps for clustering and visualization. In Zytkow JM and Rauch J (eds) Third European Conference on Principles of Data Mining and Knowledge Discovery, PKDD'99, Prague, Czech Republic, pp. 80-88, also available at http://citeseer.ist.psu.edu/flexer99use.html

[21] Fox E, Hix D, Nowell LT, Brueni DJ, Wake WC, Lenwood SH, and Rao D (1993) Users, user interfaces, and objects: Envision, a digital library. J Amer Soc Inf Sci 44: 480-491

[22] Ghaphery J (2002) My Library at Virginia Commonwealth University: Third Year Evaluation. D-Lib Magazine 8

[23] Kohonen T (1995) Self-organizing maps. Springer Verlag, Berlin

[24] Kohonen T, Kaski S, K. Lagus, Salojärvi J, Honkela J, Paatero V, and Saarela A (2000) Self organization of a massive document collection. IEEE Transactions on Neural Networks 11: 574-585

[25] Korfhage RR (1997) Information storage and retrieval. Wiley, N.Y.

[26] Kurhila J, Miettinen M, Nokelainen P, and Tirri H (2002) EDUCO - A collaborative learning environment based on social navigation. In De Bra P, Brusilovsky P and Conejo R (eds) Second International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH'2002), Málaga, Spain, May 29-31, 2002, pp. 242-252

[27] Leuski A and Allan J (2004) Interactive information retrieval using clustering and spatial proximity. User Mod & User-Adapt Interact 14: 259-288

[28] Levy AY and Weld DS (2000) Intelligent internet systems. Artif Intell 118: 1-14, also available at http://citeseer.nj.nec.com/levy00intelligent.html

[29] Mathé N and Chen J (1996) User-centered indexing for adaptive information access. User Modeling and User-Adapted Interaction 6: 225-261

[30] Mitsuhara H, Ochi Y, and Yano Y (2002) Open-ended adaptive system for facilitating knowledge construction in Web-based exploratory learning. In De Bra P, Brusilovsky P and Conejo R (eds) Adaptive Hypermedia and Adaptive Web-Based

Systems. Second International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH'2002), May 29-31, 2002, pp. 547-550

[31] Norman DA (1993) Things that make us smart. Perseus Books, Cambridge, MA

[32] Olsen KA, Korfhage RR, Sochats KM, Spring MB, and Williams JG (1993) Visualisation of a document collection: The VIBE system. IPM 29

[33] Recker MM, Walker A, and Lawless K (2003) What do you recommend? Implementation and analyses of collaborative information filtering of web resources for education. Instr Science 31: 299-316

[34] Shipman F, Furuta R, Brenner D, Chung C, and H. H (2000) Guided Paths through Web-Based Collections: Design, Experiences, and Adaptations. Journal of the American Society of Information Sciences 51: 260-272

[35] Spink A and Hurson AR (2001) Modeling interactive information retrieval (IR): An information seeking approach. In Smith M, Salvendy G, Harris D and Koubek RJ (eds) Usability evaluation and interface design. 9th International Conference on Human-Computer Interaction, HCI International'2001, New Orleans, LA, August 8-10, 2001. 1, Lawrence Erlbaum Associates, Mahwah, NJ, pp. 888-892

[36] Vesanto J (1999) SOM-based data visualization methods. Intelligent Data Analysis 3: 111-126, also available at http://citeseer.ist.psu.edu/vesanto99sombased.html

[37] Weber G and Specht M (1997) User modeling and adaptive navigation support in WWW-based tutoring systems. In Jameson A, Paris C and Tasso C (eds) User Modeling. 6th International Conference on User Modeling, Chia Laguna, Sardinia, Italy, June 2-5, 1997. SpringerWienNewYork, Wien, pp. 289-300, also available at http://www.psychologie.uni-trier.de:8000/projects/ELM/Papers/UM97-WEBER.html

[38] Wexelblat A and Mayes P (1999) Footprints: History-rich tools for information foraging. In ACM Conference on Human-Computer Interaction (CHI'99), Pittsburgh, PA, pp. 270-277

[39] Yang CC, Chen H, and Hong K (2003) Visualization of large category map for Internet browsing. Decision Support Systems 35: 89-102