

The User Interface in Text Retrieval Systems

A Letter to the Editor

Jef Raskin

Dear Editor,

I was disappointed in Offer Drori's article "The User Interface in Text Retrieval Systems" in that it purports to "define the user interface characteristic of text retrieval systems" but only discusses a particular subset of methods for text retrieval, ignoring interfaces techniques that, in many circumstances, offer significant benefits over those he presents.

Drori assumes that (1), "in text retrieval systems the search is generally defined by text (one word or more), with Boolean operators between the words" that (2), "the response ... is a general one, usually comprising the display of all system documents likely to contain a match of the request" and (3), the user first defines a detailed search criterion and then engages a search mechanism.

As Landauer has pointed out [1995] Boolean searches often are not an optimal form of target specification for text retrieval (that they are widely used is insufficient reason to promote them). Rather than presenting a list of documents as the result of a search, text retrieval can generally be improved by presenting the user with each found instance as it is found. Such a search can proceed while the user is inspecting the instance just found, making the system seem much

faster [Raskin, 1989]. An interactive, incremental search is often preferable to either of the two kinds of dialogs Drori discusses because each instance found gives feedback on the appropriateness of the search key while it is being developed, saving the user time and frustration. I have found that users tend to prefer an incremental search (e.g. that of the EMACS

text editor [Stallman 1993]) to the typical FIND dialog box as implemented in most word processors.

Drori correctly supports the concept of displaying a small portion of the document surrounding a match rather than displaying just the title or the title and the first few lines of the document; he also presents this idea as being relatively new. The approach, however, is a form of the concordance, the benefits of which have been known for centuries. Keyword-in-context indices (I implemented one for the ACM in the 1960s) are a well-established paper interface where the found word is presented in context. The Canon Cat interface [Shapiro 1989] provided near-instantaneous incremental search over collections of documents, displaying each found instance in context. Similar systems have been and will continue to be developed on the Internet.

Drori's reliance on Boolean searches, that he ignores incremental and other highly interactive search methods, and his exclusive consideration of systems that accumulate and then present multiple hits, are relics of the past and sadly limit the value of his article to readers seeking to improve interfaces to text retrieval.

Jef Raskin

References

- Landauer, Thomas K. *The Trouble with Computers*. Cambridge MA: MIT Press, 1995
- Raskin, Jef. "Systemic implications of an improved two-part cursor" *Proceedings of CHI 89: Human Factors in Computing Systems* (Austin, 30 April 1989), New York: ACM Press, 1989 pp. 167170.
- Shapiro, Ezra. "The Cat is Dead, Long Live the Interface" *Language Technology Magazine*, No. 13, 1989
- Stallman, Richard. *GNU Emacs Manual*, Ninth Edition, Cambridge, MA: Free Software Foundation, 1993

Author's Address

Jef Raskin, 8 Gypsy Hill
Pacifica CA 94044 USA
Phone: +1-650-359-8588
Fax: +1-650-359-9767
Email: jefraskin@aol.com
Web: www.cfcl.com/jef