Popularity and findability: Log analysis of search terms and queries for public services

Gilad Ravid  
Department of Industrial Engineering and Management,  
Ben-Gurion University of the Negev  
gilad@ravid.org

Judit Bar-Ilan  
Department of Information Science,  
Bar-Ilan University  
barilaj@mail.biu.ac.il

Sheizaf Rafaeli  
Center for the Study of the Information Society,  
University of Haifa  
sheizaf@rafaeli.net

Shifra Baruchson-Arbib  
Department of Information Science,  
Bar-Ilan University  
baruchs1@mail.biu.ac.il

Abstract

The Israeli Citizen’s Advice Bureau (SHIL – an acronym for its name in Hebrew: Sherut Yeutz Laezrach) is charged with providing the public with information about rights, social benefits, government and public services and civil obligations. It does so using sixty centers throughout Israel, phone hotlines and a very active website (http://shil.info). The website began operations in 1994; has undergone a number of major upgrades, and currently is a comprehensive site providing information on the above-mentioned topics. The current site contains interfaces in four languages; it addresses both national and local audiences and topics, and has an active “backoffice” that facilitates constant updating of content. The topics are arranged in a hierarchical directory-like fashion. In addition the site has a very active forum, where the citizens’ questions are answered by SHIL employees and volunteers.

The site is accompanied by an extensive set of logs and tracking mechanisms. Activity on the site approaches 10,000 pages read per day. This measure does not include the interactive use of forums. This is a level high enough to allow analysis of aggregate behavioral patterns.

A closer look at the logfiles of the SHIL website reveals surprising access patterns. About 60.8% of the external hits (requests reaching SHIL from external sites, excluding robots) are from general search engines (e.g. Google and MSN). Thus a significant source of external hits is queries submitted to general search engines, and users reach a specific page on the site linked from the search results page. This finding seems to indicate that the site is not known well enough to the public. On the other hand the site is very active, thus it seems to serve Israeli citizens well, even without being a known brand. In this study we analyzed the queries submitted to the general search engines that lead the users to the SHIL site.

The specific research questions were:

- What citizen’s information is of interest to users who arrive to SHIL on the Web from search engines?
- What are the characteristics of these queries? Are there specific problems because the site is multi-lingual?
• How do these queries compare to those submitted to general search engines and to queries submitted to local search engines?

The analysis provides an insight to the general interests of the Web users reaching the SHIL site. Most of the information sought is citizen-related, and it is surprising to see the large number of requests for SHIL pages on government offices, when these government offices have their own Web sites. Seasonal interests in some queries were also observed. The data analyzed was based on all the external hits during March-October 2005. During this period there were 438,289 external hits that were not submitted by crawlers, out of which 266,295 were external queries. This study is comparable to the recently published study by Chau, Fang and Sheng (2005) analyzing the query logs of the search engine at the Utah state government web site. The general objectives of the Utah state government web site partially overlap with the objectives of the SHIL site.

References: