

Akshay Java

<http://www.cs.umbc.edu/~aks1>

<http://socialmedia.typepad.com>

aks1@cs.umbc.edu

410-271-5695

RESEARCH INTERESTS:

Social Media, Information Retrieval, Natural Language Processing, Semantic Web, Artificial Intelligence

EDUCATIONAL BACKGROUND:

PhD: Computer Science	Sept 2008
University of Maryland Baltimore County	
M.S.: Computer Science	G.P.A.: 3.86/4.0
University of Maryland, Baltimore County	
B.E.: Computer Engineering	First class
Thadomal Shahani Engineering College, Mumbai, India.	
	May 2001

EMPLOYMENT:

NEC Laboratories America Inc. Research Assistant (May 2007-Nov 2007)

- Analysis of Microblogging phenomenon to study user intentions and communities.
- Influence models, Ranking Relationships in Social Media.

eBiquity Research Group-UMBC Research Assistant (Jan 2005-present)

- **Ph.D. research topic “Tracking Influence and Opinions in Social Media”**
- Community detection and mining structure and content in social media
- TREC Blog track: Opinion extraction and spam blog detection.
- Feeds That Matter: Clustering blog feed subscriptions for ranking and categorizing blogs.
- SemNews: Agent understandable news by converting natural language text in the news to RDF.
- SWOOGLE, a Semantic Web Search engine.

University of Sheffield, U.K Research Associate (Mar 2004-Oct 2004)

- Discovering social networks using linguistic analysis of unstructured and semi-structured web data.
- Implemented the first web crawler and search plugins for GATE NLP tool.

NEC Laboratories America Inc. Research Assistant (Jun 2003- Dec 2003)

- Meta-search and data mining for improved search result ranking.
- Developed algorithms for re-ranking results based on topic specificity, resulting in a patent application.
- Automatically predict key concepts from web search results.

CADIP Research Group-UMBC Research Assistant (Sep 2002-Dec 2003)

- Distributed Information retrieval, web crawling and distributed PageRank implementation.
- Collaborative Agent Based Information Retrieval and participation at TREC 2004.

Palo Alto Research Center (PARC) Summer Intern (May 2002-Aug 2002)

- Re-Engineered the WebCrawler for a website usability analysis toolkit
- Enhanced the performance of the parsing tools by JavaScript emulation enabled Link extraction.

UMBC- Physics department Research Assistant (Aug 2001-Apr 2002)

- Worked on an interdisciplinary project involving data mining applications in Astronomy.
 - Rule discovery and trend prediction methods for time series data analysis of Astronomy datasets.
-

SOFTWARE DEVELOPMENT:

JAVA, J2EE, JavaScript, AJAX, XML, RDF, OWL, Prolog, LISP, PERL, Shell Scripting, Lucene, RSS/ATOM, Google Maps and Search API, LingPipe, GATE, Hadoop, MATLAB, UCINet, Weka, Pajek

PATENT :

Methods and Systems for Improving Document Relevance Ranking by Discovering General and Specific Documents. **Akshay Java**, Brian Klock, Eric Glover, Vishal Shanbhag and Robert Krovetz. Provisional Patent Application Serial No. 60/523,851. [Final Patent Pending]

PUBLICATIONS :

1. *Simultaneously Clustering Graphs and Folksonomies*, Akshay Java, Anupam Joshi, Tim Finin, In preparation, May 2008.
2. *Approximating the Community Structure of the Long Tail*, Akshay Java, Anupam Joshi, Tim Finin, Submitted to ICWSM, 2008.
3. *SecondSpace: A Generative Model for the Blogosphere*, Amit Karandikar, Akshay Java, Tim Finin, Anupam Joshi, ICWSM 2008.
4. *Web2.0 Mining: Analyzing Social Media*, Anupam Joshi, Tim Finin, Akshay Java, Anubhav Kale, Pranam Kolari, Next Generation Data Mining and Cyber-Enabled Discovery for Innovation, 2007.
5. *Extracting Information from Social Media and Online Communities*, Tim Finin, Anupam Joshi, Pranam Kolari, Akshay Java, Anubhav Kale, Amit Karandikar, AI Magazine special issue on networks, Fall 2007
6. *Why We Twitter: Understanding Microblogging Usage and Communities*, Akshay Java, Xiaodan Song, Tim Finin, Belle Tseng, WebKDD Workshop on SNA, KDD 2007
7. *Feeds That Matter: A Study of Bloglines Subscriptions.*, Akshay Java, Pranam Kolari, Tim Finin, Tim Oates, ICWSM 2007
8. *Modeling Trust and Influence in the Blogosphere Using Link Polarity*, Anubhav Kale, Amit Karandikar, Pranam Kolari, Akshay Java, Tim Finin, Anupam Joshi, ICWSM 2007
9. *The UMBC/JHU BlogVox System*, Akshay Java, Pranam Kolari, Tim Finin, James Mayfield, Anupam Joshi, Justin Martineau, Proceedings of the 15th Text Retrieval Conference, Oct 2006.
10. *BlogVox: Separating Blog Wheat from Blog Chaff*, Akshay Java, Pranam Kolari, Tim Finin, James Mayfield, Anupam Joshi, Justin Martineau, Workshop on Analytics for Noisy and Unstructured Text, IJCAI 2007
11. *Blog Track Open Task: Spam Blog Classification*, Pranam Kolari, Akshay Java, Tim Finin, James Mayfield, Anupam Joshi, Justin Martineau, TREC 2006 Blog Track Notebook, Oct 2006.
12. *Characterizing the Splogosphere*, Pranam Kolari, Akshay Java, Tim Finin, WWW 2006, 3rd Annual Workshop on Weblogging Ecosystem: Aggregation, Analysis and Dynamics, May 2006
13. *Detecting Spam Blogs: A Machine Learning Approach*. Pranam Kolari, Akshay Java, Tim Finin, Tim Oates, Anupam Joshi, AAAI 2006.
14. *Modeling the Spread of Influence on the Blogosphere*, Akshay Java, Pranam Kolari, Tim Finin, Tim Oates, UMBC Tech Report TR-CS-06-03
15. *SemNews: A Semantic News Framework*. Akshay Java, Tim Finin, Sergei Nirenburg. Student Abstract and Poster Program and Intelligent Systems Demonstrations, AAAI 2006.
16. *Text Understanding agents and the Semantic Web*. Akshay Java, Tim Finin and Sergei Nirenburg. Proceedings of the 39th Hawaii International Conference on System Sciences, January 2006.
17. *Integrating Language Understanding agents on the Semantic Web*. Akshay Java, Tim Finin and Sergei Nirenburg. AAAI Fall 2005 Symposium.
18. *Swoogle: Searching for knowledge on the Semantic Web*. Tim Finin, Li Ding, Rong Pan, Anupam Joshi, Pranam Kolari, Akshay Java, Yun Peng. AAAI 05 (intelligent systems demo)
19. *UMBC at TREC 12* Srikanth Kallurkar, Yongmei Shi, R. Scott Cost, Charles Nicholas, Akshay Java, et al. in Proceedings of the Twelfth Text Retrieval Conference.
20. *Predictive Mining of Time Series Data*. Akshay Java and Eric Perlman in Proceedings of the 200th AAS meet and ADASS XII.

INVENTION DISCLOSURES

Approximate Community Detection, UMBC Office of Technology Development
PolVox: Mining the Political Blogosphere, UMBC Office of Technology Development

PRESS MIT Tech review, Wired Magazine, Network World magazine, BBC Radio podcasts (upcoming)

INVITED TALKS:

- *Social Media Analytics*. IBM Research India Jan 2007, Greylock Capital, Aug 2007, Social Web Technologies Class UMBC 2008.
- *Spam in Social Media*, Tutorial ICWSM 2007, SearchMe Inc Jul 2007.
- *Text Understanding Agents and the Semantic Web*. Systems and Information Sciences Group, Johns Hopkins Applied Physics Laboratory. Dec 2005.