SIGIR Workshop on Semantic Matching in Information Retrieval

Title: Leveraging Cold-Start Knowledge Base Population for Information Access

Speaker: Douglas W. Oard, University of Maryland (USA)

Abstract:
We might identify three canonical approaches to leveraging some degree of semantic analysis to support information access, one that starts with queries as they exist today, one that starts with existing knowledge representations, and one that starts with semantic analysis of documents. These correspond to what have been called semantic matching, semantic search, and machine reading. In this talk, I will focus on Cold Start Knowledge Base Population (KBP), an approach that we might think of as spanning semantic search and machine reading. In the Cold Start track at the Text Analysis Conference (TAC), the goal of each participating team has been to build systems to automatically "read" some coherent (but unanticipated) text collection and then to automatically generate a knowledge base with a pre-specified schema. In this talk, I will describe the evolution of the TAC Cold Start track over the three years since 2012, with a particular focus on what the levels of accuracy and coverage that have been achieved imply for the near-term usefulness of this approach as a basis for supporting information access. I will augment the presentation with a few examples of the application of a broader range of KBP techniques to additional collections. I will then finish the talk with a few speculative remarks on how these techniques might ultimately be used as one part of a complete system for supporting interactive information access.

About the Speaker:

Douglas Oard is a Professor at the University of Maryland, College Park, with joint appointments in the College of Information Studies and the Institute for Advanced Computer Studies. Dr. Oard earned his Ph.D. in Electrical Engineering from the University of Maryland. His research interests center around the use of emerging technologies to support information seeking by end users. Additional information is available at http://terpconnect.umd.edu/~oard/.