EDUCOSM - Personalized Writable Web for Learning Communities

Miikka Miettinen
Petri Nokelainen
Patrik Floréen
Henry Tirri
Complex Systems Computation Group
Helsinki Institute for Information Technology
Finland
firstname.lastname@hiit.fi

Jaakko Kurhila
Department of Computer Science
University of Helsinki
Finland
kurhila@cs.helsinki.fi

Abstract

Many of the possibilities of Web-based education are still unexplored. It seems that novel ways of thinking about both learning and technology are needed to get beyond the limitations of the traditional classroom setting. In this paper we introduce EDUCOSM, which focusses on the possibilities of collaboration and the open-ended nature of the Web. Its main features include sharing and annotation of arbitrary Web-pages, and an adaptive desktop for accessing the evolving contents of the system. EDUCOSM has been used in a real Web-based course, and the experiences are discussed along with a description of the tool. Although the approach requires both the teacher and the students to rethink their roles, the feedback received so far has been encouraging.

1. Introduction

The potential of Web-based learning environments is vast but – even today – largely unexplored. Most existing systems focus on providing a medium for the instructor to supply material and assignments to the students, who work according to a predetermined plan. Rudimentary communication tools are available, but they are rarely integrated to the actual learning process. In essence, the system constitutes a virtual classroom, where the instructor specifies and guides the activity. Obviously, this is not the only alternative. In order to realize the full potential of information technology, we need to "step out of the box" and envision new ways of organizing education that do not necessarily resemble the traditional classroom setting. Some tentative steps towards more meaningful use of technology have been taken in various areas, for example in supporting group formation and peer-helping (see [12, 4, 3] for examples).

This paper introduces a collaborative learning environment called EDUCOSM, which attempts to replace the instructor and course material –centered approach with an open-ended knowledge construction process. Both pedagogical and pragmatic considerations support this shift, as can be seen in recent research literature [10, 1, 2]. The main theoretical principles behind EDUCOSM include student-centered learning [1, 9] and evolutionary learning in communities with joint needs and goals [7]. When successful, this kind of a pedagogical approach is known to result in improved motivation and deeper understanding. It should also be noted that engagement in collaborative knowledge building excercises a much broader set of skills than traditional education. The abilities to search for, evaluate, collaborate on, synthesize, refine, and create knowledge are considered vital in the information society.

The shift in viewpoint leads to rather different requirements for the learning environment. Active participation in the collaborative process plays an essential role, and should be made as easy and natural as possible. One aspect of collaboration is to provide the results of individual work to others to utilize and elaborate on. The members of the community should be able to search, share, process, and publish information in a distributed manner, and the division of work should be encouraged rather than prohibited. On the other hand, effective tools are needed to support communication and group work. Of particular importance is the fact that these processes almost always take place in the context of certain pieces of information, and the normal separation between communication facilities and content is harmful. Therefore, the system should allow discussions to be attached to documents and document fragments, rather than adopt the popular idea of a separate discussion forum. The final important requirement is openness. From the distributed and student-centered nature of the process it follows that it cannot and should not be restricted to a predefined set of documents or even topics. The instructor has an important role in guiding the activity, but its exact direction is determined primarily by the interests of the community members. As a result, the system should allow any Web pages to be incorporated into the community's resource pool.

EDUCOSM was designed according to these principles. More specifically, it has the following features:

- A shared document pool, where the course participants can collect Web-resources for everyone to use. The numerous useful resources that are already on the Web offer a starting point for learning.

- Collaborative annotation of the documents, so that the annotations are visible to others in the system, even though the documents reside anywhere on the Web.

- Publication of the students' own work, so that it is on an equal footing with the other resources. In particular, it can be annotated and discussed effectively.

The underlying idea behind EDUCOSM is to make the learning process transparent for everyone so that it becomes a joint knowledge building effort. Effective communication tools make it natural for the students to rely on each other for help and elaboration of fruitful lines of inquiry. On the other hand, the role of teachers and tutors is very similar to everybody else. In a sense, the line between teachers and students is vague, and the responsibility for learning is shared by all members of the community.

The rest of this paper describes the EDUCOSM tool in from the user's point of view, and presents the technical details of its implementation. We also present some empirical evidence from a Web-course in computer science.

2 EDUCOSM Tool

2.1 Features and their use

EDUCOSM appears to the user as a button bar at the top of the browser window and a custom popup menu that is available on any page being accessed through the system (Fig. 1). The button bar is used for navigating between the various views, including desktop, search and filter creation views, which are described below. Functions for handling individual documents are located in the popup menu. They allow the students to add new material to the system and create annotations and newsgroups.

The basic idea is best illustrated by a scenario. Suppose that the student has become interested in a certain topic, and wants to find more information about it. From the search view he can send queries to both the system's internal search engine and Google. Internal search covers only the documents, annotations and newsgroup messages currently included in EDUCOSM, while Google can be used for searching the entire Web. The third option is to give the URL manually. The popup menu is available on all pages, allowing the student to add new material to the system with a couple of mouse clicks. The right mouse button makes the menu visible, and an option labeled ”Add to EDUCOSM” is chosen to include the visible document in the environment. In addition, the student is asked to assign the document to one of the topics of the course. The purpose of the topics is to facilitate personalization and organization of the contents of the environment. When a document is added to the system, it becomes visible on index pages and internal search, it can be annotated, and the system can recommend it to other potentially interested users. In other words, it becomes available for the entire community to collaborate on.

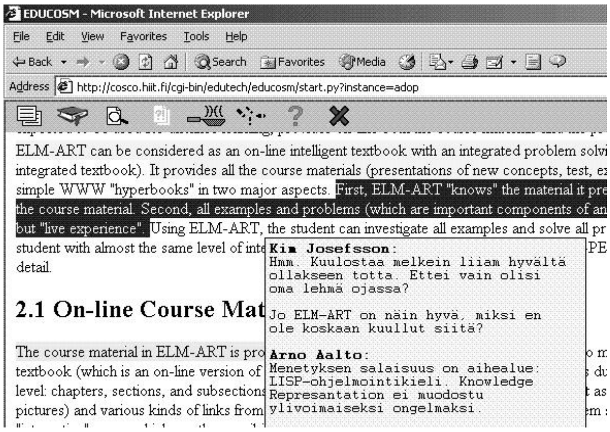

Figure 1. EDUCOSM user interface showing highlights and comments by two students.

The primary means of collaboration are annotations, hierarchical newsgroup discussions and publication of the students' own writings. There are two different types of annotations: highlights and comments. Highlights can be applied to mark important parts of the text, analogously to the way many people underline text on paper. In practice, highlighting involves selecting the text with the mouse, right-clicking the mouse to make the popup menu visible, and choosing the ”Highlight” option from the menu. Comments work the same way, except that they include the students' own reflections appearing as a tooltip when the mouse pointer is placed on top of the commented text fragment (see Figure 1). Space for longer discussions is available in newsgroups, which can be attached to individual documents. The students can publish their own work along with other resources, and all the same collaboration tools are available for discussion and feedback on their texts.

To return to our scenario, the extent to which collaboration actually takes place around the newly added document depends largely on its relevance to the interests of the course participants. Some documents are studied, annotated and discussed comprehensively, whereas others may not have much overall impact on the community. While course assignments and other requirements determined by the instructor certainly influence the activity, it is important to realize that it is the students themselves who are responsible for the actual learning process. In fact, the role of the instructor is very similar to that of a student. The instructor can participate in the discussion and perhaps provide some useful documents, but so can everybody else. One of our fundamental assumptions is that given the right kind of environment, the students do not need additional ”guidance from the top”.

Although collaboration is a crucial part of studying in EDUCOSM, it may occasionally happen that the student does not want to see all of the document's annotations simultaneously. This need can be addressed with filters that define the group of students and the time period from which annotations are shown. For example, the student may choose to see only the latest annotations of certain peers, or hide all annotations completely. Filters are created on a particular page, where the student can select the people from a list and optionally limit the time period to the last day, week or month.

Additional features include a personalized desktop and internal bookmarks. The desktop serves as the entry point to the system, and contains links to announcements, assignments, articles and newsgroup messages. The announcements and assignments are posted by the instructor, and are typically the same for everybody, although it is also possible to make them visible only to certain groups of people. The article and newsgroup links are recommendations generated for the particular user by heuristic algorithms (see the next section for details). The desktop also contains links to indexes that list all the articles and newsgroups in the system. Bookmarks work in the same way as in ordinary Web browsers, except that they are stored on the server to make them available from any client machine.

2.2 Technical implementation

From a technical point of view, EDUCOSM consists of a collection of server-side CGI-scripts and an HTML and JavaScript based client that runs inside an ordinary Web browser. The role of the server is to store data and act as a proxy between the client and the rest of the Web. The essential data includes user profiles, annotations and newsgroup messages, as well as various statistics regarding the activities of the users. The documents, on the other hand, are not copied to the EDUCOSM server. When the user requests a page, the server reads it over the network from its original location and inserts the annotations, menus and other application specific data before sending it to the client. This kind of a solution avoids copyright problems, and is in harmony with the distributed and open-ended nature of the system.

The display of the data and interaction with the user are handled by the client. Technically speaking, the annotations and menus are additional HTML elements that are incorporated into the document from JavaScript. Event handlers associated with the elements control their visibility and location on the screen, respond to user input and send updates to the server. Unfortunately, differences between various JavaScript implementations make it necessary to rely on some browser specific code, but the majority of it is isolated in just a few routines that connect directly to the browser interface.

In a sense, the key idea of EDUCOSM is to provide the community with a shared view to the Web. The server acts as the window through which everything is seen. In order to take the shared viewpoint, the user goes to the server and logs in, after which all data is routed along the same pipeline. Our current implementation of this mechanism is based on a simple manipulation of the links and other navigation paths on each page that is sent to the user. More specifically, the direct references are replaced with the URL of a CGI-script on the EDUCOSM server. The script receives the actual location of the document as a parameter, reads it over the network, and repeats the procedure to ensure that further requests are also directed to the EDUCOSM server. The next step is to check if the document has already been selected to the community's collection of resources, and add a status variable to inform the client about the correct operating mode. Other important data to be sent to the client may include annotations and the URL of the document's newsgroup. With this additional information, the client is able to supplement the publically available resource with the contributions of the community members.

Annotations are received from the server as descriptions, and are transformed to HTML elements as part of the document loading process. The application of filters as well as splitting and merging of overlapping annotations also happen at this stage. The positioning of the annotations is complicated by the fact that the original document may have changed after the annotations were created. The effect of the changes on the validity of a particular annotation is extremely hard to determine automatically, and we have ended up with a relatively simple solution. For each annotation, we save the underlying text fragment (corresponding to the selection made by the user) plus 40 words of surrounding text. The surrounding text is used for finding the right position in cases where there are several occurrences of the underlying fragment. In order for the annotation to be displayed, the underlying fragment must be completely intact, whereas changes in the surrounding text are permitted. The fragments are located with an approximate matching algorithm based on comparing sets of n-grams. It is hard to make exact predictions about the robustness of this solution, but it can be assumed that the majority of documents are stable or undergo only minor changes while being used in EDUCOSM.

The amount of content in a particular EDUCOSM environment may grow quite large over time. In order to make it easier for the users to find the most relevant information, the system provides an internal search engine, document and newsgroup indexes, bookmarks, and a personalized desktop view, which all attempt to address slightly different needs. The search engine is most useful when the user knows specifically what he is looking for, and is able to come up with an effective combination of keywords. Often this may not be the case, however, as both the accumulation of resources and collaboration take place in a completely distributed way. Looking at the index pages is a better way of getting an overall picture about the contents of the environment. The links are supplemented with flags and statistics to make it easier to identify new information. In the document index, the system marks the links that have not been visited by the particular user, and reports the number of annotations created since the last visit. Similarily, the newsgroup list shows the number of unread messages along with the total number of messages in the newsgroup. The role of the personalized desktop is to provide links to a few documents and newsgroup messages, which are expected to be particularily relevant to the user.

The current implementation of the internal search engine is based on simple word matching. The result listing consists of links to the items and sample text fragments, sorted according to the frequencies of the matching words. In addition to the documents, the search covers highlights, comments and newsgroup messages. The user can select any combination of these to be included in the results.

As an open ended and constantly evolving system, EDUCOSM is a challenging environment for personalization. Large amounts of potentially relevant data are accumulated, but straightforward implementations of techniques like collaborative or content-based filtering are not directly applicable. Instead, we resort to heuristic algorithms in choosing the document and newsgroup links for the desktop.

Documents are evaluated along two dimensions: the quality of the individual document and the relevance of its topic to the particular user. The estimation of quality is based on the number of annotations and the amount of reading time accumulated by the users. Some additional computations have to be performed to deal with the varying ages and lengths of the documents, however. In particular, new documents should be competitive against older ones and short documents against longer ones. The desired effect is achieved by using the average visit length instead of the absolute reading time, and by dividing both numbers by the word count of the document. Since the resulting quantities are by no means comparable, it is more convenient to define a relation for sorting the documents than try to combine the scores numerically. In our current implementation, the order of a pair of documents is determined by the annotation score if one document is above the average and the other below, and otherwise it is determined by the reading score. The second step is to select the recommendations from the sorted list. The relevance of each topic is estimated on the basis of the distribution of the user's activity during the past week. The activity statistics that can be attributed to topics include navigation from the document index, bookmarking, navigation from search results, annotations, and newsgroup postings. The items that appear on the desktop are unread documents chosen from the front of the sorted list according to the activity distribution of the user.

Newsgroup recommendations are similarily based on ordering the messages according to certain criteria. The highest priority is given to messages that are replies to the particular user. Next come the messages that start a new thread of discussion, and after that those which are attached to a different document than any of the predecessors. Finally, messages that are closer to the start of a discussion thread or have more replies are preferred. The set of recommendations is created by picking unread messages from the beginning of the list.

3 Empirical evaluation

3.1 Study setting

EDUCOSM was used in an advanced computer science course titled ”Adaptive educational systems”. As opposed to standard lecture courses, the course was organized as a seminar. The difference is that every student must pick a topic related to the course title, and prepare an oral presentation and a written paper on it. The course was structured to consist of separate tasks including (1) searching the Web for related material, such as research reports or other scientific resources, (2) selecting one or two articles and preparing a summary of them, (3) preparing a short oral presentation about the selected topic as an introduction for a roundtable discussion, and (4) preparing a final paper on the selected topic.

Although the course tasks were more or less traditional, the students were required to make their action (i.e. their learning process) visible to the other students by using the EDUCOSM system. Articles were brought to EDUCOSM, and they were actively annotated by the students. The students' summaries and several different versions of the final papers were also brought into the system for other students to see and comment on.

Course grading affects the way students participate. To make the process meaningful, the use of EDUCOSM was encouraged by giving part of the grade solely on the basis of the activity in the system. Specifically, the grade had four components: written paper (35% of the grade), oral presentation (15%), discussion and commenting (25%), and overall activity during the course (15% on the basis of peer-assessment). Overall activity included the number and quality of articles found and the time spent in the system.

Twenty-four students participated in the course. The duration of the course was six weeks. The students were computer science majors, both at the graduate and undergraduate level, obviously familiar with computers and various kinds of application software.

The course included only two face-to-face meetings. During the first meeting the course structure was explained and the details of taking the course, including the grading policy, were agreed upon together with the students. The use of the EDUCOSM tool was also explained. The second face-to-face meeting was the roundtable-discussion lasting four hours, giving each student roughly ten minutes for the presentation and discussion. It should be noted that the students were mostly strangers to each other, so they mainly formed their opinions of the other participants based on the activity visible in the system.

3.2 Results

After the course the students were provided with a questionnaire asking them to reflect on various aspects of their experience with EDUCOSM. On the one hand, the purpose of the questionnaire was to evaluate the usefulness of the system and gather feedback for further development. The students were asked to comment on each of the main features of the system, as well as its general value and applicability. Another objective was to obtain qualitative data for the analysis of the learning process. The main results of the study are summarized below.

Comments and highlights. The course participants greeted the annotation functionality with enthusiasm. Although they were encouraged to be active in the system, the amount of comments (total of 693) and highlights (1484) produced during the course was surprisingly high. Intuitive usability and perceived usefulness of the annotations received positive feedback:

> ”Comments seemed to replace discussion forums, and the reason to my mind was that one was able to attach comments more easily to the context. It is a very useful feature [...]”

> ”Beautiful!!! Magnificent discussion in the right context.”

> ”[In the comments] you can question your own text and write something that is not scientifically valid but you want to share it with others anyway.”

Even changes in attitudes were reported, even though the exposure to the system was relatively short:

> ”I was a bit shy with commenting. Before EDUCOSM, my attitude towards other students' highlights and annotations in course textbooks was negative. EDUCOSM changed my attitude to more positive, but I still did not write every comment that came to my mind. I could probably get rid of my self-censorship when getting used to the system.”

Common document pool. The common document pool that serves as a basis for knowledge building is in harmony with the student-centred learning principles, as it shifts the responsibility for learning to the students. This point was not missed, as seen in a quote:

> ”[The added value is that] the course contents evolve to the direction that the students want, because they can choose the material that is brought into the system.”

However, the instructor can also contribute to the document pool in the same way as the students. In this experiment, the instructor brought some documents into the system, but they did not have any special label and were therefore indistinguishable from the contributions of the students. This was done on purpose so that the instructor's choices would not affect restrictively to the actions of the students.

Annotation filters. Since the course was fairly small (24 students), there was little need to use filters to eliminate overwhelming annotations. The participants used mainly ”everybody” and ”nobody” filters, but the need for other types of filters in other courses was recognized by several students:

> ”I added highlights only to places where there were no highlights. Had there been more highlights, I would have used filters and made more highlights myself.”

> ”The only filtering I made was to eliminate all annotations from the text. When I tried to read and concentrate on a specific article, I sometimes felt that annotations by others were distracting and wanted to eliminate them.”

> ”I tried these [filters] out, but I used them very little. Comments made by others were nice to read, so I left them visible along with my own [comments]. At some point I noticed that if I had turned off other students’ comments, I probably would have commented the articles more. Now the articles were already full of comments, if the comments were left to be shown...”

Overall evaluation. The strong points of EDUCOSM compared to mainstream learning environments were also noticed. When asked how EDUCOSM changed the learning process compared to the previous seminars the students had taken, one student stated that:

> ”It supported discussion well. Usually, when I am preparing a final paper for a seminar, I discuss with my room mates. Now I was able to discuss with other course participants.”

The quote above stresses the point that even if the time and space constraints are relaxed much the same way as in conventional Web-based education, the features of EDUCOSM enable learners to discuss their on-going work more effectively. The transparency and openness of the learning process was seen as a plus in other comments as well:

> ”To let others review [my] work-in-progress is a big plus that I want to experience again.”

> ”[The added value is that] the activities take place in the system itself, i.e. not using outside links.”

4 Conclusions and future work

When requirements for education are changing due to changes in society, proper tools can be employed to make the learning process more meaningful and efficient. This paper presents EDUCOSM, a tool to search, process and publish information as a joint effort. The potential of the Web can be exploited through the collaboration of learners in a learning community. EDUCOSM is a step towards those novel ways of using information and communications technology to support learning in a way that is not ”a dull approximation of the existing school system” [8].

The first real-life trial with EDUCOSM was an encouraging experience: intuitive and valuable uses for the various features of the system were easily found. One missing ingredient is awareness of other learners present in the environment. Collaboration space awareness has been identified as an important issue (see e.g. [11]). It can be achieved in a virtual environment by using real-time social navigation [6]. Positive experiences with tools that enhance the sense of other learners in a learning community have been reported [5]. These features would fit into EDUCOSM as well.

References

[1] R. Felder and R. Brent. Navigating the Bumpy Road to Student-Centered Instruction. *College Teaching*, 44:43–47, 1996.

[2] R. Felder and R. Brent. Effective Strategies for Cooperative Learning. *Journal of Cooperation & Collaboration in College Teaching*, 10(2):69–75, 2001.

[3] J. Greer, G. McCalla, J. Vassileva, R. Deters, S. Bull, and L. Kettel. Lessons Learned in Deploying a Multi-Agent Learning Support System: The I-Help Experience. In *Artificial Intelligence in Education (AIED 2001)*, pages 410–421. IOS Press, 2001.

[4] A. Inaba, T. Supnithi, M. Ikeda, R. Mizoguchi, and J. Toyoda. How can we form effective collaborative learning groups? In G. Gauthier, C. Frasson, and K. VanLehn, editors, *Intelligent Tutoring Systems (ITS 2002)*, Lecture Notes in Computer Science 1839, pages 282–291. Springer, 2000.

[5] J. Kurhila, M. Miettinen, P. Nokelainen, and H. Tirri. EDUCO - A Collaborative Learning Environment using Social Navigation. In P. de Bra, P. Brusilovsky, and R. Conejo, editors, *Adaptive Hypermedia and Adaptive Web-Based Systems (AH 2002)*, Lecture Notes in Computer Science 2347, pages 242–252. Springer, 2002.

[6] A. Munro, K. Höök, and D. Benyon. Footprints in the snow. In A. Munro, K. Höök, and D. Benyon, editors, *Social Navigation of Information Space*, pages 1–14. Springer, 1999.

[7] M. Scardamalia and C. Bereiter. Computer-Support for Knowledge-Building Communities. *The Journal of the Learning Sciences*, 3(3):264–283, 1994.

[8] R. Schank. The Computer isn’t the Medium, It’s the Message. *Communications of the ACM*, 44(3):142–143, 2001.

[9] J. Sears and J. Marshall. *Teaching and thinking about curriculum: critical inquiries*. Teachers College Press, 1990.

[10] R. Snow, L. Corno, and D. Jackson. Individual differences in affective and conative functions. In D. Berliner and R. Calfee, editors, *Handbook of educational psycholog*. Simon & Schuster & MacMillan, 1994.

[11] H. Wang and Y. Chee. Supporting Workspace Awareness in Distance Learning Environments: Issues and Experiences in the Development of a Collaborative Learning System. In C.-H. Lee, editor, *International Conference on Computers in Education/Schoolnet 2001*, pages 1109–1116, Seoul, Korea, 2001. Incheon National University of Education.

[12] M. Wessner and H.-R. Pfister. Group Formation in Computer-Supported Collaborative Learning. In S. Ellis, T. Rodden, and I. Zigurs, editors, *ACM SIGGROUP Conference on Supporting Group Work*, pages 24–31, New York, 2001. ACM Press.