WikiGaze

October 30, 2019

[1]. It has been consistently ranked in the top ten visited sites on the Internet as per Alexa.com.

[2] This study shows that Wikipedia's content quality is comparable to traditional encyclopedias.

[3] Search engines such as Google and Bing respond to user queries by extracting structured knowledge from Wikipedia.

[4] Observes that collaborative knowledge building portals mimic the actual knowledge building process.

[5] Even with "Anyone Can Edit" feature, it dominates in terms of information provision. On wikipedia, users play a variety of role where readers are also the part of knowledge building process.

[6] Eye movement characteristic of reading is extensively studied in the psychology literature. Eye movement tracking open the door for an automated analysis of document reading.

In [7] a system is proposed for recording and analyzing web reading behavior using eye gaze.

Sanches et. al [8] show that prediction of the subjective understanding is improved by 13% if we use eye gaze instead of comprehension questions.

[9] performed on 477 Wikipedia based researches revealed that 42% of these reseachers were centered on understanding characteristics of contributors involment or article quality evaluation [10, 11, 12]. Only 20% of the studies related to readers in Wikipedia and the usage of Wikipedia. Less than 1% of the reviewed studies looked at users' reading preferences.

Lehmann et al. [13] emphasized on the importance of reading behavior analyses. The research characterizes users' reading preferences at article level i.e. it determines which articles are more "engaging" than others.

In past studies [14, 15], it was revealed that our eye gaze pattern is closely related to our thought process. Through eye gaze tracking, we can get the knowledge about the portions of an article where a user is more focused while reading.

There is evidence in research from reading psychology that eye movement patterns while reading are indeed related to textual features [16].

Personalization has been identified as being one of the grand challenges in information retrieval lately [17].

Personalized summarization [18] presents users with document extracts that are of interest to them.

Majority of the automatic personalized summarization techniques incorporate some kind of personal information for individually improving the quality of summary [19, 20, 21].

Well established devices, like Tobii [22], EyeLink [23], etc. are available for gaze tracking.

CVC ET [24, 25] is the port of Opengazer for the Linux and Mac platform and NetGazer [26] is the port of Opengazer for the Windows platform.

Xu et al. [27] talk about the relevance between the human text reading pattern and their current cognitive process. Their basic assumption is that the eyes' fixation duration on a word is directly equivalent to the user's interest in that word.

Article length is considered an effective metric for analysing an article [28, 29].

Falagas et. al [30] demonstrates the impact of article length on the number of citations that it might get in future.

Blumerstock [31] shows that article length is a critical parameter to differentiate featured articles from other articles in Wikipedia.

But it can not be said that all long articles are good. [31] shows several counterexamples that prevent using it alone as a quality measure.

References

[1] Wikipedia, "Wikipedia:Statistics — Wikipedia, the free encyclopedia." http://en.wikipedia.org/w/index.php?title=Wikipedia%3AStatistics&oldid=914609482, 2019. [Online; accessed 02-October-2019].

[2] J. Giles, "Internet encyclopaedias go head to head," 2005.

[3] T. Bergstrom and K. Karahalios, "Conversation clusters: grouping conversation topics through human-computer dialog," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2349–2352, ACM, 2009.

[4] U. Cress and J. Kimmerle, "A systemic and cognitive view on collaborative knowledge building with wikis," *International Journal of Computer-Supported Collaborative Learning*, vol. 3, no. 2, p. 105, 2008.

[5] J. Antin and C. Cheshire, "Readers are not free-riders: reading as a form of participation on wikipedia," in *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, pp. 127–130, ACM, 2010.

[6] K. Rayner, "Eye movements in reading and information processing: 20 years of research.," *Psychological bulletin*, vol. 124, no. 3, p. 372, 1998.

[7] D. Beymer and D. M. Russell, "Webgazeanalyzer: a system for capturing and analyzing web reading behavior using eye gaze," in *CHI'05 extended abstracts on Human factors in computing systems*, pp. 1913–1916, ACM, 2005.

[8] C. L. Sanches, O. Augereau, and K. Kise, "Estimation of reading subjective understanding based on eye gaze analysis," *PloS one*, vol. 13, no. 10, p. e0206213, 2018.

[9] C. Okoli, M. Mehdi, M. Mesgari, F. Å. Nielsen, and A. Lanamäki, "The peoples encyclopedia under the gaze of the sages: A systematic review of scholarly research on wikipedia," *Available at SSRN 2021326*, 2012.

[10] D. Wilkinson and B. Huberman, "Assessing the value of cooperation in wikipedia, digital libraries (cs. dl)," *Computers and Society; Physics and Society*, 2007.

[11] A. Kittur and R. E. Kraut, "Harnessing the wisdom of crowds in wikipedia: quality through coordination," in *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, pp. 37–46, ACM, 2008.

[12] B. Stvilia, M. B. Twidale, L. C. Smith, and L. Gasser, "Information quality work organization in wikipedia," *Journal of the American society for information science and technology*, vol. 59, no. 6, pp. 983–1001, 2008.

[13] J. Lehmann, C. Müller-Birn, D. Laniado, M. Lalmas, and A. Kaltenbrunner, "Reader preferences and behavior on wikipedia," in *Proceedings of the 25th ACM conference on Hypertext and social media*, pp. 88–97, ACM, 2014.

[14] A. J. Calder, A. D. Lawrence, J. Keane, S. K. Scott, A. M. Owen, I. Christoffels, and A. W. Young, "Reading the mind from eye gaze," *Neuropsychologia*, vol. 40, no. 8, pp. 1129–1138, 2002.

[15] A. Ajanki, D. Hardoon, S. Kaski, K. Puolamäki, and J. Shawe-Taylor, "Can eyes reveal interest? implicit queries from gaze patterns," *User Modeling and User-Adapted Interaction*, vol. 19, no. 4, pp. 307–339, 2009.

[16] K. Rayner, "Eye movements in reading and information processing.," *Psychological bulletin*, vol. 85, no. 3, p. 618, 1978.

[17] N. J. Belkin, "Some (what) grand challenges for information retrieval," in *ACM SIGIR Forum*, vol. 42, pp. 47–54, ACM, 2008.

[18] S. Berkovsky, T. Baldwin, and I. Zukerman, "Aspect-based personalized text summarization," in *International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*, pp. 267–270, Springer, 2008.

[19] R. Móro *et al.*, "Personalized text summarization based on important terms identification," in *2012 23rd International Workshop on Database and Expert Systems Applications*, pp. 131–135, IEEE, 2012.

[20] X. Wu, J. Li, Y. Zhang, S. Tang, and S.-Y. Neo, "Personalized multimedia web summarizer for tourist," in *Proceedings of the 17th international conference on World Wide Web*, pp. 1025–1026, ACM, 2008.

[21] C. Kumar, P. Pingali, and V. Varma, "Generating personalized summaries using publicly available web documents," in *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology-Volume 03*, pp. 103–106, IEEE Computer Society, 2008.

[22] A. Olsen, "The tobii i-vt fixation filter," *Tobii Technology*, 2012.

[23] F. W. Cornelissen, E. M. Peters, and J. Palmer, "The eyelink toolbox: eye tracking with matlab and the psychophysics toolbox," *Behavior Research Methods, Instruments, & Computers*, vol. 34, no. 4, pp. 613–617, 2002.

[24] O. Ferhat, F. Vilarino, and F. J. Sanchez, "A cheap portable eye-tracker solution for common setups," 2014.

[25] "Cvc eye tracker." `https://github.com/tiendan/OpenGazer`. Accessed: 2016.

[26] "Netgazer." `http://sourceforge.net/projects/netgazer/`. Accessed: 2016.

[27] S. Xu, H. Jiang, and F. Lau, "User-oriented document summarization through vision-based eye-tracking," in *Proceedings of the 14th international conference on Intelligent user interfaces*, pp. 7–16, ACM, 2009.

[28] F. Chevalier, S. Huot, and J.-D. Fekete, "Wikipediaviz: Conveying article quality for casual wikipedia readers," in *2010 IEEE Pacific Visualization Symposium (PacificVis)*, pp. 49–56, IEEE, 2010.

[29] Q. V. Dang and C.-L. Ignat, "Quality assessment of wikipedia articles without feature engineering," in *Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries*, pp. 27–30, ACM, 2016.

[30] M. E. Falagas, A. Zarkali, D. E. Karageorgopoulos, V. Bardakas, and M. N. Mavros, "The impact of article length on the number of future citations: a bibliometric analysis of general medicine journals," *PLoS One*, vol. 8, no. 2, p. e49476, 2013.

[31] J. E. Blumenstock, "Size matters: word count as a measure of quality on wikipedia," in *Proceedings of the 17th international conference on World Wide Web*, pp. 1095–1096, ACM, 2008.