A Turn of a Phrase Folds into Fields

[Friday], Sept 23  |  12:00  |  West Hall 105

Speaker: Mr. Guy Divita
Research Associate
University of Utah’s Medical School

Abstract: Natural Language Processing (NLP) use cases involve the identification of one or more features from a corpus of clinical documents. The found features get aggregated and analyzed along with structured data. Processing clinical records pose unique challenges. Common feature extraction techniques such as sentence identification, dictionary lookup and assertion assignment get employed as part of the process. Some example projects from VA data are reviewed to highlight aspects of challenges and novel techniques. These projects include homelessness prediction, symptom extraction, colonoscopy quality metrics.

Biography: Mr. Divita is associated with the University of Utah’s Medical School and the Salt Lake City VA Health Sciences R&D division as an NLP evangelist by day, and research associate by title. Known to a few as the wise guy, he is known for naming his tools after favorite characters, including “Sophia: an expedient information extraction tool”, “Nora, a vocabulary building tool”, his professional interests include document sectionizing, template identification, de-identification, vocabulary curation, and building NLP tools for clinician’s use. Prior to his association with the VA and the University of Utah, Mr. Divita played at the National Library of Medicine’s Lister Hill Center for Biomedical Communications for 16 years. It’s still not believable that they paid him for the privilege! He’s aided the creation of well-known tools and resources as the UMLS, the SPECIALIST Lexicon, MetaMap, MMTx, LVG, and long forgotten tools like gspell. Mr. Divita holds a Bachelor’s degree in Statistics from The George Washington University, and a Master’s degree in Computer Science from George Mason University, another fine Virginia institution.