Gill Library Publications   Gill Library

1998

Keyword/Subject: Finding a Middle Path

Terry Ballard

Gill Library, College of New Rochelle, tballard@cnr.edu

Recommended Citation

The Systems Librarian

by Terry Ballard

Keyword/Subject: Finding a Middle Path

This hybrid search approach dramatically improves patron success

I t's a long-running debate: How do you deal with the fact that library patrons fail so often when they search controlled subject fields such as Library of Congress or NLM's Medical Subject Headings?

Do you give up and blame the patrons because they won't learn to search the official way, even though the three-volume set of LC Subject Headings is right next to the OPAC terminal? Do you encourage keyword searching and subtly steer the users away from subject searches? Do you put in extra links from their search concepts to the official ones?

The Indisputable Facts

No issue in systems librarianship that I know of has generated more heated discussion. That is interesting, because these arguments are based on a long set of facts that nobody debates. Here are the facts:

- Patrons have, so far, failed to learn the syntax in controlled vocabulary subject searches. After studying this weekly for 8 years, I have consistently found that 45-50 percent of patron subject searches fail to match anything in the catalog. Studies at other schools turn up the same numbers. Furthermore, studies of the hit rate at the reference desk show that even reference librarians fail about 25 percent of the time in subject searching.
- Keyword searches, on the other hand, tend to generate a much higher percentage of hits. The problem is that there are more erroneous matches than controlled subject users encounter. That is because the titles of books are rarely an accurate indication of what is inside. For instance, *Gone with the Wind* doesn't tell you that it is a novel about the Civil War.
- Placing the keyword option higher in the OPAC menu generates a greater percentage of keyword search usage. Labeling subject searches "Library of Congress Subject Headings" brings down the usage of this option.
- Leaving volumes of Library of Congress Subject Headings near the OPAC terminals gives users a place to prop their el-

bows while they go ahead and second guess what the subjects should be.
- Adding "See References" from the searches the patrons make to the actual subject headings helps a little. In other cases, catalogers have lost track of reality. How many of us need to be told not to search by the subject "Keaton, John Frank," but search by the more popular "Keaton, Buster"?

The Disputable Points

What librarians want to do, given these facts, says a lot about their philosophy of what they are doing in the first place. Some say that the subject headings are fine—the problem is with the users. You do as much

as you can with bibliographic instruction, and if they still fail, then it's their own fault for not paying attention. Emphasizing keyword searching to give them more hits is just dumbing down to their level.

Other librarians feel that it is their job to get people to the information they need as smoothly as possible. A failure rate of 50 percent is not a good sign, so the system should be improved somehow. Librarians should point users to search options that will not unnecessarily frustrate them with counter-intuitive searches like "Eliot, T. S. (Thomas Stearns), 1888-1965—Knowledge—Language and languages."

Even though I was recently cited in print as a keyword enthusiast, I am all too aware of the deficiencies of keyword searches of title, notes, and corporate author fields. I have always yearned to provide a keyword index of the subject field, which I'll refer

to as a keyword/subject search option. This would eliminate the syntax constraints in subject searching, while keeping the powerful and useful controlled vocabulary. Earlier this year, I was able to provide this in the catalog that I maintain. I expected improved searching, but the results surprised even me.

The Keyword/Subject Approach

We substituted the new index for the old one and added the LC subject searches as an additional option on the menu. The users adopted the keyword/subject searching by a margin of about 4 to 1. The failure rate settled in at between 12 and 13 percent, and stayed that way for months. One

> ## "
> *Some librarians may object to rewarding a patron with good information when he or she misspelled a word.*
> ### "

interesting thing we found was that the people who sought out LC subject searches after the changeover had a worse hit rate than the overall users before.

Why does this result in a 75 percent reduction in the failure rate? There are a lot of answers to that. I saved a file of the failed subject searches that were performed in the last 10 days before we changed to the new system. First of all, I noticed that keyword/subject searching will return hits in name searching even if the user puts the first name first. Eight failed searches from before the changeover would have worked in the new scheme. One patron had dutifully tried to think like a technical services librarian and entered the failed search "Victoria Queen, Biography." Most of us wouldn't have thought of typing "Victoria, Queen of Great Britain, 1819-1901," but the keyword/subject search picked out the

patron's words there as well as the word biography in "Queens—Great Britain—Biography" to return six pertinent hits. That's what would have happened if the patron had made the search 10 days later. As it is, the evidence shows that he or she went away frustrated.

Even the search "house unmerican [sic] activities committee," which failed in the traditional search mode, brought in eight perfect hits in keyword subject searching. The system matched "house, activities, and committee" and threw out the misspelled word.

Some librarians may object to rewarding a patron with good information when he or she misspelled a word. My response would be along the lines of something that Mark Twain wrote to an overly fussy copy editor, "Take out your mind and dance on it." Of those who object, how many could say without looking that the proper subject heading is "United States. Congress. House. Committee on Un-American Activities"?

The Ideal Marriage

The usage data in the months since we added keyword/subject seems to show that users are learning how to "subject surf," particularly if they are using the Web-based catalog. This means that they put in a request in their own language, find a perfect match, and then link to everything else with that subject heading.

This is the ideal marriage of keyword and formal subject heading searching in my opinion. The new system is not perfect. The browse screen displays do not line up in subject heading order—they simply display an alphabetical list of all titles containing the matched words in the subject fields. Also, we have given up the use of see and see also references.

You can try it out for yourselves on the Web. The catalog is at http://qcat.quinnipiac.edu. The formal subject headings are there, but you need to click on "Advanced Searches" to get there.

Terry Ballard is automation librarian at Quinnipiac College in Hamden, Connecticut. He is the author of INNOPAC: A Reference Guide to the System, *published by Information Today, Inc. He can be reached by e-mail at ballard@quinnipiac.edu; his Web page is located at http://www.geocities.com/Athens/Delphi/3632.*

The Research Investment's Tri-Doc Division Joins OCLC ILL Document Supplier Program

OCLC Online Computer Library Center, Inc. has announced that The Research Investment, Inc., a full-service business and market research firm providing business information worldwide for more than 25 years, has joined the OCLC ILL Document Supplier Program.

Tri-Doc, the document delivery division of The Research Investment, provides articles, books, reports, studies, conference proceedings, patents, company literature,

FDA-mandated clean copies, standards, specifications, and regulations in a variety of languages with a fill rate of 97 percent, according to OCLC. Translations in any language are completed by a network of 15 subject specialists.

"Tri-Doc's expertise in finding the impossible—whether it's unpublished statistics and texts, papers that were presented but not published, program information, memoranda, or in-house documentation—

will be of particular interest to OCLC users," said Kate Vranich, president of The Research Investment. "We are delighted at the prospect of finding the 'tough stuff' for OCLC users."

"The Research Investment is a welcome addition to the OCLC ILL Document Supplier Program," said Marda Johnson, director of OCLC collections and technical services. "OCLC users now have more options for delivery of documents—especially doc-

uments that can be more difficult to find."

Libraries can access The Research Investment document delivery service using the OCLC symbol RSK.

More information on The Research Investment, which is based in Cleveland, Ohio, can be found at its Web site (http://www.researchinvest.com).

The OCLC ILL Document Supplier Program increases access to special collections through commercial vendors and public and private libraries to meet the needs of OCLC users.

Source: OCLC Online Computer Library Center, Inc., Dublin, OH, 614/764-6000; Fax: 614/764-6096; http://www.oclc.org.