

# Eugene Agichtein

Mathematics and Computer Science Department  
Emory University  
eugene@mathcs.emory.edu

1 October 2009

400 Dowman Drive, Suite W401  
Atlanta, GA 30322

Web: <http://www.mathcs.emory.edu/~eugene/>  
Telephone: 1-404-727-7962

---

## RESEARCH INTERESTS

**Information retrieval, natural language processing and text mining, and medical informatics.** My research has been focused on two complementary areas of information retrieval: *natural language processing* and *text mining* over the web, social media, and biomedical literature, and *modeling user behavior and interactions* for search tasks, with applications to web search, question answering, and text classification.

## EDUCATION

- Ph.D.** Computer Science, **Columbia University**, New York, NY May 2005  
Advisor: Professor Luis Gravano  
Dissertation: *Extracting Relations from Large Text Collections*
- M.S.** Computer Science, **Columbia University**, New York, NY 2000
- B.S.** Engineering (Interdisciplinary), **The Cooper Union**, New York, NY 1998

## HONORS AND AWARDS

- HP Labs Open Innovation Award** 2009  
For the proposal “Robust Knowledge Acquisition from Social Media,” Eugene Agichtein and Hongyuan Zha.
- Microsoft New Faculty Fellow Finalist** (one of 10 selected across N. America) 2009
- Microsoft Research “Beyond Search” Award** 2007  
For the proposal “Inferring Commercial Intent: Taxonomy Construction, Intent Classification, and Applications,” Eugene Agichtein and Charles L. Clarke.
- “Best Paper” Award, ACM International Conference on Management of Data (SIGMOD)** 2006  
For the paper “To Search or to Crawl? Towards a Query Optimizer for Text-Centric Tasks,” Panagiotis Ipeirotis, Eugene Agichtein, Pranay Jain, and Luis Gravano.
- “Best Student Paper” Award, International Conference on Data Engineering (ICDE)** 2003  
For the paper “Querying Text Databases for Efficient Information Extraction,” Eugene Agichtein and Luis Gravano.
- Award for Exemplary Service to the Computer Science Department** 2002  
Computer Science Department, Columbia University, for service as the “Research Project Liaison,” responsible for helping undergraduate students become involved in research projects.
- Award for Excellence in Teaching** 2001  
School of Engineering and Applied Science, Columbia University, for performance as the head teaching assistant of the “Database Systems” course.
- Full Tuition Scholarship, The Cooper Union** 1994-1998

**GRANTS AND GIFTS** (not including startup funds and cost sharing)

- Robust Knowledge Acquisition from Social Media** (\$75,000) 2009  
Joint PI with H. Zha (Georgia Tech), HP Labs Open Innovation Award
- Yahoo! Research Faculty Grant**, Search and Advertising group, unrestricted gift. (\$18,000) 2009
- Automatic Web-Based System for Detection of Mild Cognitive Impairment** (\$25,000) 2009  
Joint PI with E. Buffalo (Neurology), Emory Alzheimer’s Disease Research Center, Pilot Award
- Modeling User Interactions in Social Media**, unrestricted gift. (\$30,000) 2009  
PI, Cisco Research “Improving the Human Network” Open RFP award.
- Enriching Feature Representation for Product Search Ranking** (\$50,000) 2009–2010  
PI. Shopping.com/EBay company.
- Yahoo! Research Faculty Grant**, Search and Advertising group, unrestricted gift. (\$25,000) 2008
- Motorola Applied Research Center**, Software Platforms group, unrestricted gift. (\$20,000) 2008
- Inferring Commercial Intent** (\$82,000) 2007  
PI (Co-PI: Charles L. Clarke, Waterloo). Microsoft Research “Beyond Search” award.
- Adaptive Extraction of Medical Information from Text Data** (\$30,000) 2007–2008  
PI. Emory University Research Committee.
- Inferring Accuracy and Utility of Medical Literature** (\$42,000) 2007  
PI. Emory College Seed Fund
- Knowledge Discovery Fueled by Biomedical Literature** (\$93,000) 2007  
Co-PI (PI: E.V. Garcia, Radiology). Emory Information Technology Consortium
- Linking Emory Health Research with caBIG Initiative** (\$95,000) 2007  
Co-PI (PI: J.J. Lu). Emory Information Technology Consortium

**PROFESSIONAL EXPERIENCE**

- Assistant Professor, Emory University, Mathematics & Computer Science Dept.** 2006–present  
Responsible for establishing and leading a world-class information retrieval group focusing on: modeling user behavior and interactions in web search and online social networks; large-scale text mining and information extraction; and on applications of these techniques to public health and medical informatics. Additional information available at: <http://ir.mathcs.emory.edu/>.
- Visiting Researcher, Yahoo! Research** Summer 2007  
Search and Advertising group.  
Developed methods for identifying high-quality content in social media.
- Postdoctoral Researcher, Microsoft Research** 2004–2006  
Text Mining, Search, and Navigation Group. Manager: Dr. Eric Brill.  
Web search and web question answering over semi-structured web content. Developed data-driven methods for modeling searcher behavior to improve web search quality.
- Research Assistant, Columbia University** 1998–2004  
Database Group. Advisor: Professor Luis Gravano.

Developed the partially supervised, domain-independent *Snowball* information extraction system, which learns to extract structured relations from unstructured text documents. Also developed the *QXtract* system, to improve the efficiency of information extraction.

**Research Intern, Microsoft Research**

Summer 2003

Mentor: Dr. Venkatesh Ganti; Manager: Dr. Surajit Chaudhuri.

Developed a robust unsupervised system for segmenting text strings into structured records by learning the internal structure of target attributes.

**Research Intern, IBM Almaden Research Center**

Summer 2001

Mentors: Dr. Shiv Vaithyanathan and Dr. C.T. Howard Ho.

Explored partially-supervised machine learning methods to extract relations from text.

**Research Intern, NEC Research Institute**

Summer 2000

Mentor: Dr. Steve Lawrence.

Developed *Tritus*, a system for searching the web for answers to natural language questions.

**Assistant Research Scientist, New York University (part-time)**

1997–1998

Mentor: Professor Ralph Grishman.

Developed utilities to support information extraction efforts within the *Proteus* project.

**TEACHING EXPERIENCE****Emory University**

|  |             |
|--|-------------|
| “CS170: Introduction to Computer Science I,” 30 students                 | Fall 2009   |
| “CS190: The Web” (undergraduate, new course), 11 students                | Spring 2009 |
| “CS572: Information Retrieval” (graduate), 12 students                   | Fall 2008   |
| “CS171: Introduction to Computer Science II,” 7 students                 | Fall 2008   |
| “CS171: Introduction to Computer Science II,” 9 students                 | Spring 2008 |
| “CS571: Natural Language Processing” (graduate, new course), 11 students | Fall 2007   |
| “CS171: Introduction to Computer Science II,” 14 students                | Spring 2007 |
| “CS572: Information Retrieval” (graduate, new course), 4 students        | Fall 2008   |

**Student Advising**

**Current Ph.D. students:** Qi Guo (2007-), Abulimiti Aji (2008-), Qiaoling Liu (2009-), Yu Wang (2009-)

**Current M.S. and other graduate students:** Dmitry Lagun 2009-, Alvin C. Grissom 2007-, Julia Kiseleva 2009-

**Directed study/rotation project supervision:** Thomas Przybylinski (Spring 2009), Ryan P. Kelly (2008-2009), Samrat Jeyaparakash (2009)

**Graduated M.S. students:** Yandong Liu (2009)

**Graduated Honors advisees:** Walter Askew (2009), Matthew Hausknecht (2009)

**Other Educational Activities at Emory**

**Department Graduate Committee** Fall 2007-present

**CS Graduate Admissions Committee** Fall 2006-present

**Ph.D. thesis committee:** Dawid Kurzyniec, Emory University Math & CS Department Feb. 2007  
Saurav Sahay, Georgia Tech, College of Computing, Spring 2009

**Ph.D. Qualifier examiner:** James Gardner, 2009

**Columbia University (as graduate student)**

Head Teaching Assistant, “Database Systems” course, 133 students Fall 2001

Supervisor, Independent Research Projects, 5 students 2000–2004

Teaching Assistant, “Advanced Database Systems” course, 42 students. Spring 1999

**PROFESSIONAL SERVICE****Conference and Workshop Organization**

- 2010 **Area Chair (IR)**: NAACL Conference on Human Language Technologies (NAACL-HLT)  
**Senior PC Member**: AAAI Conference on Weblogs and Social Media (ICWSM)  
**Co-Chair**: Third Annual Workshop on Searching Social Media (SSM 2009) at WSDM 2010
- 2009 **Area Chair (Web IR)**: ACM Conference on Research and Development in Information Retrieval (SIGIR)  
**Area Vice-Chair (IR)**, ACM Conference on Information and Knowledge Management (CIKM)  
**Co-Chair**: SIGIR 2009 Workshop on Searching Social Media (SSM 2009)
- 2008 **Tutorials Co-Chair (IR)**: Annual Meeting of the Association for Computational Linguistics (ACL)  
**Co-Chair**: CIKM 2008 Workshop on Searching Social Media (SSM 2008)  
**Co-Chair**: SIGIR 2008 Workshop on Information Retrieval in Advertising (IRA 2008)  
**Senior PC Member**: ACM Conference on Research and Development in Information Retrieval (SIGIR)
- 2007 **Senior PC Member**: ACM Conference on Research and Development in Information Retrieval (SIGIR)

**Conference Technical Program Committee Member**

- 2010 ACM International Conference on Web Search and Data Mining (WSDM)  
International World Wide Web Conference (WWW)  
AAAI Conference on Artificial Intelligence (AAAI), AI&The Web Track  
ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)
- 2009 NAACL Conference on Human Language Technologies (NAACL-HLT)  
ACM International Conference on Web Search and Data Mining (WSDM)  
AAAI International Conference on Web Logs and Social Media (ICWSM)  
European Conference on Information Retrieval (ECIR)  
Conference on Intelligent User Interfaces (IUI)  
The Fifth International Conference on Knowledge Capture (K-CAP)
- 2008 ACM International Conference on Web Search and Data Mining (WSDM)  
European Conference on Information Retrieval (ECIR)  
AAAI Conference on Artificial Intelligence, Web track (AAAI)  
Conference on Empirical Methods in Natural Language Processing (EMNLP)  
ACM International Conference on Knowledge Discovery and Data Mining (KDD)  
Annual Meeting of the Association for Computational Linguistics (ACL)  
IEEE International Conference on Data Engineering (ICDE)  
International Joint Conference on Natural Language Processing (IJCNLP)
- 2007 NAACL Conference on Human Language Technologies (NAACL-HLT)  
Conference on Empirical Methods in Natural Language Processing (EMNLP)  
The Fourth International Conference on Knowledge Capture (K-CAP)  
Pacific Symposium on Biocomputing (PSB)
- 2006 ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD)  
ACM International Conference on Research and Development in Information Retrieval (SIGIR)  
Joint COLING/ACL Conference on Computational Linguistics (COLING/ACL)  
European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD)  
Intelligent Systems in Molecular Biology Conference (ISMB)
- 2005 IEEE International Conference on Data Mining (ICDM)  
Intelligent Systems in Molecular Biology Conference (ISMB)  
Third International Conference on Knowledge Capture (K-CAP)
- 2004 Intelligent Systems in Molecular Biology Conference (ISMB)

**Workshops Program Committee Member** (since 2006)

WSDM 2009 Workshop on Mining Click Data (WSCD09), ICDE 2008 Workshop on Ranking in Databases (DBRank 2008), SIGIR 2007 Workshop on Learning to Rank for Information Retrieval, WWW 2007 Workshop on Query Log Analysis: Social and Technological Challenges (QueryLogs 2007), ICDE 2007 Workshop on Text Data Mining and Management (TDMM 2007), IJCAI-2007 Workshop on Analytics for Noisy Unstructured Text Data (AND 2007), Bar Ilan Symposium on Foundations of Artificial Intelligence (BISFAI 2007), CIKM 2006 International Workshop on Health Information and Knowledge Management (HIKM 2006).

**Journal and Conference Referee**

Proceedings of the National Academy of Science (PNAS), Foundations and Trends in Information Retrieval, ACM Transactions on Information Systems (ACM TOIS), ACM Transactions on Database Systems (ACM TODS), International Journal on Very Large Data Bases (VLDB Journal), IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE), IEEE Internet Computing (IEEE IC), IEEE International Conference on Data Engineering (IEEE ICDE), ACM International Conference on Management of Data (ACM SIGMOD), International Conference on Very Large Data Bases (VLDB).

**INVITED TALKS AND KEYNOTES** (Since 2006)**2009 Modeling User Interactions In Social Media:**

Keynote at the IJCAI 2009 Workshop on Wikipedia and AI (WikiAI'09)  
 Johnson & Johnson "Blue Ribbon" Lecture, Rutgers University  
 University of Pennsylvania, Computational Linguistics Lunch, Philadelphia, PA  
 Los Alamos National Laboratory, Center for Nonlinear Studies, Los Alamos, NM  
 Georgia Institute of Technology, CSE Seminar, Atlanta, GA

**Searching Collaboratively Generated Content:**

Enterprise Search Summit (plenary session/keynote), New York, NY

**2008 Modeling User Interactions In Social Media:**

Carnegie Mellon University/Google joint seminar, Pittsburgh, PA  
 Georgia Institute of Technology, School of Computational Science and Engineering, Atlanta, GA  
 Shopping.com, Burbank, CA  
 Emory Libraries Symposium on Scientific Collaboration, Atlanta, GA  
 Yahoo! Research, New York, NY

Centers for Disease Control (CDC), Atlanta, GA

SIAM Data Mining Conference (SDM 2008), Atlanta, GA

**2007 Patterns in Search: Mining Web Search User Behavior:**

University of Waterloo  
 Georgia Institute of Technology, GVVU Center brownbag  
 Yahoo Research, Santa Clara

**TUTORIALS**

2009 E. Agichtein, Modeling Searcher Behavior, Russian Summer School on Information Retrieval (RuSSIR'09)

2007 E. Agichtein, Towards Web-Scale Information Extraction, ACM SIGKDD webcast (invited)

2006 E. Agichtein and S. Sarawagi, Scalable Information Extraction and Integration  
 at the ACM International Conference on Knowledge Discovery and Data Mining (KDD)

**PATENTS**

P1. C. Burges and E. Agichtein, Training a learning system with arbitrary cost functions, U.S. Patent 7,472,096 (filed 2005, issued 2008).

P2. E. Agichtein and S. Lawrence, Method for retrieving answers from an information retrieval system, U.S. Patent 7,269,545 (filed 2001, issued 2007).

**INVITED JOURNAL PAPERS**

I1. 1. E. Agichtein, E. Gabrilovich, and H. Zha, "The Social Future of Web Search: Modeling, Exploiting, and Searching Collaboratively Generated Content," in the IEEE Data Engineering Bulletin issue on "New Directions in Web Search," June 2009

I2. E. Agichtein, "Web Information Extraction and User Information Needs: Towards Closing the Gap, in the IEEE Data Engineering Bulletin issue on "Web Scale Data, Systems, and Semantics," Dec. 2006

- I3. E. Agichtein, Scaling Information Extraction to Large Document Collections, in the IEEE Data Engineering Bulletin Special Issue on “Searching and Mining Digital Libraries,” Dec. 2005

### **REFEREED JOURNAL PAPERS**

- J1. E. Agichtein, Y. Liu, and J. Bian, Modeling Information Seeker Satisfaction in Community Question Answering, ACM Transactions on Knowledge Discovery from Data (TKDD), special issue on “Social Computing”, 2009.
- J2. S. Sahay, S. Mukherjea, E. Agichtein, E. V Garcia, S. Navathe, and A. Ram, Discovering Semantic Biomedical Relations utilizing the Web, in the ACM Transactions on Knowledge Discovery from Data (TKDD), special issue on Bioinformatics, 2008
- J3. P. Ipeirotis, E. Agichtein, P. Jain, and L. Gravano, Towards a Query Optimizer for Text-Centric Tasks, in the ACM Transactions on Database Systems (TODS), vol. 32, no. 4, 2007
- J4. E. Agichtein, S. Lawrence and L. Gravano, Learning to Find Answers to Questions on the Web, in the ACM Transactions on Internet Technology (TOIT) Special Issue on “Machine Learning for the Internet”, 2004
- J5. H. Yu and E. Agichtein, Extracting Synonymous Gene and Protein Terms from Biological Literature, in Bioinformatics, 2003 (also in Proc. of ISMB 2003)
- J6. F. M. Torres, E. Agichtein, L. Grinberg, G. Yu, and R. Q. Topper, A note on the application of the “Boltzmann simplex”-Simulated Annealing algorithm to global optimizations of argon and water clusters, Journal of Molecular Structure (THEOCHEM), 1997

### **REFEREED CONFERENCE PAPERS**

- C1. J. Bian, Y. Liu, D. Zhou, E. Agichtein, and H. Zha, Learning to Recognize Reliable Users and Content in Social Media with Coupled Mutual Reinforcement, in Proc. of the International World Wide Web Conference (WWW), 2009 (12% accepted)
- C2. Q. Guo, E. Agichtein, C. Clarke and A. Ashkan, “In the Mood to Click? Towards Inferring Searcher Receptiveness to Search Advertising,” (short paper) in Proc. of IEEE/WIC/ACM International Conference on Web Intelligence (WI), 2009 (25% accepted).
- C3. A. Ashkan, C. Clarke, E. Agichtein, and Q. Guo, “Estimating Ad Clickthrough Rate through Query Intent Analysis,” in Proc. of IEEE/WIC/ACM International Conference on Web Intelligence (WI), 2009 (18% accepted)
- C4. A. Ashkan, C. Clarke, E. Agichtein, and Q. Guo, Classifying and Characterizing Query Intent in Sponsored Search, in Proc. of the 31st European Conference on Informational Retrieval (ECIR), 2009 (30% accepted)
- C5. Y. Liu, J. Bian, and E. Agichtein, Predicting Information Seeker Satisfaction in Community Question Answering, in Proc. of the ACM SIGIR International Conference on Research and Development in Information Retrieval (SIGIR), 2008 (17% accepted)
- C6. B. Li, Y. Liu, and E. Agichtein, CoCQA: Co-Training Over Questions and Answers with an Application to Predicting Question Subjectivity Orientation, in Proc. of Conference on Empirical Methods in Natural Language Processing (EMNLP), 2008 (21% accepted)
- C7. Y. Liu and E. Agichtein, You’ve Got Answers: Towards Personalized Models for Predicting Success in Community Question Answering (short paper), in Proc. of the Annual Meeting of the Association for Computational Linguistics (ACL), 2008 (25% accepted)
- C8. J. Bian, Y. Liu, E. Agichtein and H. Zha. Finding the Right Facts in the Crowd: Factoid Question Answering over Social Media, in Proc. of the International World Wide Web Conference (WWW), 2008 (11% accepted)

- C9. E. Agichtein, C. Castillo, D. Donato, A. Gionis, G. Mishne, Finding High Quality Content in Social Media, with an Application to Community-Based Question Answering, in Proc. of the ACM Web Search and Data Mining Conference (WSDM), 2008 (16% accepted)
- C10. C. Clarke, E. Agichtein, S. T. Dumais, and R. W. White, The Influence of Caption Features on Click-through Patterns in Web Search, in Proc. of the ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), 2007 (18% accepted)
- C11. E. Agichtein, C. Burges, and E. Brill, Question Answering over Implicitly Structured Web Content, in Proc. of the IEEE/WIC/ACM Conference on Web Intelligence (WI), 2007 (17% accepted)
- C12. P. Jurczyk and E. Agichtein, Discovering Authorities in Question Answer Communities Using Link Analysis (short paper), in Proc. of the ACM Conference on Information and Knowledge Management (CIKM), 2007 (26% accepted)
- C13. E. Agichtein, E. Brill, and S. T. Dumais, Improving Web Search Ranking by Incorporating User Behavior Information, in Proc. of the ACM SIGIR Conference on Research and Development on Information Retrieval (SIGIR), 2006 (19% accepted)
- C14. E. Agichtein, E. Brill, S. T. Dumais, and R. Ragno, Learning User Interaction Models for Predicting Web Search Result Preferences, in Proc. of the ACM SIGIR Conference on Research and Development on Information Retrieval (SIGIR), 2006 (19% accepted)
- C15. P. Ipeirotis, E. Agichtein, P. Jain, and L. Gravano, To Search or to Crawl: Towards a Query Optimizer for Text-Centric Tasks, in Proc. of the ACM Conference on Management of Data (SIGMOD), Best Paper Award, 2006 (13% accepted)
- C16. E. Agichtein and Z. Zheng, Identifying "Best Bet" Web Search Results by Mining Past User Behavior (short paper), in Proc. of the ACM International Conference on Knowledge Discovery and Data Mining, (KDD), Industrial Applications track, 2006 (24% accepted)
- C17. E. Agichtein, Confidence Estimation Methods for Partially Supervised Relation Extraction (short paper), in Proc. of the SIAM Conference on Data Mining (SDM), 2006 (30% accepted)
- C18. E. Agichtein and S. Cucerzan, Predicting Accuracy of Extracting Information from Unstructured Text Collections, in Proc. of the ACM Conference on Information and Knowledge Management (CIKM), 2005 (18% accepted)
- C19. E. Agichtein and V. Ganti, Mining Reference Tables for Automatic Text Segmentation, in Proc. of the ACM International Conference on Knowledge Discovery and Data Mining (KDD), 2004 (12% accepted)
- C20. E. Eskin and E. Agichtein, Combining Text Mining and Sequence Analysis to Discover Protein Functional Regions, in Proc. of the Pacific Symposium on Biocomputing (PSB), 2004 (28% accepted)
- C21. E. Agichtein and L. Gravano, Querying Text Databases for Efficient Information Extraction, in Proc. of the International Conference on Data Engineering (ICDE), Best Student Paper Award, 2003 (14% accepted)
- C22. H. Yu and E. Agichtein, Extracting Synonymous Gene and Protein Terms from Biological Literature, in Proc. of the Conference on Intelligent Systems for Molecular Biology (ISMB), 2003 (15% accepted)
- C23. E. Agichtein, S. Lawrence and L. Gravano, Learning Search Engine Specific Query Transformations for Question Answering, in the 10th World Wide Web Conference (WWW), 2001 (20% accepted)
- C24. E. Agichtein and L. Gravano, Snowball: Extracting Relations from Large Plain-Text Collections, in the 5th ACM International Conference on Digital Libraries (ACM DL), 2000 (33% accepted)

## **REFEREED WORKSHOP, POSTER AND DEMO PAPERS**

- W1. Q. Guo and E. Agichtein. Beyond Session Segmentation: Predicting Changes in Search Intent With Client-Side User Interactions (poster), in Proc. of the ACM SIGIR International Conference on Research and Development in Information Retrieval (SIGIR), 2009 (34% accepted)
- W2. Y. Liu, N. Narasimhan, V. Vesudehvan and E. Agichtein. Is This Urgent? Exploring Time-Sensitive Information Needs in Community Question Answering (poster), in Proc. of the ACM SIGIR International Conference on Research and Development in Information Retrieval (SIGIR), 2009 (34% accepted)
- W3. Q. Guo, R. Kelly, S. Deemer, A. Murphy, J. A. Smith, and E. Agichtein. EMU: The Emory User Behavior Modeling System for Automatic Library Search Evaluation (poster), in Proc. of Joint Conference on Digital Libraries (JCDL), 2009
- W4. Q. Guo, E. Agichtein, C. Clarke and A. Ashkan. Understanding "Abandoned" Ads: Towards Personalized Commercial Intent Inference via Mouse Movement Analysis, in Proc. of the SIGIR 2008 Workshop on Information Retrieval in Advertising (IRA), 2008
- W5. A. Ashkan, C. Clarke, E. Agichtein and Q. Guo. Characterizing Query Intent From Ad Clickthrough Data, in Proc. of the SIGIR 2008 Workshop on Information Retrieval in Advertising (IRA), 2008
- W6. J. Bian, Y. Liu, E. Agichtein and H. Zha, A Few Bad Votes Too Many? Towards Robust Ranking in Social Media, in Proc. of the WWW Workshop on Adversarial Information Retrieval (AIRWeb), 2008
- W7. Q. Guo and E. Agichtein, Exploring Client-Side Instrumentation for Personalized Search Intent Inference: Preliminary Experiments, in Proc. of the AAAI 2008 Workshop on Intelligent Techniques for Web Personalization and Recommender Systems (ITWP), 2008
- W8. B. Li, Y. Liu, A. Ram, E. V. Garcia, and E. Agichtein, Subjectivity Analysis for Questions in QA Communities (poster), in Proc. of the ACM SIGIR International Conference on Research and Development in Information Retrieval (SIGIR), 2008
- W9. Y. Liu, E. Agichtein, On the Evolution of the Yahoo! Answers QA Community (poster), in Proc. of the ACM SIGIR International Conference on Research and Development in Information Retrieval (SIGIR), 2008
- W10. Q. Guo, E. Agichtein, Exploring Mouse Movements for Inferring Query Intent (poster), in Proc. of the ACM SIGIR International Conference on Research and Development in Information Retrieval, 2008
- W11. P. Jurczyk and E. Agichtein. HITS on Question Answer Portals: an Exploration of Link Analysis for Author Ranking (poster), in Proc. of the ACM SIGIR 2007
- W12. L. Xiong and E. Agichtein. Towards Privacy-Preserving Query Log Publishing, in Proc. of the Query Log Analysis: Social and Technological Challenges Workshop at WWW 2007
- W13. S. Sahay, E. Agichtein, E.V. Garcia, B. Li, and A. Ram. Semantic Annotation and Inference for Medical Knowledge Discovery, in Proc. of NSF Symposium on Next Generation Data Mining (NGDM), 2007
- W14. E. Agichtein and S. Cucerzan, Predicting Extraction Performance by Using Context Language Models, in Proc. of the SIGIR Workshop on Methodologies and Evaluation of Lexical Cohesion Techniques in Real-World Applications (SIGIR ELECTRA), 2005
- W15. E. Agichtein, S. Cucerzan, and E. Brill, Analysis of Factoid Questions for Effective Relation Extraction (poster), in Proc. of the ACM SIGIR 2005
- W16. E. Agichtein, P. Ipeirotis, and L. Gravano, Modeling Query-Based Access to Text Databases, in Proc. of the Sixth International Workshop on the Web and Databases (WebDB), 2003 (25% accepted)
- W17. E. Agichtein, C.T. H. Ho, V. Josifovski, and J. Gerhardt. Extracting Relations from XML Documents, in Springer Lecture Notes in Computer Science (LNCS), Volume 2814, "Conceptual Modeling for Novel Application Domains"; also in the Workshop on XML Schema and Data Management (XSDM), 2003

- W18. E. Agichtein and Luis Gravano, QXtract: A Building Block for Efficient Information Extraction from Plain-Text Databases (demo), in the ACM International Conference on Management of Data (SIGMOD), 2003
- W19. E. Agichtein, L. Gravano, J. Pavel, V. Sokolova, A. Voskoboynik. Snowball: A Prototype System for Extracting Relations from Large Text Collections (demo), in the ACM International Conference on Management of Data (SIGMOD), 2001
- W20. A. Borthwick, J. Sterling, E. Agichtein, and Ralph Grishman. Exploiting Diverse Knowledge Sources via Maximum Entropy in Named Entity Recognition, in the Sixth Workshop on Very Large Corpora, 1998
- W21. E. Agichtein, E. Eskin and L. Gravano. Combining Strategies for Extracting Relations from Text Collections, in the ACM SIGMOD Workshop on Data Mining and Knowledge Discovery (DMKD), 2000

### **OTHER PUBLICATIONS**

- O1. E. Agichtein, M. Hearst, and I. Soboroff, The Search and Social Media Workshop at SIGIR 2009: Workshop Report, to appear in SIGIR Forum, 2009., December issue
- O2. E. Agichtein, W. Askew, and Y. Liu, Combining Lexical and Semantic Features for Textual Entailment Classification, in Proc. of TAC 2008
- O3. S. Sahay, B. Li, E. V. Garcia, E. Agichtein, and A. Ram, Domain Ontology Construction from Biomedical Text, in the Proc. of the 2007 International Conference on Artificial Intelligence (ICAI), 2007
- O4. S. Cucerzan and E. Agichtein, Factoid Question Answering over Unstructured and Structured Content on the Web at TREC 2005, in the Proc. of the TREC 2005 conference
- O5. E. Agichtein, Extracting Relations From Large Text Collections, Ph.D. Thesis, Columbia University, 2005
- O6. A. Borthwick, J. Sterling, E. Agichtein, and R. Grishman, NYU: Description of the MENE Named Entity System as used in MUC-7, in the proceedings of the 7th Message Understanding Conference (MUC-7)

### **MISCELLANEOUS**

Citizenship: United States of America.

Languages: English: fluent, Russian: fluent (mother tongue).