

Title: Mining and Learning from Context in Web Search

Speaker: Paul Bennett

When: Friday, February 15, 1:30pm

Where: GSB 3.130

Abstract:

User and behavioral modeling can play a critical role in a variety of online services such as web search, advertising, e-commerce, and news recommendation. With regard to web search, while information retrieval has made significant progress in returning relevant results for a single query, much search activity is conducted within a richer context of a current task focus, recent search activities as well as longer-term preferences. For example, our ability to accurately interpret the current query can be informed by knowledge of the web pages a searcher was viewing when initiating the search or recent actions of the searcher such as queries issued, results clicked, and pages viewed. We develop a framework that enables representation of a broad variety of context including the searcher's long-term interests, recent activity, current focus, and other user characteristics. We then demonstrate how that can be used to improve the quality of search results. We describe recent progress on three key challenges in this domain: enriching information retrieval via automatically generated metadata; mining contextual signals from large scale logs; and using contextual representations in learning to improve both standard ad hoc and personalized retrieval.

This talk will present joint work with Fedor Borisyuk, Jinyoung Kim, Nam Nguyen, Lidan Wang, Filip Radlinski, Ryen White, Kevyn Collins-Thompson, Wei Chu, Susan Dumais, Peter Bailey, Emine Yilmaz, Xiaoyuan Cui, David Sontag, Sebastian de la Chica, and Bodo von Billerbeck.

About the speaker:

Paul Bennett is a Researcher in the Context, Learning & User Experience for Search (CLUES) group at Microsoft Research where he focuses on the development, improvement, and analysis of machine learning and data mining methods as components of real-world, large-scale adaptive systems. His research has advanced techniques for ensemble methods and the combination of information sources, calibration, consensus methods for noisy supervision labels, active learning and evaluation, supervised classification (with an emphasis on hierarchical classification) and ranking with applications to information retrieval, crowdsourcing, behavioral modeling and analysis, and personalization. He completed his dissertation on combining text classifiers using reliability indicators in 2006 at Carnegie Mellon where he was advised by Profs. Jaime Carbonell and John Lafferty.