

The Development of an Educational Digital Library for Human-Centered Computing

Edward Clarkson, Jason Day and Jim Foley

College of Computing and GVU Center, Georgia Institute of Technology

801 Atlantic Drive, Atlanta, GA 30332-0280

{edccclark, dayja, foley}@cc.gatech.edu

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 the *efficiency* of education by affording the reuse of stored objects. To that end, we have created the Georgia Tech Human-Centered Computing (HCC) Education Digital Library (HCC EDL) [12] for use by the worldwide HCC and Human-Computer Interaction (HCI) communities.

This paper focuses two aspects of this work:

- A taxonomy of selected online repositories, which reveals a lack of narrowly focused educational digital libraries that provide rich affordances for content browsing.
- Our prototype design and implementation of an HCC EDL, including our requirements gathering methods and results. This process includes the development of an HCI topic taxonomy.

Our prototype repository addresses the relatively sparse space in the digital library design space exposed by our taxonomy and does so in conjunction with the user requirements we have collected. It targets two primary user groups: students, for self-study; and instructors, for course preparation. Such a repository has the ability to improve educational practice both at institutions with current HCI faculty and ease the barrier to entry at institutions without any existing HCI faculty or curriculum. It also can serve as a valuable resource for students interested in the subject outside of any classroom structure or for those who are interested in alternative presentations of classroom materials. Finally, the HCC EDL provides a platform for future research in a range of areas, such as digital library usability issues or facilitating alternative pedagogies [27].

MOTIVATION AND RELATED WORK

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; however, a prerequisite of such recycling is the availability of such

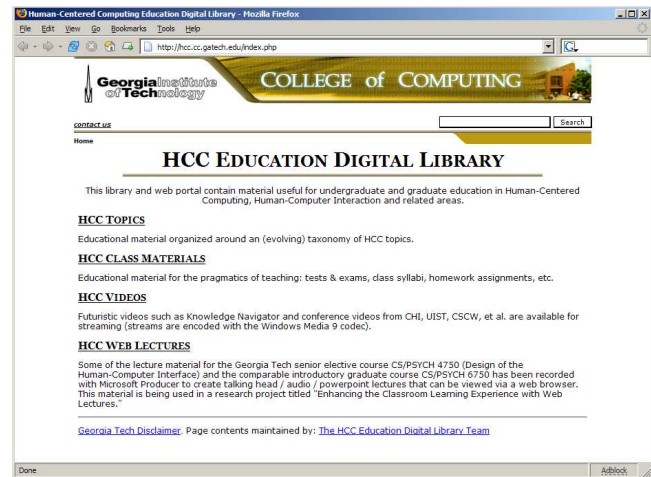


Figure 1 – The HCC Education Digital Library.

resources. One common method of accessibility is through colleagues; it is rare to encounter a practicing educator who has not borrowed at least some portion of his or her curriculum from a trusted associate. Likewise, students (self-motivated or otherwise) depend on available resources for their education. By definition, students are new to the field; as such they are forced to rely on external resources for educational materials. In many cases, professors or experienced peers fulfill this niche.

But what of students or educators who lack such social networks? More explicitly, what about students who are interested in HCI, but whose institution has no courses on the topic? Likewise, what of professors interested in starting a new HCI curriculum at a school without one (e.g., in a smaller liberal-arts college faculty)? In either case, the web is potentially their primary resource for finding and evaluating HCI educational materials.

Unfortunately, current resources are not ideal for such situations. There are HCI-specific resources, but their contents are generally not educational materials. And although general repositories have extensive breadth of coverage, the depth of their material in narrower subjects such as 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 with a specific focus and substantial depth of material. Such a facility supports the improvement of existing HCI education programs and encourages the development of new curricula, as has been suggested previously [37].

Existing projects have put considerable effort into their contents (in the form of, for example, building sizable collections and rich metadata schemes and standards). However, with few exceptions there is relatively little variety in access beyond standard search or simple browsing mechanisms. Researchers have long noted differences between searching and browsing activities, which users employ variously depending on their goals, environmental capabilities and related factors [33]. However, while current efforts generally have extensive search features, their facilities and affordances for browsing behaviors are limited—for example, single-level ordered listings of all materials that meet some specific metadata characteristic (e.g., alphabetical by author name). Even when more detailed structures are employed, the broad focus of most online repositories is such that even high browsing granularity does not reach groupings of materials that are sufficiently specific for advanced studies (the MERLOT collection [18], for example). Our work attempts to address these needs via several approaches, including a more detailed topic classification and structured collections of documents that are relevant to user tasks (such as course syllabi).

Other research supports the idea of creating more specialized browsing constructs. Sumner and Dawe have examined the effects of educational digital library content presentation on reuse [44] in the context of the Digital Library for Earth System Education (DLESE) project [9]. They found that instructors’ use of digital libraries could be divided into two primary behaviors (class vs. course preparation). By looking at these tasks from a reuse perspective, Sumner and Dawe contend that contextual information about library resources and the composition of those resources are important considerations for library design. Sumner and Marilino have also found that even teachers can have difficulty understanding how characteristics of particular resources are connected with broader aspects of the field [42]. Our detailed topic hierarchy and our syllabus documents address each of these concerns: both provide the data within the context of related resources and support the task of composing resources because of those groupings.

Researchers recognize the value of multiple content display mechanisms. In their evaluation of different types of hypermedia architectures [39], two of Salampanis and Diamantaras’ conclusions are to 1) offer affordances for multiple information seeking strategies and 2) support parallel, interleaved use of those strategies. In earlier work, Xie came to similar conclusions in her study of library user behaviors [45], identifying (among others) support for opportunistic interaction and information-

seeking strategy shifts as ways to improve information retrieval systems.

Finally, the practice of linking to content (versus storing it locally) has its own pragmatic consequences in the form of broken links. Though there has been extensive research on various approaches to mitigating this problem ([32,35,40], to name only a few) none of them can guarantee 100% effectiveness. Conversely, Fogg’s survey of web users indicated that even small numbers of broken links (or other indicators of amateurism) have a significant negative impact on users’ subjective impressions of site credibility [30]. As such, link repair mechanisms cannot be considered a panacea for resource maintenance.

Survey of Selected Online Digital Repositories

As a means of both motivating our approach and to survey other similar efforts, we have organized a taxonomy of selected digital repositories (see Table 2). The sheer number of researchers working on digital libraries and related issues makes it difficult to discuss related research in a comprehensive manner; we omit a large number of significant projects (such as the Alexandria Digital Library [3]) in order to present repositories that are more closely related to our own work in some way (e.g., audience or content characteristics). Likewise, the number of possible dimension combinations of such a taxonomy is nearly infinite; we have chosen a few that we believe are interesting or informative. These dimensions focus on three aspects of repository design: characteristics of the data itself (storage location, breadth), target audience (purpose, educational level) and content browsing facilities (organization, granularity). More detailed descriptions of each category are found in Table 1.

These dimensions reveal several patterns in existing work. Educational contents are by far the most popular type of materials in our survey collection: 16 of the 21 repositories

Dimension	Description
<i>Purpose</i>	Intended task supported by the repository contents. Education indicates support for actual pedagogy; Research refers to assisting academic investigations (e.g., literature searches); Reference combines elements of education and research—for example, educational materials not intended for classroom use a la encyclopedia-style resources.
<i>Level</i>	The level of education for which the repository contents are intended.
<i>Location</i>	Where repository contents are stored (i.e., on local servers or elsewhere on the internet).
<i>Breadth</i>	The range of subjects covered by the repository.
<i>Organization</i>	Methods for structuring content display without requiring user searching.
<i>Granularity</i>	How finely divided the organizational structure is where applicable; entries are by example.

Table 1 – Descriptions of taxonomy dimensions from Table 2.

focus exclusively on objects targeted for use directly in the classroom. The type of classroom, however, varies considerably. Twelve of the 21 repositories have contents suitable for higher education settings, while another 7 have no specific focus on educational level. Conversely, focusing on a relatively narrow subject area is relatively rare. Although it is difficult to quantify the breadth of a body of knowledge, 9 of the 21 surveyed repositories include content from any discipline, while only 4 (including

the HCC EDL) target HCI or similarly-scoped topics.

In addition, relatively few repositories employ browsing systems that present their contents in relatively small chunks, especially when those systems take the form of hierarchical classifications. This has significant consequences especially for higher education repositories; in combination with the tendency toward including content from any or broad subject areas, even fine-grained hierarchies (or other means for browsing content) still leave

Name	Target Audience		Content		Browse Access	
	Purpose	Level	Location	Breadth	Organization	Granularity
ACM Digital Library [1]	Research/Reference	Hi-Ed/Post-Grad	Local	Computer Science	Publication Hierarchy	"Proceedings::CHI::CHI 2005"
AESharenet [2]	Education	All	Local	All		
AIS SIGHCI Teaching Resources [4]	Education	Hi-Ed	Remote	HCI	By Subfield	"Information Visualization"
Ariadne (SILO) [5]	Education	All	Local	All		
CAREO [6]	Education	Hi-Ed	Remote	All	By Topic	"Science"
CLOE [7]	Education	Hi-Ed	N/A	N/A	N/A	N/A
Connexions [8]	Education	All	Local	All	By Title/Author/Keyword/Popularity	"Central limit theorem"
Digital Library for Earth System Education [9]	Education	All	Remote	Earth Science	By Topic/Grade Level/Type	"Atmospheric Science"
eCourses Web Portal [10]	Education	Hi-Ed	Local	Engineering	By Course Topic	"Rigid Body Equilibrium"
The Gateway to Educational Materials [11]	Education	All	Remote	All	By Subject/Type/Level/etc.	"Educational Technology"
HCC EDL [12]	Education	Hi-Ed/Post Grad	Local	HCI/HCC	Hierarchical topic taxonomy	"Prototyping the UI::Prototyping Methods::Software Prototyping::Tools"
HCI Bibliography [13]	Research/Reference	Hi-Ed/Post-Grad	Remote	HCI	2-level topic taxonomy	"HCI Education::Programs"
IEEE Xplore [14]	Research/Reference	Hi-Ed/Post-Grad	Local	Engineering	Alphabetical by Publication	"IEEE Symposium on InfoVis"
iLumina [15]	Education	All	Remote	Science and Mathematics	3-level topic taxonomy	"Computer Science::Human-Computer Interaction::Graphical User-Interface Design"
Interaction-Design.org [16]	Reference	Hi-Ed/Post-Grad	Local	Interaction Design	Alphabetical by Keyword	"Affordances"
MathDL [17]	Education	9-12+	Remote	Mathematics	By type and 2-level topic taxonomy	"Analysis::Complex Analysis"
MERLOT [18]	Education	Hi-Ed	Remote	All	3-level topic taxonomy	"Science and Technology::Computer Science::Human-Computer Interaction"
MIT OpenCourseWare [19]	Education	Hi-Ed	Local	All	Courses by Department	"UI Design and Implementation"
NEEDS [20]	Education	All	Remote	Engineering		
Open Video Project [21]	Research	Hi-Ed/Post-Grad	Local	All	By Genre/Duration/etc.	"Documentary"
SMETE [22]	Education	All	Remote	Science, Engineering, Math and Technology	By Topic	"Mathematics"
Wise-Online [23]	Education	Vocational	Local	All	By Topic	"Economics"

Table 2 – Taxonomy of selected online repositories of educational materials, categorized by characteristics of repository contents and access mechanisms. Adapted from <http://elearning.utsa.edu/guides/LO-repositories.htm>.

groupings at a relatively high level. For example, the MERLOT [18] repository has one of the more detailed topic classifications in our survey; however, due to its topic breadth its finest-grained grouping is the entire field of HCI.

A number of these systems illustrate some of the issues raised in the literature discussed in the previous section. By virtue of their missions to publish complete course data, the MIT OpenCourseWare [19] and OU eCourses [10] projects present all of their contents within the contexts of course syllabi. Similarly, the Connexions library [8], affords greater reuse of learning objects by grouping objects into ‘modules’ and courses in addition to individually accessible components. Groupings aside from syllabi also present library contents in contexts outside the usual search query results view. The Walden Path [31] and Panorama [34] systems are similar in this respect, in which context is expert- (Walden) or system-created (Panorama).

SYSTEM REQUIREMENTS

Existing research and our own backgrounds in HCI highlighted the relative lack of HCI-focused educational digital libraries. They also amplified our interest in the issues surrounding browsing mechanisms for such repositories. However, this constitutes only the barest outline of a design for our system. As a result, we have conducted a series of activities to gather more precise user requirements. As we have mentioned, our two primary audience segments are students and educators—groups that have very different constraints and motives with respect to an HCI repository. In keeping with the principles of user-centered design, we have employed a number of requirements-gathering techniques with potential members of our user population.

Focus Groups

In the Fall of 2003, we conducted a series of focus groups with Georgia Tech students and faculty on the characteristics 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 was substantial agreement between faculty and TAs and students about the types of data that would be most useful: lecture notes, sample tests and exams, videos, software tools, etc. Not surprisingly, students tended to rate content most 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 a particular course, such as an introductory HCI course, would be especially useful. This matches what we might expect [33]: browsing within an organized scheme lowers cognitive demands, especially for novice users. The faculty groups had more discussion over efficiency issues. Faculty stressed that the ability to

navigate library facilities (such as browsing interfaces or an online document submission process) in as few steps or mouse clicks as possible was important because of pressure on faculty time.

Online Surveys

We conducted two online surveys of HCI faculty at other institutions; the first was conducted in July 2004 (S1) and the second in November 2004 (S2). Subjects were solicited via two email lists: the ACM SIGCHI education mailing list and a large list of the last author’s personal contacts. We received 15 responses to S1 and 20 to S2. The surveys allowed for either anonymous or personalized responses; 9 of 35 total subjects (1/15 for S1 and 8/20 for S2) were anonymous. The responses had a moderate international component: 11 of 35 total respondents reported email addresses with non-U.S. top-level domains. The texts of the questions from both surveys are listed in Table 3. Ratings on S2 were taken on a 5-point Likert scale, and both surveys collected free-form responses.

The results of both surveys strongly 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 87.5% and 89.5% on S1 and S2, respectively). Conference videos, class assignments and example projects were also rated highly (requested by 63+%) by both response groups. Half of the S2 responses mentioned a sense of community as a main incentive for contributing material to the library:

1. What should be in an HCC educational library to make it useful to you as a teacher in preparing courses and individual lectures?
2. What methods for accessing library materials would be most useful to you?
3. Please rate the following levels of granularity for an organizational scheme:
4. What would motivate you to contribute your own material to such a library?
5. Please answer the following questions concerning review and feedback mechanisms:
 - Please rate the following mechanisms for screening material **prior** to its inclusion in the library:
 - Please rate the following comment/review mechanisms for material **already** included in the library:
6. Who should comprise a formal peer-review committee for submissions as described in question 5?
7. To whom should review/feedback results be available?
8. If such a library existed, with content and granularity and search means meeting your desires as state above, would you be likely to use the library as a resource?
9. If yes, what about the library is attractive to you; If not, why not?
10. What other issues should we think about with this project, from the perspective of a teacher or student?

Table 3 – Questions from online surveys. Survey #1 included all questions shown above; #2 omitted questions 2 and 9.

for example, “When creating my own course materials, I have borrowed and benefited from others. Thus I am happy to give back to the community.” Responses also favored structured contents, highly rating course-level organization 3.90¹ ($\sigma = 0.55$) and narrowly-scoped subtopics 4.10 ($\sigma = 0.91$). A number of free-form responses underscored our approach to the problem:

“I would like to see the detailed level of granularity 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.”

A major portion of the surveys concerned inquiries about potential review and feedback mechanisms. S2 yielded 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, $\sigma = 0.77$). This emphasis on quality contents mirrors the priorities found in a study of educator attitudes towards digital library contents [43]. Conversely, there was little agreement on the details of such processes, either pre- or post-inclusion in the repository. Neither limited Amazon.com-style reviews (3.00, $\sigma = 1.12$) nor formal peer-review assessments gained favorable marks (3.35, $\sigma = 1.11$) as screening systems. Similarly, both unrestricted (2.74, $\sigma = 0.99$) and author-only Amazon-style (2.72, $\sigma = 0.89$) received negative scores. Peer-only Amazon-style review was the only positively-viewed system (3.75, $\sigma = 0.85$).

Though responders to the survey were self-selected, it was somewhat encouraging to find that 100% of the subjects indicated they would be likely to use an HCC EDL if it met their requirements. Other anecdotal responses bolstered our belief in the utility of such a facility:

“I’m a new teacher, and when I began teaching HCI I was completely lost as to which material I should include in an [undergraduate] course and which to leave out. This [HCC EDL] would help immensely, if it existed a year or two ago. Also, I’m always trying to update and revise my courses, so this resource would help in giving me ideas on what other people are doing, and whether it works or not.”

Intellectual property (IP) was another common issue in free form responses. Most subjects’ central question was concern about copyright permission of library contents, from the perspective of both the donating author and the person using the material. However, there were few detailed suggestions on how these matters might be addressed.

¹ All ratings are on a 5-point Likert scale with 5.0 as the high rating, and are taken from S2 responses.

HCI Education Workshop [29]

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 can be found in [25]. Many of the issues raised in the discussion were, by this point, familiar to us. As it turned out, so were many of the group’s conclusions (although it should be noted that some of group participants may have participated in the online surveys, perhaps biasing our findings towards consistency).

As with the online surveys, participants noted the challenge of IP, especially the tension between author rights and reasonable access to their works. The Creative Commons licensing system [24] is one solution, though it does not address common situations such as the inclusion of embedded copyrighted works (e.g., a Far Side[®] cartoon used in a PowerPoint lecture).

Likewise, participants also favored a screening process over a post-submission ranking as a means of creating and maintaining a high-quality corpus. The group noted that such a vetting process entails some kind of organization to perform those duties, and that creating one that is self-sustaining is a significant challenge (from both a funding and a social perspective). The Wikipedia² community model is one such model, but there are significant differences between the community participants. It is not clear how well that model might work for a project like the HCC EDL given the time-sensitivity of many university faculty members (the source of our library contents). Similarly, contributors noted the problem of breaking links (especially for resources like class web pages which tend to be relatively short-lived). Forcing contributors (or another group) to maintain such links is a significant burden, but local mirroring is a resource-intensive practice. A method to balance the two concerns is necessary—one suggestion was to employ a resource cache³ similar to that employed by Google.

Finally, the group emphasized the importance of context for library materials in several ways. First, the context of use of a pedagogical artifact (such as lecture notes or a video) is as important as the artifact itself. A video or picture embedded in a lecture may be inscrutable without knowledge of an instructor’s commentary or motivation. On a larger scale, the workshop echoed findings from our previous surveys and focus groups, emphasizing the need for showing contents in the context of related items via facilities such as a topic taxonomy, syllabi, or ‘trails’ (a la

² <http://wikipedia.org/>

³ <http://www.google.com/help/features.html#cached>

the idea of Walden Paths [31]. Syllabi also mitigate what Sumner calls the ‘granularity problem’ [44], in which library resources do not match the scope desired by users.

Web Server Log Analysis

Soon after our initial focus groups with students and faculty, we created some rough prototypes of the HCC EDL, which we have refined over time according to what we have learned from our users. The site has also been in active use since Spring 2004 as part of a related project on using web-based lectures as part of a project to enhance classroom experience [27, **Error! Reference source not found.**]. Using this site, we had the capacity to investigate user requirements in a more quantitative manner, in contrast to the more qualitative processes we have described thus far.

Beginning in January 2005, we instrumented our prototype repository to log customized data about its usage. Several aspects of our audience and their usage environment affected our design for our logging scheme. We are interested in tracking user access patterns over time; furthermore, a substantial portion of our user base are students, who commonly access the repository from multiple physical locations (e.g., personal desktops in dorms, laptops, different campus labs). As a result, we wanted our web server to differentiate only users who were truly distinct—and conversely, be able to identify a user over the course of an entire semester, regardless of his or her location.

To those ends, we configured the HCC EDL web server (Apache/1.3.29) to check for the presence of a special cookie when any page request was made. If it was not present, it presented a splash page to the user requesting that he or she provide an anonymous, semi-unique identifying code⁴. This page explained the reasons we were asking for this information and informed users that they could bypass this step if desired. If the user did not enter a code, Apache recorded the standard session data.

Analysis of this data is ongoing; we hope to acquire data about user access patterns that will be informative for future design iterations.

Requirements Summary and Discussion

We have used a variety of approaches to requirements gathering as a form of triangulation, increasing the confidence we have in our findings that are consistent across methods. Regularly occurring requirements include:

- A method to ensure that access to contributed materials is not affected by link degradation.

⁴ We requested the last four digits of the user’s phone number followed by the day of the user’s birth. This code is stable over time and unlikely to have collisions between a relatively small group (< 500) of users.

- A screening mechanism to ensure a significant baseline quality, including a sustainable, scalable organization framework to support such a system.
- Extensive support for placing items of interest in useful contexts, such as in class syllabi, topic hierarchies or paths, or along with pedagogical tips.
- An explicit and clear explanation of IP issues, including the rights and permissions of both authors and users.

These characteristics of a digital library are not unique individually—but their *combination* is interesting. Moreover, the requirements are noteworthy as having been derived from a specific user population. The need for such a repository in this user population has been identified and the repository represents a focused subject area rather than a general-interest; as such, its requirements are significant on their own.

This is not to say that we have concluded the process of collecting such data. Though we have had input from a variety of teachers and professors, that sample has been heavily biased toward faculty at larger research-oriented universities. Considering that (as we have pointed out) instructors at smaller teaching colleges and universities are a prime audience for this resource, we need to acquire more data about if and how their needs differ from their larger-school colleagues. Likewise, professionals are another potential audience which has been neglected to a large degree. Finally, we have concentrated to a large extent on faculty users because they not only represent a large user base, but the likely source of nearly all our content. However, student needs are clearly distinct, and deserve a more thorough treatment than what we have completed thus far.

PROTOTYPE REPOSITORY

As we have mentioned, we implemented the first versions of a prototype repository in early 2004. This development has proceeded in an iterative fashion as we have refined our requirements and augmented our content collection.

HCI Topic Taxonomy

Very early in the process we identified a topic hierarchy as a useful navigational tool. However, no such hierarchy is in common use—so we developed one. After developing a preliminary version for early editions of the repository, we subsequently used textbooks as model resources, since this hierarchy is specifically for educational use (and classifies educational materials). We chose five textbooks commonly used in undergraduate and graduate HCI survey courses [26, 28, 36, 38, 41]. We synthesized their overall contents as well as their organization and structure with our own experience and expertise to form the hierarchy partially shown in Figure 2.

The taxonomy has evolved over time as we have collected ill-fitting documents and found inconsistencies in its

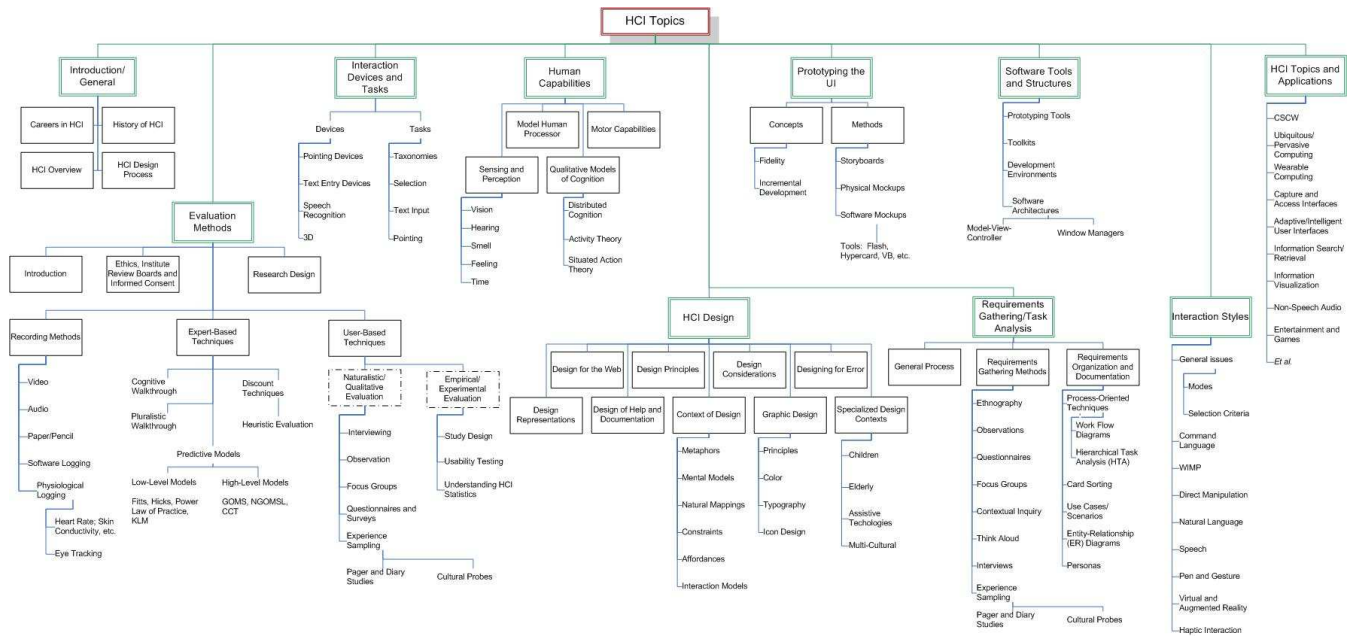


Figure 2 – Diagram of a hierarchical HCI topic taxonomy (parts omitted). Developed using [26, 28, 36, 38, 41].

structure. We expect this process will continue, especially as the repository evolves from what is practically an HCI library to a more truly interdisciplinary HCC facility. The taxonomy has been specified as an XML topic map (XTM), which supports such modifications in a relatively seamless manner.

We do not intend for this hierarchy to be exclusive; our goal is merely for an adequate representation of the field. Card-sorting exercises with people knowledgeable in HCI will provide evidence that it meets those goals or data from which we can adjust the taxonomy.

Prototype HCC EDL Implementation

Our initial prototype repository consists of a static set of hand-edited web pages serving over 500 documents from over 30 authors at more than 15 institutions (both academic and industrial). The types of documents reflect the demands of our users: lecture slides, classic HCI videos, example homework and project assignments, tests and exams, web lectures, etc. In addition to cataloging these items individually, the repository also contains a number of syllabus documents (taken from real courses) from various parts of the HCI field. A template ensures a consistent look-and-feel for the syllabus documents, which contain links to the component documents in the library related to that class as well as other data about the class (e.g., textbooks).

We have acquired library material via personal solicitations to authors for permission to distribute their materials. Such solicitations have been arisen either from searching or general appeals for contributions to HCI education mailing lists or personal contacts. Though labor-intensive, this process has allowed us to tightly control content quality.

Copies of the documents are served from servers under our control so we can guarantee their availability.

We have applied our HCI topic taxonomy to our contents by classifying each document under at least one (but perhaps more) node of the hierarchy. Any level of the hierarchy can contain documents (not just the leaf nodes), but the current collection does not completely fill the hierarchy; unpopulated nodes are not shown in the browsing interface. Where nodes are sparsely populated, the repository collapses the nodes into the appropriate parent: for example, there separate nodes for low- and high-level predictive models in the evaluation section (see Figure 2 lower left). However, their parent ‘predictive models’ node has subsumed both categories since there are relatively few documents on either sub-topic.

Conversely, when there are enough documents to warrant a separate page on a particular sub-topic, that link is shown on the parent node along with a few representative documents from the sub topic (see Figure 3). At present, those nodes are selected by us as ‘editor’s picks’, but could easily be selected on the basis of number of downloads or user ratings. Like topic subsumption, the representative documents minimize the depth of the hierarchy, which in turn reduces the average number of page transitions necessary to access the library corpus.

To this point, we have handled IP issues on an *ad hoc* basis—permissions to use material have been handled informally over email. Likewise, our site documentation directs authors who believe their material is being improperly used or who would like previously included materials removed simply to email us for resolution (though this has not occurred to date). Nevertheless, IP issues are

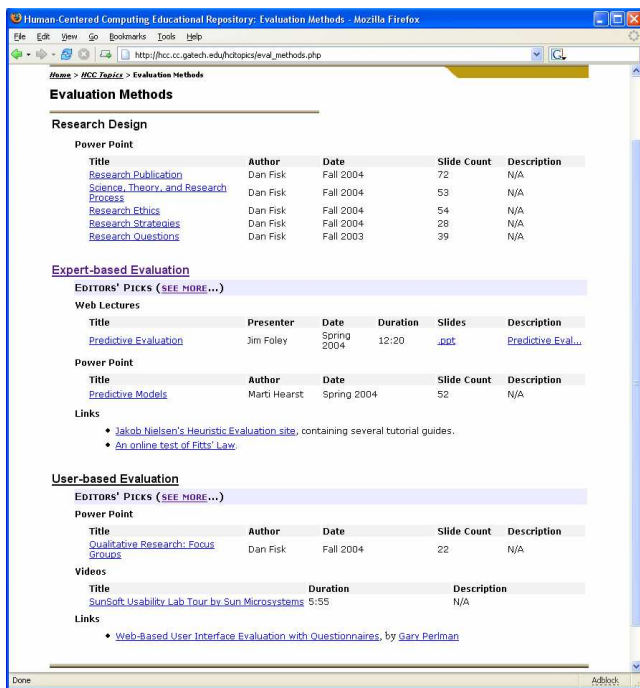


Figure 3 – HCI topic taxonomy implementation.

confusing and important enough that a more explicit policy is desirable. Like many online facilities, the Creative Commons licensing policy is a likely candidate for future iterations of our design.

The features we have described are a direct result of user requests and requirements. We have paid careful attention to screening submissions to create a high-quality corpus containing the types of documents people have requested. Those documents are mirrored locally so that we obviate the need for a system to detect or correct broken links. We have attempted to afford document reuse by providing materials at multiple granularity scales (on their own and within larger-scale syllabi). Those syllabi, along with the HCI topic taxonomy, provide informative and useful context for library materials.

FUTURE WORK

The current version of the repository only partially satisfies our requirements. We are in the process of implementing a more robust database-backed, dynamically generated version, which will dramatically decrease the maintenance costs associated with the current static pages. It will also allow for improved UI features, such as document table layouts that can be sorted, reordered and customized at the user's preference. We will also be able to store a much wider variety of metadata that will be more easily searchable.

The current library home page (see Figure 1) segments the contents into video, web lecture, class materials and the HCI topic taxonomy (all of the above plus everything else). These divisions are somewhat arbitrary. We would like to

explore other approaches based more directly on user needs, such as segmenting the contents based on audience type (e.g., student, teacher or professional) or task (e.g., finding lectures on a particular topic, constructing a course syllabus, etc.). This segmentation involves exploring other ways of putting documents in contexts aside from our basic taxonomy. Other high-priority features include the ability to submit new materials online, an online syllabus building tool, and an infrastructure for mirroring documents a la the Google cache.

There are also a number of broader research agendas of interest. As we mentioned, our current system involves a great deal of involvement on our part: there are a variety of research possibilities in terms of organizational systems for rating content, especially in the specialized context of educational digital libraries. Corresponding to Sumner and her collaborators ideas about granularity mismatch issues, we would also like to investigate techniques (aside from simple syllabus aggregation) to minimize this problem. In a general sense, we are interested in the interaction of user experience levels (i.e., novice vs. expert) and various methods of providing context.

Additionally, we are interested in using our repository as a platform for related research. We are already using it as a delivery mechanism for web lectures. We hope that it will prove useful in examining tangential issues such as automatic hierarchy reorganization based on user behavior or the effects of broken links on subjective user impressions of quality.

CONCLUSIONS

This work has presented our work in determining and defining the need for a digital library in the HCI/HCC educational community. A survey of similar digital library projects shows that though there is a diversity of educational digital libraries, many are broadly focused or lack detailed or contextual interfaces for browsing their contents. Moreover, the synthesis of these features is rare indeed. Motivated by this gap in the literature, we have produced a set of user requirements for such a resource specifically for the field of HCC and HCI. These requirements for what we call the HCC EDL have been collected and triangulated by using an array of requirements gathering methods. Furthermore, those methods have repeatedly shown the utility of an HCC EDL meeting those requirements.

We have also presented our initial work on implementing those requirements, including the creation of an HCI topic hierarchy. Our prototype repository embeds repository content within our topic hierarchy, supporting browsing activities by placing items within a useful context. We have also screened contents and mirrored them locally to assure a high-quality, always-available corpus. Finally, we have identified both specific research questions and more general areas for additional work. Like many researchers, we believe educational digital libraries hold a great deal of

promise for improving the efficiency and quality of higher education. We believe this work is a step toward a better understanding of the processes by which that promise can be realized and a useful facility for HCC/HCI education.

ACKNOWLEDGMENTS

Thanks to past developers on the HCC EDL: Remco Groeneweg, Aarjav Trivedi and Jonathan Woodbridge. Credit is due to the many individuals worldwide, too numerous to list here, who have contributed their time and materials to build the HCC EDL. This work is supported by the Stephen J. Fleming Chair in Telecommunications at the Georgia Institute of Technology.

REFERENCES

1. ACM Digital Library, <http://portal.acm.org/>.
2. AShareNet, <http://www.aSHAREnet.com.au/>.
3. Alexandria Digital Library, <http://www.alexandria.ucsb.edu/>.
4. AIS SIGHCI Teaching Resources
5. Ariadne (SILO)
6. Campus Alberta Repository of Educational Objects (CAREO), <http://www.careo.org/>.
7. Co-operative Learning Exchange (CLOE), <http://cloe.on.ca/>.
8. Connexions, <http://cnx.rice.edu/>.
9. Digital Library for Earth System Education (DLESE), <http://www.dlese.org/>.
10. eCourses Web Portal, Gramoll, K. ed., <http://www.ecourses.ou.edu/>.
11. The Gateway to Educational Materials (GEM), <http://www.thegateway.org/>.
12. Human-Centered Computing Education Digital Library (HCC EDL), <http://hcc.cc.gatech.edu/>.
13. HCI Bibliography, Perlman, G. ed., <http://hcibib.org/>.
14. IEEE Xplore, <http://ieeexplore.ieee.org/>.
15. iLumina, <http://www.ilumina-dlib.org/index.asp>.
16. Interaction-Design.org, Soegaard, M., ed. <http://interaction-design.org/>.
17. The MAA Mathematical Sciences Digital Library, <http://www.mathdl.org/>.
18. Multimedia Educational Resource for Learning and Online Teaching (MERLOT), <http://www.merlot.org/>.
19. MIT OpenCourseWare, <http://ocw.mit.edu/>.
20. National Engineering Education Delivery System (NEEDS), <http://www.needs.org/>.
21. Open Video Project, <http://www.open-video.org/>.
22. Science, Mathematics, Engineering and Technology Education (SMETE) Open Federation, <http://www.smete.org/smete/>.
23. Wisconsin Online Resource Center (Wisc-Online), <http://www.wisc-online.com/>.
24. Creative Commons, 2005. Creative Commons home page: <http://creativecommons.org/>.
25. Bederson, B., et al. HCI Graduate Education Workshop Report, CHI '05. <http://hcc.cc.gatech.edu/chi2005workshop.htm>
26. Carroll, J., ed. *HCI Models, Theories, and Frameworks: Toward a Multidisciplinary Science*. Elsevier, 2003.
27. Day, J., Foley, J., Groeneweg, R., van der Mast, C. Enhancing the Classroom Learning Experience with Web Lectures. GVU Tech Report GVU-04-18.
28. Dix, A., Finlay, J., Abowd, G., and Beale, R. *Human-Computer Interaction*, 3rd ed. Prentice Hall, 2003.
29. Foley, J. et al. Graduate Education in Human-Computer Interaction. In *Extended Abstracts of CHI '05 (Workshops)*, pp. 2113-2114.
30. Fogg, et al. What Makes Web Sites Credible? A Report on a Large Quantitative Study. In *Proc. of CHI '01*, pp. 61-68.
31. Futura, R. et al., Hypertext Paths and the World-Wide Web: Experiences with Walden Paths.
32. Ingham, D., Little, M., Caughey, S. and Shrivastava, S. Fixing the Broken Link Problem: The W3Object Approach. In *Proc. of WWWC '96*.
33. Marchionini, G. and Shneiderman, B. Finding Facts vs. Browsing Knowledge in Hypertext Systems. *Computer* 21 (1), pp. 70-80.
34. Pant, G., Tsioutsoulouklis, K., Johnson, J. and Giles, C. Panorama: Extending Digital Libraries with Topical Crawlers. In *Proc. of JCDL '04*.
35. Phelps, T. and Wilensky, R. Robust Hyperlinks: Cheap, Everywhere, Now. In *Proc. of DDEP '00*.
36. Preece, J., Rogers, Y., Sharp, H. *Interaction Design: Beyond Human-Computer Interaction*. Wiley & Sons, 2002.
37. Kotzé, P., Lorés, J., Oestreicher, L., and Palmer, L. *A Framework for the Development of a Database of Educational Material for Human-Computer Interaction*. Report of Workshop on Developing Educational Material for HCI, NordiCHI 2000.
38. Rosson, M., Carroll, J. *Scenario-Based Usability Engineering*. Elsevier, 2001.
39. Salamapasis, M and Diamantaras, K. Experimental User-Centered Evaluation of an Open Hypermedia System and Web Information Seeking Environments.
40. Shafer, K., Weibel, S., Jul, E. and Fausey, J. Introduction to Persistent Uniform Resource Locators. In *INET '96*.
41. Shneiderman, B., Plaisant, C. *Designing the User Interface*, 4th ed. Addison-Wesley, 2005.
42. Sumner, T. and Marlino, M. Digital Libraries and Educational Practice: A Case for New Models.
43. Sumner, T. Understanding Educator Perceptions of "Quality" in Digital Libraries.
44. Sumner, T. and Dawe, M. Looking at Digital Library Usability from a Reuse Perspective.
45. Xie, H. Shifts of Interactive Intentions and Information-Seeking Strategies in Interactive Information Retrieval.