SWAPNA SOMASUNDARAN

Curriculum Vitae

CONTACT: ssomasundaran@ets.org

EDUCATION

Ph.D. Computer Science, University of Pittsburgh, Pittsburgh, PA, August 2010
Advisor: Janyce Wiebe
Dissertation: Discourse-level Relations for Opinion Analysis

M.S.E. Computer Science, Johns Hopkins University, Baltimore, MA. October 2002

B.E. Electronics Engineering, Mumbai University, Mumbai, June 1999

PROFESSIONAL EMPLOYMENT HISTORY

Research Scientist, Research & Development, Educational Testing Service, Princeton, NJ (2013-present)

INDIVIDUAL CONTRIBUTION AND RESEARCH

- Acquired project funding as Principal Investigator (6 projects) or collaborator (3 projects) by writing proposals to internal research allocation.
- Led research in multiple NLP areas such as Sentiment Analysis, Discourse, Narratives, and Graph-methods while building applications for Automated Writing Evaluation. Built new capabilities in narrative quality assessments, discourse coherence, persuasiveness in argumentative writing, evaluating idea development using graphs, scoring source-based writing using topic signatures, scoring picture-based items.
- Provided technical expertise across various groups at ETS. Advised K-12 Business Unit for responding to Request for Proposals (RFP). Served as NLP expert on an international research grant (MEWS) and Speech Construct Relevance and Validity (CRV) initiative.
- Authored Publications and Patents
- Disseminated research outcomes via external conference presentations and internal seminars

MANAGEMENT AND LEADERSHIP

- Co-led the “Text and Speech Systems” sub-initiative (FY2015, FY2016) with a goal to advance research towards building resources and tools that support an array of NLP and Speech systems for aiding item generation, improving instruction, ensuring test security, in addition to creating foundational capabilities that can be applied to assessments. Monitored budgets, managed proposals, advised project leads for 5 projects in FY2015 and 8 projects in FY 2016.

- Managed multi-year (4 projects) and single-year (3 projects) projects under ETS research allocation. Co-ordinated teams (team size varied from 2 to 30), external consultant, staff and resources for reaching milestones and project deliverables.
- Managed requests for e-rater from external researchers by mentoring proposal writing, and acting as a liaison between legal and research.
- Mentored student intern and other ETS staff on research projects, resulting in conference/workshop publications, and patent activity. Supervised Research Assistants and Associates on ETS projects.
- Assisted in the recruitment of full-time employees.

Research Scientist, Siemens Corporate Research, Princeton, NJ, USA (2010-2012)

- Research and Development: Led research in the areas of relationship mining and word spaces. Developed technologies for projects in domains such as fraud detection, personalized healthcare, patent invalidity search and fleet monitoring.
- Project Acquisition and Management: Played leading and supporting roles in the acquisition of customer projects via prototype development, customer workshop, presentations and discussions, managed several customer projects ranging from small-scale prototype development to large-scale projects. Hired, trained and supervised three software consultants.
- Proposal Writing: Led the writing of government proposals, involving multiple universities and departments within Siemens. Assisted in shaping proposals for internal funding.
- Mentoring: Advised student interns on research projects, resulting in journal, conference and workshop publications, and patent applications. Co-advised interns from other groups.
- University Collaborations: Led an internally funded collaboration with UC Berkeley and ICSI FrameNet on developing frame semantics for encoding knowledge in the medical domain. Represented Siemens on the advisory board for SALTS (Rutgers Laboratory for the Study of Applied Language Technology and Society). Led a DARPA proposal with CMU and University of Pittsburgh.
- Miscellaneous: Initiated weekly reading club. Collaborated cross-department and cross-program on various projects via module development and discussions. Assisted in the recruitment of full-time employees, software consultants and interns.

Software Consultant, Parity Teltech, IBM TJ Watson Center, New York, USA (2002-2004)

- Assisted in the development of Unstructured Information Management Architecture (UIMA) architecture. The UIMA architecture is the foundation of IBM Watson, the NLP system that won Jeopardy in 2011.

Senior Systems Engineer, Cognitive Systems Research Lab, Tata Infotech Limited, Mumbai, India (2000-2001)

- Developed a Question Answering system. The system was demoed at NAACL 2001.
- Managed and coordinated a team of interns and outside experts.

Associate Systems Engineer, Tata Infotech Limited, Mumbai, India (1999-2000)

- Implemented touch-screen info kiosks for the Central Railways in Mumbai.

Undergraduate Internship, Bhabha Atomic Research Center, Mumbai, India (1998-1999)

- Developed a system to automate the testing process for bus-based embedded system I/O cards.

AWARDS and DISTINCTIONS

- SPOT award at ETS 2014, for outstanding team achievement in responding to a difficult RFP.
- Best Reviewer Award, EMNLP 2010: Conference on Empirical Methods in NLP (2010)
- Andrew Mellon Pre-Doctoral Fellowship, University of Pittsburgh (2009-2010)
- Best Graduate Student Research Award, Computer Science Department, University of Pittsburgh (2009)
- Outstanding Paper Presentation Award, School of Arts and Sciences Grad Exposition, University of Pittsburgh (2009)
- Popular Press, Research featured in Department of Homeland Security Network Newsletter (2007) (http://www.dhsnetwork.org/Somasundaran.html)
- Incoming Doctoral Student Fellowship, Faculty of Arts and Sciences (FAS), University of Pittsburgh (2004)
- Outstanding Employee Contribution Award, Tata Infotech Ltd, India (2001)

PUBLICATIONS

"" indicates that papers were first-authored by a student who Swapna mentored*

Michael Flor, Swapna Somasundaran (2017). Sentiment Analysis and Lexical Cohesion for the Story Cloze Task, *to appear* in LSDSem EACL Workshop, Valencia, Spain *(This publication was accompanied by an oral presentation)*

Swapna Somasundaran, Brian Riordan, Binod Gyawali, Su-Youn Yoon. (2016). Evaluating Argumentative and Narrative Essays using Graphs, in Proceedings of COLING 2016, Osaka, Japan. *(This publication was accompanied by a poster presentation)*

Lei Chen, Ben Leong, Gary Feng, Chong Min Lee, and Swapna Somasundaran (2015) Utilizing Multimodal Cues to Automatically Evaluate Public Speaking Performance , in Proceedings of International Conference on Affective Computing and Intelligent Interaction (ACII), Xi'an, China, 2015 *[Google Scholar Citation as of 01/18/2017: 5] (This publication was accompanied by a poster presentation)*

Swapna Somasundaran, Chong Min Lee, Martin Chodorow and Xinhao Wang (2015) Automated Scoring of Picture-based Story Narration , The 10th Workshop on Innovative Use of NLP for Building Educational Applications, NAACL 2015, Denver, CO, USA *[Google Scholar Citation as of 01/18/2017: 4] (This publication was accompanied by an oral presentation)*

* Noura Farra, Swapna Somasundaran and Jill Burstein (2015) Scoring Persuasive Essays Using Opinions and their Targets . The 10th Workshop on Innovative Use of NLP for Building Educational Applications, NAACL 2015, Denver, CO, USA *[Google Scholar Citation as of 01/18/2017: 5] (This publication was accompanied by a poster presentation)*

Swapna Somasundaran Jill Burstein and Martin Chodorow, (2014) Lexical Chaining for Measuring Discourse Coherence Quality in Test-taker Essays , COLING 2014, Dublin, Ireland *[Google Scholar Citation as of 01/18/2017: 10] (This publication was accompanied by an oral presentation)*

Swapna Somasundaran and Martin Chodorow, (2014) Automated Measures of Specific Vocabulary Knowledge from Constructed Responses ("Use These Words to Write a Sentence Based on this Picture"), In Proceedings of the Ninth Workshop on Innovative Use of NLP for Building Educational Applications, Baltimore, Maryland, ACL 2014 *[Google Scholar Citation as of 01/18/2017: 5] (This publication was accompanied by an oral presentation)*

Jill Burstein, Swapna Somasundaran and Martin Chodorow, (2014) Finding your "inner-annotator": An experiment in annotator independence for rating discourse coherence quality in essays. 8th Linguistic Annotation Workshop (LAW VIII 2014) , Dublin, Ireland. *[Google Scholar Citation as of 01/18/2017: 0] (This publication was accompanied by an oral presentation)*

Beata Beigman Klebanov, Nitin Madnani, Jill Burstein and Swapna Somasundaran (2014) Content Importance Models for Scoring Writing From Sources. , In Proceedings of ACL 2014 (short papers) Baltimore, Maryland *[Google Scholar Citation as of 01/18/2017: 5] (This publication was accompanied by a poster presentation)*

* Kateryna Tymoshenko, Swapna Somasundaran, Vinodkumar Prabhakaran, Vinay Damodar Shet (2012) Relation Mining in the Biomedical Domain using Entity-level Semantics ECAI 2012: 20th European Conference on Artificial Intelligence, Montpellier, France, August 27-31 *[Google Scholar Citation as of 01/18/2017: 2](This publication was accompanied by an oral presentation)*

* Dingcheng Li, Swapna Somasundaran and Amit Chakraborty (2012) ERD-MedLDA: Entity relation detection using supervised topic models with maximum margin learning Journal of Natural Language Engineering (JNLE), 18 , pp 263-289 doi:10.1017/S1351324912000058 *[Google Scholar Citation as of 01/18/2017: 2]*

* Dingcheng Li, Swapna Somasundaran and Amit Chakraborty (2011) A Combination of Topic Models with Max-margin Learning for Relation Detection In Proceedings of TextGraphs-6: Graph-based Methods for Natural Language Processing, Workshop at ACL-HLT 2011, Portland, Oregon *[Google Scholar Citation as of 01/18/2017: 6] (This publication was accompanied by an oral presentation)*

Swapna Somasundaran (2010), Discourse-level relations for Opinion Analysis, PhD Thesis, University of Pittsburgh. *[Google Scholar Citation as of 01/18/2017: 27]*

Swapna Somasundaran and Janyce Wiebe, (2010), Recognizing Stances in Ideological On-line Debates. , In Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and

Generation of Emotion in Text, pages 116-124, Los Angeles, CA. Association for Computational Linguistics, 2010. *[Google Scholar Citation as of 01/18/2017: 152] (This publication was accompanied by a poster presentation)*

Swapna Somasundaran and Janyce Wiebe, (2009), Recognizing Stances in Online Debates , ACL 2009: Joint conference of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, August 2-7, 2009, Singapore. *[Google Scholar Citation as of 01/18/2017: 158] (This publication was accompanied by an oral presentation)*

Swapna Somasundaran, Galileo Namata, Janyce Wiebe and Lise Getoor (2009) Supervised and Unsupervised Methods in Employing Discourse Relations for Improving Opinion Polarity Classification , EMNLP 2009: conference on Empirical Methods in Natural Language Processing August 6-7, 2009, Singapore. *[Google Scholar Citation as of 01/18/2017: 73] (This publication was accompanied by an oral presentation)*

Swapna Somasundaran, Galileo Namata, Lise Getoor and Janyce Wiebe, (2009) Opinion Graphs for Polarity and Discourse Classification , TextGraphs-4: Graph-based Methods for Natural Language Processing, 7th August 2009, Singapore *[Google Scholar Citation as of 01/18/2017: 24](This publication was accompanied by an oral presentation)*

Swapna Somasundaran, Janyce Wiebe and Josef Ruppenhofer (2008) Discourse Level Opinion Interpretation, COLING, Manchester, 18-22 August, 2008 *[Google Scholar Citation as of 01/18/2017: 83] (This publication was accompanied by an oral presentation)*

Swapna Somasundaran, Josef Ruppenhofer and Janyce Wiebe (2008) Discourse Level Opinion Relations: An Annotation Study, SIGdial Workshop on Discourse and Dialogue, Columbus, Ohio, June 2008 *[Google Scholar Citation as of 01/18/2017: 28] (This publication was accompanied by an oral presentation)*

Josef Ruppenhofer, Swapna Somasundaran and Janyce Wiebe (2008) Finding the Sources and Targets of Subjective Expressions. LREC 2008, Marrakech, Morocco. *[Google Scholar Citation as of 01/18/2017: 66] (This publication was accompanied by an oral presentation)*

Swapna Somasundaran, Josef Ruppenhofer and Janyce Wiebe (2007) Detecting Arguing and Sentiment in Meetings SIGdial Workshop on Discourse and Dialogue, Antwerp, Belgium, September 2007 *[Google Scholar Citation as of 01/18/2017: 57] (This publication was accompanied by an oral presentation)*

Swapna Somasundaran, Theresa Wilson, Janyce Wiebe and Veselin Stoyanov (2007) QA with Attitude: Exploiting Opinion Type Analysis for Improving Question Answering in On-line Discussions and the News, International Conference on Weblogs and Social Media March 26-28, 2007, Boulder, Colorado, U.S.A. *[Google Scholar Citation as of 01/18/2017: 64] (This publication was accompanied by an oral presentation)*

Swapna Somasundaran, Janyce Wiebe, Paul Hoffmann, Diane Litman (2006).Manual Annotation of Opinion Categories in Meetings. ACL Workshop: Frontiers in Linguistically Annotated Corpora

(COLING/ACL 2006), Sydney, Australia *[Google Scholar Citation as of 01/18/2017: 15] (This publication was accompanied by an oral presentation)*

Theresa Wilson, Paul Hoffmann, Swapna Somasundaran, Jason Kessler, Janyce Wiebe, Yejin Choi, Claire Cardie, Ellen Riloff, Siddharth Patwardhan (2005). OpinionFinder: A system for subjectivity analysis *Demo* in Human Language Technologies Conference/Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP 2005), Vancouver, Canada. *[Google Scholar Citation as of 01/18/2017: 377]*

PATENT ACTIVITY

- J. Burstein, S. Somasundaran, M. Chodorow (2015). Computer-Implemented Systems and Methods for Measuring Discourse Coherence. US Patent 20,150,248,397 (*Patent Granted;* Docket No. 011948-1016-999, Serial No.: 14/633,200)
- S. Somasundaran, V. Prabhakaran, V. Damodar Shet, & K. Tymoshenko (2012). A System and Method for Generating Knowledge Base for Medical Concepts. (*Patent Granted;* Publication number: US8639678 B2).
- B. Beigman Klebanov, N. Madnani, J. Burstein, S. Somasundaran (2015). Computer-Implemented Systems and Methods for Evaluating Use of Source Material in Essays. US Patent 20,150,254,229
- S. Somasundaran, M. Chodorow, J. Tetreault (2015). Systems and Methods for Automated Scoring of Textual Responses to Picture-Based Items. (US Patent 20,150,243,181)
- S. Somasundaran, D. Li, A. Chakraborty. (2012). Relationship Mining Using Latent Dirichlet Allocation. (Patent application filed with USPTO: 2010P20118EP, 2010E19409 US).

TALKS AND PRESENTATIONS

- Swapna Somasundaran, Yoko Futagi, Nitin Madnani, Nancy Glazer, Matt Chametsky and Cathy Wendler (2016). Prediction of Passage Acceptance/Rejection Using Linguistic Information. NCME 2016, Washington DC.
- Invited Talk: Discourse-level Opinion Analysis (2009). University of Maryland, College Park
- Invited Talk: Discourse-level Opinion Analysis (2009). Indian Institute of Information Technology (IIIT), Hyderabad, India
- Recognizing Stances in Online Debates (2009). AI forum, Intelligent Systems Program, University of Pittsburgh
- Opinion is not only word deep: Discourse-level Relations for Opinion Analysis (2009). Arts and Sciences Graduate Student Expo, University of Pittsburgh (*won the outstanding paper presentation award*)
- Discourse-level relations for opinion analysis (2009). Computer Science Research Competition, University of Pittsburgh (*won best graduate student research award*)
- Discourse Level Opinion Interpretation (2008). AI forum, Intelligent Systems Program, University of Pittsburgh

- Detecting Arguing and Sentiment in Meetings (2007). AI forum, Intelligent Systems Program, University of Pittsburgh

ETS SEMINARS AND CLIENT PRESENTATIONS

- Presented story-cloze system "Sentiment Analysis and Lexical Cohesion for the Story Cloze Task" at the LTA Brown Bag Presentation (2017)
- Presented R&D work on Source-Use at the TOEFL Symposium (to the TOEFL Committee of Examiners) (2015)
- Presented Passage Finding work at the SELT Symposium (2015)
- Presented Narrative Quality Assessment work “Towards Assessing Narrative Essays” at NLP Turnbull Brown Bag (2015)
- Presented Lexical Chaining for Discourse coherence work at NLP Brown bag (2014)

PROFESSIONAL ACTIVITIES

Organizing Committee, Program Committee

- Organizer (TextGraphs Workshop, 2017)
- Area Chair for Sentiment Analysis and Argument Mining (COLING 2016)
- Handbook Chair (EMNLP 2016)
- Organizer (TextGraphs Workshop, 2010)
- Vice President of Women in Computer Science at University of Pittsburgh (WiCS) (2009-2010)

Editorial Board

- Permanent member of the editorial board of the Journal of Writing Analytics (2016-present).

Advisory Board

- SALTS (Rutgers Laboratory for the Study of Applied Language Technology and Society) (2011).

PC member or reviewer

Journals

- Journal of Writing Analytics (2016-present)
- Journal of AI in Education (AIED) (2016
- Journal of Computational Linguistics (CL) (2014, 2013)
- IEEE Transactions on Affective Computing (TAFC) (2013, 2011)
- IEEE Intelligent Systems (2010)

Book Chapter (reviewer)

- Emerging Applications of Natural Language Processing: Concepts and New Research (2010)

Conferences (PC member/reviewer)

- ACL (2017, 2014 2013, 2011, 2010, 2009, 2008)
- EMNLP (2016, 2015, 2014, 2011, 2010, 2009, 2007)
- NAACL (2013, 2010)

- EACL (2014)
- COLING (2016, 2014, 2010)
- LREC (2016, 2014)
- IJCNLP (2013)

Workshops (PC member)

- LSDSem (2017, 2015)
- Uphill Battles in NLP (2016)
- BEA (2016, 2015, 2014)
- SAAIP (2016, 2013, 2011)
- Student Research Workshop (SRW) (2015)
- TextGraphs (2011)
- MLG (2010)
- NESCAI (2008, 2007)

Misc (reviewer)

- COLING Demo (2014)
- ETS Research Report Series (2013-present)

TEACHING

- 2002 (Summer) Instructor, Intermediate Programming, Johns Hopkins University
- 2006 (Fall) Teaching Assistant, Introduction to NLP, University of Pittsburgh
- 2005 (Fall) Teaching Assistant, Introduction to NLP, University of Pittsburgh
- 2006 (Fall) Teaching Assistant, Introduction to Artificial Intelligence, University of Pittsburgh
- 2005 (Fall) Teaching Assistant, Discrete Structures for Computer Science, University of Pittsburgh
- 2001 (Fall) Teaching Assistant, Database Systems, Johns Hopkins University

TRAINING and SKILLS

TECHNICAL AND PROGRAMMING SKILLS: Experience in Java, Perl, Python, Shell scripts, UIMA (Unstructured Information Management Architecture), Eclipse, NetBeans, Linux, Mac OSX, Windows, Git, SVN, SQL server management studio, Latex

DATA SCIENCE SKILLS: Experience with machine learning algorithms such as logistic regression, linear regression, Support Vector Machines, Naïve Bayes, Elastic Nets; graph algorithms, and tools such as Weka, Gephi, NetworkX, cytoscape

MANAGEMENT: Completed Rutgers mini MBA program at ETS. Completed Project Management Training at Siemens.

LANGUAGES: English, Hindi, Malayalam, Marathi