

**Pavan Kapanipathi**  
Greater New York City Area  
USA

+1 713 435 9982 (phone)  
pavan046[at]gmail.com  
<http://pavan046.github.io/>

## Research Interests

Semantics, Natural Language Understanding, Machine Learning, Knowledge Graphs, Artificial Intelligence, Recommendation Systems, User Modeling.

## Education

**Phd, Computer Science**, Wright State University, (2012 - 2016)  
**MS, Computer Science**, Wright State University, (2009 - 2012)  
**BE, Computer Science**, Visvesvaraya Technological University, (2003 - 2007)

## Experience

**IBM TJ Watson Research Center**, Research Staff Member, (June 16 - Present)  
**Samsung Research America**, Research Intern, (May 14 - December 14)  
**IBM TJ Watson Research Center**, Research Intern, (May 13 - August 13)  
**Digital Enterprise Research Institute**, Research Intern, (April 11 - August 11)  
**Accenture**, Programmer, (July 07 - March 09)  
**Robert Bosch**, Intern, (February 07 - June 07)

## Research Projects

**Watson Conversation:** The focus of the project was on innovating and developing new methodologies in the areas of ML and NLP (NLU) to improve performance of cognitive applications, primarily Watson conversation. The work involves unsupervised/semi-supervised learning of short-textual data to recommend actions that facilitate better performance of cognitive applications.

**Domain Specific Knowledge Graphs:** In this project we address the challenge of extracting domain-specific knowledge graphs from large knowledge graphs such as Wikipedia. For instance a movie recommendation system would not require all the information on Wikipedia but may require knowledge related to only movies. We develop techniques to extract such domain specific knowledge.

**User-specific information extraction from social data:** User modeling comprises of acquiring necessary information of users for personalization and recommendation systems. I have focused on extracting *hierarchical interests of users*, *home location of a user*, and *users' activities of interest* from users' tweets. These techniques are unsupervised and harness knowledge graphs for semantic processing of Twitter data. They are compared against information retrieval, machine learning, and probabilistic modeling approaches.

## Selected Publications, Patents, and Grant Proposals

S.Lalithsena, **P. Kapanipathi**, A. Sheth. Harnessing relationships for domain-specific subgraph extraction: A recommendation use case. IEEE BigData2016.  
**P. Kapanipathi**, R.Krishnamurthy, A. Sheth, K. Thirunarayan. Knowledge Enabled Approach to Predict the Location of Twitter Users. ESWC2015.  
**P. Kapanipathi**, P. Jain, C. Venkataramani, A. Sheth. User Interests Identification on Twitter Using a Hierarchical Knowledge Base. ESWC2014  
**P. Kapanipathi**, J. Anaya, A. Sheth, B. Slatkin, A. Passant. Privacy-Aware and Scalable Content Dissemination in Distributed Social Network. ISWC2011.  
P. Mendes, A. Passant, **P. Kapanipathi**, A. Sheth. Linked Open Social Signals. WI2010. **(16.6%)**  
**Patent:** Joakim Soderberg, Edwin A. Heredia, **Pavan Kapanipathi**, Glenn Algie, Alan Messer. Semantic Enrichment of Trajectory Data.  
**Grant Proposal:** One of the contributors of award winning proposals to NSF and NIH during PhD. NIH: \$1.5M, 2014, NSF: \$2.0M, 2015, and NSF: \$200K, 2015

## Skills, Achievements, and Service

**Programming:** Python, Java, SQL, PIG, SPARQL.  
**Tools, & Softwares:** Database (MySQL), Graphs (Virtuoso, Jena, Gephi), Machine Learning (scikit-learn, gensim), Natural Language Processing (nltk, Stanford CoreNLP), Version Management (SVN, GIT).  
**Open Source Contributions:** Twarql, SMOB.  
**Awards:** Triplication Challenge 2010, ISWC2011 Travel Grant.  
**Invited Talks:** EMC invited talk, 2015. Frontiers of Cloud Computing and Big Data Workshop, 2014 (IBM Research).  
**PC Member:** WWW2017, ISWC2017, ESWC(2016, 15, 14), WI(2017, 16), HT2015, DERIVE2012.