

# How successfully do users search the Web?

One real-time site allows you to “spy” on searchers

by Julia K. Nims and Linda Rich

It has long been a desire of academic librarians everywhere to look over the shoulders of library users and observe how they actually search for information. Without blatantly violating our users' privacy and sacrificing our professional ethics, we have been unable to do this with any great success.

This was our predicament until we learned about the McKinley Search Voyeur Web site (<http://voyeur.mckinley.com/cgi-bin/voyeur.cgi>), a real-time site that allows people to view a sample of searches currently entered by users of the McKinley Magellan search engine.

With great anticipation, we visited the site, and it captivated us immediately; first, perhaps, by the large number of sex-related searches, but secondly, because of the large number of obviously poorly planned searches.

As librarians who regularly teach university students search strategies, both for electronic databases and the Web, we began to consider what we could learn from visiting this site. What types of mistakes do people make when searching the Web? Could we generalize that our university's students were conducting similarly poor searches when they searched the Web? If we could make this jump from Magellan users to our local users, how could what we learned from the Web Voyeur help us when we instructed our

students? Ideally, of course, we would like to study the Web searches performed by Bowling Green State University students alone. At this point, such examination is not possible and perhaps not desirable with all the ethical considerations. However, we are not averse to using a “spy” service freely available on the Web.

## Magellan and the “Web Voyeur”

Magellan (<http://www.mckinley.com>) describes itself as “an online guide to the Web that includes original editorial content, a directory of rated and reviewed Web sites, a vast database of yet-to-be reviewed sites, and a powerful search engine . . .” ([http://www.mckinley.com/feature.cgi?faq\\_bd](http://www.mckinley.com/feature.cgi?faq_bd) March 22, 1997). As an online guide to the Web, users can browse the Magellan database, going from one general topic, such as art, to more specific topics such as museums. Additionally, they could use the Magellan search engine to search a huge Web database.

Search Voyeur, a Magellan service, allows people to peek in on searches currently conducted on the Magellan site. The page shows a random selection of current searches and automatically refreshes itself every 15 seconds. If a search seems interesting, voyeurs can click on the search and jump to the results screen. By examining the searches appearing on the “Search Voy-

---

### About the authors

Julia K. Nims is coordinator of library user education and Linda Rich is reference librarian at Bowling Green State University; e-mail: [jk\\_nims@bgnnet.bgsu.edu](mailto:jk_nims@bgnnet.bgsu.edu); [lrich@bgnnet.bgsu.edu](mailto:lrich@bgnnet.bgsu.edu)

eur" screen, we made some observations about the search strategies of the public.

### **Our study**

Before going further, let's clarify what we cannot ascertain from this site. First, as we viewed the searches we kept in mind that we do *not* know the intent of the searcher, and we do *not* know if the searcher was satisfied with the results of the search. Second, we do not know how faithful the Search Voyeur mirrors the searches entered in the Magellan search engines. For example, we could not determine if searches displayed on Search Voyeur were marked to search the entire Web or whether they were marked to only include Magellan's "reviewed sites," an option Magellan gives on its front search screen. We tried to do a search of our own in which we had marked "reviewed sites only." After numerous tries, we found what appeared to be our search on the Search Voyeur site. When we followed the link to the results screen, the screen stated that it had searched the entire Web, not just the reviewed sites. Of course, it is remotely possible that the search we followed on the Search Voyeur was not our search. However, it is important to realize that this occurrence suggests that the Search Voyeur may not always mirror searches entered from the Magellan search page.

We encountered another problem while studying Search Voyeur—the Magellan search engine changed drastically. Early in our study, Magellan used its own search engine software with its own rules for entering precise searches. Toward the end of our examination of Search Voyeur, Magellan switched to new search engine software. This change in software radically changed how Web pages are retrieved since the new search engine software searches for concepts, not only key words. With this change in search engine software, we observed a severe change in the content focus of the searches listed on the Search Voyeur site. When we first visited the site, a large percentage of the searches were sex-related. After the switch in software, very few of the searches had to do with sex at all. Either the users of Magellan suddenly started searching for topics other than sex-related ones or the administrators at Magellan started

filtering some searches from Search Voyeur. (We contacted Magellan and asked for an explanation of these changes. They did not reply to our requests for information.)

Because of these unknowns with the Search Voyeur, we chose to focus our questions on observed search strategies and not on content. We examined the Search Voyeur pages during four, 30 minute sessions over a six-week period, coming up with 1,090 search strings from which we could make generalizations. So what are the mistakes that are being made? We categorized the following "mistakes" as we see them:

- one-word searches
- inclusion of stop words
- typing errors
- inclusion of entire or partial URLs
- exclusion of Magellan's "fine tuning"

suggestions for conducting productive searches.

### **One-word searches**

After studying the Voyeur site for approximately 240 minutes and examining 1,090 individual search strings, we counted 353 one-word searches that appeared on the voyeur screen. While some one-word searches might have been specific enough to retrieve relevant pages, most of the one-word searches would be considered ineffective on any database. For example, we observed searches for each of the following terms: "computer," "women," "sex," and "automobiles." One search that appeared on the Voyeur site was "network." What was this searcher trying to find? Information on television networks or, perhaps, computer networks? Or was the searcher looking for techniques for employee networking within a corporation? We followed the search link, and indeed it had retrieved sites dealing with a great variety of "networks," including computer networks, research networks, and television networks. Even without knowing the searcher's intent, it is safe to deduce that this was a poor search. Perhaps the searcher waded through the large number of unrelated sites to find the needed information, but it could not have been an efficient search.

### **Inclusion of stop words**

Sixty-two of the observed searches included

## Tips for successful Web searching

### In general:

- *Always* check for typos!
- Try searching for multiple terms or phrases rather than a single word.
- *Read the search engine's help screens!*

### Web Search Engine Features

Search Engine	AND, OR, NOT—(+/-)	Truncation	Phrase Searching	Upper/LowerCase	URL Searching
<b>AltaVista</b> www.altavista.digital.com	+/- and, or, and not	use "*"	quotation marks	upper case retrieves exact match	yes (ex. u:image)
<b>Excite</b> www.excite.com	+/- AND, OR AND NOT (must be in caps)	no	quotation marks	matches capital letters	no
<b>HotBot</b> www.hotbot.com	+/- and, or, not	no	quotation marks	insensitive (except for "interesting cases")	no
<b>Lycos</b> www.lycos.com	+/- and, or, not	automatic	quotation marks	insensitive	no
<b>Magellan</b> www.mckinley.com	+/- AND, OR, AND NOT (must be in caps)	no	quotation marks	matches capital letters	no
<b>Yahoo!</b> www.yahoo.com	+/- and, or, not	use "*"	quotation marks	upper case retrieves exact match	yes (ex. u:image)

For more comparisons of Web search engine features, point to:

**Search Engine Comparison Chart—Kansas City Public Library**  
<http://www.kcpl.lib.mo.us/search/chart.htm>

**Search Engine Reviews**  
<http://searchenginewatch.com/reviews.htm>

**Search Engine Tips**  
<http://www.hamline.edu/library/bush/handouts/slatable.html>

stop words, such as *the*, *of*, *in*, and *an*. The Magellan help screens state that the search engine ignores these words, so theoretically it should not matter whether or not users include them. However, some searches appeared to want these words in them, such as the search "help for hire." This search could have been for a company by that name, or for this particular phrase on a page. Without determining how to do a phrase search to include the stop word, the searcher probably had a frustrating time trying to find the desired page from the thousands that appeared with the words "help" and "hire" in them.

### Typing errors

The number of searches containing typographical errors did not surprise us. That is a common mistake all library researchers make at one point or another. Disregarding searches for proper names, possible company or product names, and terms in languages other than English, we noted 86 searches that included blatant typing errors (*sumbarine* instead of *submarine*, for example). What is alarming about these typing errors is that they retrieve results! For instance, the search for the term "mathmetics" retrieved more than 2,000 results. (The creators of the pages had misspelled the word, too.) When discussing search strategies with students in an instructional session, we always suggest that a search on a database that retrieves no results be retried immediately, checking the spelling. When searching the Web, this general principle does not hold true. Virtually any search will bring back some Web pages, and it is entirely possible that the searcher erroneously assumed that he or she had conducted a good search (after all, it did retrieve lots of results).

### Entire or partial URLs

Of the nearly 2,000 searches we observed, 25 contained all or part of a URL. Over the weeks we observed the Search Voyer, the types of results that searches obtained when searching for a URL differed dramatically. In the early weeks of our study, searches for all or partial URLs did not retrieve the page with that URL. Instead, it located pages that had the URL as a link somewhere on

the page. However, in the latter weeks of our study, the Magellan search engine changed, and searches for a URL did jump to pages with that address. Despite the change in the Magellan search engine, we are still concerned about searches for URLs, which imply that the searcher wanted to jump to that page. Rather than typing the URL in the location box, the searcher typed it into the search engine, suggesting an unfamiliarity with navigating the Web.

### Exclusion of Magellan search suggestions

Of all the searches we observed, few of them included Magellan search options designed to retrieve precise results. For example, none of the searches employed the "+" in front of a word instructing Magellan that the word must be in all pages returned or a "-" preceding a word directing Magellan to exclude pages with that word. Furthermore, only 64 of the searches used quotation marks to find specific phrases and none of the searches included any nesting. Less than 13 percent (135 searches) used any Boolean operators. Of the searches that did, none used the operator "not" and only one used the "or" operator. This could be an instance where the Search Voyer does not accurately mirror the Magellan search strings. However, it is even more probable that Web searchers are not reading the help documentation before doing a search on a Web search engine.

### What does this mean to librarians?

None of these errors were a surprise to us. We've observed them in library instruction sessions and at our reference desk. Searches are not thought out properly, errors are made, erroneous assumptions are rampant. What we did find alarming was that the problems poor search strategies create are magnified when they are conducted on the Web. The most obvious problem is that almost any search, no matter how badly constructed, will retrieve something on the Web. The next leap of logic that states that what was retrieved must be good information, is unfortunately all too common an occurrence with researchers. In future sessions on how to search the Web, we will stress these common search strategy errors and suggest ways that our students can avoid them. ■

BLACKWELL'S

# *Collection Manager*

## Collection Development Services

- ✔ Intuitive presentation of approval profiles
- ✔ Web-based review of new title announcements
- ✔ Flexible selection & ordering features
- ✔ Eliminates double-keying of orders
- ✔ Adaptable to library workflows
- ✔ Developed in partnership with librarians

*Collection Manager* — A Complete Workflow Solution

For more information contact • Matt Nauman • 1-800-547-6426 • [mktg@blackwell.com](mailto:mktg@blackwell.com) • Lake Oswego, Oregon