

Eugene Agichtein

Associate Professor
Mathematics and Computer Science Department, Emory University

400 Dowman Drive, Suite W401
Atlanta, GA 30322

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

RESEARCH SUMMARY

Eugene Agichtein founded and leads the Emory Intelligent Information Access Laboratory (IR Lab): <http://ir.mathcs.emory.edu/>. The IR Lab works on information retrieval and text and data mining, 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 medical informatics. He collaborates closely with researchers in Psychology, Neuroscience, and Biomedical Informatics at Emory, and as well as with scientists at Georgia Tech, Google, Yahoo! Labs, and Microsoft Research. Dr. Agichtein and colleagues have produced an influential body of work resulting in 5 awarded U.S. Patents and more than 100 publications, including 4 best paper awards, with over 8,500 citations (h-index=38) according to Google Scholar.

EDUCATION

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

SELECTED HONORS AND AWARDS

- | | |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| Karen Sparck Jones Award from the British Computer Society | 2014 |
| For “significant contribution to research on information retrieval and natural language processing”. Presented invited keynote lecture at the ECIR 2014 Conference. | |
| “Best Student Paper” Award, ACM Conf. on Web Search and Data Mining (WSDM) | 2014 |
| For the paper “Discovering Common Motifs in Cursor Movement Data for Improving Web Search”, by Dmitry Lagun*, Mikhail Ageev*, and Eugene Agichtein. | |
| * indicates a visitor or student supervised by me. | |
| A. P. Sloan Research Fellowship | 2012 |
| Honored among about 100 researchers chosen from across all areas of natural sciences and mathematics. | |
| “Best Paper” Award, ACM Conf. on Information Retrieval (SIGIR) | 2011 |
| For the paper “Find It If You Can: Modeling Different Types of Web Search Success Using Interaction Data”, by Mikhail Ageev*, Qi Guo*, Dmitry Lagun*, and Eugene Agichtein. | |
| DARPA Computer Science Study Group Fellow | 2011 |
| Microsoft New Faculty Fellow Finalist (one of 10 selected across North America) | 2009 |
| Microsoft Research “Beyond Search” Award | 2007 |
| “Best Paper” Award, ACM Conf. on Management of Data (SIGMOD) | 2006 |
| “Best Student Paper” Award, IEEE Conf. on Data Engineering (ICDE) | 2003 |

PROFESSIONAL EXPERIENCE

Associate Professor, Emory University, Mathematics & Computer Science Dept. 2012–present
Director of Graduate Studies, Computer Science & Informatics Program 2016–present

Assistant Professor, Mathematics & Computer Science Dept. 2006-2012

Founded and leading the Emory Information Retrieval lab (IR Lab) 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 about the IR Lab is at: <http://ir.mathcs.emory.edu/>.

Visiting Principal Research Scientist, Yahoo! Labs 2014-2015

Search and Question Answering Research (while on leave from Emory University)
 Developed new methods for automated question answering for web search queries.

Visiting Researcher, Microsoft Research Summer 2012

Context, Learning and User Experience Group.
 Explored searcher intention and interaction for multimedia retrieval.

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

Research Intern, IBM Almaden Research Center Summer 2001

Research Intern, NEC Research Institute Summer 2000

Assistant Research Scientist, New York University (part-time) 1997–1998

TEACHING**Emory University**

“CS325: Artificial Intelligence,” 28,35,25 students respectively Spring 2014, 2016,2017
 “CS584: Intelligent Assistants (Graduate, new course),” 12 students Spring 2017
 “CS557: Artificial Intelligence (Graduate),” 22 students Fall 2016
 “CS572: Information Retrieval” (graduate, new course), 12-17 students Fall 2008, Spring 2010, 2016
 “CS571: Natural Language Processing” (graduate, new course), 11 students Fall 2007, Spring 2014
 “CS190: The Web” (new freshman seminar), 11-16 students Spring 2009, Spring 2010, Spring 2013

“CS370: Software Engineering Practicum,” 20 students	Spring 2012, 2013
“CS171: Introduction to Computer Science II,” 14-26 students	Spring 2007, Fall 2008, Fall 2011
“CS170: Introduction to Computer Science I,” 28-30 students	Fall 2009, Fall 2010
“CS730: Text Mining for Social Media” (advanced graduate seminar), 5 students	Fall 2010
“CS190: The Web: Concepts and Technologies” (freshman seminar), 16 students	Spring 2010, 2013
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

STUDENT AND POSTDOC SUPERVISION

Graduated Ph.D. Students:

Qi Guo: 2007-2012. Ph.D. awarded November 2012. First position: Applied Scientist, Microsoft.
 Dmitry Lagun: 2009-2014. Ph.D. awarded May 2015; First position: Research Scientist, Google.
 Qiaoling Liu: 2009-2014. Ph.D. awarded May 2015; First position: Data Scientist, CareerBuilder.com.
 Yu Wang: 2009-2015. PhD awarded December 2015. First position: High tech startup.
 Denis Savenkov: 2011-2017. May 2017; First Position: Research Scientist, Facebook.

Current Ph.D. students:

Payam Karisani: 2016-
 Zihao Wang: 2017-

Current Undergraduate and other students advised:

None

Ph.D. thesis committee:

Shuguang Han, University Of Pittsburgh, Spring 2016-2017
 Anna Shtok, Technion, Computer Science, Spring 2015
 Shuang Hong Yang, Georgia Tech, Spring 2012
 Jeff Huang, University of Washington, Spring 2013
 James Gardner, Emory Math&CS Department, Spring 2011
 Steven Crane, Georgia Tech, Spring 2011
 Dawid Kurzyniec, Emory Math & CS Department, Spring 2007
 Saurav Sahay, Georgia Tech, Fall 2011
 Jiang Bian, Georgia Tech, Fall 2010

Postdocs supervised:

Dr. Baoli Li (2007-2008): currently Associate Professor in Henan University of Technology, China
 Dr. Alexander Kotov (Fall 2011-2013): Currently Assistant Professor at Wayne State University
 Dr. JongHo Shin (Fall 2011-2012): Web-based Automated Early Detection of Alzheimer’s Disease (NIH)

Ph.D. Qualifier examiner: Ablimit Aji (Emory Math&CS, Spring 2012)

Graduated M.S. students: Yandong Liu (2009), Alvin Grissom (2009), Akshatha K. Pai (2010), Noah Adler (2014).

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

Other Students and Visitors Supervised:

Pavel Braslavsky from Ural State University (2015): Community question answering (**Fulbright Scholar**)
 Dr. Mikhail Ageev from Moscow State University (Winter 2011, 2013): Search success models (NSF)

Nikita Zhiltsov from Kazan State University (2012-2013): Robust Entity Search (**Fulbright Scholar**)
 Dr. Tianyong (Tony) Hao from City University of Hong Kong (Spring 2010): Collaborative Question Answering
 Julia Kiseleva from St. Petersburg State University (2009-2010): Query Segmentation (Shopping.com)

Directed study/rotation project supervision: Thomas Przybylinski (Spring 2009),
 Ryan P. Kelly (2008-2009), Samrat Jeyaparakash (2009), Brandon Lock (2010-2011), Haojian Jing (2012)

Other Educational Activities

Marshall Gregory Pedagogy Seminar Participant Winter 2011
 A 5-day intensive workshop on teaching methods and approaches, Emory CFDE

CS Graduate Admissions Committee Fall 2006-present

PROFESSIONAL SERVICE

Editorial Board

Journal of Information Retrieval (JIR), 2010-present

Conference and Workshop Organization

- 2017 **Program Committee Chair:** The 2017 World Wide Web Conference (WWW'17):
 WWW is the most visible CS conference for all areas of Web-related research
Area Chair 2017 ACM SIGIR Conference on Research and Development in Information Retrieval
Area Chair 2017 ACL Conference of Association of Computational Linguistics
- 2016 **Area Chair:** 2016 ACM SIGIR Conference on Research and Development in Information Retrieval
Organizer: 2016 TREC competition on Live Question Answering (LiveQA), with Yahoo Labs
Senior PC member: 2016 ACM International Conference on Web Search and Data Mining (WSDM)
Co-Chair: SIGIR 2016 Workshop on Web Question Answering
- 2015 **Area Chair:** ACM Conference on Research and Development in Information Retrieval (SIGIR)
Organizer: TREC competition on Live Question Answering (LiveQA), with Yahoo Labs
Senior PC member: ACM Conference on Information and Knowledge Management (CIKM)
Senior PC member: ACM International Conference on Web Search and Data Mining (WSDM)
Workshops Co-Chair: ACM International Conference on Web Search and Data Mining (WSDM)
Co-Chair: SIGIR 2015 Workshop on Web Question Answering
- 2014 **Area Chair:** ACM Conference on Research and Development in Information Retrieval (SIGIR)
Senior PC member: AAAI Conference on Weblogs and Social Media (ICWSM)
- 2013 **Co-Chair:** ACM Workshop on Health Search and Discovery at SIGIR 2013.
Area Chair: ACM International Conference on World Wide Web (WWW)
Short Papers Chair: European Conference on Information Retrieval (ECIR)
Area Chair: North American Conference on Computational Linguistics (NAACL)
Area Chair: ACM Conference on Research and Development in Information Retrieval (SIGIR)
Senior PC member: ACM Conference on Information and Knowledge Management (CIKM)
Senior PC member: AAAI Conference on Weblogs and Social Media (ICWSM)
- 2012 **Program Committee Co-Chair:** ACM International Conference on Web Search and Data Mining (WSDM)
Workshops Co-Chair: ACM Conference on Research and Development in Information Retrieval (SIGIR)
Area Chair: ACM Conference on Research and Development in Information Retrieval (SIGIR)
Area Chair: Conference on Empirical Natural Language Processing (EMNLP)
Co-Chair: WWW 2012 Workshop on Collaborative Question Answering (CQA'2012)
Senior PC member: AAAI Conference on Artificial Intelligence (AAAI)
Senior PC member: AAAI Conference on Weblogs and Social Media (ICWSM)
Senior PC member: ACM Conference on Information and Knowledge Management (CIKM)
Prize Review Committee Member: Overstock.com \$1M Recommendation Challenge, USA
Steering Committee Member: Recommendation track (RecLab) at CLEF 2012, EU/Spain.

- 2011 **Area Chair:** ACM Conference on Research and Development in Information Retrieval (SIGIR)
Senior PC Member: AAAI Conference on Weblogs and Social Media (ICWSM)
Workshops Co-Chair: International World Wide Web Conference (WWW)
- 2010 **Area Chair (IR):** NAACL Conference on Human Language Technologies (NAACL-HLT)
Area Chair (Web IR): ACM Conference on Research and Development in Information Retrieval (SIGIR)
PC Vice Chair: IEEE International Conference on Data Mining (ICDM)
Senior PC Member: AAAI Conference on Weblogs and Social Media (ICWSM)
Co-Chair: Third Annual Workshop on Searching Social Media (SSM 2010) 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 (SSM2008)
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, Selected listing (Since 2006)

ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD): 2006,2008-2012
 International World Wide Web Conference (WWW): 2010-2011
 Annual Meeting of the Association for Computational Linguistics (ACL): 2008,2010,2011
 IEEE International Conference on Data Engineering (ICDE): 2008, 2011
 ACM International Conference on Web Search and Data Mining (WSDM): 2008,2009,2010
 AAAI Conference on Artificial Intelligence (AAAI), AI&The Web Track: 2008, 2010
 NAACL Conference on Human Language Technologies (NAACL-HLT): 2007,2009
 AAAI International Conference on Web Logs and Social Media (ICWSM): 2008, 2009
 Conference on Intelligent User Interfaces (IUI): 2009
 International Conference on Knowledge Capture (K-CAP): 2007,2009
 Conference on Empirical Methods in Natural Language Processing (EMNLP): 2007, 2008
 IEEE International Conference on Data Engineering (ICDE)
 Conference on Empirical Methods in Natural Language Processing (EMNLP)

Journal 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).

INVITED TALKS AND KEYNOTES (Since 2006)

- 2016 **No Contribution Left Behind: User Feedback as (Implicit) Contribution to Wikipedia**
 Wiki2016: Workshop on Collaborative Creation at WWW 2016.
- 2015 **Answering Searcher Questions:**
 Microsoft Research DB/IR day, invited talk
 Technion: Israel Institute of Technology; Haifa University; Bar-Ilan University.
- 2014 **Inferring Searcher Attention and Intention by Mining Behavior Data:**
 Opening Keynote at the ECIR 2014 Conference in Amsterdam, Netherlands
- 2013 **Looking Beyond the Clicks: Acquiring and Mining Searcher Examination and Interaction Data:**
 2013 Workshop on Web Search Click Data at WSDM conference
- 2012 **Extracting Meaning from Online Behavior Data:**
 University of Texas at Dallas, TX
 University of Maryland, CLIP Colloquium
 Microsoft Research, Redmond, WA

- 2011 **Extracting Meaning from Online Behavior Data:**
 Microsoft Research, Redmond, WA
 Yahoo! Research, Santa Clara, CA
 Army Research Laboratories (ARL)
 GVU “Brown Bag” at Georgia Institute of Technology, Atlanta, GA
- 2010 **Inferring Searcher Intent from Interaction Data:**
 Invited talk at the SIGIR 2010 workshop on Understanding Query Intent
 Invited talk at the SIGIR 2010 workshop on Feature Generation for Search
 Google Tech Talk at Google Research, Zurich
- 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
 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, GVU Center brownbag
 Yahoo Research, Santa Clara

TUTORIALS

- 2012 E. Agichtein and E. Gabrilovich, Organizing and Searching Collaboratively Generated Content, at the ACM WSDM 2012 and at the WWW 2012 Conferences.
- 2011 E. Agichtein and E. Gabrilovich, Organizing and Searching Collaboratively Generated Content, at the ACM SIGIR 2011 and at the AAAI 2011 Conferences.
- 2010 E. Agichtein, Inferring Searcher Intent, at the International World Wide Web Conference (WWW 2010)
 E. Agichtein, Inferring Searcher Intent, at the AAAI Conference on Artificial Intelligence (AAAI 2010)
- 2009 E. Agichtein, Modeling Searcher Behavior, at the 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. E. Agichtein, Q. Guo, P. Wolff, Automatic client-side user-behavior analysis for inferring user intent, U.S. Patent 8,606,725 (filed 2009, issued 2013)
- P2. E. Agichtein, C. Burges, and E. Brill, Question Answering over Structured Content on the Web, U.S. Patent 7873624 (filed 2005, issued 2011)
- P3. V. Ganti, T. Vassilakis, and E. Agichtein, Segmentation of Data into Structured Records, U.S. Patent 7,627,567 (filed 2004, issued 2009).
- P4. C. J. Burges and E. Agichtein, Training a learning system with arbitrary cost functions, U.S. Patent 7,472,096 (filed 2005, issued 2008).

- P5. E. Agichtein and S. Lawrence, Method for retrieving answers from an information retrieval system, U.S. Patent 7,269,545 (filed 2001, issued 2007).

PUBLICATIONS

INVITED JOURNAL PAPERS

- I1. Eugene Agichtein, David Carmel, Dan Pelleg, Yuval Pinter, and Donna Harman. "Overview of the TREC 2015 LiveQA Track. Proc. of The Twenty-Fourth Text REtrieval Conference (TREC 2015)". NIST Special Publication, February 2016.
- I2. 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
- I3. 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
- I4. 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 (* indicates students supervised or co-advised by me)

- J1. *T. Hao and E. Agichtein. Finding Similar Questions in Collaborative Question Answering Archives: Toward Bootstrapping-based Equivalence Pattern Learning. Journal of Information Retrieval (JIR), Spring 2012
- J2. *D. Lagun, C. Manzanares, S. M. Zola, E. A. Buffalo, and E. Agichtein, Detecting cognitive impairment by eye movement analysis using automatic classification algorithms, Journal of Neuroscience Methods, Volume 201, Issue 1, September 2011.
- J3. 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.
- J4. *S. Sahay, S. Mukherjea, E. Agichtein, E. V Garcia, S. Navathe, and A. Ram, Discovering Semantic Biomedical Relations utilizing the Web, ACM Transactions on Knowledge Discovery from Data (TKDD), special issue on Bioinformatics, 2008
- J5. P. Ipeirotis, E. Agichtein, P. Jain, and L. Gravano, Towards a Query Optimizer for Text-Centric Tasks, ACM Transactions on Database Systems (TODS), vol. 32, no. 4, 2007
- J6. E. Agichtein, S. Lawrence and L. Gravano, Learning to Find Answers to Questions on the Web, ACM Transactions on Internet Technology (TOIT) Special Issue on "Machine Learning for the Internet", 2004
- J7. H. Yu and E. Agichtein, Extracting Synonymous Gene and Protein Terms from Biological Literature, in Bioinformatics, 2003 (also in Proc. of ISMB 2003)
- J8. 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 TOP-TIER CONFERENCE PAPERS (* indicates students/visitors supervised or co-advised by me)

- C1. *D. Savenkov and E. Agichtein, EviNets: Neural Networks for Combining Evidence Signals for Factoid Question Answering (short paper), in Proc. of ACL 2017
- C2. *A. Vtyurina, *D. Savenkov, E. Agichtein, and C. Clarke. Exploring Conversational Search With Humans, Assistants, and Wizards (Late-breaking work). in Proc. of CHI 2017
- C3. *P. Braslavski, *D. Savenkov, E. Agichtein, and A. Dubatovka. What Do You Mean Exactly?: Analyzing Clarification Questions in CQA. In Proc. of CHIIR 2017.
- C4. *D. Savenkov and Eugene Agichtein, When a Knowledge Base Is Not Enough: Question Answering over Knowledge Bases with External Text Data, in Proc. of ACM SIGIR 2016
- C5. D. Pelleg, O. Rokhlenko, I. Szpektor, E. Agichten and Ido Guy, When the Crowd is Not Enough: Improving User Experience with Social Media through Automatic Quality Analysis, in Proc. of ACM CSCW 2016
- C6. *D. Lagun, E. Agichtein. Inferring Searcher Attention by Jointly Modeling User Interactions and Content Salience, in Proc. of SIGIR 2015
- C7. *Y. Wang, *D. Fink, and E. Agichtein, SEEF: Planned Social Event Discovery and Attribute Extraction by Fusing Twitter and Web Content, in Proc. of ICWSM 2015.
- C8. Xin Chen, Yu Wang, Fusheng Wang, and Eugene Agichtein. A Comparative Study of Demographic Attribute Inference in Twitter, In Proc. of ICWSM 2015
- C9. Kotov A., Rakesh V., Agichtein E. and Reddy C. Geographical Latent Variable Models for Microblog Retrieval. In Proc. of ECIR 2015
- C10. *D. Lagun, *M. Ageev, *Q. Guo, and E Agichtein, Discovering common motifs in cursor movement data for improving web search, in Proc. of the 7th ACM international conference on Web search and Data Mining (WSDM) 2014 (**best student paper award**)
- C11. C.H. Tan, E. Agichtein, P. Ipeirotis, E. Gabrilovich, Trust, but Verify: Predicting Contribution Quality for Knowledge Base Construction and Curation, in Proc. of the 7th ACM international conference on Web search and Data Mining (WSDM) 2014
- C12. *N. Zhiltsov, E Agichtein, Improving entity search over linked data by modeling latent semantics, In Proc. of the 22nd ACM international conference on Conference Information and Knowledge Management (CIKM), 2013
- C13. A. Kotov, E. Agichtein, The importance of being socially-savvy: quantifying the influence of social networks on microblog retrieval, in In Proc. of the 22nd ACM international conference on Conference Information and Knowledge Management (CIKM), 2013
- C14. *Q. Guo, *H. Jin, *D. Lagun, *S. Yuan, E. Agichtein, Mining touch interaction data on mobile devices to predict web search result relevance. In Proc. of the 36th international ACM SIGIR conference, 2013. (**selected as a “best of 2013” by ACM Computing Reviews**).
- C15. *M. Ageev, *D. Lagun, E. Agichtein, Improving search result summaries by using searcher behavior data, In Proc. of the 36th international ACM SIGIR conference, 2013. (**best paper runner-up**)
- C16. D. Pelleg, *D. Savenkov, E. Agichtein, Touch Screens for Touchy Issues: Analysis of Accessing Sensitive Information from Mobile devices, In Proc. of the 7th International AAAI Conference on Weblogs and Social Media (ICWSM), 2013.

- C17. *Q. Guo and E. Agichtein, Beyond Dwell Time: Estimating Document Relevance from Cursor Movements and other Post-click Searcher Behavior, in Proc. of WWW 2012 (12% accepted)
- C18. *Q. Liu, E. Agichtein, G. Dror, Y. Maarek, and I. Szpektor. When web search fails, searchers become askers: understanding the transition. In Proc. of SIGIR 2012 (18% accepted)
- C19. E. Agichtein, R. White, S. Dumais, P. Bennett, Search, interrupted: understanding and predicting search task continuation, in Proc. of SIGIR 2012 (18% accepted)
- C20. *Y. Wang, E. Agichtein and M. Benzi, TM-LDA: Efficient Online Modeling of Latent Topic Transitions in Social Media, in Proc. of KDD 2012 (11% accepted)
- C21. E. Agichtein, E. Buffalo, *D. Lagun, A. Levey, C. Manzanares, *J. Shin, S. Zola, Developing and Validating Automated Web-Based Behavioral Diagnostics for Early Detection of Alzheimers Disease, in Medicine-X Conference, 2012
- C22. *Q. Guo, *D. Lagun, and E. Agichtein, Predicting web search success with fine-grained interaction data (short paper), in Proc. of CIKM 2012
- C23. *M. Ageev, *Q. Guo, *D. Lagun, and E. Agichtein, Find It If You Can: Modeling and Predicting Different Types of Web Search Success with Interaction Data, in Proc. of SIGIR 2011 (19% accepted) **“Best Paper” award**
- C24. *D. Lagun and E. Agichtein, ViewSer: Enabling Large-Scale Remote User Studies of Web Search Examination and Interaction, in Proc. of SIGIR 2011 (19% accepted)
- C25. *Q. Liu, E. Agichtein, G. Dror, E. Gabrilovich, Y. Maarek, D. Pelleg, I. Szpektor, Predicting Web Searcher Satisfaction with Existing Community-based Answers, in Proc. of SIGIR 2011 (19% accepted)
- C26. *D. Lagun and E. Agichtein, ViewSer: A Tool for Large-Scale Remote Studies of Web Search Result Examination, in Proc. of the 29th ACM Conference on Human Factors in Computing Systems (CHI) Extended Abstracts, 2011.
- C27. *Q. Liu and E. Agichtein, Modeling Answerer Behavior in Web-Scale Collaborative Question Answering, with Implications for Question Recommendation, in Proc. of ECIR 2011 (20% accepted).
- C28. K. Radinsky, E. Agichtein, E. Gabrilovich, S. Markovitch: A word at a time: computing word relatedness using temporal semantic analysis. in Proc. of WWW 2011 (12% accepted).
- C29. *A. Aji, *Y. Wang, E. Agichtein, and E. Gabrilovich, Using the past to score the present: Extending Term Weighting Models with Revision History Analysis, CIKM 2010 (13% accepted).
- C30. *Q. Guo and E. Agichtein, Ready to Buy or Just Browsing? Detecting Web Searcher Goals from Interaction Data, In Proc. of the ACM SIGIR International Conference on Research and Development in Information Retrieval (SIGIR), 2010 (17% accepted)
- C31. *Q. Liu, *Y. Liu, and E. Agichtein. Exploring Web Browsing Context for Collaborative Question Answering, in Proc. of IiX 2010.
- C32. *Q. Guo and E. Agichtein, Towards Predicting Web Searcher Gaze Position from Mouse Movements. In Proc. of the 28th ACM Conference on Human Factors in Computing Systems (CHI) Works-In-Progress, 2010.
- C33. *Y. Wang and E. Agichtein, Query Ambiguity Revisited: Clickthrough Measures for Distinguishing Informational and Ambiguous Queries (short paper). In Proc. of NAACL-HLT, 2010 (33% accepted).
- C34. *A. Aji and E. Agichtein, Deconstructing Interaction Dynamics in Knowledge Sharing Communities, Proc. of the 2010 International Conference on Social Computing, Behavioral Modeling, and Prediction (SBP), 2010.

- C35. *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)
- C36. *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).
- C37. 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)
- C38. 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)
- C39. *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)
- C40. *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)
- C41. *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)
- C42. *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)
- C43. 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)
- C44. C. Clarke, E. Agichtein, S. T. Dumais, and R. W. White, The Influence of Caption Features on Clickthrough Patterns in Web Search, in Proc. of the ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), 2007 (18% accepted)
- C45. 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)
- C46. *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)
- C47. 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)
- C48. 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)
- C49. 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)

- C50. 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)
- C51. 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)
- C52. 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)
- C53. 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)
- C54. 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)
- C55. 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)
- C56. 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)
- C57. 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)
- C58. 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 PAPERS AND POSTERS (* indicates students supervised by me)

- W1. E. Kravi, E. Agichtein, I. Guy, Y. Kanza, A. Mejer, D. Pelleg, Searcher in a Strange Land: Understanding Web Search from Familiar and Unfamiliar Locations (poster), in Proc. of SIGIR 2015
- W2. *Y. Wang, T. Clark, E. Agichtein, J. Staton. Towards Tracking Political Sentiment through Microblog Data, In Proc. of ACL 2014 Workshop on Social Dynamics, 2014
- W3. *D. Lagun, E. Agichtein, Effects of Task and Domain on Searcher Attention, Effects of Task and Domain on Searcher Attention (poster), in Proc. of SIGIR 2014
- W4. *D. Savenkov, E. Agichtein, To hint or not: exploring the effectiveness of search hints for complex informational tasks (poster), in Proc. of SIGIR 2014
- W5. *D. Lagun, E. Agichtein, Re-Examining Search Result Snippet Examination Time for Relevance Estimation (poster), in Proc. of SIGIR 2012
- W6. *B. Lock, E. Agichtein, K. Holmes, and P. Wolff. Using Web Corpus Statistics to Infer Conceptual Structure (Poster). Proc. of the Cognitive Science Conference (COGSCI) 2011
- W7. *Q. Guo, *S. Yuan, and E. Agichtein, Pinch, Zoom, and Slide: Touch Screens Shown Helpful for Detecting Success in Mobile Search (poster), Proc. of SIGIR 2011
- W8. *Y. Wang and E. Agichtein, Temporal Latent Semantic Analysis for Collaboratively Generated Content: Preliminary Results (poster), Proc. of SIGIR 2011

- W9. *D. Lagun, E. Agichtein, E. Buffalo, C. Manzanares, S. Zola, VPW: An Interactive Prototype of a Web-Based Visual Paired Comparison Cognitive Diagnostic Test (demo paper), in Proc. of the ACM International Health Informatics Symposium (IHI) 2010
- W10. *Q. Guo, and E. Agichtein. Exploring Searcher Interactions for Distinguishing Types of Commercial Intent (poster). in Proc. of the 19th International World Wide Web Conference (WWW), 2010
- W11. *J. Kiseleva, Q. Guo, E. Agichtein, D. Billsus and W. Chai. Unsupervised Query Segmentation Using Click Data: Preliminary Results (poster). in Proc. of the 19th International World Wide Web Conference (WWW), 2010
- W12. *Q. Liu, *Y. Liu and E. Agichtein. Collaborative Question Answering from the Answerers Perspective: Exploring Answerer Context in Question Recommendation. in Proc. of the Third Annual Workshop on Search in Social Media (SSM 2010)
- W13. *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)
- W14. *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)
- W15. *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
- W16. *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
- W17. 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
- W18. *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
- W19. *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
- W20. *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 Int. Conf. on Research and Development in Information Retrieval (SIGIR), 2008
- W21. *Y. Liu, E. Agichtein, On the Evolution of the Yahoo! Answers QA Community (poster), in Proc. of the ACM SIGIR Int. Conf. on Research and Development in Information Retrieval (SIGIR), 2008
- W22. *Q. Guo, E. Agichtein, Exploring Mouse Movements for Inferring Query Intent (poster), in Proc. of the ACM SIGIR Int. Conf. on Research and Development in Information Retrieval, 2008
- W23. *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 Int. Conf. on Research and Development in Information Retrieval, 2007
- W24. 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

- W25. *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
- W26. 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
- W27. E. Agichtein, S. Cucerzan, and E. Brill, Analysis of Factoid Questions for Effective Relation Extraction (poster), in Proc. of the ACM SIGIR Int. Conf. on Research and Development in Information Retrieval, 2005
- W28. 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)
- W29. 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 (XSMD), 2003
- W30. E. Agichtein and Luis Gravano, QXtract: A Building Block for Efficient Information Extraction from Plain-Text Databases (demo), in ACM Int. Conf. on Management of Data (SIGMOD), 2003
- W31. E. Agichtein, L. Gravano, J. Pavel, V. Sokolova, A. Voskoboynik. Snowball: A Prototype System for Extracting Relations from Large Text Collections (demo), in ACM Int. Conf. on Management of Data (SIGMOD), 2001
- W32. 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
- W33. 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 (* indicates students or postdocs supervised by me)

- O1. E. Agichtein, E. Buffalo, *D. Lagun, A. Levey, C. Manzanares, *J. Shin, S. Zola, Developing and Validating Automated Web-Based Behavioral Diagnostics for Early Detection of Alzheimers Disease, presented at the Stanford Medicine-X Conference, September 2012
- O2. Katz, Daniel Martin, Bommarito, Michael James, Seaman, Julie, Candebub, Adam and Agichtein, Eugene, Legal N-Grams? A Simple Approach to Track the Evolution of Legal Language, in Proceedings of JURIX 2011: The 24th International Conference on Legal Knowledge and Information Systems, December 2011.
- O3. E. Agichtein, E. Buffalo, *D. Lagun, A. Levey, C. Manzanares, S. Zola, VPW: a Web-based visual paired comparison task for early detection of amnesic mild cognitive impairment (abstract), The Neuroscience 2010 Conference
- O4. *A. Pai, E. Agichtein, A. Post, and J. Saltz. The Emory System for Extracting Medical Concepts at i2b2 Challenge: Combining Knowledge-Based and Machine Learning Techniques. in Proc. of the i2b2 NLP Challenge Workshop, 2010.
- O5. E. Agichtein, M. Hearst, and I. Soboroff, The Search and Social Media Workshop at SIGIR 2009: Workshop Report, SIGIR Forum, 2009., December issue
- O6. E. Agichtein, *W. Askew, and *Y. Liu, Combining Lexical and Semantic Features for Textual Entailment Classification, in Proc. of TAC 2008

- O7. *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
- O8. 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
- O9. E. Agichtein, Extracting Relations From Large Text Collections, Ph.D. Thesis, Columbia University, 2005
- O10. 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)

GRANT PROPOSALS FUNDED

- Improving Search Ranking of Biomedical Datasets with Machine Learning. (\$76,000)** 2017
Emory University PI, subcontract from UCSD.
- Yahoo! Labs Faculty Research Engagement Program Award. (\$20,000)** 2015
Sole PI, competitive process each year.
- Discovering Relations from Question and Answer Archives (\$80,000)** 2014
Sole PI, Google Research Award
- Extracting Causal Relations from Natural Text (\$34,500)** 2013
Co-PI (PI: Phillip Wolff). Google Research Award.
- Automated Web-Based Behavioral Diagnostics of Cognitive Impairment (\$1,370,000)** 2011-2015
Role: PI. Co-Investigators: E. Buffalo, A. Levey, S. Zola.
National Institutes of Health, R01, Nat. Inst. of Biomedical Imaging & Bioinformatics
- III: Modeling and Inferring Searcher Intent by Mining User Interactions (\$500,000)** 2010-2014
Role: sole PI
National Science Foundation, CISE, Division of Information and Intelligent Systems
- Computer Science Study Group (CSSG) (\$499,000)** 2011-2014
Role: sole PI. \$99,520 Phase 1 + \$399,353 Phase 2).
Defence Advanced Research Projects Agency (DARPA)
- Sloan Research Fellowship (\$50,000)** 2012-2014
Alfred P. Sloan Foundation
- Yahoo! Labs Faculty Research Engagement Program. (\$112,000)** 2008-2013
Sole PI, award competitively renewed each year.
- Yahoo! Equipment Donation. (about \$26,000 value)** 2012
Sole PI, 512-node Hadoop cluster, servers donated to the department by Yahoo! Labs
- Google Research Award, unrestricted gift. (\$27,500)** 2011-2012
Sole PI, competitive open RFP program.
- HP Labs Open Innovation Award (\$75,000)** 2009

Role: joint PI with H. Zha (Georgia Institute of Technology); Emory portion: \$40,000.

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