



The **HCI** *Bibliography:*
**Ten Years Old,
But What's It
Done for Me Lately?**

GARY PERLMAN

E-mail: director@hcibib.org

Web: <http://www.acm.org/~perlman/>

A Brief History of the HCI Bibliography

The Idea

The HCI Bibliography grew out of my experiences in 1988 while writing a curriculum module on user interface development for the Software Engineering Institute [2]. I was delighted to have work-study students type in the bibliographic information for about 200 references; I especially liked having the abstracts and/or tables of contents online. I could search the file, reorder records, import them into bibliography management tools, and so on. I thought, "If a couple of students can put online hundreds of records in a few weeks, what could hundreds or thousands of people do?" In a short time, their efforts could be merged into a bibliography to be used by thousands. Maybe authors could donate records of their own publications!

How It Really Got Started

Of course, that was a naive view. Authors wrote long messages explaining why they could or would not provide an abstracted entry. People donated bibliographic data that had several errors per abstract. The Executive Committee of the Special Interest Group on Computer-Human Interaction (SIGCHI EC) expressed a lack of interest in helping because of the likelihood of failure of such a project. Fortunately, work-study students at The Ohio State University signed up to do data entry, and some people on the Net were willing to validate entries. Also, publishers were willing to give permissions to have their materials online, free of charge. In 1991, after 2 years of getting started, the first paper on the HCI Bibliography (HCIBIB) appeared in the *SIGCHI Bulletin* [3], boasting of more than 1,000 entries and promising that "eventually, all of HCI will be online and freely accessible around the world." Once started, the SIGCHI EC became a consistent supporter and sponsor of the project.

Related Efforts

During the compilation of the HCI Bibliography, other projects "competed." ACM Press had published a book with the title *Resources in Human-Computer Interaction* [1] that was actually a printout of a query performed on

ACM publications. Although it had several indexes, I could not help but think that any printed index would be a relic of outdated ideas. Around the same time, I received in the mail a printed bibliography on hypertext, ordered by author and nicely bound as a report. I could not help but think how ironic it was that a bibliography on hypertext in particular would be (1) on paper and (2) in one organization (and one that was least useful, except perhaps for the authors). I was convinced that online information was the only long-term option and that a format that identified the different parts of entries would allow a variety of search and display options. The UNIX Refer format was chosen because it was simple enough to explain to non-experts.

Another project, HILITES [4], had broader coverage and more features, but was costly to maintain and therefore costly to provide. I surveyed several hundred registered HCIBIB users and concluded that HILITES was beyond the financial reach of most people in HCI and that, by being provided as a CD-ROM, did not serve the needs of many.

Becoming Established

To simplify coverage of the major sources of HCI publications, an early decision was to cover whole journal volumes and conference proceedings that were substantially if not primarily about HCI. Each of these modules was kept in one file (or several files with related names for large conferences). During the early 1990s, the backlog of modules proceedings was added to the HCIBIB database, going as far back as the first volume (1969) of the *International Journal of Man-Machine Studies* (renamed *International Journal of Human-Computer Studies* in 1994). In more recent years, OCR scanning of entries has proved more accurate than having volunteers type, especially when supplemented by hundreds of automated checks.

Although HCIBIB was recognized as the primary source of HCI bibliographic information, it was a database and not a search service. A variety of search services were provided via e-mail and, later, via the Web, which appeared to be very popular, if judged only on the number of requests I received about these

unaffiliated services. None of these services was authorized, and they typically lagged in their coverage, in some cases having less than half the released records. In 1997, the HCIBIB moved to its own domain, hcibib.org, offering Web and FTP access. In April 1998, its search service started.

Services Provided by the HCI Bibliography

Data in the HCI Bibliography

As of September 1998, the HCI Bibliography

- ◆ Covered more than 15 major HCI conferences (www.hcibib.org/confer.html)
- ◆ Covered more than 10 major HCI journals (www.hcibib.org/journal.html)
- ◆ Had more than 18,100 entries (about 400 files, almost 19 megabytes) including
 - About 400 entries on books, many with tables of contents;
 - More than 4,000 links to full text online, most requiring a subscription;
 - About 800 entries on Internet resources, categorized to create the link indexes on the SIGCHI Web site:
 - HCI-Sites (478 links in 13 categories), www.acm.org/sigchi/hci-sites/
 - Publications (59 links in 4 categories), www.acm.org/sigchi/publications/
 - Education (55 links in 3 categories), www.acm.org/sigchi/education/
 - Intercultural (109 links in 12 categories), www.acm.org/sigchi/intercultural/
 - Kids and Computers (60 links in 6 categories), www.acm.org/sigchi/kids/
 - The link index for ACM SIGCAPH: Accessibility (62 links in 5 categories), www.acm.org/sigcaph/links/

All these pages have links to forms to suggest new resources.

One principle of the HCI Bibliography has been that currency of coverage is not as important as affordability, correctness, or portability. Once online in a portable format, materials online will remain online indefinitely. The HCI Bibliography usually has had a backlog of materials to put online and, once online, a backlog of materials to validate. As of

September 1998, both backlogs were relatively low.

Web Site

The HCI Bibliography (<http://www.hcibib.org/>) was redesigned in April 1998 and since then until September 1998 has had more than 17,500 visitors (over 100 per day). Several pages discuss how the project is run (e.g., publisher permissions, data collection and validation, support) and pages about what's new in the database and the search service. Users can browse pages by, for example, publication type, publication date, or date of release. The site includes a list of the most frequent authors (those with 10 or more authored entries in the HCI Bibliography), with links to retrieve all the publications by each author.

Search Service

The HCIBIB search service started in April 1998 and between then and September 1998 has processed more than 30,000 searches (about 6,000 a month). Monthly counts show the service growing from an initial 150 a day to about 300 a day during that period. On September 3, 1998, the service handled more than 1,000 searches for the first time.

The HCI Bibliography search service is based on the glimpse search engine (glimpse.cs.arizona.edu), a free, not-for-profit tool that runs on the server for hcibib.org. Ironically, the search system for the HCI Bibliography has serious usability problems. Compound those with the generally unplanned nature of searches for a free Web service, and it is clear (from the server logs) that many searches miss a lot of what is desired. The search service provides extensive advice about how to improve a search, using knowledge of (1) commonly misspelled author names in HCI, (2) common (non-discriminating) terms in HCI, (3) differences between British and American spelling, and (4) methods to broaden or narrow a search, and by providing relative search term frequencies in the database and frequencies of terms in results. The best searches for various topics are maintained in the HCIBIB database as Internet resources. They can be modified with terms to further restrict a search. For example,

the following search yields more than 1,400 records on hypertext OR hypermedia:

`{hypertext,hypermedia}`

A comma means OR, and braces imply grouping. The search could be modified to find more than 60 records on books:

`{hypertext,hypermedia};isbn`

(semicolon means AND) or almost 500 records with links to full text:

`{hypertext,hypermedia};http`

There are search options to control whether the search is case-sensitive, or whether whole words must be matched. There is an option for an approximate match that will allow one error per search term. These options can have large effects: a whole word search for AI yields 100 records; but anywhere within words, it matches most of the records in the database.

Results can be viewed in HTML or raw Refer format, in brief or detailed views, and search terms can be highlighted in the text. Records contain bookmarks that are actually links to search for a record's identifier; they can be saved for future reuse. International standard book numbers (ISBNs) are displayed as links to amazon.com, from which any royalties are donated to the Central Ohio local ACM SIGCHI chapter, BuckCHI.

Voltaire wrote, "The best is the enemy of the good." So, irony aside, the glimpse search engine lets people search the HCI Bibliography from the convenience of their browsers. Over time, a more usable front end may be provided for the queries. It might make a good project for a user interface course.


Conclusions

I am occasionally asked how the HCI Bibliography project is managed (e.g., how to get per-

missions for materials, how to get them online, how to validate them). These procedures (and their motivations) are documented online, but they have often discouraged people from creating a bibliography service for a field that could use it. It turns out that high-quality bibliographic data are expensive, at least 10 minutes per record when all is counted, but often more. Given that there are too many publications for most people to browse, good bibliographic records provide a reasonable point of access (especially when titles, abstracts, and keywords are well done by authors, which is unfortunately infrequent).

I'm often asked why I put hundreds of hours per year into the HCI Bibliography. Besides the obvious compulsive disorder the work satisfies, I've found that I personally get a lot out of doing work that has lasting value and that is used by other people. It has also been a great source of data (and now a platform) for exploring uses of hypertext for doing research (especially with a search service on the Web). More recently, it has been a reason for me to learn more about how to provide Web-based services, which I can apply in other contexts. And, occasionally, people thank you.

References

1. ACM. *Resources in Human-Computer Interaction*. New York, ACM Press, 1990.
2. Perlman, G. *User Interface Development*. Curriculum Module CM-17. Pittsburgh, Software Engineering Institute, 1989.
3. Perlman, G. "The HCI Bibliography Project." *SIGCHI Bulletin* 23, 3 (1991): 15–20.
4. Shackel, B. et al. "HILITES—The Information Service for the World HCI Community." *SIGCHI Bulletin* 24, 3 (1992): 40–49. 

PERMISSION TO MAKE DIGITAL OR
HARD COPIES OF ALL OR PART OF THIS
WORK FOR PERSONAL OR CLASSROOM
USE IS GRANTED WITHOUT FEE
PROVIDED THAT COPIES ARE NOT
MADE OR DISTRIBUTED FOR PROFIT OR
COMMERCIAL ADVANTAGE AND THAT
COPIES BEAR THIS NOTICE AND THE
FULL CITATION ON THE FIRST PAGE.
TO COPY OTHERWISE, TO REPUBLISH,
TO POST ON SERVERS OR TO REDIS-
TRIBUTE TO LISTS, REQUIRES PRIOR
SPECIFIC PERMISSION AND/OR A FEE.
© ACM 1072-5220/99/0300 \$5.00