



Evidence Summary

Study of Search Engine Transaction Logs Shows Little Change in How Users use Search Engines

A review of:

Jansen, Bernard J., and Amanda Spink. "How Are We Searching the World Wide Web? A Comparison of Nine Search Engine Transaction Logs." Information Processing & Management 42.1 (2006): 248-263.

Reviewed by:

David Hook

Manager, Operations Information and Configuration Management, MDA

Brampton, Ontario, Canada

E-mail: david.hook@mdacorporation.com

Received: 02 June 2006

Accepted: 18 June 2006

© 2006 Hook. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Objective – To examine the interactions between users and search engines, and how they have changed over time.

Design – Comparative analysis of search engine transaction logs.

Setting – Nine major analyses of search engine transaction logs.

Subjects – Nine web search engine studies (4 European, 5 American) over a seven-year period, covering the search engines Excite, Fireball, AltaVista, BWIE and AllTheWeb.

Methods – The results from individual studies are compared by year of study for percentages of single query sessions, one-

term queries, operator (and, or, not, etc.) usage and single result page viewing. As well, the authors group the search queries into eleven different topical categories and compare how the breakdown has changed over time.

Main Results – Based on the percentage of single query sessions, it does not appear that the complexity of interactions has changed significantly for either the U.S.-based or the European-based search engines. As well, there was little change observed in the percentage of one-term queries over the years of study for either the U.S.-based or the European-based search engines. Few users (generally less than 20%) use Boolean or other operators in their queries, and these percentages have remained relatively stable. One area of noticeable change is in the

percentage of users viewing only one results page, which has increased over the years of study. Based on the studies of the U.S.-based search engines, the topical categories of 'People, Place or Things' and 'Commerce, Travel, Employment or Economy' are becoming more popular, while the categories of 'Sex and Pornography' and 'Entertainment or Recreation' are declining.

Conclusions – The percentage of users viewing only one results page increased during the years of the study, while the percentages of single query sessions, one-term sessions and operator usage remained stable. The increase in single result page viewing implies that users are tending to view fewer results per web query. There was also a significant difference in the percentage of queries using Boolean operators between the US-based and the European-based search engines. One of the study's findings was that results from a study of a particular search engine cannot necessarily be applied to all search engines. Finally, web search topics show a trend towards information or commerce searching rather than entertainment.

Commentary

The authors have undertaken a study of some very interesting and important questions: as the Internet and search engines evolve, and as more and more people use the Internet, how are users' search patterns and techniques changing?

One thing that the researchers of this study do particularly well is to clearly identify their research questions – searching for trends and differences in the number of one-query sessions, one-term queries and results pages viewed, and the changes in the search topics themselves. Their methodology is equally well described, complete with definitions and full explanations. The authors detail an extensive background

literature search and identify the three different types of methodologies used to study web searching in the past. Although there have been many previous studies of search engine use, the authors point out that to date there have been few studies that compare the findings across the different studies. Thus, this study fills a void in the research, complementing existing research.

It is unfortunate that the number of search engines covered by the various studies was very limited. This appears to have been necessary, due to the limited amount of usable data from the other search engines, but meant that trend analysis of U.S.-based search engines was based on only two search engines. Notably missing from the search engines was Google, not only because it is currently the most popular search engine, but also because its development has changed user expectations of search engines and web searches. The limited number of studies and search engines included puts into question the usefulness of the results, but it appears that the researchers were working with what they had available to them.

For the most part, the authors do a thorough job of analyzing and explaining the results of their comparison studies. There were two cases (percentage of one-query sessions and percentage of users viewing only one search results page) where the results of the 1998 study of the Alta Vista search engine had to be discounted because the user sessions were artificially limited to 5 minutes, whereas studies have shown that the average user session is typically 15 minutes in length. This factor presumably affected the way the search engine was used and therefore the results were skewed. While the authors do identify this variable, one wonders if there are other variables in the studies that may have also had an impact on the study results. It is unclear whether or not the authors have investigated the

impacts of other potential variations among the different studies, as this example is one of the few described.

In some cases, the conclusions drawn from the results do not appear to have a base. The article states that “[u]sers have a low tolerance of viewing any results past the first page. They prefer to reformulate the Web query rather than wade through result listings.” Yet the percentages of single query sessions are almost as high as those of the single-page views.

This study does shed some light on the changes in how users are using search engines. In some cases, the authors have found little change in the patterns of search engine use, which, although not as exciting as discovering a change, is just as important to report. Although more research is needed to better understand the variables and comparisons among different search engine studies, this study presents research that is useful to the library and information science profession. Certainly, the trends in search topics are useful for reference and collection planning, but the use of one-term queries among search engine users might also give some insight as to how library users approach reference questions.

Knowledge of users’ search techniques can also be beneficial for those providing library instruction. Knowing that users are typically viewing only the first page of search results, for example, indicates that instructors should be educating users on how to identify and discern the legitimate sources that may be contained on the page. As well, knowing that few users are using advanced queries, should instructors put more emphasis on teaching the use of Boolean operators - or should Boolean searching be taught at all? Similarly, for those designing library search interfaces, knowing that few users use Boolean operators may mean that there is less need to build advanced search capabilities into their features, or perhaps a greater need to simplify them.