Keynote: Apache clinical Text Analysis and Knowledge Extraction System (cTAKES)

Abstract

The presentation will focus on methods and software development behind the cTAKES platform. An overview of the modules will set the stage, followed by more in-depth discussion of some of the methods and evaluations of select modules. The second part of the presentation will shift to software development topics such as optimization and distributed computing including UIMA integration, UIMA-AS, as well as our plans for UIMA-DUCC integration. A live demo of cTAKES will conclude the talk.

About the speakers

Pei Chen is a Vice President of Apache Software Foundation, leading the top-level cTAKES project[1]. He is also a lead application development specialist at the Informatics Program at Boston Childrens Hospital/Harvard Medical School. Mr. Chen's interests lie in building practical applications using machine learning techniques. He has a passion for the end-user experience and has a background Computer Science/Economics. Mr. Chen is a firm believer in the open source community contributing to cTAKES as well as other Apache Software Foundation projects.

Guergana Savova, Ph.D. is member of the faculty at Harvard Medical School and Childrens Hospital Boston. Her research interest is in natural language processing (NLP), especially as applied to the text generated by physicians (the clinical narrative) focusing on higher level semantic and discourse processing which includes topics such as named entity recognition, event recognition, relation detection, and classification including co-reference and temporal relations. The methods are mostly machine learning spanning supervised, lightly supervised, and completely unsupervised. Her interest is also in the application of the NLP methodologies to biomedical use cases. Dr. Savova has been leading the development and is the principal architect of cTAKES. She holds a Masters of Science in Computer Science and a PhD in Linguistics with a minor in Cognitive Science from University of Minnesota.

[1] `http://ctakes.apache.org/`