Arabic Users' Satisfaction with the Online Information as Obtained from Google

Azzah Al-Maskari, Mark Sanderson and Paul Clough
Department of Information Studies
University of Sheffield
Sheffield, UK
Emails: {lip05aaa, m.sanderson, p.d.clough}@shef.ac.uk

Abstract

This paper reports an experimental study that evaluates the retrieval effectiveness of Google in response to 104 Arabic queries. The study investigates Arabic users' satisfaction with the accuracy and coverage of search results. Analysis of results indicates that Arabic users are not highly satisfied with the results of Google.

Introduction

Search engines are among the most popular and useful services on the web. Ingwersen and Järvelin (2005) recommend evaluating an IR system based on how much it helps the user achieve their task effectively and efficiently. Therefore, to evaluate a search engine holistically, the users' rating of the results should be taken into account. Many previous studies have evaluated the effectiveness of web search engines (i.e., Su 1998; Spink, 2002, Su, 2003; Bar-Ilan & Gutman, 2003; Griesbaum, 2004). However to the best of our knowledge there is no previous evaluation of an Arabic speakers' satisfaction with Google.

Methodology

The study asked 26 Arabic speakers to search on four queries from a pool of 104 Arabic queries. Topics were in four categories (Religion, Art, Health, and Politics); each category contained two types of topics (Arabic-specific and general). The motivation for having Arabic-specific topics and general topics stems from the fact that many Arabic people are interested in finding general information which is not specifically related to the Arabic world. The tasks were designed to emulate a simple information-finding task (e.g., find web pages that contain relevant information about Mozart). Users searched directly in Google[1], which was chosen for its popularity and high effectiveness as identified by Hawking et al. (2001). Users were required to save five relevant pages within twelve minutes for each topic. Users were also asked to judge the first ten pages resulting from the best query they issued. They rated the effectiveness of each page at three levels of relevancy: *highly-relevant*[2]; *reasonably relevant*[3]; *not*

[1] http://www.google.co.uk/

[2] The page directly addresses the core issue of the topic

[3] The page only points to the topic, but it does a not discus the themes of the topic thoroughly

relevant[4]. They also rated their satisfaction[5] with the "accuracy[6]", "coverage[7]" and "ranking[8]" of the results.

Findings

We have compared the effectiveness of Google in retrieving information for the Arabic-specific and general topics based on user effort and time spent in completing the tasks. User effort is measured by the number of viewed pages (either reading the snippet or opening the actual page). Results show that there is no significant difference in the number of pages viewed between the Arabic-specific and general topics across all categories. Similarly, no significant difference is found in the time taken to complete the tasks.

Table 1 lists the measures used to evaluate the effectiveness of Google: Precision (obtained from users rating of pages relevancy), users' satisfaction with accuracy, coverage and ranking of the results. Arabic topics are significantly (p=0.01) more precise than general topics; consequently users are significantly more satisfied with accuracy (p=0.02) and ranking (p=0.009) of Arabic-specific topics than the general topics.

	Arabic-specific topics	General topics
Precision @ 10	0.69	0.59
Users' sat. accuracy	0.64	0.47
Users' sat. coverage	0.61	0.50
Users' sat. ranking	0.60	0.44

Table 1: List of measures utilized in this study

Users rated the easiness[7] of finding five relevant pages per topic. As anticipated, Arabic-specific topics were significantly easier (p= 0.02) than the general topics. This can be explained by two factors: (1) the wider availability of online information that discusses Arabic-specific topics than general topics, and (2) the deeper knowledge shown by Arabic users for Arabic-specific topics. Surprisingly, we found low Pearson's correlation between easiness and precision of the results (0.31), easiness and users' satisfaction with accuracy (0.47), easiness and users' satisfaction with coverage (0.33) and easiness and completion time (0.34).

Upon completion of the search tasks, users were asked to give their feedback about Google. This included how Google helped them in their search, and their satisfaction with the coverage, accuracy and overall satisfaction of results. The effectiveness of Google was rated on a five-point scale from highest (5) to worst (1). The results are shown in Table 2 which indicates that most users are satisfied with only 32% of Google's results.

[4] The page does not contain any information about the topic.

[5] very satisfied=1, partially satisfied=0.5, not satisfied=0

[6] Precision of the results; [7] Completeness of the results; [8] Order of the results

[7] very easy=1, partially easy=0.5, not easy=0

	5 (highest)	4	3	2	1 (least)
Coverage	19.23%	34.62%	30.77%	11.54%	3.85%
Accuracy	15.38%	38.46%	34.62%	11.54%	0.00%
Overall satisfaction	15.38%	30.77%	34.62%	11.54%	7.69%

Table 2: Overall users' satisfaction of Google

Conclusion

This small research study quantified the effectiveness of Google search results in response to 104 Arabic queries. Analysis of results indicates that Arabic information is not precise, thus Arabic users are not highly satisfied with the results. This is likely due to a lack of valid and reliable online information in Arabic. Therefore, we would like to carry out future research by comparing Google versus Araby[8], a new Arabic search engine. A further study will repeat the 104 Arabic queries in English and compare the results of the two sets of languages (Arabic and English).

Acknowledgement

Work partially supported by European Community under the Information Society Technologies (IST) programme of the 6th FP for RTD - project MultiMatch contract IST-033104, and the Ministry of Manpower, Oman. The authors are solely responsible for the content of this paper. It does not represent the opinion of supporting funding agencies, and the supporters are not responsible for any use that might be made of data appearing therein.

References

- Bar-Ilan, J. & Gutman, T. (2003) How do search engines handle non-English queries?-A case study. *WWW2003 Proc. of the Twelfth International World Wide Web Conference,* 415.
- Hawking, D., Craswell, N., Bailey, P. & Griffihs, K. (2001). Measuring Search Engine Quality. *Information Retrieval,* 4, 33-59.
- Ingwersen, P. & Järvelin, K. (2005). The turn: integration of information seeking and retrieval in context, Springer.
- Gurrin, C. & Smeaton, A. (2003). Improving the evaluation of Web search systems. In F. Sebastiani (Ed.), *Advances in information retrieval: 25th European Conference on IR Research, ECIR 2003, Pisa, Italy, April 14-16, 2003. Proceedings*, (pp.25-40) Berlin; New York: Springer. (Lecture Notes in Computer Science **2633**)
- Spink, A. (2002). A user centered approach to evaluating human interaction with Web search engines: an exploratory study. *Information Processing & Management*, **38**(3), 410-426.
- Su, L. T., Chen, H. & Dong, X. (1998) Evaluation of Web-Based Search Engines from the End-User's Perspective: A Pilot Study. *Proceedings of the ASIS Annual Meeting,* 35**,** 348-61.
- Su, L. T. (2003) A comprehensive and systematic model of user evaluation of Web search engines: I. Theory and backgr. *Journal of the American Society for Information Science and Technology. Hoboken,* 54**,** 1175, 18 pages.

[8] http://araby.com/