How Medical Expertise Influences Web Search Interaction

Ryen W. White
Microsoft Research
Redmond, WA 98052
ryenw@microsoft.com

Susan Dumais
Microsoft Research
Redmond, WA 98052
sdumais@microsoft.com

Jaime Teevan
Microsoft Research
Redmond, WA 98052
teevan@microsoft.com

ABSTRACT
Domain expertise can have an important influence on how people search. In this poster we present findings from a log-based study into how medical domain experts search the Web for information related to their expertise, as compared with non-experts. We find differences in sites visited, query vocabulary, and search behavior. The findings have implications for the automatic identification of domain experts from interaction logs, and the use of domain knowledge in applications such as query suggestion or page recommendation to support non-experts.

Categories and Subject Descriptors
H.3.3 [**Information storage and retrieval**]: Information Search and Retrieval – *query formulation*; *search process*.

General Terms
Measurement, Experimentation, Human Factors.

Keywords
Domain expertise, Web search.

1. INTRODUCTION
The amount of knowledge a person has about a particular domain can be an important determinant of their search behavior and ultimately, their search success. Domain expertise has been studied extensively in the information science community [6]. It differs from search expertise, in that it concerns subject-matter knowledge rather than search knowledge. Studies of domain expertise have highlighted several differences between experts and novices, including: site selection and sequencing [3], task completion time [2], vocabulary and search expression [1], and the number and length of queries, and search effectiveness [7]. These studies involved small numbers of subjects with carefully controlled tasks, making it difficult to generalize their findings.

This poster presents a large-scale log-based study of Web search behavior. It contrasts the search strategies of domain experts with those of domain non-experts through analysis of naturalistic interaction log data over a three-month period of time. This large-scale analysis allows us to identify greater diversity in vocabulary, site visits, and user tasks than with smaller-scale studies, *e.g.*, [3]. We focus on the medical domain because of the complexity of the subject matter and potential benefit to novices of identifying effective search strategies. In addition to highlighting differences in the search behavior of experts and non-experts, we describe the potential benefits of being able to identify domain experts and leverage their querying strategies and source selection abilities.

2. IDENTIFYING DOMAIN EXPERTS
To study the effects of domain expertise on Web search, we first selected a set of people interested in the medical domain. From this group of interested people we separated experts from non-experts based on visitation to a specialist medical search engine.

SIGIR'08, July 20–24, 2008, Singapore.
ACM 978-1-60558-164-4/08/07

2.1 Data Sources
We used the interaction logs of over 500 thousand consenting users from a three-month period from May 2007 to July 2007. From these logs we extracted: (i) *browser trails*, a temporally-ordered set of URLs comprising all pages viewed per Web browser instance or browser tab, and (ii) *search sessions,* subsets of browser trails, beginning with a query to a search engine such as Google, Yahoo!, or Live Search, and terminating with a period of user inactivity of 30 or more minutes. This threshold has been used previously to demarcate search sessions in logs [4].

2.2 Identifying Topical Interest
To identify users interested in the medical domain we classified pages in the browser trails into the topics from a Web directory. Given the large number of pages involved, the classification needed to be automatic. Our classifier assigned labels to pages based on the top level of the Open Directory Project (ODP) in a similar way to [5], by starting with URLs that were in the ODP and backing-off to cover other URLs. Using this classifier we identified the proportion of pages that each user visited that were medically related (*i.e.*, those in the "Health" category of the ODP). While the goal was to identify people with some degree of interest in medical topics, we also wanted to remove outliers who viewed mainly medical pages (*e.g.*, data-entry workers employed to process prescriptions), or viewed few medical pages, since both groups may skew our analysis. Thus, we selected people whose page views contained between 5% and 50% medically-related pages. This resulted in the selection of 16,658 medical users. The logs contain 1,129,398 search sessions for these users.

2.3 Separating Experts from Non-Experts
To identify experts and non-experts, we divided users based on whether they had ever visited or queried for the PubMed search engine.[1] PubMed is used primarily by medical researchers and physicians, and provides access to citations and abstracts of biomedical research articles. PubMed visitation therefore seemed to be an appropriate expertise filter. This approach classified 15.6% of the people interested in the medical domain as experts and 84.4% as non-experts. We also classified each search session from these two groups as medical or non-medical based on whether the session contained a page tagged with the ODP label "Health" by our classifier. This provides us with medical experts and non-experts, engaged in medical and non-medical sessions.

3. EXPERTS VERSUS NON-EXPERTS
We compared the characteristics of the pages visited (source selection), search queries, and search sessions of experts and non-experts. Table 1 shows the mean (M), standard deviation (SD), and median values for search and browsing features for each group during medical and non-medical search sessions. Given the large sample sizes, all observed differences in the means between groups were statistically significant (*i.e.*, all $p < .001$, independent

[1] http://www.pubmedcentral.nih.gov

Table 1. Features of Web search interaction for experts and non-experts in medical and non-medical sessions.

Session	Features			User					
				Expert			Non-expert		
Medical	Number of sessions / Number of queries			39,599 / 502,511			90,035 / 804,012		
				M	SD	Median	M	SD	Median
	Query	Length	Tokens	3.57	2.04	3	3.32	1.89	3
			Characters	28.60	11.16	25	22.39	10.31	20
		% of queries w/ technical terminology		**48.32**	**38.11**	**44**	**15.29**	**14.85**	**14**
	Session	Length	Pages	36.10	45.11	22	35.94	46.67	22
			Queries	12.69	13.68	7	8.93	11.38	5
			Seconds	1686.05	1849.49	962	1444.23	1634.41	925
		Branches		8.91	11.36	5	5.71	8.74	3
		Unique domains		8.15	6.52	6	6.13	4.68	5
Non-medical	Number of sessions / Number of queries			211,599 / 969,123			788,165 / 3,373,346		
				M	SD	Median	M	SD	Median
	Query	Length	Tokens	3.10	2.27	3	2.96	2.06	3
			Characters	21.04	14.38	18	19.89	11.88	17
		% of queries w/ technical terminology		4.01	10.03	3	3.11	7.76	2
	Session	Length	Pages	18.07	29.53	8	19.85	33.12	9
			Queries	4.58	8.08	2	4.28	7.76	2
			Seconds	734.69	1211.79	276	762.03	1251.29	296
		Branches		4.25	7.31	2	4.5	7.81	2
		Unique domains		4.09	3.98	3	4.27	4.04	3

measures t-tests). We applied Cohen's d-tests to determine the effect size for each between-group comparison and highlight the pertinent effect sizes in the following findings summary. We begin by summarizing interaction patterns for medical sessions.

Source selection: Experts visited different sites than non-experts. Domain experts primarily visited technical sites (*e.g.*, nih.gov, mayoclinic.com, nejm.org), and non-experts visited consumer-oriented sites about health and wellbeing rather than only medical issues (*e.g.*, about.com, calorie-count.com, webmd.com).

Queries: Experts issued longer queries than non-experts (d= 0.62). We also analyzed the technical nature of the query vocabulary used by the two groups. To quantify this we computed the proportion of queries that contained complex clinical terminology from parts of the Unified Medical Language System (UMLS) MetaThesaurus.[2] As can be seen in Table 1 (bolded row), experts used many more technical query terms than non-experts (d=1.14).

Sessions: Experts exhibited different search behaviors than non-experts. Per session, experts: issued more queries (d=0.30), branched more (*i.e.*, revisited an earlier page in the session and then browsed to a new page) (d=0.32), visited more unique domains (d=0.36), and spent longer searching (d=0.14).

While there were substantial differences between medical experts and novices for medical sessions, the differences were relatively small for non-medical sessions. This suggests that our participants were similar in their general search behaviors, and differed mainly in sessions where domain expertise was relevant.

One challenge in interpreting these results is that experts and non-experts may be searching for different things rather than searching differently – that is, the observed differences may be task differences rather than expertise differences. To address this concern we developed two methods to identify comparable tasks. For one, we identified search sessions that began with the same query. There were 12 queries describing medical conditions (*e.g.*, pregnancy, diabetes, HIV), issued at least ten times by both experts and non-experts. We also identified sessions that ended with the same URL. For the matched queries and sessions, the between-group differences noted earlier in this section held. Thus, it seems that even for similar information needs experts and non-experts still search differently.

[2] http://www.nlm.nih.gov/pubs/factsheets/umlsmeta.html

4. CONCLUSIONS

We have described the first large-scale log-based study of domain expertise in Web search. Our findings show that we can identify domain experts (and comparable non-experts) from logs. In line with previous work (*e.g.*, [6]), we showed that domain experts search differently than non-experts in terms of the sites they visit, the query vocabulary they use, and their patterns of interaction. The biggest differences were in their queries and source selection. We believe that these features can be used to automatically identify and quantify domain expertise from search and browse logs. Identifying domain experts makes it possible to provide expert query suggestions and site recommendations to non-expert users, and to personalize search results based on expertise. Our future work will focus on the development of such applications and on the extension of the methods described to other domains.

5. REFERENCES

[1] Allen, B.L. (1991). Topic knowledge and online catalog search formulation. *Lib. Quart.*, 61(2): 188-213.

[2] Bhavnani, S.K. (2001). Important cognitive components of domain-specific search knowledge. *Proc. TREC*, 571-578.

[3] Bhavnani, S.K. (2002). Domain-specific search strategies for the effective retrieval of healthcare and shopping information. *Proc. SIGCHI*, 610-611.

[4] Downey, D., Dumais, S. & Horvitz, E. (2007). Models of searching and browsing: Languages, studies and application. *Proc. IJCAI*, 2740-2747.

[5] Shen, X., Dumais, S. & Horvitz, E. (2005). Analysis of topic dynamics in Web search. *Proc. WWW*, 1102-1103.

[6] Wildemuth, B. M. (2004). The effects of domain knowledge on search tactic formulation. *JASIST*, 55(3): 246-258.

[7] Zhang, X., Anghelescu, H.G.B. &Yuan, X. (2005). Domain knowledge, search behavior, and search effectiveness of engineering and science students: An exploratory study. *Inf. Res.*, 10(2): p 217.