NATIONAL UNIVERSITY OF SINGAPORE

School of Computing

C S S E M I N A R

Title: **Knowledge Base Population from Conversational Media**

Speaker: Professor Douglas W. Oard
University of Maryland and HLTCOE
USA

Date/Time: 23 May 2016, Monday, 03:00 PM to 04:15 PM

Venue: Executive Classroom, COM2-04-02

Chaired by: Dr Chua Tat Seng, KITHCT Chair Professor, School of Computing (chuats@comp.nus.edu.sg)

Abstract:

The usual setup for the so-called "cold start" Knowledge Base Population task is that we receive some collection of documents and we are asked to produce a "knowledge graph" in which entities, attributes of those entities and relationships between those entities are identified based on information attested in that collection. I'll begin this talk by describing our participation in the cold start KBP task of the Text Analysis Conference. I'll then spend the majority of the talk exploring what we should do differently when what we received are not isolated documents, but rather some form of conversational media. I'll focus principally on email as one computationally tractable conversational medium, and on person and organization entities and their attributes, presenting results for both the Enron and the Avocado email collections. I'll then conclude the talk with a preview of some work that we are doing now with recorded phone calls made by Enron energy traders, and with a few thoughts on next steps. This is joint work with Mark Dredze, Tim Finin, Ning Gao, Dawn Lawrie, Jim Mayfield, and Paul McNamee.

Biodata:

Douglas Oard is a Professor at the University of Maryland, College Park, with joint appointments in the College of Information Studies (Maryland's iSchool) and the University of Maryland Institute for Advanced Computer Studies (UMIACS). He is also affiliated with at the Human Language Technology Center of Excellence (HLTCOE) at The Johns Hopkins University. 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/.