Bos, N., Drabenstott, K., Krajcik, J., Soloway, E., Talley, M., Woolridge, S., & Miller, J. (2000). Students' Searching and Evaluating with the Artemis Interface to a Digital Library. In B. Fishman & S. O'Connor-Divelbiss (Eds.), *Fourth International Conference of the Learning Sciences* (pp. 150-156). Mahwah, NJ: Erlbaum.

Students' Searching and Evaluating with the Artemis Interface to a Digital Library

Nathan Bos, Karen Drabenstott, Joseph Krajcik, Elliot Soloway, Melanie Talley, Steven Woolridge, and John Miller
CREW 701 Tappan St suite C2420 Ann Arbor, MI 48109
(734) 647-7730
nathan.bos@umich.edu

Abstract: This paper describes the design and evaluation of the Artemis interface to the University of Michigan Digital Library, which was built to assist students in searching on-line collections. Noteworthy features of Artemis include the ability to store search results in an on-line workspace, browse collections as well as keyword-search, and access diverse collections (both Web pages and proprietary collections) with a single agent-based architecture. To evaluate Artemis, we conducted in-situ videotaped sessions with students using Artemis in a project-based curriculum. We supplemented this with analysis of log file data collected at the server level. Analysis of videotaped sessions show that students spend a considerable percentage of their in Artemis evaluating search results on a number of levels. Log files show that students also developed fairly sophisticated search strategies, using both browsing and searching, and constructing multiple-keyword searches. Shortcomings and future directions of the Artemis system are also discussed.

Keywords: digital-library, middle-school, high-school, science education

Perspective on Searching and Inquiry

Problems and Promises of Searching on the World-Wide Web

Research on students' searching with electronic databases identifies a number of common problems. Students typically have trouble formulating keyword searches in unfamiliar content areas, sorting through long lists of search results, and organizing and storing results of their searches. (Marchionini, 1995; Moore, 1995; Wallace and Kupperman, 1997). For students, search sessions in the library are often disorienting and unproductive.

Searching Should be part of an Integrated Process of Inquiry

Searching is too often treated by students (and researchers) as an isolated task to be done once at the beginning of a research project. When assigned to search for research materials, students sometimes treat the search as an end in itself, and do not evaluate search results in the light of their broader research questions and information needs. But searching is more effective when it is part of an ongoing, cyclical research process (Kulthau, 1996) which includes identifying questions, searching, experimenting, writing, and raising of more research questions and additional searching as needs arise. This need for a cyclical research process is particularly apparent in inquiry-based curriculum models where students have considerable freedom to design and purse their own questions over time. Teaching students to *evaluate* search results is a key focal point for making searching more process-oriented. An expert searcher monitors the process and evaluates results all along as to their relevance to questions at hand. Evaluation can also be done at a number of levels: students should know how to evaluate individual documents, abstracts, titles in a list of search 'hits', or entire collections.

Problems and Promises of the Web

The challenges of searching and evaluating are exacerbated when the venue being searched is large, distributed collection such as the World-Wide Web. The Web is vast and uncontrolled. Unlike proprietary databases, on the Web no single agency takes on the jobs of screening content, writing abstracts, or indexing by keywords. Instead, the Web is dynamically updated by many different players, with different perspectives, levels of knowledge, etc. Some high-quality sites are well-organized and offer useful search features, but there is no standardization between these sites. Internet search engines offer a range of clever solutions to searching and indexing problems, but none of these offers features such as quality filtering and abstract-writing.

Even as it presents new problems, though, the Web also offers stunning new possibilities. The Web could be a central information-repository like no other, bringing together publishers, government agencies, non-profit and for-profit corporations, and individual information-providers in ways never before possible. The Web can allow full-text access to content more easily than other media, which is important for resource-thin K-12 settings. The Web is accessible from any networked computer, allowing learners access in the classroom (instead of the school library) or away from school. And the Web may offer a place for students to share their knowledge with each other, becoming information-providers as well as information-consumers.

Design of Artemis Addresses Needs of Student Information-seekers

The Artemis interface, and the digital library that is provides access to, were designed to address a set of known challenges with student searching, and to support information-seeking as part of an ongoing research process. Visible features of the Artemis interface are labelled in figure 1. (See Wallace, et al. 1998 for a longer description of Artemis' design.)

Students Need Help Generating Search Terms

Because they are usually not experts in the domains where they are working, students often have not mastered the basic vocabulary of a scientific sub-discipline. Without help, a student searching for, say, information on the asteroid collision suspected to kill the dinosaurs will often generate keyword searches that are far too broad (e.g. 'space', 'death', or 'extinction') or are off the mark (e.g. 'space rocks', 'deadly collisions'). A good solution to this problem is to give the option of *browsing* a list of search terms rather than generating new keywords. Given a list of candidate terminology, students can often select more appropriate search terms. Also, browsing is an opportunity for tacit learning, as students pick up the vocabulary of a sub-discipline by working with it. Artemis features a browsable list of terms and phrases in a hierarchical list, taken from the widely-used Sears terms (Miller, 1994).

Students Need Ways to Save Information from Searches

Students often do not have good ways of saving information that results from searching. Because students avoid spending the time for physically finding books and magazines in the library, and the cost of photocopying and printing that may or may not prove to be useful, much quality information is lost. Artemis provides persistent bins, or 'Driving question folders' where students can store search results. Students simply drag-and-drop results from a list of search 'hits' into one of these named bins, and a pointer to this found information is stored permanently, or until the student chooses to delete these items. Because only pointers to the information are stored, the cost of storing many search hits on the server is trivial, and students are encouraged to make heavy use of the driving question folders.

Students Need Support in Staying Focused and Organizing their Work

Even when students save information in paper or note form, they sometimes do not organize material in logical ways, or fail to keep track of papers and notes, and material is unavailable when needed. Artemis supports sorting and storing of information in logical partitions because it allows creation of unlimited numbers of Driving Question folders. Students are encouraged to create multiple folders with informative names (e.g. 'When did dinosaurs live?' and 'What killed the dinosaurs?') and can drag relevant search hits into the appropriate folder immediately. Students are encouraged to name these folders with question names rather than topics, fitting with the inquiry-based philosophy in which Artemis is used.

Students Need Help in Quickly Identifying Appropriate Content, and 'Vetting' the Vast Internet Collections

Especially when searching on the Internet, there may be vast collections available, but a small fraction of it may be appropriate for students at a particular level. Much Internet-based information is either too technical, or too shallow for student use. Also there is a great deal of information that is unreliable (e.g. personal opinion expressed as fact) or uninformative (e.g. a search for 'dinosaurs' yields dozens of museum home pages with driving directions and exhibit hours). These problems also exist, to a lesser extent, in all other on-line collections. The problem of vetting appropriate information led Artemis' designers to employ a team of 'Cybrarians' (mostly master's degree students from the University's School of Information) who searched available Web collections and identified useful and appropriate resources. Good resources were categorized with appropriate search terms (matching with the available browse terms) and abstracts were written to aid student evaluation. These vetted Web collections were presented in search results as a separate 'Web collection', distinct from Web sites returned from at-large, unvetted searching. One question for our evaluation research would be to find out whether students used these abstracts, and whether they distinguished between different collections.

Students Need to be Able to Access Full-text Articles from Multiple Locations

K-12 school libraries do not have the budgets to keep vast collections of specialized scientific resources on-hand. Students need to be able to access full-text of many resources over the Web, and need to be able to access these wherever they are (not just in the school library). Artemis' Java-based interface makes full-text search results available from any networked computer without special client software. The agent-based architecture of the UMDL makes it a true 'distributed' system, allowing it to flexibly integrate multiple collections from different sources. Artemis could search collections from ProQuest, Elsevier, Grolier's Encyclopedia Americana, from the indexed Web collections, and use HotBot to search the Web at large.

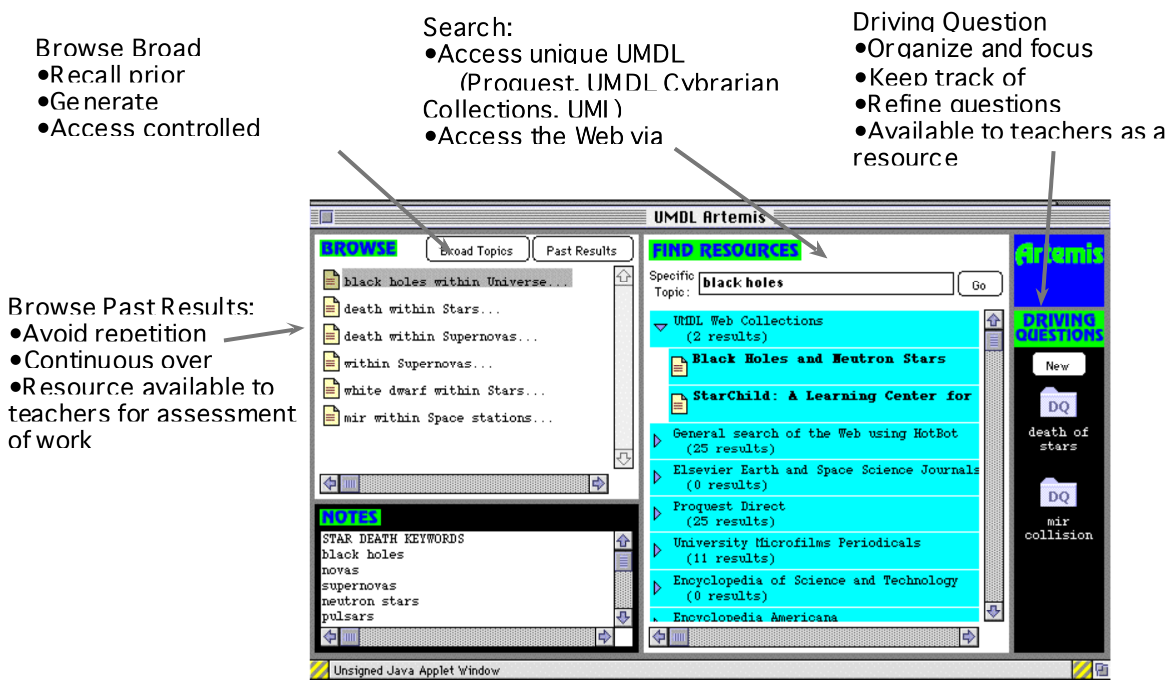

Figure 1: Artemis Interface and visible features

Settings for Evaluation

Students in two settings used Artemis during the 1997-98 school year. The first setting was a high school where the science curriculum was entirely project-based. All sophomore students in the second year of the three-year 'Foundations of Science' curriculum were introduced to Artemis during class, and then a small number were recruited to be videotaped doing some of their personal project research on Artemis after school or during lunch. Teachers were asked to recommend students who were fairly verbal (to ensure good verbal protocols) and were a cross-section of class achievement levels. The second setting for Artemis usage were middle school students. These students were taking part in structure inquiry-based curriculums using Artemis. These students were not the primary targets for the Artemis evaluation, but since we have detailed data on their search behaviors in the log file analyses, they are included as subjects. Both sets of students were attending public schools in a medium-size midwestern University town.

Space prevents a fuller explanation of the exact curriculums students worked within on these projects, or of project-based philosophy behind these curriculums. The Hi-CE web site (www.hi-ce.org) is also a good online resource about Artemis, project-based learning, and the digital library project.

Note on Combination of Methods

Our evaluation of Artemis included two main data sources: process video of selected students using the interface, and log file reports of overall system usage. Videotaped session of students using Artemis allowed fine-grained analysis of students behaviors, goals, and problems. Log files collected on the server side supplemented this, providing coarse-grained data on user actions that happen sporadically, over time, and in the field--types of actions that are difficult and expensive to capture with videotape or observation data. We are enthusiastic about the potential of combining these data sources-- one detailed but narrow, the other thinner but broader.

Process Video Data

We recorded eight student sessions with Atemis using process video equipment kits. This setup captures output from the students' computer screen and prints it onto hi-8 videotape, and records an audio recording of conversation between students onto the audio tracks of the same tape. If two students are present for the session, they each wear microphones, and their voices are captured separately onto left and right channels of the tape. Students were asked to talk aloud when they were working, and researchers used a prompting protocol (prompts such as 'what are you doing now?' and 'why did you just do that?') to keep them talking and focus their comments on the kind of information we needed. Even though these sessions took place outside of class time, it was important that they be in context. Students were not asked to do artificial tasks, but rather to search for information related to their current curriculum projects.

Eight students sessions with Artemis were coded for types of events and duration of events. Events were given a primary code, start and end times, and in some cases a secondary code. Primary codes divided sessions into states, including Log-in (time spent logging in), Demonstrate (researcher demonstrates an aspect of the system), Browse, Search, Collection (student is looking at collections results of a search), Title (students looking at titles of search results), Abstract, Document (student is looking at a specific document), Web (user has followed a link into the World-Wide Web) Problem and Question states. Secondary codes were assigned to more specifically designate types of events within broad categories. For example, within the Document category, students could be in the states of Document-Evaluate, Document-Scroll, Document-Problem (system problem of some kind within document state), Document-Comment (discussing something about the document with each other, not directly related to evaluating), Document-General (actions do not fit into other sub-category.) Single events that were not assigned a durations, such as Document-Open, were also coded. Reliability of this coding scheme was checked by comparing two independently-rated sessions and satisfactory agreement was achieved.

Log File Data

We captured data from all Artemis logins for the entire 1997-1998 school year. We removed all known logins by Artemis staff and developers during this time, yielding a database of 6,642 Artemis session. (Multiple logins by the same person on the same day were counted as single sessions.) Log files captured most interactions within the system, including searches performed, hits returned, and saves made into driving question folders.

All Artemis log files for the school year were captured, parsed, and written to a relational Filemaker Pro database (*Logger* system developed by John Miller). From there, they are searchable on any field using standard database query tools available in Filemaker. All log file analyses were done using queries on this database.

Results

Process video and log files allowed us to study how students *evaluated* search hits in Artemis, how they *constructed searches*, and how high school students used the *persistent workspace* features over time.

Evaluation of Search Results

Because evaluation was identified as perhaps the most important component of a process-oriented view of searching, we performed detailed analyses of how long students spent time evaluating search results as compared to other activities during the search session. We hoped to see a relatively large proportion of time spent searching, as opposed to formulating searches through browsing topics or performing keyword queries. This is not because we thought formulating searches was a waste of time, but rather because we wanted to see an integrated process of searching and thoughtful evaluation, rather than a string of poorly-evaluated searches.

We could not find previous research on how much time searchers might be expected to spend evaluating meta-data, so it would be a somewhat subjective judgement of how much was enough time. Further research could profitably compare meta-data use in different systems (e.g. Artemis versus a Web search engine), compare meta-data evaluation time with some criterion (e.g. search effectiveness) or do more detailed qualitative studies of how students go about this evaluation. For this study, we were satisfied to examine the raw percentages, gleaning what we could about our system's usage. These data may also be useful as baseline data for future research.

Eight process video sessions (as described in methods) lasting an average of 40 minutes each were analyzed. Figure 1. Shows the proportion of time students spent in five states. Time spent on the Web at large (which was uncoded as to whether or not it was evaluation or 'surfing') was excluded because it was of less interest to this study.

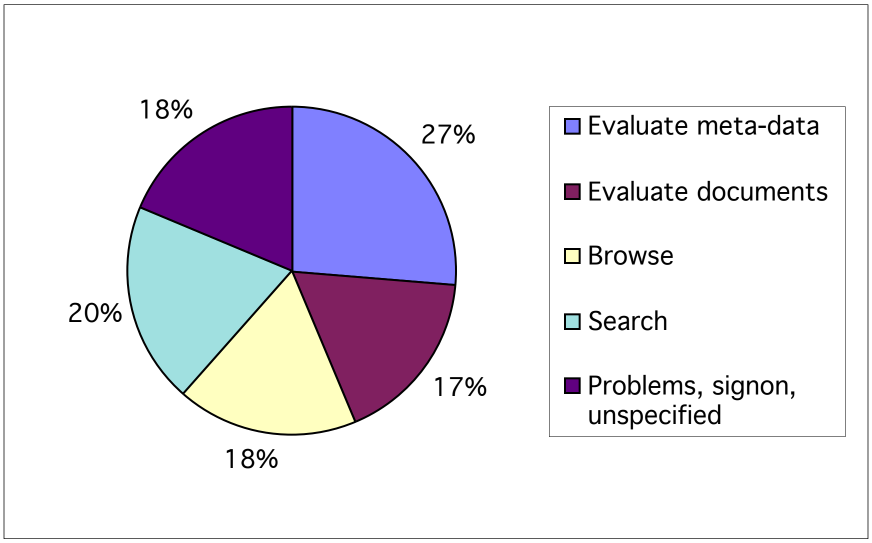

Figure 2 Percentage of time spent in five Artemis search states

Figure 2 shows that students spent a large amount of time evaluating meta-data and documents. (Meta-data is information about content, such as titles and abstracts, that are not actual content.) Evaluating meta-data includes activities such as evaluation of lists of collections (choosing which collection to open) lists of titles (choosing which title to look at first), and evaluation of document abstracts. This encouraging finding about evaluation of meta-data (27% of time spent) was somewhat of a surprise. It would have been easy and not unexpected for students to quickly skip over this data in a rush to get to the documents. Instead, however, we saw students spending time in thoughtful evaluation at every level. The verbal protocols give anecdotal examples of students looking at where information is coming from (collections-evaluate), making inferences about the usefulness of titles, and using abstracts to decide whether or not to proceed to accessing full documents.

Students also spent considerable time reading and evaluating the full-text of documents (17% of time spent). Again, we might have seen students hurriedly collecting all search hits, relevant or not, in an effort to 'finish' one search and get on to the next. Instead, we saw students reading and evaluating the usefulness of full documents, and making decisions about future searches on the basis of what they had and had not found.

Search Query Analysis

Using log files we examined how Artemis users performed searches during the 1997-98 school year. We had log file records of 17,169 Artemis searches over 6,642 sessions. In the average Artemis session, students performed 2.5 searches, and looked at 2.9 documents from their search results. This give a baseline measure of the success of the system at allowing large numbers of searches and document retrievals. Data shows that fairly complex searches were performed using this system. In 55% of sessions (3691/ 6642), students did at least one keyword search term (rather than relying solely on the browse function.) In 22% of session (1522/6642) students did multiple-keyword searches. Inspection of the set of complex searches shows that multiple-keyword searches were not the sole domain of the older students; most were performed by middle school students. This data suggests that given the proper support (e.g. browse tools, and instruction) students can do keyword searches. It may be appropriate to introduce more complex boolean search tools to students at the middle school level.

Persistent Information Usage

Log files also allowed us to examine student use of Artemis' information-storage capability. Students saved documents to their Driving Question folders 3,395 times, or an average of .51 documents per session. This means that 17.7% of all documents accessed were deemed worthy of being saved for later retrieval.

Most of these DQ folder saves were done by middle school students, as part of a structured curriculum that required them to do so. Data from high school students was not so encouraging. Login records show that high school students rarely used Artemis outside of the class sessions where it was explicitly required. The precise numbers were not deemed worth getting (because they were known to be low and would require considerable time), but HS student accesses to Artemis outside of class sessions were fairly rare. Instead, given free access to the Web, these students continued to rely on the Internet search engines they were accustomed to using. During structured class sessions students did use Artemis and saved items to the Driving Question folders. However, log files show that in most instances students did not access the system later to retrieve them, but rather printed out the materials they wanted to keep during the search session.

Why didn't Artemis become the search engine of choice for these high school students, given the advantages described earlier in this paper? Why didn't students use Artemis as a place to store their research? In-class observation and informal interviews with students yield some clues:

1) Established habits. Students in these classes were used to using Internet search engines, and had developed their own procedures with them.

2) Printing preferred to electronic storage. The persistent storage features of Artemis were not used because students had in-class access to laser printers and were in the habit of printing whatever they wanted to keep. Students preferred their paper-based system to Artemis' persistent storage feature because they could take paper versions home more easily, could highlight and annotate them, and could keep them in the same place as all their other (paper) class materials.

3) Artemis was not available everywhere. Because of licensing agreements with proprietary publishers, Artemis was only accessible from within the high school.

4) The Artemis Java applet had a long startup time. Artemis was built with a very early version of Java, and took some time to download (30 seconds to 1 minute). This delay was enough to discourage some students from starting it up, when they could access familiar Web search engines more quickly.

Despite these shortcomings of the first version of Artemis in providing a preferred venue for high school students, we are optimistic about this model for the future. Each item in the above list is potentially solvable, given the relentless advances in networking speed, computer power, and computer accessibility. There are good reasons, then, to continue to push the model or persistent, integrated interfaces to distributed collections.

Conclusion

The Artemis interface, with its underlying digital library, represent a better approach to student inquiry in electronic collections. Artemis allows keyword browsing as well as searching, allows full-text access to a variety of on-line collections through a single interface, and allows storage and retrieval of search hits.

Evaluation showed that students using Artemis spent a good portion of their time evaluating search results and search meta-data such as document titles, collection titles, and abstracts. These results are encouraging in that they represent a process approach to searching. Log files also show that students commonly performed multiple-keyword searches, suggesting a certain level of sophistication in search strategies. High school students did not use Artemis as often as we had hoped for their in-class research, as shown by inspection of the log files. Most of the apparent reasons for this lack of usage are things that can be corrected with better and faster technology, and so the prospects for future use of an Artemis-like model are bright.

A note on methods: combination of fine-grained video analysis and coarse-grained transaction log analysis was useful in getting multiple lenses into usage. This combination would be even more useful in situations where good repeated-usage data is available through log files, and can then be matched more closely with observations of individual students.

References

Bos, N.D. (In press.) Giving back to the Web: High School students' social filtering of scientific resources on the World-Wide Web. *Journal of Science Education and Technology.*

Kulthau, C.C. (1996). The process of learning from information. In C. Kulthau (Ed.) *The Virtual school library: gateway to the information superhighway.* 95-104 Englewood, CA: Libraries Unlimited.

Marchionini, G. (1995). *Information seeking in electronic environments.* New York: Cambridge University Press.

Miller, J. (Ed.) (1994). *Sears list of search headers.* New York: H.W. Wilson.

Moore, P. (1995). Information problem solving: A wider view of library skills. *Contemporary Educational Psychology,* 20, 1-31.

Wallace, R. & Kupperman, J. (1997). Online search in the science classroom. in E. Soloway (symposium chair) *Using online digital resources to support sustained inquiry learning in K-12 science classrooms.* Symposium presented at the annual meeting of the American Educational Research Association, Chicago, IL, March, 1997. [on-line] http://mydl.soe.umich.edu/papers

Wallace, R., Bos, N., Hoffman, J., Eccleston Hunter, C., Krajcik, J., Soloway, E., Kiskis, D., Klann, E., Peters, G., Richardson, D., Ronen, O. (1998). ARTEMIS: Learner-Centered Design of an Information Seeking Environment for K-12 Education. *Conference proceedings of CHI98, Human Factors in Computing Systems.* New York: Association for Computing Machinery

.

Acknowledgements

The authors wish to acknowledge the contributions of many more members of the UMDL and MYDL research groups than could be named here. This work was supported by the National Science Foundations Digital Libraries Initiative.