

# Pavan Kapanipathi

1449 Newgate Ct Apt C Fairborn Ohio 45324  
pavan@knoesis.org • +1 (713) 435-9982 • <http://knoesis.org/researchers/pavan>

**INTERESTS** User Modeling; Personalization; Recommendation; Knowledge Graphs; Linked Data; Semantic Web; Text Mining; Information Retrieval; Applied Machine Learning; Social Data Analysis

**EDUCATION** **Kno.e.sis Center, Wright State University**, Dayton, Ohio, USA  
Doctor of Philosophy (Ph.D.) in Computer Science Jan 2012 – Apr 2016 (Expected)  
• Thesis: Semantic personalized filtering on social media  
• Adviser: Dr Amit P Sheth  
Master of Science (M.S.) in Computer Science Aug 2009 – Aug 2012

**EXPERIENCE** **Samsung Research America**, San Jose, California, USA  
Research Intern, Advanced Technology Labs May 2014 – Dec 2014  
Developed a semantic enrichment engine for trajectory data (patent pending); Predicted location specific activities of interest based on tweets.  
**IBM T J Watson Research Center**, Yorktown heights, New York, USA  
Research Intern, Infosphere Streams Group May 2013 – Aug 2013  
Modeled Twitter users' preferences as a hierarchy, inferred from a knowledge graph (Hierarchical Interest Graph); Implemented an adaptation of spreading activation algorithm to score the interests in the hierarchical interest graph. [\[paper\]](#)  
**Digital Enterprise Research Institute**, Galway, Ireland  
Research Intern, Social Software Unit Apr 2011 – Aug 2011  
Collaborated with Google for extending their Pubsubhubbub protocol to a privacy aware Semantic Hub; Semantic Hub is used by SMOB (an open source, distributed, semantic microblogging framework) for content dissemination. [\[paper\]](#)  
**Accenture**, Bangalore, India  
Software Programmer Jul 2007 – Mar 2009  
Undergone training in C++; Clients: Drugstore and SFR; Technologies used: VC++, C#, & Java.  
**Bosch**, Bangalore, India  
Research Intern Feb 2007 – Jun 2007  
Developed a Test-program compiler in C# used for calibration process of fuel injection pumps.

**PROJECTS** **User Modeling and Recommendation:** Understanding users by determining their attributes and interests from the content they generate on the social web. (Spreading activation, SVM, LDA)  
• Modeled Hierarchical Interest Graphs from users' social data; Prototyped a recommendation system that harnesses the hierarchical interest graph for recommending tweets to users.  
• Predicted home location of users from their social data with a novel, knowledge-base driven, unsupervised methodology. [\[paper\]](#)  
• Built a probabilistic model for determining user activities of interest from social data.  
**Twarql:** Information overload on the social web is increasing, making it harder to track relevant information. Twarql is a semantic filtering system that allows flexible, complex querying of social data in near real-time. (Java, SPARQL, RDF, Storm, Hadoop) [\[paper\]](#)  
**Twitris:** A Semantic Social Web platform with real-time monitoring and multi-faceted analysis of social signals to provide insights and event analysis (Storm, MySQL, PHP, Servlet)  
• Integrated Twarql pipeline into Twitris to perform semantic analysis of social data.

**SKILLS** **Programming:** Java, Python, R, SQL, PIG, SPARQL. **Tools & Softwares:** Database (MySQL), Graphs (Virtuoso, Jena, Gephi), Machine Learning (Weka, scikit-learn), Information Retrieval (Lucene), Natural Language Processing (Stanford CoreNLP), Version Management (SVN, GIT), Big Data (Hadoop, Storm), open to learning new skills. **Open source:** Twarql, SMOB. **Proposals:** Contributed to 3 funded proposals from NSF and NIH.

**HONORS & AWARDS** Invited to present my research at venues that include Big Data tutorial [\[ppt\]](#), EMC CTO Office [\[ppt\]](#), and Frontiers of cloud computing workshop at IBM TJ Watson Research Center [\[ppt\]](#); Winner of the Triplication Challenge for our submission on Twarql [\[paper\]](#); PC Member and External reviewer for more than 15+ conference and workshops.

**PUBLICATIONS** Published 15+ papers, cited 170 times, venues: ISWC, ESWC, WWW, WI, and Websci.  
[\[Google Scholar\]](#) [\[LinkedIn\]](#) [\[Twitter\]](#)