

SHORT TERM SCIENTIFIC MISSION (STSM) SCIENTIFIC REPORT

This report is submitted for approval by the STSM applicant to the STSM coordinator

Action number: IC1302

STSM title: Search as Learning: experiments with keyword-based search in large scale online learning

STSM start and end date: 24/09/2019 to 04/10/2019

Grantee name: Leif Azzopardi

PURPOSE OF THE STSM:

DelftX delivers Massively Open Online Courses to hundreds of thousands of students - a key challenge faced is to personalize the experience and ensure that students find relevant answers and information regarding the course material. Each MOOC provides lectures, notes, videos, Q&As, discussions, various data, and other related resources - yet the search functionality is limited (Davis et al, 2016). Thus the aim of this STSM is to implement various keyword based approaches (query suggestion, query auto-completion, faceting, etc) to improve student's access to information within the courses - and help enhance their learning experience. This will be considered in the context of MOOCs and University Site Search.

To this end, we propose to implement a number of innovations regarding keyword based approaches, and undertake A/B testing to determine the influence of such access on the overall search and learning experience (as done in previous work by Hiemstra, Hauff and Azzopardi, 2017). Currently the A/B testing and search infrastructure is being developed at Delft by Hauff and her team based on the work of Buivyas and Azzopardi (2016). This will provide the critical infrastructure to deploy the experiments. During the proposed visit, Azzopardi, Hauff and her team will implement a number of methods typically used for web search and knowledge extraction - and adapt them to the novel context of MOOC search. To provide and enable better suggestion and recommendation through their keyword based searching, we will first develop methods for extracting tasks and sub-tasks which is used to support complex search tasks (Awadallah et al, 2017; Mehrotra & Yilmaz, 2015). Using query logs supplied by DelftX from their MOOCs will, therefore, be conducted and analysis and extract out semantic knowledge relating to the needs of users. This analysis and extraction will provide the underlying suggestions for the query auto-completion methods to be implemented and evaluated for A/B testing. The results from the evaluation will be used to write publication(s) contributing to the replication and validation of lab based methods, as well as novel contributions to keyword based search methods within the context of search and learning.

Buivyas, M., & Azzopardi, L. (2016). Pienapple search: an integrated search interface to support finding, refining and sharing. Proceedings of the Association for Information Science and Technology, 53(1), 1-5.

Hiemstra, D., Hauff, C., & Azzopardi, L. (2017). Exploring the Query Halo Effect in Site Search: Leading People to Longer Queries. To appear in the Proceedings of ACM SIGIR.

Davis, D., Chen, G., van der Zee, T., Hauff, C., & Houben, G. J. (2016). Retrieval Practice and Study Planning in MOOCs: Exploring Classroom-Based Self-Regulated Learning Strategies at Scale. In

European Conference on Technology Enhanced Learning (pp. 57-71). Springer International Publishing.

Hassan Awadallah, A., White, R. W., Pantel, P., Dumais, S. T., & Wang, Y. M. (2014, November). Supporting complex search tasks. In Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management (pp. 829-838).

Mehrotra, R., & Yilmaz, E. (2015, May). Towards hierarchies of search tasks & subtasks. In Proceedings of the 24th International Conference on World Wide Web (pp. 73-74).

DESCRIPTION OF WORK CARRIED OUT DURING THE STSMS

The visit consisted of several parts: the first part of the visit involved a number of offline activities to ensure that data was available. This consisted of the development of infrastructure for facilitating, collecting and logging search in the MOOC context, and implementing and running autocomplete experiments in the university site search context. During the visit, an overview of the contexts and the data were discussed and described to decide on what and how to analyse the data, as well as to plan further experiments given the data. This then led to: development of Jupyter notebooks to perform statistical analysis of query logs in site search, which required developing and testing our evaluation metrics and that the log was being processed appropriately. An analysis of the A/B experiment on autocomplete was then performed (see below). Two subsequent A/B tests were then designed and the code written to deploy the experiments, and then the experiments were kicked off. These experiments will continue to explore factors that affect the length of key word based queries.

In the MOOC Context space, only limited data was collected, which provided interesting insights into how people search (or not) when learning online. This then led to designing a number of experiments to analyse how people learn about a particular topic via searching and to measure the change in knowledge, and the learner's efficiency.

DESCRIPTION OF THE MAIN RESULTS OBTAINED

In the autocomplete A/B experiment logs, over a 3 week period, 1332 queries were issued in the autocomplete condition, while 1334 queries were issued in the non-autocomplete condition. The mean term length was 2.44 and 2.12 and the mean reciprocal block rank was 0.67 and 0.69 respectively. The analysis found that in the autocomplete condition, queries were significantly longer when autocomplete was enabled. This was observed with respect to both term length and character length. However, contrary to previous findings, users in the autocomplete condition did not submit queries which were significantly better. Two subsequent A/B experiments were set up on: (1) query box length, with (a) short query box and (b) a long query box, and (2) device type, with (a) mobile and (b) desktop. These experiments in conjunction with our previous work on glowbox will contribute to a journal article on factors that affect the length of site search queries.

FUTURE COLLABORATIONS (if applicable)

In the context of site search, as mentioned above, two additional A/B tests have been designed, and are currently running. These experiments will complement the Autocomplete experiments (and the prior experiments conducted) and contribute to a journal article submission on "Experiments on Key Word Based Queries in SiteSearch" with Claudia Hauff (Delft University) and Djoerd Hiemstra (University of Twente).

In the context of the MOOC, a learning based search experiment has been designed to evaluate how people use search to learn in a MOOC and non-MOOC context – this will enable us to examine the efficiency and effectiveness of the learner – and the learning environment. We hypothesize that active search to learn is likely to decrease the efficiency of learners, but will lead to greater effectiveness. These experiments will also provide an implicit measure of the quality of the learning resource/method.