

Findability in Health Information Websites

Hamman W. Samuel
Department of Computing
Science
University of Alberta
Edmonton, Alberta, Canada
hwsamuel@cs.ualberta.ca

Osmar R. Zaiane
Department of Computing
Science
University of Alberta
Edmonton, Alberta, Canada
zaiane@cs.ualberta.ca

Jane Robertson Zaiane
School of Library and
Information Studies
University of Alberta
Edmonton, Alberta, Canada
jane.zaiane@ualberta.ca

Abstract—In this study, we investigate how health information consumers locate content on health information websites. Preliminary results show that there is room for improvement in terms of finding specific content on health websites, that is, findability. We focus on and identify usability issues with three key aspects of health websites: search box, navigation menu, and home page. Results are based on a population sample of users with varied backgrounds, familiarity with medical terms, and a diversified range of question types. Consumer trends in looking up information demonstrate that using the search box is the method of choice, while navigation menus and links on the home page are not effectively being utilized. Ultimately, we propose possible solutions aimed at improving the overall quality of health information websites, such as faceted search, metaphor exploration, multi-dimensional views, and trending topics.

I. INTRODUCTION

The Pew Internet and American Life survey recently reported that as of 2011, 80% of Internet consumers look up healthcare-related information online[9]. Compared with the 2000 figure of 55%, it can be seen that the Internet is playing an increasingly important role in finding health information[16]. Despite the popularity of health information websites, it is unclear whether health information consumers are able to easily locate the information they are seeking. Keselman et al. point out that consumers often find it difficult to locate health information online[11]. In other words, the degree of findability needs to be examined.

The term ‘findability’ is well-known in the area of information architecture, and means the ease of locating information on a website[13]. When applied to health information websites, findability is a measure of how easily specific health-related content can be found by an information consumer. Findability can be viewed from two perspectives of the consumer: the visibility of a website from the Internet via search engines, or visibility of specific content on the website itself. Our focus is on the latter of these. Consequently, we assume that the consumer is already on a particular website and will use facilities available on the website itself to find pertinent information.

There has been previous research on the topic of findability within health websites. Fisher et al. looked at improving the consumer’s search experience on health information portals[8]. They investigate usability on five health websites, but their focus is on improving search capabilities using

ontologies. Also, Bentley has recently looked into enhancing findability of healthcare portals by incorporating semantic search[1]. We extend these approaches by including two other key features that are used in locating information, namely navigation menus and the home page. Madle et al. recently carried out a survey of the WHO Labresources portal involving twenty-one participants. Their results showed that browsing via the navigation menu was more popular than using search box[12]. However, since the invited participants were all public health professionals, they were expectedly more familiar with medical terminology and categorization of medical topics. We look at other profiles of health information consumers based on different literacy parameters: information, technological, and health literacy.

In our study, we look at findability in health information websites by briefly examining trends in how consumers seek information on three top health information websites: WebMD.com, Yahoo! Health, and MayoClinic.com. We use the device of an online survey with tree testing. Tasks are modeled around key features of health websites that are meant to be helpful in locating information: search box, navigation menu, home page. We note that these features are cross-cutting and independent of the technology platforms, whether desktop, laptop, tablet, or mobile. We carried out the survey in two phases: trial phase and pilot phase. Our results show that using the search box is the most popular method, while navigation menus alone and links on the home page are seldom used. Ultimately, based on the observations, we propose possible solutions with faceted search, metaphor exploration, multi-dimensional views, and trending topics.

II. METHODOLOGY

The survey study was carried out in two phases: an initial trial phase, followed by a broader pilot phase.

A. Target Websites Selection

For the trial phase, we arbitrarily chose five health information websites with varying popularity based on their traffic: WebMD, Yahoo! Health, PatientsLikeMe, HealthPost, and Doctissimo. For the pilot phase, selection of the target websites was a three-step process. First, we looked at the top health websites as reported by CAPHIS[4] and eBizMBA[6]. Next, we checked the number of unique visitors to these sites from Compete.com and sorted them. Finally, we selected

the top three popular websites, which were WebMD, Yahoo! Health, and Mayo Clinic. A summary of rankings of the top ten of these health websites is provided in Figure 1.

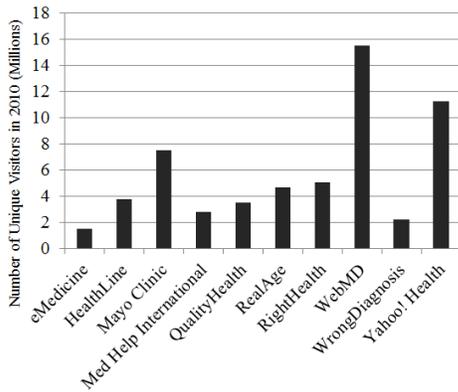


Fig. 1. Top Health Information Websites by Traffic [Data: Compete.com]

B. Survey Formulation

We used an online survey to get feedback results. Our survey was modeled as a questionnaire with elements of tree testing incorporated. In a tree test, consumers are asked to group a given subject under the correct menu category on a website[21]. The questions were based on one task per website, which included one tree test task. Consumers performed the task by navigating to the health website and then returned to the questionnaire to give feedback about the task. The results of the survey comprised of feedback on these tasks from the respondents. The nature of the task was uniform and involved giving participants a question related to health and asking them to visit a particular health website to locate the answer.

In the trial phase, five questions from top ten health questions of 2010 searched on Ask.com were arbitrarily chosen and assigned to each of the websites. For the pilot phase, the health questions for the tasks were chosen randomly from a mixed selection of three main sources: 1) top ten health questions of 2010 searched on Ask.com[10], 2) top health questions of 2010 reported on CNN.com[3], and 3) top questions of 2010 and 2011 asked of Alberta Health Services, Canada[20], [19]. Selection criteria was based on variety in the question types: image search, multiple-choice, binary choice, and multiple-answer questions. Also, the use of popular questions allows our study to be more in line with real-world scenarios. Each health question was then assigned to a health website, which led to the formulation of the survey tasks. A follow-up question after the respondent completes the task was asked to determine the completion of the task. The three health questions eventually selected, along with response types for the follow-up questions are listed in Table I.

Feedback on the tasks included four feedback questions and measured responses for each feedback question. The first question evaluated the respondent's completion of the task by asking a question related to the task. The second question

was based on tree testing and asked the respondent to identify under which menu title the answer to the task was located at. Next, two subjective questions were asked about 1) the ease of completing the task, and 2) the preferred method for finding answers: search box, menu navigation, home page links, or others. There was also allocation for free-form text responses about general experience with each health website.

C. Survey Administration

In the trial phase, the survey was administered through general invitations on social networking websites like Facebook. However, in the pilot phase, the environment was more controlled based on literacy levels. Petch's profiling of health information consumers shows that the level of literacy, health literacy, and technological literacy affect their approach to finding information[15]. Norman and Skinner also suggest a profiling of consumers based on six essential literacies that contribute to overall e-health literacy. In addition to traditional literacy and health literacy, Norman and Skinner include scientific literacy, and sub-divide technological literacy into information, media, and computer literacy[14]. Cultural literacy is also an important aspect, because consumers in different countries would use the same terms differently in the context of their culture.

The levels of information, computer, and health literacy were 1) controlled, and 2) measured in the pilot phase. The literacy levels were controlled by inviting select categories of users to take the survey. This included two types of users with educational or vocational backgrounds in 1) library or information science, 2) computing science. These categories of users matched with expected high information and computer literacy levels. A third category of users was general users, who represented user profiles most likely found in real-world scenarios. Their literacies were not pre-gauged. We received a total of seventy-two responses. These responses underwent a clean-up process in which incomplete responses were removed, leading eventually to fifty tabulated responses. Among these, thirteen respondents were from the high information literacy category, 24 were from the high computer literacy group, and thirteen were from the unknown literacies category. About 68% of the respondents were female, while 32% were male. Also, a 72% majority of the respondents were in the 20-35 years age group, while 24% were in the 36-55 years category, and a minority 4% were less than 20 years. About 10% identified their area as medicine/health care, while only 24% out of all the respondents were non-students.

Measurements of literacy were done as follows. In an initial page of the questionnaire, respondents were asked to rate their familiarity with medical terms on a three-point scale, i.e. health literacy. In addition, respondents were asked the frequency of Internet usage on a three-point scale, i.e. computer literacy. They were also asked to rate their own ability to find/locate information on a website using a three-point scale, i.e. information literacy. Gender, occupation, and age category were also asked of the respondents. Figure 2

TABLE I
LISTING OF HEALTH QUESTIONS AND HEALTH WEBSITES

Health Question	Follow-up Question	Response Type	Health Website
What are the symptoms/signs of breast cancer?	Which of the following symptom(s) is/are for breast cancer?	Multiple Answers	WebMD
What is the treatment for chicken pox?	Which of the following is a treatment for chicken pox?	Multiple Choice	Yahoo! Health
What does herpes look like? Look for images.	Were you able to view any images/pictures of what herpes looks like?	Binary Response	Mayo Clinic

shows the measured literacy distributions of various users based on feedback on an initial page in the survey.

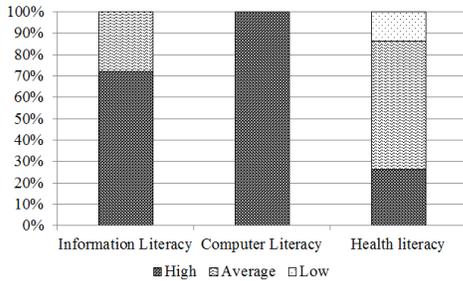


Fig. 2. Measured Literacy Levels of Respondents

All the respondents identified themselves as having a high proficiency with computers, while the majority were of high or average information literacy. A minority of respondents were either highly proficient with medical terms or had low health literacy, with most having average health literacy. Based on these profiling details, the average respondent who took the survey can be described as follows: female student between 20-35 years old who is very good at using computers and finding information on websites, but has an average knowledge of medical terms.

III. RESULTS

In this pilot survey, the measured responses to each feedback question were counted and summed, and an average of each measured response was taken across the three websites and the different literacy levels. The first question related to each task was used as an accuracy checker. Results of the survey showed that most participants found information they were looking for fairly easily and quickly. The results also showed that using the search box was the more popular method of locating information as shown in Figure 3, with all the results as percentages. Whether the allocated task was completed with a correct answer or not is referred to in Figure 3 as ‘completeness’, while ‘categorization’ refers to the respondent being able to identify the menu item under which their answers were located. The respondent’s subjective responses to ‘ease’ of finding the results, and their preferred ‘method’ for locating information are also shown.

Respondents generally found it hard to identify the menu categorizations across all the literacy levels. It should be noted that these respondents all had a high level of computer literacy, and on average information literacy levels were also

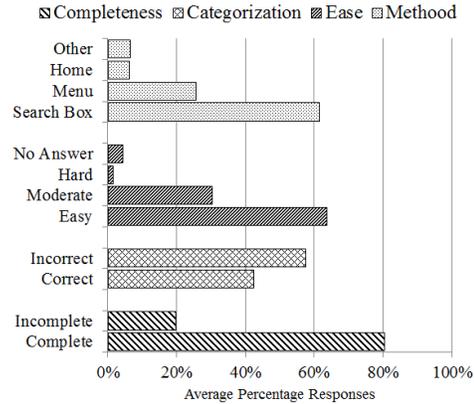


Fig. 3. Survey Feedback Summary with Average Percentage Responses

good. Other methods users identified that helped them in locating the required information included the ‘find’ function of the browser, and external search like Google. One user commented in reference to the question on locating an image that they prefer to use Google Image Search and then redirect to the appropriate website from there.

In the trial phase of the survey, similar results were obtained. Figure 4 shows a summary of the trial phase, in which the percentage of respondents using navigation only (N), search only (S), or home page only (H) were analyzed. The trial phase also looked into how popular combinations of these methods were, such as search and navigation (SN), search and home page (SH), navigation and home page (NH), and all three methods together as well (SNH). Even in the trial phase, search was the method of choice.

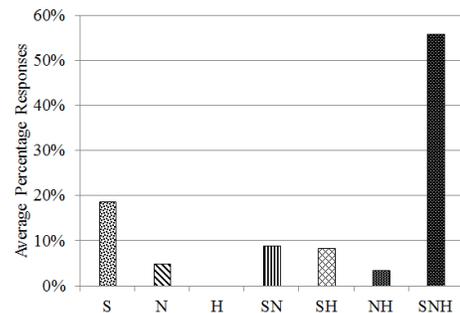


Fig. 4. Use of Search box (S), Navigation menu (N), and Home page (H) in the Trial Phase

IV. DISCUSSION

The home page is the first point of contact of the consumer. Our pilot study shows that the home page's link on its own are not useful in overall findability of a website. Almost 6% of users indicated they used home page links, while about 7% used other methods, such as the Google Search 'site:' hack. We propose that the home page links can be improved and contribute to the findability of the website if the trending topics and frequently asked questions are shown on the home page. Health websites can dynamically refresh their home page content based on consumers' search and click-through habits.

Keselman et al. pointed out that navigation menus are commonly not effective because of a disparity between consumer and the website terminology[11]. Navigation menus require consumers to have a high-level idea of what they are looking for in the domain language, which may not always be the case. Consequently, consumers may find using the navigation menu counter-intuitive. In the field of information architecture, metaphor exploration is a method that can be used to create better and more intuitive navigation menus. Metaphor exploration is focused on coming up with creative relationships by relating familiar ideas with new ones[17]. Health websites can re-think and leverage appropriate metaphors to make the navigation menu more intuitive.

Given that the search box is popular, it should be highlighted on the home page. In line with [8], [1], we propose that providing additional features for a search box will enhance the findability of the health website. In line with the need for navigability, faceted search is ideal. Faceted search allows search results to be additionally grouped under various categories and taxonomies[23]. In addition, instead of aggregating search results from the whole website, search results can be grouped within a category, topic or web page. By increasing the granularity of search results, consumers could be able to locate information more efficiently. Also, when looking for new or unknown medical terms, consumers are likely to misspell words. In order to provide a mapping between lay and technical terms, we propose incorporating multi-dimensional views into health information websites similar to [22], [24], [18].

V. CONCLUSION

We looked at findability in health information websites by briefly examining consumer trends in seeking information. Our results showed that using the search box is the most popular method, while navigation menus alone and links on the home page are rarely used. We proposed possible improvements to findability with faceted search, metaphor exploration, multi-dimensional views, and trending topics. We emphasize that this is preliminary work, and the results serve as a base for further research and investigation. The results have provided suggestions and hypotheses that can be tested in future work. We intend to carry out even more controlled surveys with larger participation and varied literacies.

REFERENCES

- [1] J. Bentley. How Can Healthcare Portals Boost Their 'Findability'? <http://www.healthcare-digital.com/blogs/editor/how-can-healthcare-portals-boost-their-findability>, Mar 2011. Digital Healthcare.
- [2] G. K. Berland, M. N. Elliott, L. S. Morales, J. I. Algazy, R. L. Kravitz, M. S. Broder, D. E. Kanouse, J. A. Munoz, J. A. Puyol, M. Lara, K. E. Watkins, H. Yang, and E. A. McGlynn. Health Information on the Internet: Accessibility, Quality, and Readability in English and Spanish. *Journal of the American Medical Association*, 285(20), 2001.
- [3] E. C. CNN. Your Top Health Searches, Asked and Answered. <http://www.cnn.com/2010/HEALTH/10/21/top.health.searches.answered/index.html>, Nov 2010.
- [4] Consumer and P. H. I. S. (CAPHIS). Top 100 List: Health Websites You Can Trust. <http://caphis.mlanet.org/consumer/>, 2010.
- [5] Z. Dou, S. Hu, K. Chen, R. Song, and J. R. Wen. Multi-dimensional search result diversification. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, pages 475–484. ACM, 2011.
- [6] eBizMBA. Top 15 Most Popular Health Websites. <http://www.ebizmba.com/articles/health-websites>, 2011.
- [7] G. Eysenbach and C. Khle. How Do Consumers Search For And Appraise Health Information On The World Wide Web? Qualitative Study Using Focus Groups, Usability Tests, And In-Depth Interviews. *British Medical Journal*, 324(7337):573–577, 2002.
- [8] J. Fisher, F. Burstein, R. Manaszewicz, , and K. Lazarenko. Health Information Portals: How Can We Improve the User's Search Experience. In *Proc. 17th European Conference on Information Systems*, pages 1–13, 2009.
- [9] M. Freudenheim. Health Care Is High Among Web Searches. <http://www.pewinternet.org/Media-Mentions/2011/NYT-Health-Care-Is-High-Among-Web-Searches.aspx>, Feb. 2011. Pew Internet and American Life Project.
- [10] D. Goodwin. Ask.com's Top Questions of 2010: Users Prefer Dunkin' Coffee. <http://blog.searchenginewatch.com/101201-213649>, Dec 2010. Search Engine Watch.
- [11] A. Keselman, A. C. Browne, and D. R. Kaufman. Consumer Health Information Seeking as Hypothesis Testing. *Journal of the American Medical Informatics Association*, 15(4):484–495, 2008.
- [12] G. Madle, A. Berger, S. Cognat, S. Menna, and P. Kostkova. User Information Seeking Behaviour: Perceptions and Reality. An Evaluation of the WHO Labresources Internet Portal. *Informatics for Health and Social Care*, 34(1):30–38, 2009.
- [13] P. Morville. *Ambient Findability*. O'Reilly Media, 2005.
- [14] C. D. Norman and H. A. Skinner. eHealth Literacy: Essential Skills for Consumer Health in a Networked World. *Journal of Medical Internet Research*, 8(2), 2006.
- [15] T. Petch. Content Analysis of Selected Health Information Websites: Final Report. Technical report, Action for Health, Simon Fraser University, 2004.
- [16] L. Rainie and S. Fox. The Online Health Care Revolution. <http://www.pewinternet.org/Reports/2000/The-Online-Health-Care-Revolution.aspx>, Nov 2000.
- [17] L. Rosenfeld and P. Morville. *Information Architecture for the World Wide Web*. O'Reilly Media, 1998.
- [18] SerialSolutions. Aqua Browser. <http://www.serialssolutions.com/aquabrowser/>, Mar 2011.
- [19] A. H. Services. 2010 Top Health Questions. <http://www.albertahealthservices.ca/3664.asp>, Nov 2010.
- [20] A. H. Services. Top Health Questions. <http://www.albertahealthservices.ca/580.asp>, Nov 2011.
- [21] D. Spencer. Card-Based Classification Evaluation. http://www.boxesandarrows.com/view/card_based_classification_evaluation, Apr 2003. Boxes and Arrows.
- [22] TagGalaxy. Tag Galaxy. <http://www.taggalaxy.de>, Mar 2011.
- [23] D. Tunkelang. *Faceted Search*. Morgan and Claypool, 2009.
- [24] VisualThesaurus. Think Map. <http://www.visualthesaurus.com>, Mar 2011.
- [25] P. Younger. Internet-based Information-seeking Behaviour Amongst Doctors and Nurses: A Short Review of the Literature. *Health Information and Libraries Journal*, 27(1):2–10, 2010.