

NOTICE: Author's Post-Print Version (Final Draft Post-Refereeing)

This is the author's version of a work that was accepted for publication in:

The 28th Australian Conference on Computer-Human Interaction (OzCHI '16)

Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication.

The author acknowledges that the copyright of the final version is owned by the publisher. The final version can be accessed with the following link:

<http://dx.doi.org/10.1145/3010915.3011843>

What Are Health Website Visitors Doing? Insights from Visualisations Towards Exploratory Search

Patrick Cheong-Iao Pang Mitchell Harrop Karin Verspoor Jon Pearce Shanton Chang

Department of Computing and Information Systems

The University of Melbourne, Parkville, Victoria, 3010, Australia

mail@patrickpang.net, {mharrop, karin.verspoor, j.pearce, shanton.chang@unimelb.edu.au}

ABSTRACT

Understanding the navigation flows of health website visitors can lead to strategies for improved support for health information-seeking. In this study, we have obtained the navigation data of visitors to Better Health Channel – one of the largest consumer health information websites in Australia, and visualised the data to compare the different patterns of visitors' activities, as accessed on desktop and mobile devices. These visualisations provide insights about the preferences for search strategies, and the traffic flows patterns of visitors. We then discuss these insights with the existing work of health information-seeking behaviour. This paper extends our observations to design recommendations, specifically to facilitate exploratory search in health information-seeking, and establishes directions for future research.

Author Keywords

Health information-seeking behaviour; information needs; navigation flows; information visualisation

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

In Health Information-Seeking Behaviour (HISB), people often adopt exploratory search (Pang et al., 2015a), which causes them to extensively look into different sources for making sense about a loose search target (Marchionini, 2006). While health information websites play an important role in HISB, many do not support exploratory search (Pang et al., 2015b). Recent HCI research suggests that the design of consumer health information websites should support different information needs and the use of exploratory search during health information-seeking (Pang et al., 2014; Lee et al., 2015). In this paper, we analysed the traffic of a health website and we will propose some navigation patterns about these different types of searches in the health setting.

In this study we collaborated with Better Health Channel (BHC – <http://www.betterhealth.vic.gov.au>). BHC is one

of the most popular consumer health websites in Australia. Our sample data were drawn from a Google Analytics account during a three-month period from May to July 2016. With these data, we have created two interactive visualisations (one for the desktop and another for the mobile version of the website) to show the categories of health information read by visitors, as well as the number of users travelling from and to these categories. We used the analysis of this website as a case study to inform the design of future websites that support various types of HISB.

The visualisations give insights into the information needs of BHC website users and reveal navigation patterns corresponding to exploratory search in HISB. These insights include the starting and ending points of their information-seeking journey, the subsequent information that the users pursue in their sessions, and their information-seeking behaviours in the desktop and the mobile counterparts. From these observations, we present recommendations to build health websites that satisfy users' needs and information-seeking behaviours.

RELATED WORK

Health information seekers not only demand different types of information but also use different approaches to look for the information, namely: focused and exploratory search (Lee et al., 2014; Pang et al., 2016). Some seekers prefer reading little relevant information with focused search, while others are keen to learn every aspect about a health issue comprehensively with an exploratory one (Pang et al., 2016). Therefore, understanding the web of traffic health websites can help to design for users' diverse needs, search approaches and behaviours.

In this study we used the data captured by Google Analytics. Prior research has investigated its usefulness for scientific research (Plaza, 2011; Pakkala et al., 2012). In addition, Google Analytics has been specifically used for studying general information-seeking behaviour (Clark et al., 2014), and evaluating a web-based medical intervention (Crutzen et al., 2012).

Information visualisation is a useful technique to make sense of the large amount of data captured by Google Analytics. Among types of visualisations for illustrating connections and flows (Fruchterman and Reingold, 1991; Biuk-Aghai et al., 2011; Biuk-Aghai et al., 2014), we chose Sankey diagrams (Tufte, 1983) for displaying our data. This algorithm does not have performance problems with large datasets, and is able to clearly illustrate the origin and destination relationship found in web traffic (Riehmman et al., 2005; Wongsuphasawat and Gotz, 2012).

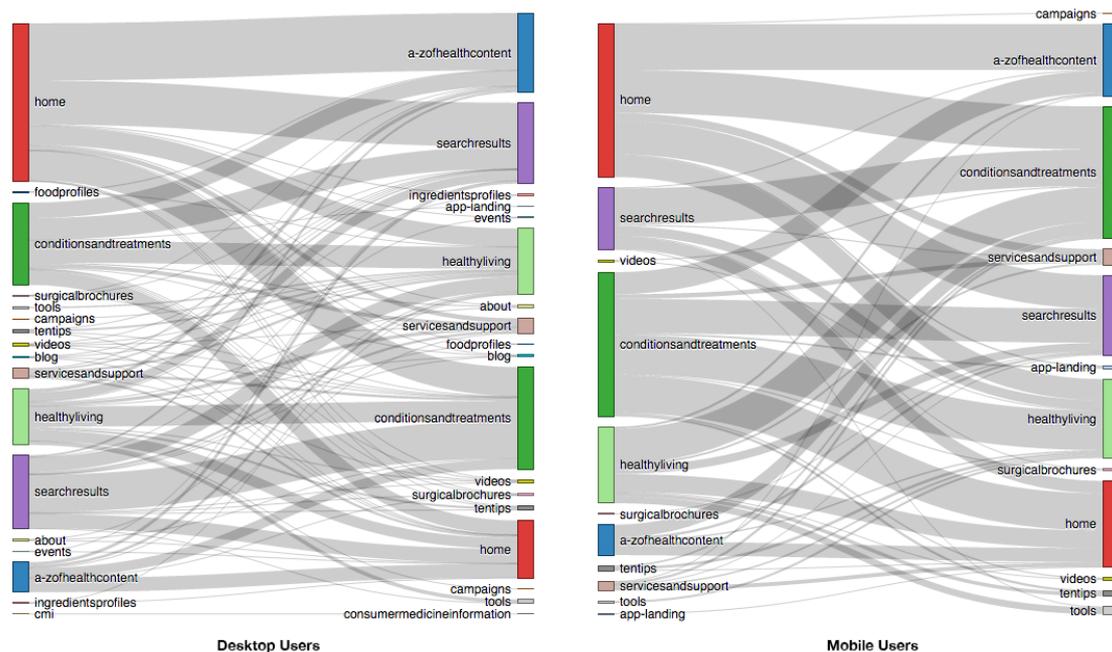


Figure 1. Visualisations of navigation flows of BHC visitors (traffic flows from left to right).

THE DATASET AND THE VISUALISATION

We obtained the visitor data from May to July 2016 of the BHC website. In this research period, the website recorded 2,041,395 unique users who viewed more than one web page, consisting of 1,149,024 (56%) desktop and 892,371 (44%) mobile users. The categories of origin and destination web pages, which are extracted from the URL path, are visualised in two interactive Sankey diagrams using HTML5 and D3 library. The visualisations for both desktop and mobile environments are shown in Figure 1. The items in the visualisations are positioned to reduce the lengths of the flows and to minimise overlaps. The interactive version is located at <http://bit.ly/2bAVgLj>

OBSERVATIONS AND DISCUSSIONS

Seeing which categories users access and move between can give us interesting insights about the information they need. As seen in Figure 1, the traffic from “conditions and treatments” and “healthy living (styles)” is dominant among the website content (other than the home page and the search results, which are typical entry points to the site). This indicates that these two categories of health information are often accessed and health websites should focus on enhancing the content in these two categories.

The destinations of outgoing traffic after reading a given page in the BHC website are also worth examining, as they indicate users’ subsequent information needs. We can see that outgoing traffic from the “conditions and treatments” and “healthy living (styles)” pages flows to a mix of different categories, or returns back to the home page/search results. In other words, a number of users continue to look for additional information after the first read, by navigating to other pages, browsing the home page, and using the search function. Also, the scope of the seeking of information is not limited to a single type of information. This implies the existence of exploratory search, in which users find, investigate and learn information in different parts of the site (Marchionini, 2006; Pang et al., 2014; Pearce and Chang, 2014).

We identify that both browsing and searching were the techniques used to find information in BHC. The browsing activities usually involve reading web pages and using menus and links for navigation. These activities are reflected in the significant flows that travel from the home page to “A-Z health content” (i.e. the menu of all health issues) and directly to the “conditions and treatments” categories. On the other hand, substantial traffic leaves from “search results”, which indicates that this group of users continued to read in the website after performing their searches.

In addition, we found that visitors do more than simply read static information in BHC. The visualisations illustrate several flows towards “tools” and “services and support”, which originate from the major content categories of the site. The “tools” category provides interactive content such as body mass index (BMI) calculators, while “services and support” offers information about health service providers and supporting communities. This observation is consistent with other research that suggests the needs of self-management (Lee et al., 2014), and making informed decisions of using the healthcare system after HISB (Hersh, 2009).

CONCLUSIONS AND FUTURE WORK

This paper presents visualisations of the navigation flows of both desktop and mobile users of a health website (BHC) in Australia. The visualisations reveal insights about how people find health information using the exploratory search approach. While both searching and browsing are used, people tend to explore and obtain a variety of health information, instead of just focusing on a single topic. In addition, we found that the interactive tools for enabling self-management and accessing support services are vital in health information-seeking. A major limitation is that the data were obtained only from BHC, and cannot necessarily be applied to other health websites. Future work would focus on generalising our findings and further investigating these observations.

ACKNOWLEDGMENTS

We acknowledge Better Health Channel for their data and the support from the Department of Health and Human Services of the Victorian State Government, Australia in this research.

REFERENCES

- Biuk-Aghai, R.P., Pang, C.-I., Cheang, F.H.H., 2011. Visualization of Large Category Hierarchies, in: Proceedings of VINCI '11. p. 2:1--2:10. doi:10.1145/2016656.2016658
- Biuk-Aghai, R.P., Pang, C.-I., Si, Y.-W., 2014. Visualizing large-scale human collaboration in Wikipedia. *Futur. Gener. Comput. Syst.* 31, 120–133. doi:http://dx.doi.org/10.1016/j.future.2013.04.001
- Clark, D.J., Nicholas, D., Jamali, H.R., 2014. Evaluating information seeking and use in the changing virtual world: the emerging role of Google Analytics. *Learned Publishing*, 27(3), pp.185–194.
- Crutzen, R., Roosjen, J.L., Poelman, J., 2012. Using Google Analytics as a process evaluation method for Internet-delivered interventions: an example on sexual health. *Heal. Promot. Int.* doi:10.1093/heapro/das008
- Fruchterman, T.M.J., Reingold, E.M., 1991. Graph drawing by force-directed placement. *Softw. Pract. Exp.* 21, 1129–1164. doi:10.1002/spe.4380211102
- Hersh, W., 2009. *Information Retrieval: A Health and Biomedical Perspective*, 3rd ed. Springer-Verlag, New York. doi:10.1007/978-0-387-78703-9
- Lee, K., Hoti, K., Hughes, J.D., Emmerton, L., 2014. Dr Google and the Consumer: A Qualitative Study Exploring the Navigational Needs and Online Health Information-Seeking Behaviors of Consumers With Chronic Health Conditions. *J. Med. Internet Res.* 16, e262. doi:10.2196/jmir.3706
- Lee, K., Hoti, K., Hughes, J.D., Emmerton, L.M., 2015. Consumer Use of “Dr Google”: A Survey on Health Information-Seeking Behaviors and Navigational Needs. *J. Med. Internet Res.* 17, e288.
- Marchionini, G., 2006. Exploratory search: from finding to understanding. *Commun. ACM* 49, 41. doi:10.1145/1121949.1121979
- Pakkala, H., Presser, K., Christensen, T., 2012. Using Google Analytics to measure visitor statistics: The case of food composition websites. *Int. J. Inf. Manage.* 32, 504–512. doi:10.1016/j.ijinfomgt.2012.04.008
- Pang, P.C.-I., Chang, S., Pearce, J., Verspoor, K., 2014. Online Health Information Seeking Behaviour: Understanding Different Search Approaches. Proceedings of PACIS 2014.
- Pang, P.C.-I., Chang, S., Verspoor, K., Pearce, J., 2016. Designing Health Websites Based on Users' Online Information Seeking Behaviours: A Mixed-method Observational Study. *J. Med. Internet Res.* 18, e145. doi:10.2196/jmir.5661
- Pang, P.C.-I., Verspoor, K., Chang, S., Pearce, J., 2015a. Conceptualising health information seeking behaviours and exploratory search: result of a qualitative study. *Health Technol. (Berl.)* 5, 45–55. doi:10.1007/s12553-015-0096-0
- Pang, P.C.-I., Verspoor, K., Pearce, J., Chang, S., 2015b. Better Health Explorer: Designing for Health Information Seekers, Proceedings of OzCHI '15. pp. 588–597. doi:10.1145/2838739.2838772
- Pearce, J. & Chang, S., 2014. Exploration Without Keywords: The Bookfish Case. Proceedings of OzCHI '14, pp.176–179. doi:10.1145/2686612.2686639.
- Plaza, B., 2011. Google Analytics for measuring website performance. *Tour. Manag.* 32, 477–481. doi:10.1016/j.tourman.2010.03.015
- Riehmann, P., Hanfler, M., Froehlich, B., 2005. Interactive Sankey diagrams. Proceedings of INFOVIS 2005. doi:10.1109/INFVIS.2005.1532152
- Tufte, E.R., 1983. *The visual display of quantitative information*. Graphics Press, Cheshire, CT.
- Wongsuphasawat, K., Gotz, D., 2012. Exploring Flow, Factors, and Outcomes of Temporal Event Sequences with the Outflow Visualization. *IEEE Trans. Vis. Comput. Graph.* doi:10.1109/TVCG.2012.225
- Wu, Y., Liu, S., Yan, K., Liu, M., Wu, F., 2014. OpinionFlow: Visual Analysis of Opinion Diffusion on Social Media. *IEEE Trans. Vis. Comput. Graph.* doi:10.1109/TVCG.2014.2346920