
An Educational Digital Library for Human-Centered Computing

Edward Clarkson

College of Computing and GVV
Center, Georgia Institute of
Technology
801 Atlantic Drive, Atlanta, GA
30332-0280
edclark@cc.gatech.edu

James D. Foley

College of Computing and GVV
Center, Georgia Institute of
Technology
801 Atlantic Drive, Atlanta, GA
30332-0280
foley@cc.gatech.edu

Jason A. Day

College of Computing and GVV
Center, Georgia Institute of
Technology
801 Atlantic Drive, Atlanta, GA
30332-0280
dayja@cc.gatech.edu

Abstract

Digital libraries have great potential to improve the educational experience. As a result, there are a wide variety of such repositories, especially those that focus specifically on education. But relatively few focus on topics as specific as Human-Computer Interaction (HCI) or Human-Centered Computing (HCC). In addition, support for browsing behavior, with a few exceptions, is both weak and not suitable for user needs. This paper presents our work to create a repository of educational materials for a relatively narrowly-targeted field (HCC/HCI), including our requirements gathering methods and results. Finally, we discuss the HCC Education Digital Library (HCC EDL) as a platform for investigating broader digital library research questions, such as exploring alternative designs for content browsing mechanisms.

Keywords

HCC, HCI, human-centered computing, digital libraries, education, requirements gathering.

ACM Classification Keywords

H.5.4. Hypertext/Hypermedia: User issues.

Copyright is held by the author/owner(s).

CHI 2006, April 22-27, 2006, Montréal, Québec, Canada.

ACM 1-59593-298-4/06/0004.

Introduction

Technology has the potential to improve both the efficacy and efficiency of the educational experience from the perspective of both students and teachers. Specifically, educational repositories can enhance the quality of education through the provision of a diverse set of learning materials, and enhance efficiency by affording the reuse of stored objects. To that end, we have created the Human-Centered Computing (HCC) Education Digital Library (HCC EDL) for use by the worldwide HCC and Human-Computer Interaction (HCI) communities (accessible at <http://hcc.cc.gatech.edu>). The HCC subject area is related to the HCC doctoral program at Georgia Tech, which incorporates the cognitive and learning sciences more centrally with that of HCI.

This paper focuses on our repository design and implementation process, particularly our requirements gathering methods and results. This process includes the development of an HCI topic taxonomy in which we structure library contents. We also discuss how our experience can serve as a case study for digital library development in other subject areas and its use in digital library research topics.

Motivation

The improvement of HCI education at both the graduate and undergraduate levels is an implicit interest of anyone involved in the field. Educators have long recognized the value of the reuse of educational materials (by which we mean resources used directly for teaching—lecture notes, assignments, videos, etc.). However, a prerequisite is the ready availability of such resources, and current online resources are not ideal. There are HCI-specific online resources, but their

Space constraints obviate a full treatment of related work as it relates to our motivations. A fuller discussion is available in the technical report GIT-GVU-05-33 (http://www.cc.gatech.edu/gvu/research/tr/tr05_33.html).

Reader's Note

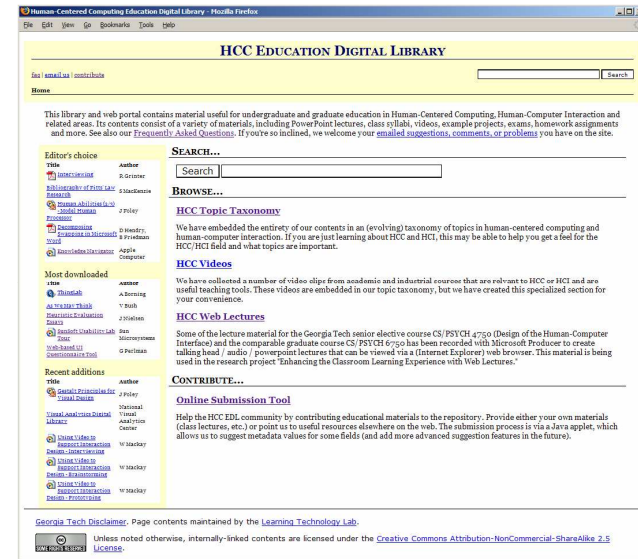


Figure 1. The Human-Centered Computing Education Digital Library (as of February 2006).

contents generally are not educational materials (e.g., the HCI bibliography¹). More all-purpose repositories have extensive breadth of coverage, but the depth of their material in narrower subjects, like HCI, is insufficient to support a course curriculum. Moreover, any education at the senior undergraduate or graduate level is specific enough that general resources (e.g., materials on “chemistry”) are of limited use. It is in this gap that the HCC EDL resides—a repository for a relatively advanced, narrow subject with the requisite depth of material to make it useful to teachers and students in higher education.

¹ <http://www.hcibib.org>.

Q: What should be in an HCC educational library to make it useful to you as a teacher in preparing courses and individual lectures?

- Syllabi Tests
 ... etc ... Other: _____

Example question from S1 and S2.

Requirements Gathering

Existing research and our own backgrounds in HCI highlighted the relative lack of HCI-focused educational digital libraries. However, this constitutes only the barest outline of a requirements list. Thus, we have conducted a series of activities to gather more precise user requirements. Our two primary audience segments are students and educators—groups with very different constraints and motives with respect to HCI education. In that regard, we view our use of multiple data gathering techniques as particularly important: users respond to various methods differently (e.g., faculty with limited time are more likely to respond to a short survey than make room on their calendar for a 60+ minute focus group). Likewise, methods like online surveys can reach wider audiences, which can highlight previously under-served issues (for example, internationalization problems).

Focus Groups

We conducted a series of focus groups with Georgia Tech students and faculty with a focus on the attributes of a valuable educational repository. These groups involved 4 faculty and graduate teaching assistants (TAs) and 5 senior undergraduate HCI students. The focus groups concentrated on two aspects of our library design: content type and organization. There has been substantial agreement between faculty and TAs and students about what kinds of data types would be most useful: lecture notes, sample tests and exams, videos, software tools, etc. Students have, however, tended to rate content directly related to classroom evaluation (e.g., sample tests) more highly than faculty or TAs. The student groups in particular also suggested that the ability to access content relevant to specific courses would be especially useful.

Online Surveys

We conducted two online surveys of HCI faculty outside Georgia Tech; the first (S1) received 15 responses and the second (S2) 20. There were 11 total participants from outside the United States across both surveys. The formatting and some of the questions differed between S1 and S2, but in general included both objective and free-form responses.

The results of both surveys supported the findings of our focus groups with respect to the basic kinds of content. Lectures and course syllabi were the two most desired types of content on both surveys (requested by 87%+ of S1 and S2 responses). At least 63% of S1 and S2 respondents favored conference videos, class assignments and example projects as well. S2 also had a more refined section on potential review and feedback mechanisms, which included 5-point Likert scale response templates (all Likert scale statistics are out of 5 points, where 5 is most positive and 1 is most negative, and all are from S2 only). The responses on S2 showed a strong consensus for some kind of screening process to ensure high-quality contents (4.15, $\sigma = 0.67$), and some preference for post-inclusion review (3.80, $\sigma = 0.77$).

Both Likert and free-form responses also favored strongly-structured content. Organization into course-level structures was rated favorably (3.90, $\sigma = 0.55$), mirroring participants' opinions of syllabus documents in general. Participants were likewise favorable about organization via a narrowly-scoped topic hierarchy (4.10, $\sigma = 0.91$). As one response summarized:

"I would like to see the detailed level of granularity [in a topic hierarchy] because many instructors of

HCI are thrown into the task rather than being trained in it. Many in the CS field have little or no psychological intuition and therefore do not teach in a way to convey this to the students. More detailed analysis of the topic material will help.”

HCI Education Workshop [1]

Given the level of interest in the repository and HCI education in general from members of the HCI community, the last author organized a workshop on graduate HCI education at the ACM CHI 2005 conference. Among the discussion topics was the idea of a repository for HCI educational materials. The format of the discussion was an unstructured group dialogue, a summary of which is available². The results broadly corroborated findings in our previous work with respect to content screening and review mechanisms, content types, and the importance of structure in contextualizing content. Intellectual property (IP) concerns were also a topic of discussion, with the consensus that the Creative Commons licensing scheme³ is flexible and promising approach.

Requirements Implementation

Given our requirements gathering process and results, we concluded that a properly designed and implemented repository of HCI/HCC educational materials would be a welcome resource for the CHI community. To those ends, we have developed the HCC EDL, which can be found at <http://hcc.cc.gatech.edu>. It has undergone several

² <http://hcc.cc.gatech.edu/chi2005workshop.htm>

³ <http://creativecommons.org/about/licenses/meet-the-licenses>

major revisions to date and is under active development.

A theme of contextualization should be apparent from our discussion of user requirements. Throughout the process, from a variety of sources and methods, users have repeatedly emphasized the importance of placing items of interest in useful contexts, such as in class syllabi, topic hierarchies or paths, or along with pedagogical tips. The HCC EDL currently attempts to satisfy those requirements primarily in two ways: an infrastructure for class syllabi and an HCI topic taxonomy.

A visual representation of our most recent revision of the taxonomy is shown in Figure 2. We have created it by synthesizing the contents and structure of a number of prominent HCI textbooks, and forming it into a hierarchy with medium depth/breadth levels [3]. We have yet to perform any rigorous validation on our structure, but we discuss our plans for doing so in our future work.

All documents within the library are embedded within at least one of these taxonomy categories, which also place related documents in close proximity. Documents are inclusive of syllabi as well; there is a standard syllabus template, which is treated as a first-class document, which in turn connects appropriate resources (lectures, homework assignments, etc.) from the repository. The role of resource granularity has been identified as a major factor in document reuse [5]; syllabi (and to a lesser extent the taxonomy topic groupings) address this issue by providing a granularity alternative to individual documents.

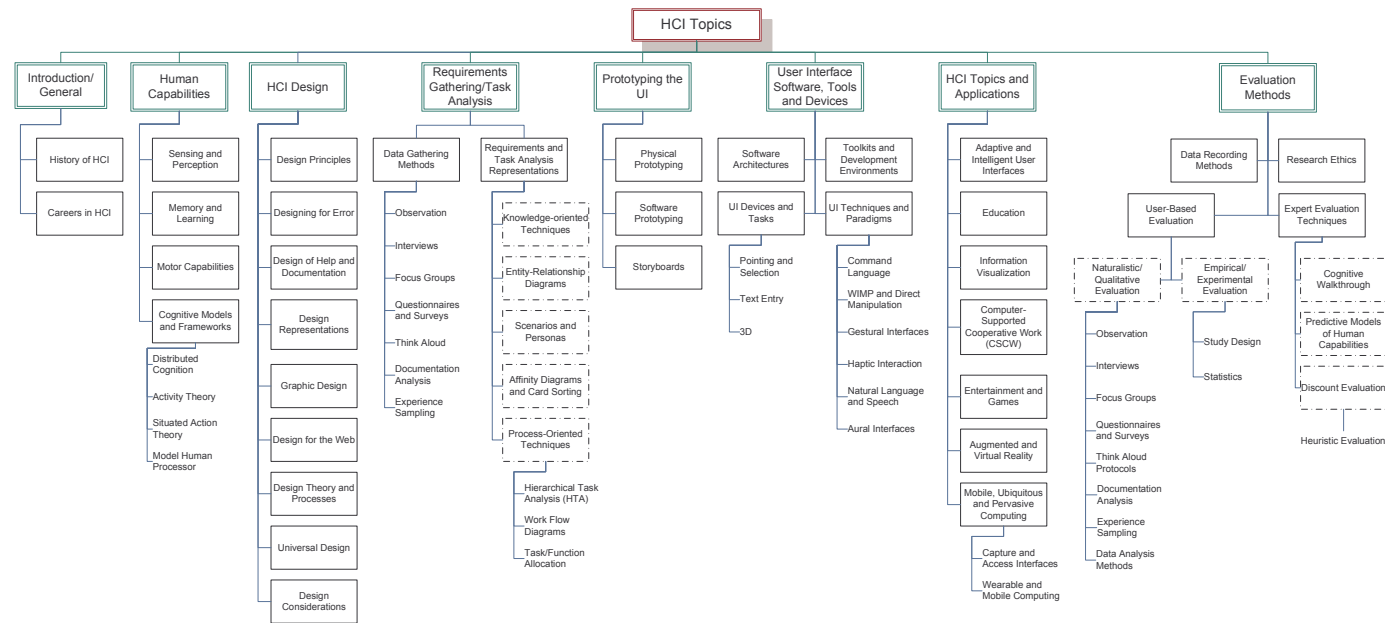


Figure 2. A taxonomy of Human-Computer Interaction topics.

Including these syllabi, the HCC EDL includes over 500 documents from over 30 authors at more than 15 institutions. Its front page is shown in Figure 1. It is built on a MySQL database backend with a dynamically generated DHTML UI, and includes an online contribution system.

Current and Future Work

Our research agenda generally focuses on leveraging the HCC EDL to explore more general issues in digital library design and use. We are currently building a syllabus generation tool that lets users construct course syllabi from library contents (and their own materials). We hope to leverage this tool both to generate new

content for the repository and to provide data on relationships between documents. We are also augmenting our online submission system with more automated features in order to streamline the contribution process. The goal is to use these systems to quantify the combinatorial effects of small design improvements. For example, if we use a variety of simple automated techniques to extract different metadata (e.g., TF-IDF keyword extraction; document title extraction), how do they affect the contribution process’ ease-of-use individually vs. in combination?

We are also currently exploring the impact of broken resource links on users’ subjective impressions of

digital library quality. Libraries that point to remote content always take on the risk of stale links. We should be able to provide quantitative data on the impact of those choices by varying the amount of ‘brokenness’ between users and surveying their opinions of library quality. Thus, DL designers have a more complete picture of the impact of their design choice (i.e., remote linking vs. local storage).

Other significant interests are methods to support interleaved search and browse behaviors; visualization techniques for repository contents; and how they can complement each other. There is existing work on both subjects (e.g., [2], [4]). We hope to leverage the HCC EDL—along with its structured and controlled environment, in comparison to the web at large—to create novel interaction designs for DL users.

Other work includes validating the HCI taxonomy. We believe a card-sorting exercise with a group of HCI experts may yield a satisfactory procedure for such a task. By having such experts categorize a relatively small group of documents into a given hierarchy, we can determine the amount of agreement within the group with respect to the appropriate categorization. We speculate that a high degree of agreement is indicative of an acceptable taxonomy. Such a validation might then yield both a confirmation of our specific hierarchy as well as a generalizable methodology for the digital library community as a whole.

Conclusions

A resource that facilitates the reuse of educational materials could be a valuable addition to the HCI/HCC community. The HCC EDL attempts to satisfy this niche

and do so via a thorough, user-centered process. We have presented this process as a case study for those interested in building narrowly-targeted digital repositories. As we have discussed, we look forward to using the library as a platform for advancing work in the DL user interface/experience field.

Acknowledgements

Building an educational resource for HCC and HCI education would not be possible without the many individuals worldwide, too numerous to list here, who have contributed their time and materials to build the HCC EDL. Thanks also to past and present developers on the HCC EDL project: Andy Cox, Remco Groeneweg, Aarjav Trivedi, Jonathan Woodbridge and Kevin Yang. This work is supported by the Stephen J. Fleming Chair in Telecommunications at the Georgia Institute of Technology.

References

- [1] Foley, J. *et al.* Graduate Education in Human-Computer Interaction. In *Extended Abstracts of CHI '05* (Workshops), pp. 2113-2114.
- [2] Kampanya, N., Shen, R., Kim, S., North, C. and Fox, E. Citiviz: A Visual User Interface to the CITIDEL System. In *Proceedings of ECDL '04*, pp. 122-133.
- [3] Larson, K. and Czerwinski, M. Web Page Design: Implications of Memory, Structure and Scent for Information Retrieval. In *Proceedings of CHI '98*, pp. 25-32.
- [4] Olston, C. and Chi, E. H. ScentTrails: Integrating Browsing and Searching on the Web. *ACM Transactions on Computer-Human Interaction*, 10(3), pp. 177-197.
- [5] Sumner, T. and Dawe, M. Looking at Digital Library Usability from a Reuse Perspective. In *Proceedings of JCDL '01*, pp. 416-425.