



ONLINE DATABASES

Visualize the Perfect Search

By Carol Tenopir

STUDIES SHOW THAT PEOPLE START with a search engine like Google when they need information. As early as 2004, in a focus group for one of my research studies, a college freshman bemoaned, "Why is Google so easy and the library so hard?"

Making the library easy

The Google interface works so well because of the powerful algorithm that lies behind the simple interface and the massive amount of content it covers. Google shows that bringing together millions of indexed web sites and their relative importance based on human judgment (inbound and outbound links) can be a powerful relevance indicator.

The latest to try to get it right by making complexity seem simple is Summon from Serials Solutions (owned by ProQuest), introduced at the American Library Association Midwinter Meeting this year. Though currently only in beta tests at Oklahoma State University and Dartmouth College libraries, Summon is already getting lots of attention.

Summon aims to gather all materials in a library's collection for searching, including the library catalog, e-journals, databases, institutional repositories, and digital library collections. Unlike federated search systems, it will search preharvested indexes of content with weighting applied to aid relevance ranking. (See LibraryJournal.com/summon for more details.)

Finely tuned machine

As with Google, a good search system combines effective relevance ranking with a wealth of content. According to Jane Burke, VP of Serials Solutions, Summon's relevance fine-tuning will be "based on content, machine, and human feedback."

A further critical component will be the weighting of variables based on content type, taking into account individual metadata elements. For example, the information used in fine-tuning the relevance ranking of journal articles includes quality factors such as how many

times an article has been cited, as well as the impact factor and peer review status of the journal in which an article appears. For other types of materials, such as magazine and newspaper articles, currency also will influence relevance.

While it's too soon to tell how well the Summon service will work for its primary audience (students and faculty), Serials Solutions is positioning itself as a starting place for research instead of Google.

For years ProQuest, under the leadership of John Law, has conducted extensive studies on how college students interact

with research with the process of search. The lifetime work of Ben Shneiderman, founder of the Human-Computer Interaction Lab (HCIL) at the University of Maryland, suggests that combining visualization with interactivity allows researchers to discover, not just retrieve.

Visualization approaches use the content of the database or digital library to guide interface choices. For example, the International Children's Digital Library was designed by a team that included kids; it provides search filters for things like color of covers, types of stories,

What is missing [from simple search] is a deep element of interactive discovery

with information for their coursework. These studies have shown that students recognize that the library has high-quality and authoritative content but still choose Google owing to its ease of use. By making it easy to search high-quality content, Serials Solutions hopes to minimize the conflict between the need for quality and the desire for simplicity.

A different tack: visualization

But is the search box solution really the best for facilitating research? Certainly Google and Summon combine both ease of searching and power in retrieval. What is missing is a deep element of interactive discovery that com-

happy or sad books, length of books, and more. PhotoFinder (now Photomesa) displays thumbnails of all photos in a collection, allowing users to mouse over images to see larger versions, zoom in to regroup photos, and restrict a search using criteria from captions such as people's names or subject categories.

Shneiderman and the HCIL conduct numerous user studies to develop and test the visualization approach to information system design and to keep up with changes in user expectations. In the end, they promote a drastically different approach to the search box.

Visual interfaces that display the entire contents of a database and allow manipulation by multiple criteria are much more complex and visually appealing than a simple search box. In fact, they are almost the opposite approach—colorful, cluttered, dynamic, and content-rich.

These extensive research projects seem to point us in different directions. Do we go for visualization and intense interactivity, or for the elegance of a simple search box? The jury is still out.

Carol Tenopir (ctenopir@utk.edu) is Professor at the School of Information Sciences, University of Tennessee, Knoxville

LINK LIST

Human-Computer Interaction Lab
www.cs.umd.edu/hcil

International Children's Digital Library
childrenslibrary.org

PhotoFinder
www.cs.umd.edu/hcil/photolib

Photomesa
photomesa.com

Summon
serialsolutions.com/summon