Slow Search: Improving Information Retrieval Using Human Assistance

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

ABSTRACT

We live in a world where the pace of everything from communication to transportation is getting faster. In recent years a number of "slow movements" have emerged that advocate for reducing speed in exchange for increasing quality, including the slow food movement, slow parenting, slow travel, and even slow science. We propose the concept of "slow search," where search engines use additional time to provide a higher quality search experience than is possible given conventional time constraints. While additional time can be used to identify relevant results within the existing search engine framework, it can also be used to create new search artifacts and enable previously unimaginable user experiences. This talk will focus on how search engines can make use of additional time to employ a resource that is inherently slow: other people. Using crowdsourcing and friendsourcing, it will highlight opportunities for search systems to support new search experiences with high quality result content that takes time to identify.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval – *search process.*

Keywords

Slow search, crowdsourcing, speed, information retrieval.

1. INTRODUCTION TO SLOW SEARCH

Substantial research and engineering effort has been devoted to achieving low latency in large, complex computing systems like search engines. Search engines target speed for good reason. Research suggests that people perceive results that are delivered quickly as higher quality and more engaging than those delivered more slowly [10]. To achieve near-instantaneous speed, search engines make a number of compromises and limit the complexity of the features and models used to identify relevant documents. The resulting fast, word-oriented matching used by search engines ignores the rich semantics of text but is an efficient way to capture some aspects of the similarity between queries and documents.

Although searchers have grown accustomed to rapid responses to their queries, recent advances in our understanding of how people search suggest there are scenarios where a search engine could take significantly longer than a fraction of a second to return relevant content [9]. People often invest minutes, hours, or even days in more complex or exploratory search tasks. A person planning a vacation or researching a medical diagnosis, for example, may be willing to wait for better results or insights. By building slow search user experiences [2], search engines can make use of additional time to identify the most relevant content [4, 6], load balance responses over time, and incorporate slow resources [1, 3].

One slow resource that is particularly interesting to explore is people. Although computers can consider billions of documents much faster than a person can, a person can better understand each document. Crowdsourcing research suggests paid crowd workers can be incorporated into the search process in a way that is personalized [8] and accounts for privacy considerations [5]. This enables underperforming search engine components, such as those for query understanding and result processing, to be done by crowd workers [4]. Initial explorations show that while incorporating crowd workers into existing search pipelines increases robustness, the increase in performance is not significant enough to warrant the extra time and cost involved. Instead, approaches that use crowd workers to drastically change the search experience [1, 3] or that incorporate feedback from highly relevant people (e.g., friends) [3, 7] appear most likely to be successful.

Our hope is that slow search will inspire new and creative research into how search experiences can be enriched with a more nuanced notion of relevant search resources and time constraints.

CIKM '15, October 19–23, 2015, Melbourne, VIC, Australia.
ACM 978-1-4503-3794-6/15/10.
http://dx.doi.org/10.1145/2806416.2806417

2. REFERENCES

[1] Bernstein, M., Teevan, J., Dumais, S.T., Libeling, D. and Horvitz, E. Direct answers for search queries in the long tail. CHI 2012.

[2] Eickhoff, C., Teevan, J., White, R. and Dumais, S.T. Lessons from the journey: A query log analysis of within-session learning. WSDM 2014.

[3] Jeong, J., Morris, M., Teevan, J. and Liebling, D. A crowd-powered socially embedded search engine. ICWSM 2013.

[4] Kim, Y., Collins-Thompson, K. and Teevan, J. Crowdsourcing for robustness in web search. TREC 2013.

[5] Lasecki, W., Teevan, J., and Kamar, E. Information extraction and manipulation threats in crowd-powered systems. CSCW 2014.

[6] Livne, A., Gokuladas, V., Teevan, J., Dumais, S.T. and Adar, E. CiteSight: Supporting contextual citation recommendation using differential search. SIGIR 2014.

[7] Morris, M.R., Teevan, J. and Panovich, K. A comparison of information seeking using search engines and social networks. ICWSM 2010.

[8] Organisciak, P., Teevan, J., Dumais, S.T., Miller, R.C. and Kalai, A.T. A crowd of your own: Crowdsourcing for on-demand personalization. HCOMP 2014.

[9] Teevan, J., Collins-Thompson, K., White, R.W. and Dumais, S.T. Viewpoint: Slow search. CACM, 57(8), 2014.

[10] Teevan, J., Collins-Thompson, K., White, R.W., Dumais, S.T. and Kim, Y. Slow search: Information retrieval without time constraints. HCIR 2013.