Interactive Learning with a Digital Library in Computer Science

Edward A. Fox
Department of Computer Science
Virginia Tech
Blacksburg, VA 24061-0106

Abstract

We have developed tools, expanded our digital library systems and content, and built almost 5000 "pages" of WWW-accessible courseware, increasing the interactivity and quality of learning about computer science. Evaluation has shown that learning practices have changed, most students are happy with the emerging infrastructure and pedagogy, and there is growth in both local and remote access to our http://ei.cs.vt.edu server.

Introduction

Key concepts of our 1993-1996 NSF-funded Education Infrastructure (EI) project [6] are to improve CS education by increasing interactivity and use of a digital library. The main objectives/accomplishments were to:

1. expand the content and software (especially interfaces [13, 15]) initially developed with NSF support of our "Envision" digital library project, "A User-Centered Database from the Computer Science Literature" [10],
2. develop/apply algorithm visualization tools [14],
3. incorporate use of specialized digital library systems like Netlib [3] into related courses,
4. add new courses related to human-computer interaction (HCI), multimedia (MM), and a freshman level introduction to Networked Information (NetInfo or NI for short),
5. significantly change courses like "Computer Professionalism" (CompProf), to make use of interactivity (e.g., asynchronous online debates) and digital library support (e.g., adding to a large History collection), and
6. apply the key concepts to improve other courses.

Resulting changes have far exceeded our expectations.

Changing Infrastructure

Over the course of this project, the infrastructure required for computer science education at Virginia Tech, and the way students learn the field, has dramatically changed. *Students now expect course materials to be online.* Late in 1995, I was threatened by a student in another department who thought I was preventing him from gaining access to (then not yet online) courseware!

In Fall 1993, we made Information Retrieval courseware available with Gopher, and began usability testing of Mosaic. By Fall 1994, our first "paperless" courses were in place, with all course materials online, and lectures given in Internet-capable rooms with LCD projection, typically of a World Wide Web browser. By early 1995 we began careful monitoring of servers and networks, so that statistics in this paper are based on 18 months of collected data. Dependence on and use of our primary server, a DEC Alpha, has expanded to such an extent that a second, mirroring server will be installed by 1997, with a 155Mbps ATM card and connection to support high bandwidth communications. For reliability, that machine will be placed in the Computing Center, several miles from the UPS-protected CS machine room that houses our main "EI" project server, ei.cs.vt.edu, accessible at http://ei.cs.vt.edu. During the same time, our sister-institution in this project, Norfolk State University (NSU), has upgraded from a single 56 Kbps to a T1 line, and an ATM-based state Wide Area Network (WAN) is undergoing pilot testing. NSU and Virginia Tech both have significantly expanded their undergraduate lab space and upgraded the computers and networking involved. Virginia Tech student network access has dramatically expanded, with faster modems in the campus pool, the addition of a local commercial internet service provider, and rapidly increasing numbers of ethernet connections in dormatories and local apartment complexes. Since all CS majors have a workstation running UNIX and/or Windows 95, those with 10BaseT jacks in their rooms are able to benefit from much better network performance and faster access to educational information, which they claim significantly enhances their learning experience. By late 1997 most residential students will have an ethernet connection.

Changing Pedagogy

With our new infrastructure becoming available, we have made use of labs and networking to better accomodate student preferences and to add interactivity to the learning experience. We have shifted certain courses from lecture-only, to having occasional sessions in a lab, or half of the classes in a lab, or (in select cases) all sessions in a lab. Even though our majors have bought workstations as entering freshmen since 1985 and use them extensively in both standalone and networked modes, students find the labs to be of great value, and benefit from the interactive learning that takes place in both "open" and "closed" labs. Results from a survey of students in the spring 1996 course on multimedia (MM) indicate a strong consensus that more time in lab is preferrable, and less time is desired for lectures. Similar comments came from many students in the freshman Operating System Tools (OS Tools). A survey administered to both the MM and OS Tools classes, as well as focus groups conducted for those courses, confirmed our earlier informal observation that students vary widely in their beliefs regarding where they learn best (e.g., in lecture, lab, or at home). We also found that though students differ in their rankings of the value of online resources and services (e.g., old quizes, email, listservs, newsgroups), most wanted information with clear practical use (e.g., lecture notes, assignments) plus some combination of passive and active communication mechanisms (e.g., WWW pages, email). Table 1 shows the amount of online material developed in four courses; seven others also have 175 or more HTML pages.

Course/Files	*HTML*	*GIF*
NetInfo	175	300
CompProf	600	350
Multimedia	525	250
InfoRetrieval	350	125

Table 1. Counts of Active Files for Selected Courses

Starting with InfoRetrieval and continuing with the Multimedia and NetInfo courses, we have adapted Keller's Personalized System of Instruction [11] to our networked environment. Students proceed at their own pace, study on their own, get help through asynchronous communication with peers and instructors, and in general have much greater flexibility learning. Many students prefer this type of course, and in the case of NetInfo we simply could not accomodate the demand any other way, in this time of scarce resources. However, students have requested that we add interim deadlines, since they tend to procrastinate and require help with time management.

Changing Evaluation

Our evaluation involves typical traditional methods, e.g., pre- and post tests, surveys, and focus groups. We performed usability studies of tools we developed or applied, and used formative evaluation methods to refine both our tools and courseware. Yet, our project still requires additional approaches to evaluation.

The investigators in our project are instructors who changed their allocation of time, behavior, pedagogy, course materials, and tools. To understand the effects of these changes, ethnographic practices are of great value - especially regarding use of asynchronous communication (i.e., online debates) in the CompProf course [12].

Another shift in our evaluation has been to rely on network monitoring, logging, and analysis. Here we draw upon special tools for this purpose [2]. Part of this work has helped improve our quality of service through caching [1]. The rest has helped us understand what students really do, what course materials are accessed, how use of multimedia effects network traffic, and how both remote and local accesses increase over time. Figure 1 shows recorded accesses of the EI server from January 1995 through May 1996. Robot accesses and indexing runs have been removed from the data, except at the time of the first peak shown. There has been a gradual increase in both remote and total access counts, if we ignore the valleys occuring during mid-semester, summer, and end-of-year breaks. We believe that the remote accesses reflect growing interest in and use of our repository.

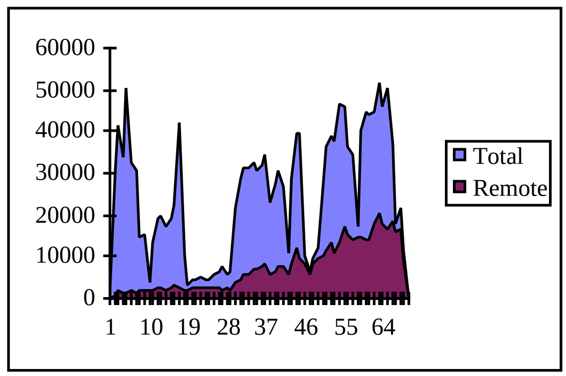

Figure 1. Weekly Accesses 1/95 - 6/96

Figure 2 gives details, focusing on Spring 1996 and two of our courses with extensive online use. The Multimedia course had about 50 students, as opposed to over 100 for NetInfo. It also followed the Keller Plan more strictly, so the peak at the end is relatively more severe, reflecting a higher level of procrastination.

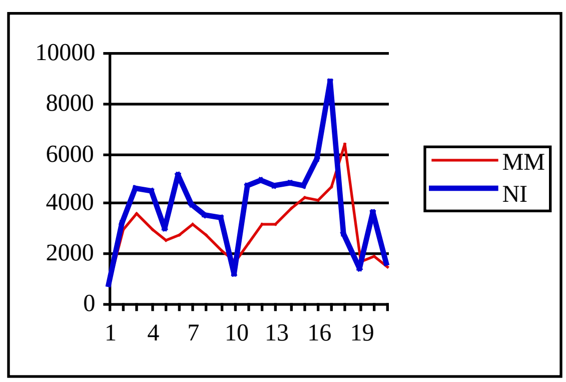

Figure 2. Weekly Accesses Spring 1996

Figure 3 gives a clearer relative comparison over time, based on access counts at the end of the Fall 1995 and Spring 1996 semesters. We note that, except for NetInfo, which, in these, its 4th and 5th offerings, showed little change, there was significant increase in accesses. While local (student) access grew, there was a sharper increase in remote access, confirming the observation from Figure 1. The History collection, which is used in the CompProf course as well as by those interested in the History of Computing, also grew in popularity.

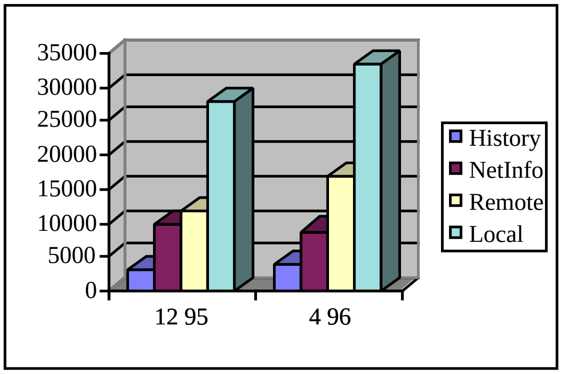

Figure 3. Accesses for Last Week of Semester

Figures 4 and 5 illustrate how many accesses are due to our course materials, and in particular, courses developed by project investigators. Figure 5, instead of counting accesses, records the number of bytes transmitted by the server, which grows large when images, PostScript files, and multimedia content are involved.

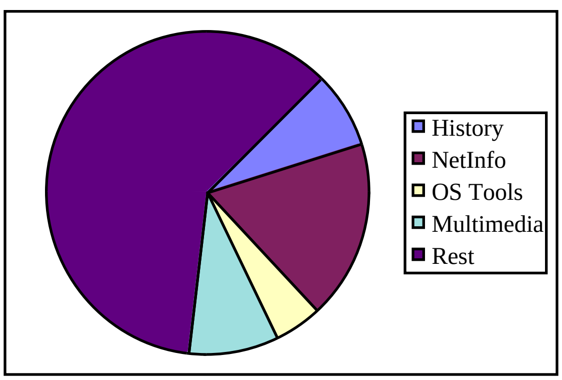

Figure 4. Accesses for Last Week of Spring 96

Changing Tools

In addition to the traditional tools used when teaching about computer science, we have applied or developed others to help with interactive learning and digital library access.

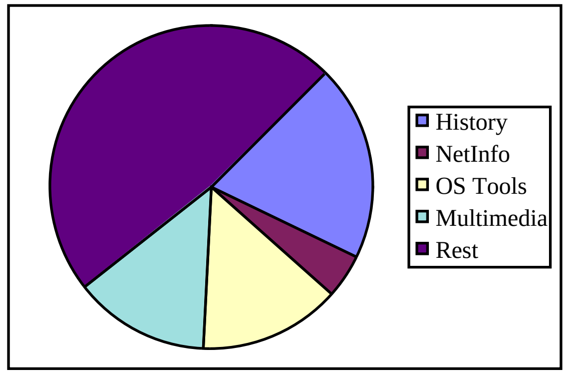

Figure 5. Bytes for Last Week of Spring 96

One focus of our project has been to help students undertand algorithms. These are often difficult to comprehend, being both abstract and complicated. Our Swan tool [14] supports algorithm visualization, has been used in one class in Spring 1996, and will be used in at least three classes in Fall 1996. We have begun evaluation work to test its utility, including for students to annotate and visualize their own programs.

Hypertext is an important technology for our project. There is use of KMS, and some of Hyper-G (now called HyperWave), but much more of Netscape [5]. A recently developed Chat-like tool, Virtual Q&A, allows us to automatically prepare an FAQ archive, to help instructors who communicate asynchronously to avoid repeatedly answering questions. This supplements several tools to support online debates [12].

Building and Applying Digital Libraries

The field of digital libraries has expanded rapidly in recent years, with whole conferences exploring the key issues [7], some of which deal with content, while others concern systems and their use.

Building digital library content typically follows one or more basic approaches. First, one can develop a digital library around existing electronic materials, as was done in the Netlib case [3]. Second, one can build upon a publisher's collection, as we do with ACM. Third, one can apply hypertext technology (e.g., WWW), as we have done in constructing our on-line courseware. Fourth, one can collect scattered materials into a unified resource, as is done with CS reports [4]. Though undergraduate student use of reports is limited, there is expanded use by graduate students and professionals. Our early work on the WATERS system [8] helped lead to the Networked CS Technical Report Library (NCSTRL - accessible at http://www.ncstrl.org) which is growing rapidly.

Building digital library systems requires identifying requirements [9], adopting a good architecture, applying library and information science concepts, and following human-computer interaction guidelines. We began this with the Envision system [10], and continued that effort with new work on interfaces [13, 15]. The main Envision interface [13] has recently been connected to our MARIAN on-line public access catalog (http://opac3.cc.vt.edu) - see the three screen dumps that accompany this paper on the conference CD-ROM. To broaden the utility of the ACM content we converted and organized, we installed an RS/6000 system at ACM headquarters with about 2 gigabytes of page image files, and are applying the same software used with NCSTRL. Further discussion with ACM and with potential partners such as IBM are scheduled, to explore issues of intellectual property rights and other matters that must be resolved before broader access to the collection is allowed.

Conclusion

In conclusion, we have worked on a wide range of objectives around the theme of improving learning by increasing interactivity and by applying digital library concepts, content and systems. As a result, a new infrastructure has emerged, our pedagogy has been transformed, utilization of the courseware we developed has grown rapidly both locally and remotely, and many tools have been constructed. Students learn new topics, often in new ways, and we have continued to progress in developing digital library content, systems, and interfaces.

Acknowledgments

Special thanks go to NSF (grant CDA-9312611) and ACM for their support of this project. I am grateful to the Virginia Tech Dept. of CS, College of Arts and Sciences, and Computing Center for their generous cost sharing. My co-investigators on the Envision and EI projects, and especially the many students involved, deserve most of the credit for our accomplishments.

References

1. M. Abrams, C. Standridge, G. Abdulla, S. Williams and E. Fox. Caching Proxies: Limitations and Potentials, in *Proc. 4th International World-Wide Web Conference,* Boston, Dec. 1995.

2. M. Abrams, S. Williams, G. Abdulla, S. Patel, R. Ribler and E. Fox. Multimedia Traffic Analysis Using Chitra95. In *Proc. 3rd Int. Multimedia Conf. and Exhibition, Multimedia '95,* San Francisco, Nov. 5-9, 1995, 267-276.

3. J. Dongarra and E. Gross. Distribution of Mathematical Software via Electronic Mail. *Commun. ACM,* May 1987, 30(5):403-407.

4. E. Fox. World-Wide Web and Computer Science Reports. *Commun. ACM,* Apr. 1995, 38(4): 43-44.

5. E. Fox and D. Barnette. Improving Education through a Computer Science Digital Library with Three Types of WWW Servers. In *Proc. Second International WWW '94: Mosaic and the Web, WWW'94,* Chicago, IL, Oct. 17-20, 1994.

6. E. Fox, N. Barnette, C. Shaffer, L. Heath, W. Wake, L. Nowell, J. Lee, D. Hix, and H. Hartson. Progress in Interactive Learning with a Digital Library in Computer Science. In *Proc. ED-MEDIA 95, World Conference on Educational Multimedia and Hypermedia,* Graz, Austria, June 17-21, 1995, 7-12.

7. E. Fox and G. Marchionini, eds., *Proceedings of the First ACM International Conference on Digital Libraries, DL'96,* Bethesda, MD, Mar. 20-23, 1996.

8. J. French, E. Fox, K. Maly, and A. Selman. Wide Area Technical Report Service --- technical reports online. *Commun. ACM,* Apr. 1995, 38(4): 45.

9. H. Gladney, E. Fox, Z. Ahmed, R. Ashany, N. Belkin, and M. Zemankova. Digital Library: Gross Structure and Requirements: Report from a March 1994 Workshop. In *Proc. Digital Libraries '94,* June 19-21, 1994, College Station, TX, ed. J. Schnase, J. Leggett, R. Furuta, T. Metcalfe, 101-107.

10. L. Heath, D. Hix, L. Nowell, W. Wake, G. Averboch, and E. Fox. Envision: A User-Centered Database from the Computer Science Literature. *Commun. ACM,* Apr. 1995, 38(4): 52-53.

11. F. Keller. Goodbye, teacher ... J. of Applied Behavioral Analysis, 1(1):79-89, Spring 1968.

12. S. Laughton. An ethnographic study of Internet-based applications in education. Ph.D. Dissertation, Dept. of Computer Science, Virginia Tech, Blacksburg, VA, May 1996.

13. L. Nowell, R. France, D. Hix, L. Heath, and E. Fox. Visualizing Search Results: Some Alternatives to Query-Document Similarity, in *Proc. SIGIR'96,* Zurich, Switzerland, Aug. 18-22, 1996, in press.

14. C. Shaffer, L. Heath, J. Nielsen, and J. Yang. SWAN: A Student-Controllable Data Structure Visualization System. In *Proc. ED-MEDIA 96, World Conference on Educational Multimedia and Hypermedia,* Boston, June 17-22, 1996, 632-637.

15. W. Wake and E. Fox. SortTables: A Browser for a Digital Library. In *Proc. 4th Int. Conf. on Information and Knowledge Management, CIKM '95,* Baltimore, MD, Nov.28-Dec.2, 1995, 175-181.