

SARA ROSENTHAL

15 Fairway Oval, Spring Valley, NY 10977
(845) 304-8685 (cell)
www.cs.columbia.edu/~sara
sara@cs.columbia.edu

EDUCATION

Columbia University Graduate School of Arts and Sciences, New York, NY

Doctor of Philosophy, Computer Science, Expected Spring 2015 (GPA 4.05/4.0)

Masters of Philosophy, Computer Science, May 2012

Completed Thesis Proposal on Detecting Personal and Social Influence in Threaded Discussions in November 2012

Completed Candidacy Exam in June 2011 on Sociolinguistics in Blogs

Completed Breadth Requirements in May 2010

Natural Language Processing Group, Advisor: Kathleen McKeown

Columbia University, School of Engineering and Applied Sciences, New York, NY

MS in Computer Science; Emphasis: Machine Learning and Natural Language Processing, May 2008 (GPA 3.8/4.0)

Ramapo College of New Jersey, School of Theoretical and Applied Science, Mahwah, NJ

BS in Computer Science, Minor in Mathematics, August 2005 (GPA 3.83/4.0)

Rockland Community College, Monsey, NY

AAS in Computer Information Systems, January 2003

RESEARCH EXPERIENCE

Columbia University, Department of Computer Science, New York, NY

Spring 2008-Present

Graduate Research Assistant

Advisor: Kathleen McKeown

- DARPA DEFT and IARPA SCIL Project:
 - Research in this area has included predicting the age of bloggers, opinion detection, and identifying agreement, persuasion, claims, and influencers in weblogs, micro-blogs, and discussion forums.
- DARPA GALE Project:
 - Maintained the pipeline for a Question Answering System
 - Created system tools using Natural Language Processing techniques

Columbia University, Center for Computational Learning Systems, New York, NY

Fall 2007, Summer 2009

Graduate Research Assistant

Advisor: Rebecca Passonneau

- Con Edison Project: Analyzed Con Edison trouble tickets to predict problematic events and modified the Machine Learning Weka code to give more priority to trouble class.

Ramapo College of NJ, Department of Computer Science, Mahwah, NJ

Spring 2005-Spring 2006

Undergraduate Research Assistant

Advisor: Amruth Kumar

- Problettes - A CS Tutor Software: Created a rule base expert system to improve the grading system of the tutor to give the users helpful feedback.

RELATED WORK EXPERIENCE

IBM Research, Yorktown Heights, NY
Research Intern

May 2013 - August 2013

- Worked with Ashish Jagmohan in the Services Group
- Research area was in analyzing virtual discussions, resulting in one publication

Google Inc., New York, NY
Software Engineer Intern

May 2007 - August 2007

- Worked with Kerry Hannan on the Sentiment Analysis Team
- Created a Google iGadget graphing sentiment over time for sports, celebrities and politicians

RustyBrick, Inc, Suffern, NY
Web Developer

January 2006 - January 2007

- Developed the front and backend of client websites
- Communicated with other programmers, graphic designers, and clients on various projects
- Select websites: canvascomplete.com, exceldiamonds.com, soleconnections.com

TEACHING EXPERIENCE

Columbia University
Mentor

2009-Present

Alisa Krivokapic, Undergrad (Present)
Project: Generating demographic features for politics and religion through Twitter

Ethan Robert Grant, Undergrad (Spring 2014 – Summer 2014)
Projects: Annotating Influence, Comprehensive analysis of sentiment lexicons

Suvarna Bothe, Masters Student (Summer 2014)
Projects: Building a sentiment Knowledge Base

Gregory Barber, Undergrad (Summer 2013 – Fall 2013)
Projects: Finding the sentiment of an Entity
Research resulted in one publication

Lakshmi Vikraman, Masters Student (Summer 2012 – Fall 2013)
Projects: Domain Adaptation in Sentiment Analysis

Chris Kedzie, Undergrad (Summer 2012 – Fall 2013)
Projects: Gathering Online Documents related to the 2012 Presidential Election

Sahil Yakhmi, Undergrad (Spring 2011, Fall 2011, Spring 2012)
Projects: Sentiment Analysis in Urdu in Blogs

Jacob Andreas, Undergrad (Fall 2009, Spring 2009, Spring 2011)
Projects: Agreement in Blogs, Sentiment Analysis in Blogs, PP Attachment in Blogs
2011 CRA award research finalist, research resulted in two publications

Elizabeth Kierstead, Undergrad (Fall 2010)
Projects: Emoticon statistic calculator, Wiktionary reader

Mukund Jha, Masters Student (Fall 2009)
Projects: PP Attachment in Blogs
Research resulted in one publication

Yousuf Hauke, High School Student (Summer 2009)

Projects: Research blogs in Urdu
William Lipovsky, Undergrad Student (Summer 2009)

Projects: PP Attachment using Mechanical Turk
Research resulted in one publication
Coleman Moore, Undergrad Student (Summer 2009)

Projects: Sentence Fusion using Mechanical Turk
Research resulted in one publication

Columbia University 2007-2011
Teaching Assistant

Natural Language Processing, Michael Collins (Fall 2011)
Natural Language Processing, Kathleen McKeown (Fall 2009)
Network Theory, Dragomir Radev (Spring 2008)
Data Structures in Java, David Elson (Fall 2007, Spring 2008)
Guest lectured class on AVL trees and Hash Tables
Data Structures in Java, Shlomo Hershkop (Spring 2007)

Ramapo College of NJ Spring 2006
Tutor
Tutored single student in Organization of Programming Languages course

PUBLICATIONS

"Detecting Influencers in Social Media Discussions"

Sara Rosenthal
In proceeding of XRDS: Crossroads, The ACM Magazine for Students, 2014.

"Enterprise Discussion Analytics"

Sara Rosenthal and Ashish Jagmohan
To appear in proceedings of the 23rd ACM International Conference and Knowledge Management. 2014. Shanghai, China.

"Semeval-2014 Task 9: Sentiment Analysis in Twitter"

Sara Rosenthal, Preslav Nakov, Alan Ritter, and Veselin Stoyanov
In Proceedings of the 8th International Workshop on Semantic Evaluation. 2014, Dublin, Ireland

"Columbia NLP: Sentiment Detection of Sentences and Subjective Phrases"

Sara Rosenthal, Apoorv Agarwal, and Kathleen McKeown
In Proceedings of the 8th International Workshop on Semantic Evaluation. 2014, Dublin, Ireland

"A Rule-Based Expert System for Diagnosing Student Errors in a Tutor on Counter-Controlled Loops"

Sara Rosenthal and Amruth Kumar
In proceeding of the Second Workshop on AI-supported Education for Computer Science. June 2014, Honolulu, Hawaii

"Columbia NLP: Sentiment Slot Filling"

Sara Rosenthal, Gregory J. Barber, Kathleen McKeown
In proceedings of the TAC KBP 2013 Workshop, February 2014, Gaithersburg, Maryland

"Semeval-2013 Task 2: Sentiment Analysis in Twitter"

Preslav Nakov, Zornitsa Kozareva, Alan Ritter, Sara Rosenthal, Veselin Stoyanov, and Theresa Wilson
In Proceedings of the 7th International Workshop on Semantic Evaluation. Association for Computational Linguistics. June

2013, Atlanta, Georgia

"Columbia NLP: Sentiment Detection of Subjective Phrases in Social Media"

Sara Rosenthal and Kathleen McKeown

In Proceedings of the 7th International Workshop on Semantic Evaluation. Association for Computational Linguistics. June 2013, Atlanta, Georgia

"Detecting Opinionated Claims in Online Discussions"

Sara Rosenthal and Kathleen McKeown

In proceedings of the Sixth IEEE International Conference on Semantic Computing (IEEE ICSC2012) Special Session on Semantics and Sociolinguistics in Social Media, September 2012, Palermo, Italy.

"Detecting Influencers in Written Online Conversations"

Or Biran, Sara Rosenthal, Jacob Andreas, Kathleen McKeown and Owen Rambow

In proceedings of the Language and Social Media Workshop in the 13th Annual North American Meeting of the Association of Computational Linguistics (NAACL-HLT), June 2012, Montreal, Quebec.

"Annotating Agreement and Disagreement in Threaded Discussion"

Jacob Andreas, Sara Rosenthal, and Kathleen McKeown

In proceedings of the 8th International Conference on Language Resources and Computation (LREC), May 2012, Istanbul, Turkey.

"Age Prediction in Blogs: A Study of Style, Content, and Online Behavior in Pre- and Post-Social Media Generations"

Sara Rosenthal, Kathleen R. McKeown

In proceedings of ACL-HLT 2011, 26% acceptance rate, 18% acceptance rate as Oral Presentation

"Corpus Creation for New Genres: A Crowdsourced Approach to PP Attachment"

Kapil Thadani, Mukund Jha, Jacob Andreas, Sara Rosenthal and Kathleen R. McKeown

In Workshop on Creating Speech and Language Data With Amazon's Mechanical Turk of NAACL-HLT, June 2010.

"Time-Efficient Creation of an Accurate Sentence Fusion Corpus"

Kathleen R. McKeown, Sara Rosenthal, Kapil Thadani and Coleman Moore

In proceedings of NAACL-HLT, June 2010.

"Semi-Automated Annotation for Prepositional Phrase Attachment"

Sara Rosenthal, William J. Lipovsky, Kathleen R. McKeown, Kapil Thadani and Jacob Andreas

In proceedings of LREC, May 2010.

"Who, What, When, Where, Why? Comparing Multiple Approaches to the Cross-Lingual 5W Task."

Kristen Parton, Kathleen R. McKeown, Robert Coyne, Mona T. Diab, Ralph Grishman, Dilek Hakkani-Tur, Mary Harper, Heng Ji, Wei Yun Ma, Adam Meyers, Sara Stolbach, Ang Sun, Gokhan Tur, Wei Xu and Sibel Yaman

In proceedings of ACL-IJCNLP 2009.

INVITED TALKS

Detecting Influencers in Social Media Discussions

IBM Thomas J. Watson Research Center

<http://www.research.ibm.com/labs/watson/index.shtml>

January 19, 2015

Detecting Influencers in Social Media Discussions

IHMC Ocala
<http://www.ihmc.us>
September 15, 2014

Detecting Influencers in Online Conversations

NYC Media Lab Research Summit 2012
<http://www.nycmedialab.org/>
September 12, 2012

Sentiment Detection using Subjective Phrases and Influence Detection in Social Media

Harmony Institute
<http://harmony-institute.org/>
July 24, 2012

Search in Structured Networks

Columbia University, Dragomir Radev's Network Theory Class
November 3, 2010

PROFESSIONAL ACTIVITIES

Semeval 2015 Task 10: Sentiment Analysis in Twitter

<http://alt.qcri.org/semEval2015/task10/>
Co-organizer, Primary Contact

Semeval 2014 Task 9: Sentiment Analysis in Twitter

<http://alt.qcri.org/semEval2014/task9/>
Co-organizer, Primary Contact

Semeval 2013 Task 2: Sentiment Analysis in Twitter

<http://www.cs.york.ac.uk/semEval-2013/task2/>,
Co-organizer

Reviewer: NAACL, TAACL, AAAI, IPM, LREV, INRT, TOIS, SemEval

Board Member, Women in Computer Science Club (WICS), Columbia University
Web Administrator (2009-2010), Corporate Liaison (2011-2012), Secretary (Fall 2012)

FELLOWSHIPS AND AWARDS

IBM Open Collaboration Research (OCR) with Kathleen McKeown and Ashish Jagmohan (2014)
Computer Science Service Award (2012, 2013)
MSTA Fellowship, Columbia University (Spring 2008)
Magna cum Laude, Ramapo College of NJ (Spring 2005)

TECHNICAL SKILLS

Programming Languages: Java, Perl, JSP, PHP, MySQL, HTML, JavaScript, CSS, C++

Applications: MATLAB, Eclipse, LaTeX, Subversion