Beyond the Search Box:
Helping Users Find Health Information on the Web

Kevin Duh
University of Washington
Seattle, WA
kevinduh@u.washington.edu

Shawn Medero
Healia
Bellevue, WA
shawn@healia.com

Mike Schultz, Tom Eng
Healia
Bellevue, WA
{mikes,tom}@healia.com

ABSTRACT
Internet users are increasingly relying on the Web for health information. Their information needs can often be quite complex, ranging from researching a personal illness to comparing the pros and cons of various treatments. We believe that a search interface beyond the traditional search box is necessary to support users in making informed health decisions. In this paper, we describe the search interface of Healia, a consumer health search engine, which contains advanced search features such as personalization, faceted browsing, and query suggestion. We present some analyses of the query logs to seek to understand how users interact with our search interface.

Categories and Subject Descriptors
H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces.

General Terms
Measurement, Design, Human Factors.

Keywords
Health Vertical Search, Search Interface Design, User Behavior

1. INTRODUCTION

eHealth is an industry of growing importance. The Internet provides opportunities for users to seek health advice from potentially millions of online peers and experts at any time of the day. Forrester Research found that as many as 84% of American Internet users have researched health information online in 2006 [1], and that the majority of these searches involve questions relating to specific medical conditions of the searcher or searcher's family/friends [2]. As the amount of health content proliferates on the web, there is an increasing demand for search engines and portals to organize and filter information in a personalized fashion.

Information need for health-related questions may be quite complex and varied, but we can categorize users into two general groups. In the first group, users may have been just diagnosed by a health professional with a certain illness, and is motivated to understand specific issues related to the illness in detail. Queries such as "What are the treatments for a 5 year-old with strep throat?" or "Clinical trials for diabetes in African American women" indicate the need for **highly personalized** (e.g. 5 year-old, African American women) as well as **highly specific** (e.g. treatment, clinical trials) results. In the second group, users may be attempting to self-diagnose prior to a hospital visit.[1] In this case, queries may be **underspecified** as users may not have the medical expertise to know what to search for, and an interactive interface may be needed to help users explore the options. In fact, a user study reported on the Journal of the American Medical Association [3] has concluded that "using search engines and simple search terms is not efficient."

Our goal is to develop a better search engine and search interface to support users in understanding health information and making health decisions. This work examines the search interface deployed by Healia, a health-related vertical search engine that focuses on the above challenges (i.e. highly personalized/specific results, underspecified queries).[2] The paper is divided as follows: First, we describe Healia's search interface, highlighting the features we believe are important in supporting user interaction and information finding in health. Then, we present results from query log analysis, which show how these advanced features are utilized. Finally, we present our conclusions and thoughts on future work.

2. HEALIA SEARCH INTERFACE

The Healia Search Interface, which can be accessed at http://www.healia.com (a screenshot is shown in Appendix A), consists of five main areas of user interaction: a search box, a personalization filter, faceted browsing, suggested query terms, and entry to Pubmed/Clinical Trials information. We imagine the searcher may use this interface in the following scenario:

---

[1] [2] also reports that for an increasing number of young users, the Internet is the preferred source to learn about health.

[2] Another major challenge for health search engines is to provide information that is credible and trustworthy. In this paper we focus on the interface aspects and will not discuss how we optimize the Healia search engine to achieve this.

1. Enter query term, e.g. diabetes, and see initial results.

2. Personalize the results with the filter, e.g. click on "Female" and "African American" to return results specific to a demographic. The personalization filter also allows filtering of results based on reading level and accreditation.

3. Explore the various facets of diabetes, which includes "Prevention," "Causes," "Symptoms," "Diagnosis," and "Treatment".

4. Try the suggested query terms, which proposes similar searches and more specific/general medical terms.

5. Further, if the user is determined to understand more, the entry points to Pubmed journal articles and clinical trials information provide a way to sift through expert information.

We can view user interaction with Healia as the following diagram (Figure 1), where the searcher is given one of five actions.[3] Upon choosing an action, the searcher will see a new results page and can continue interacting with the system with different actions until satisfaction.

In the following, we will study user behavior on the Healia website under the framework of these five user actions.
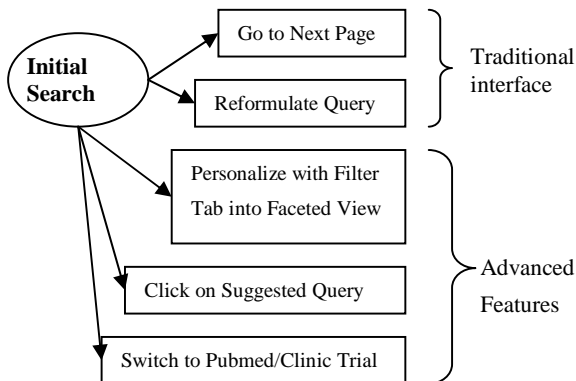
Figure 1: Five possible user actions are available to a searcher on each Healia search results page.

3. QUERY LOG STUDY

We are interested in understanding how searchers use the Healia search interface, in particular, to what extent are the advanced search features used by different types of users. To this end, we mine the query logs to obtain statistics of the five different user actions defined in Figure 1. We filter the log such that only user sessions with one or more actions after the initial search results page are used (i.e. we do not consider cases where the user

---

[3] In this study, we will not examine other Healia features, such as the entry points to the Healia Health Guides (editorial content) and the Healia Communities social support network. Also, we consider personalization filters and faceted tabs as the same type of user action since they both involve filtering the current list of search results.

session ends after a single query and there is no further interaction with the system).

Following the work of [4], we divide our users into regular users and "expert" users, where "expert" is defined by whether the user enters Healia's PubMed search interface to access scientific journal article. Manual inspection of these "expert" search queries reveal many technical terms and PubMed author names, leading us to believe that these searches are meant to pinpoint specific documents and is therefore qualitatively different from the complex and exploratory search tasks of a consumer health user. Among the 6800 unique users in our data, roughly 8% were classified as "expert" under this heuristic.

3.1 What are the most frequent actions taken by users?

First, we measured the frequency of each user action and show the results in Table 1. We observe that the traditional search interface actions of "Reformulate Query" and "Go to Next Page" consists of the majority (82.7%) of user actions and the advanced search features are utilized with less frequency (17.3%) in total. Among the advanced features, "Suggested Query" and "Personalization / Faceted Tab" are used equally often. Interestingly, many user-entered query reformulations are often achievable by personalization filters and tabs, for example:

- "strep throat"(original query) → "strep throat in children" (reformulated query, typed in by user)
- "quit smoking"→ "quit smoking methods"
- "uterine infection"→ "cause of uterine infection"

These query reformulations reflect the need to get more personalized and specific information, which is exactly what can be accomplished by the advanced features, but users often chose to type additional query terms (which is more time consuming). The reason may be that users now are used to the single box search interface.

Table 1 also shows that expert users use advance features roughly 3%-5% more than regular users.
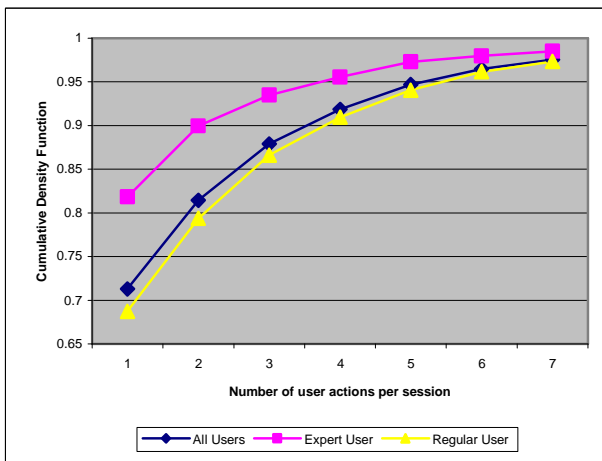
Table 1. Percentage of User Actions

User Action	ALL USERS	EXPERT USERS	REGULAR USERS
Reformulate Query	47.9	43.4	49.5
Go to Next Page	34.8	33.7	35.7
Personalization Filter / Tab into Facet	8.3	12.7	7.5
Click on Suggested Query	7.6	10.2	7.3
Switch to PubMed or Clinical Trials	1.4	-	-

3.2 How long do users interact with the search engine?

Second, we calculated the length of a user session, in terms of the number of user interactions. Long user sessions indicate an extended interaction with the search interface. Figure 2 shows the cumulative density function for user actions: 71% of all user

sessions end after one user action, 81% of all user sessions end with two or less user actions, and 91% of all user sessions end with four or less user actions. The majority of user sessions are short, but there are a significant number of extended interactions.

We also observe that the sessions of expert users are shorter than that of regular users. Two possible explanations are: (1) the search tasks of regular users are more complex and require extended interaction; (2) expert users used advanced search features more often than regular users, thus finding information faster. Further work is needed to test these hypotheses.

Figure 2: Number of user actions per session. This is a cumulative plot indicating the percentage of user sessions with less than or equal to X user actions.

3.3 How diverse are the actions in each user session?

We are also interested in seeing whether users employ a variety of actions in a user session, since a diversity of actions implies the user's sophistication with the search strategy. We found that users tend to stick to a few actions (possibly due to familiarity with its intended results): Of all the sessions that have at least three actions, 44% involve only one type of action, e.g.:

- reformulate query → reformulate query → reformulate query

- next page → next page → next page

42% of user sessions involve two types of actions, e.g.:

- reformulate query → next page → reformulate query

- suggested term → personalize → personalize

Only 12% of user sessions involve three or more types of actions.

3.4 What kinds of personalization filters and facets are being used?

Figure 3 indicates facet usage by measuring the percentage of time each facet tab is clicked on in the query log. We find that users are most interested in the "symptoms" facet of their search results, implying that users are indeed using the Internet as a tool for self-diagnosis. In fact, as many as 20% of distinct queries entered in

conjunction with faceted tabs contain the words "photo" or "picture" (e.g. "pictures of pink eye", "scabies photo").
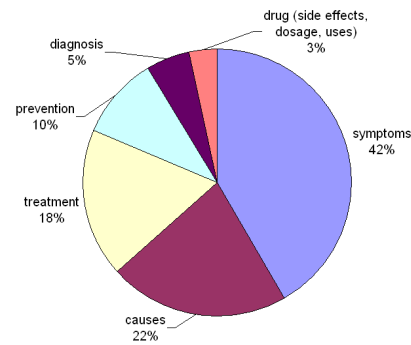
Figure 3: Facet usage. Users are most interested in finding out about "symptoms" (42%), "causes" (22%), and "treatments" (18%) of diseases.

Figure 4 shows the percentage of time each type of personalization filter is used. Users most often filter results by "gender" and "age." The more popular setting for the gender filter is "female" (68%); for the age filter, the breakdown is "kids" (57%), "teens" (26%), "seniors" (17%). These statistics may have interesting implications as to who may be the main consumers of Internet health information (i.e. women and parents).
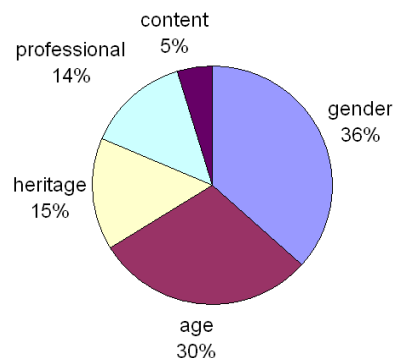
Figure 4: Personalization filter usage. Users filter results most often by gender (female/male) and age (kids/teens/senior), followed by heritage (African/Asian/Hispanic/Native), professional, and content (e.g. easy to scan, interactive tools)

4. SUMMARY AND PROSPECTS

We have advocated that a vertical search engine for health should provide features that support the complex information need of users, which can be *highly personalized*, *highly specific*, and *underspecified.* Consumer health search is an "exploratory search" problem [5] where users are "searching to learn." Our query log study of the Healia search interface found that:

1) Users sometimes opt to use the traditional single search box paradigm even when advanced features provide one-click solutions to personalization and more specific information. Nevertheless, we observe a promising ~17% usage of advanced features on Healia.

2) Expert user sessions are shorter than those of regular users. It is not yet clear whether this is due to simpler information need for technical searches, or faster task completion since experts use more advanced features.

3) User interactions with the search interface are not very diverse, with only 12% of user sessions involving three or more actions.

4) The most commonly-used facet is "symptoms", implying an audience that uses health search for self-diagnosis. Commonly-used filters are gender and age.

We are interested in the following open questions:

- How do we design search interfaces so that advanced search features can be easily learned and adopted?

- What other advanced search features are useful in helping consumer health users make informed health decisions?

Regarding the first point, it has been shown by [6] that a user who learns a good search strategy performs significantly better in retrieving domain-related information. Further, [7] presents design recommendations for making faceted search, in particular, more effective.

We have recently built a new version of the Healia search interface, which includes federated search (of the Web, PubMed, and Clinical Trials), a more streamlined presentation of filters and tabs, and significant improvements in response time for user interactions. We believe these enhancements will further improve the user experience; it would be interesting to perform a comparative study of query logs between these two versions for evaluation purposes.

5. REFERENCES

[1] Forrester Research. (2006). North American Consumer Technology Adoption Study (NACTAS) Q4.

[2] Hanson, J. (2007). How Different Generations Use Online Health Research. Forrester Research.

[3] Berland, G.K., et. al. (2001). Health Information on the Internet—Accessibility, Quality, and Readability in English and Spanish. Journal of the American Medical Association.

[4] White, R., Dumais, S., Teevan, J. (2008). How Medical Expertise Influences Web Search Interaction. SIGIR.

[5] Marchionini, G. (2006). Exploratory search: From finding to understanding. Communications of the ACM, 49 (4).

[6] Bhavnani, S.K. (2002). Domain-specific search strategies for the effective retrieval of health and shopping information. SIGCHI.

[7] Hearst, M. (2006). Design recommendations for hierarchical faceted search interfaces. SIGIR Faceted Search Workshop.

Appendix A: Healia Search Engine Interface, August 2008 (http://www.healia.com)