# Jordan Boyd-Graber

Jordan.Boyd.Graber@colorado.edu • +1 (920) 524-9464 • http://cs.colorado.edu/~jbg
111B ECCS • Computer Science • College of Engineering and Applied Science • Boulder, CO 80309

## University Appointments

University of Colorado Boulder — Boulder, CO
**Assistant Professor, Computer Science** — *2014–Present*

University of Maryland — College Park, MD
**Assistant Professor, College of Information Studies** — *2010–2014*
**Assistant Professor, Institute for Advanced Computer Studies** — *2011-2014*
**Affiliate Assistant Professor, Computer Science** — *2011-2015*
**Postdoc (Advisor: Philip Resnik)** — *2009-2010*

## Education

Princeton University — Princeton, NJ
**Ph.D., Computer Science** — *2010*
Thesis: Linguistic Extensions of Topic Models (Advisor: David Blei)
**M.A., Computer Science** — *2007*

California Institute of Technology — Pasadena, CA
**B.S., Computer Science** — *2004*
**B.S., History** — *2004*

## Other Employment

Princeton University — Princeton, NJ
**Writing Fellow, Princeton Writing Center** — *2007-2008*

Google — New York, NY
**Intern** — *2007*

University of California Los Angeles — Los Angeles, CA
**Digital Humanities Programmer** — *2004*

California Institute of Technology — Pasadena, CA
**Newsprint Researcher / Programmer, Einstein Papers Project** — *2003-2004*
**Peer Tutor, Hixon Writing Center** — *2001-2004*
**Lab Technician, Caltech Digital Media Center** — *2001-2003*

**Immigration status**: U.S. citizen

## Fellowships, Prizes, and Awards

- Best Paper Award, CoNLL 2015
- Honorable Mention, Best Student Paper, NIPS 2009
- Computing Innovation Postdoctoral Fellowship 2009 (declined)
- Richter Undergraduate Research Fellowship, 2001 and 2002
- American Association for Artificial Intelligence student award, International Science and Engineering Fair 2000
- Caltech Jorgensen Scholarship 2001-2004

## Publications

Students directly advised or co-advised in underline.

### 2015

1. Viet-An Nguyen, **Jordan Boyd-Graber**, Philip Resnik, and Kristina Miler. **Tea Party in the House: A Hierarchical Ideal Point Topic Model and Its Application to Republican Legislators in the 112th Congress**. *Association for Computational Linguistics*, 2015, 11 pages (25% Acceptance Rate).
2. He He, Alvin Grissom II, **Jordan Boyd-Graber**, and Hal Daumé III. **Syntax-based Rewriting for Simultaneous Machine Translation**. *Empirical Methods in Natural Language Processing*, 2015.
3. Forough Poursabzi-Sangdeh and **Jordan Boyd-Graber**. **Speeding Document Annotation with Topic Models**. *NAACL Student Research Workshop*, 2015.
4. Anupam Guha, Mohit Iyyer, Danny Bouman, and **Jordan Boyd-Graber**. **Removing the Training Wheels: A Coreference Dataset that Entertains Humans and Challenges Computers**. *North American Association for Computational Linguistics*, 2015, 11 pages (26% Acceptance Rate).
5. Stephen H. Bach, Bert Huang, **Jordan Boyd-Graber**, and Lise Getoor. **Paired-Dual Learning for Fast Training of Latent Variable Hinge-Loss MRFs**. *International Conference on Machine Learning*, 2015, 10 pages (20% Acceptance Rate).
6. Paul Felt, Eric Ringger, **Jordan Boyd-Graber**, and Kevin Seppi. **Making the Most of Crowdsourced Document Annotations: Confused Supervised LDA**. *Conference on Computational Natural Language Learning*, 2015, 10 pages.
7. Vlad Niculae, Srijan Kumar, **Jordan Boyd-Graber**, and Cristian Danescu-Niculescu-Mizil. **Linguistic Harbingers of Betrayal: A Case Study on an Online Strategy Game**. *Association for Computational Linguistics*, 2015, 10 pages (25% Acceptance Rate).
8. Thang Nguyen, **Jordan Boyd-Graber**, Jeff Lund, Kevin Seppi, and Eric Ringger. **Is your anchor going up or down? Fast and accurate supervised topic models**. *North American Association for Computational Linguistics*, 2015, 10 pages (26% Acceptance Rate).
9. Yi Yang, Doug Downey, and **Jordan Boyd-Graber**. **Efficient Methods for Incorporating Knowledge into Topic Models**. *Empirical Methods in Natural Language Processing*, 2015, 9 pages (25% Acceptance Rate).
10. Mohit Iyyer, Varun Manjunatha, **Jordan Boyd-Graber**, and Hal Daumé III. **Deep Unordered Composition Rivals Syntactic Methods for Text Classification**. *Association for Computational Linguistics*, 2015, 11 pages (25% Acceptance Rate).
11. Weiwei Yang, **Jordan Boyd-Graber**, and Philip Resnik. **Birds of a Feather Linked Together: A Discriminative Topic Model using Link-based Priors**. *Empirical Methods in Natural Language Processing*, 2015, 5 pages (22% Acceptance Rate).
12. Philip Resnik, William Armstrong, Leonardo Claudino, Thang Nguyen, Viet-An Nguyen, and **Jordan Boyd-Graber**. **Beyond LDA: Exploring Supervised Topic Modeling for Depression-Related Language in Twitter**. *NAACL Workshop on Cognitive Modeling and Computational Linguistics*, 2015.

### 2014

1. Viet-An Nguyen, **Jordan Boyd-Graber**, and Philip Resnik. **Sometimes Average is Best: The Importance of Averaging for Prediction using MCMC Inference in Topic Modeling**. *Empirical Methods in Natural Language Processing*, 2014, 6 pages (30% Acceptance Rate).
2. Naho Orita, Naomi Feldman, and **Jordan Boyd-Graber**. **Quantifying the role of discourse topicality in speakers' choices of referring expressions**. *ACL Workshop on Cognitive Modeling and Computational Linguistics*, 2014.
3. Yuening Hu, Ke Zhai, Vlad Eidelman, and **Jordan Boyd-Graber**. **Polylingual Tree-Based Topic Models for Translation Domain Adaptation**. *Association for Computational Linguistics*, 2014, 11 pages (26% Acceptance Rate).
4. Mohit Iyyer, Peter Enns, **Jordan Boyd-Graber**, and Philip Resnik. **Political Ideology Detection Using Recursive Neural Networks**. *Association for Computational Linguistics*, 2014, 10 pages (26% Acceptance Rate).
5. Kimberly Glasgow, Clay Fink, and **Jordan Boyd-Graber**. **Our grief is unspeakable: Measuring the community impact of a tragedy**. *The International AAAI Conference on Weblogs and Social Media*, 2014, 9 pages (20% Acceptance Rate).

6. Ke Zhai, **Jordan Boyd-Graber**, and Shay B. Cohen. **Online Adaptor Grammars with Hybrid Inference**. *Transactions of the Association for Computational Linguistics*, 2014, 12 pages.
7. Viet-An Nguyen, **Jordan Boyd-Graber**, Philip Resnik, Deborah Cai, Jennifer Midberry, and Yuanxin Wang. **Modeling Topic Control to Detect Influence in Conversations using Nonparametric Topic Models**. *Machine Learning*, 2014, 48 pages.
8. Viet-An Nguyen, **Jordan Boyd-Graber**, Philip Resnik, and Jonathan Chang. **Learning a Concept Hierarchy from Multi-labeled Documents**. *Neural Information Processing Systems*, 2014, 9 pages (25% Acceptance Rate).
9. Yuening Hu, **Jordan Boyd-Graber**, Brianna Satinoff, and Alison Smith. **Interactive Topic Modeling**. *Machine Learning*, 2014, 56 pages.
10. Ke Zhai, **Jordan Boyd-Graber**, and Shay B. Cohen. **Hybrid Online Inference with Adaptor Grammars**. *NIPS Workshop on Advances in Variational Inference*, 2014.
11. Mohit Iyyer and Jordan Boyd-Graber Hal Daumé III. **Generating Sentences from Semantic Vector Space Representations**. *NIPS Workshop on Learning Semantics*, 2014.
12. Alvin Grissom II, He He, **Jordan Boyd-Graber**, John Morgan, and Hal Daumé III. **Don't Until the Final Verb Wait: Reinforcement Learning for Simultaneous Machine Translation**. *Empirical Methods in Natural Language Processing*, 2014, 11 pages (30% Acceptance Rate).
13. Alison Smith, Jason Chuang, Yuening Hu, **Jordan Boyd-Graber**, and Leah Findlater. **Concurrent Visualization of Relationships between Words and Topics in Topic Models**. *ACL Workshop on Workshop on Interactive Language Learning, Visualization, and Interfaces*, 2014.
14. Jason Chuang, John D. Wilkerson, Rebecca Weiss, Dustin Tingley, Brandon M. Stewart, Margaret E. Roberts, Forough Poursabzi-Sangdeh, Justin Grimmer, Leah Findlater, **Jordan Boyd-Graber**, and Jeffrey Heer. **Computer-Assisted Content Analysis: Topic Models for Exploring Multiple Subjective Interpretations**. *NIPS Workshop on Human-Propelled Machine Learning*, 2014.
15. **Jordan Boyd-Graber**, David Mimno, and David Newman. **Care and Feeding of Topic Models: Problems, Diagnostics, and Improvements**. *Handbook of Mixed Membership Models and Their Applications*, 2014, 39 pages.
16. Thang Nguyen, Yuening Hu, and **Jordan Boyd-Graber**. **Anchors Regularized: Adding Robustness and Extensibility to Scalable Topic-Modeling Algorithms**. *Association for Computational Linguistics*, 2014, 10 pages (26% Acceptance Rate).
17. Mohit Iyyer, **Jordan Boyd-Graber**, Leonardo Claudino, Richard Socher, and Hal Daumé III. **A Neural Network for Factoid Question Answering over Paragraphs**. *Empirical Methods in Natural Language Processing*, 2014, 12 pages (26% Acceptance Rate).

**2013**

1. Viet-An Nguyen, **Jordan Boyd-Graber**, Jonathan Chang, and Philip Resnik. **Tree-Based Label Dependency Topic Models**. *NIPS Workshop on Topic Models: Computation, Application, and Evaluation*, 2013.
2. Yuening Hu, Ke Zhai, Vlad Edelman, and **Jordan Boyd-Graber**. **Topic Models for Translation Domain Adaptation**. *NIPS Workshop on Topic Models: Computation, Application, and Evaluation*, 2013.
3. **Jordan Boyd-Graber**, Kimberly Glasgow, and Jackie Sauter Zajac. **Spoiler Alert: Machine Learning Approaches to Detect Social Media Posts with Revelatory Information**. *ASIST 2013: The 76th Annual Meeting of the American Society for Information Science and Technology*, 2013, 9 pages.
4. Ke Zhai and **Jordan Boyd-Graber**. **Online Topic Models with Infinite Vocabulary**. *International Conference on Machine Learning*, 2013, 9 pages (20% Acceptance Rate).
5. Viet-An Nguyen, **Jordan Boyd-Graber**, and Philip Resnik. **Lexical and Hierarchical Topic Regression**. *Neural Information Processing Systems*, 2013, 10 pages (25% Acceptance Rate).
6. Thang Nguyen, Yuening Hu, and **Jordan Boyd-Graber**. **Evaluating Regularized Anchor Words**. *NIPS Workshop on Topic Models: Computation, Application, and Evaluation*, 2013.
7. Naho Orita, Rebecca McKeown, Naomi H. Feldman, Jeffrey Lidz, and **Jordan Boyd-Graber**. **Discovering Pronoun Categories using Discourse Information**. *Proceedings of the Cognitive Science Society*, 2013, 6 pages.
8. Viet-An Nguyen, **Jordan Boyd-Graber**, and Stephen Altschul. **Dirichlet Mixtures, the Dirichlet Process, and the Structure of Protein Space**. *Journal of Computational Biology*, 2013, 48 pages.

9. Yuening Hu, **Jordan Boyd-Graber**, Hal Daumé III, and Z. Irene Ying. **Binary to Bushy: Bayesian Hierarchical Clustering with the Beta Coalescent**. *Neural Information Processing Systems*, 2013, 9 pages (25% Acceptance Rate).
10. Viet-An Nguyen, Yuening Hu, **Jordan Boyd-Graber**, and Philip Resnik. **Argviz: Interactive Visualization of Topic Dynamics in Multi-party Conversations**. *North American Association for Computational Linguistics*, 2013, 4 pages (50% Acceptance Rate).

**2012**

1. Viet-An Nguyen, **Jordan Boyd-Graber**, and Philip Resnik. **SITS: A Hierarchical Nonparametric Model using Speaker Identity for Topic Segmentation in Multiparty Conversations**. *Association for Computational Linguistics*, 2012, 10 pages (19% Acceptance Rate).
2. Ke Zhai, **Jordan Boyd-Graber**, Nima Asadi, and Mohamad Alkhouja. **Mr. LDA: A Flexible Large Scale Topic Modeling Package using Variational Inference in MapReduce**. *ACM International Conference on World Wide Web*, 2012, 10 pages (12% Acceptance Rate).
3. Viet-An Nguyen, **Jordan Boyd-Graber**, and Philip Resnik. **"I Want to Talk About, Again, My Record On Energy ...": Modeling Topic Control in Conversations using Speaker-centric Nonparametric Topic Models**. *Mid-Atlantic Student Colloquium on Speech, Language, and Learning*, 2012.
4. Vladimir Eidelman, **Jordan Boyd-Graber**, and Philip Resnik. **Topic Models for Dynamic Translation Model Adaptation**. *Association for Computational Linguistics*, 2012, 5 pages (21% Acceptance Rate).
5. Yuening Hu and **Jordan Boyd-Graber**. **Suggesting Constraints for Interactive Topic Modeling**. *ICML Workshop on Machine Learning in Human Computation and Crowdsourcing*, 2012.
6. Ke Zhai and **Jordan Boyd-Graber**. **Online Topic Model with Infinite Vocabulary**. *Mid-Atlantic Student Colloquium on Speech, Language, and Learning*, 2012.
7. Yuening Hu, Ke Zhai, Sinead Williamson, and **Jordan Boyd-Graber**. **Modeling Images using Transformed Indian Buffet Processes**. *International Conference of Machine Learning*, 2012, 8 pages (27% Acceptance Rate).
8. Asad B. Sayeed, **Jordan Boyd-Graber**, Bryan Rusk, and Amy Weinberg. **Grammatical structures for word-level sentiment detection**. *North American Association of Computational Linguistics*, 2012, 10 pages (31% Acceptance Rate).
9. Yuening Hu and **Jordan Boyd-Graber**. **Efficient Tree-Based Topic Modeling**. *Association for Computational Linguistics*, 2012, 5 pages (21% Acceptance Rate).
10. **Jordan Boyd-Graber**, Brianna Satinoff, He He, and Hal Daumé III. **Besting the Quiz Master: Crowdsourcing Incremental Classification Games**. *Empirical Methods in Natural Language Processing*, 2012, 12 pages (25% Acceptance Rate).
11. Yuening Hu and **Jordan Boyd-Graber**. **Bayesian Hierarchical Clustering with Beta Coalescents**. *Mid-Atlantic Student Colloquium on Speech, Language, and Learning*, 2012.

**2011**

1. Brianna Satinoff and **Jordan Boyd-Graber**. **Trivial Classification: What features do humans use for classification?**. *Workshop on Crowdsourcing Technologies for Language and Cognition Studies*, 2011.
2. Clay Templeton, Kenneth R. Fleischmann, and **Jordan Boyd-Graber**. **Simulating Audiences: Automating Analysis of Values, Attitudes, and Sentiment**. *IEEE International Conference on Social Computing*, 2011, 4 pages (10% Acceptance Rate).
3. Kenneth R. Fleischmann, Clay Templeton, and **Jordan Boyd-Graber**. **Modeling Diverse Standpoints in Text Classification: Learning to Be Human by Modeling Human Values**. *iConference*, 2011, 2 pages.
4. Clay Templeton, Travis Brown, Sayan Battacharyya, and **Jordan Boyd-Graber**. **Mining the Dispatch under Supervision: Using Casualty Counts to Guide Topics from the Richmond Daily Dispatch Corpus**. *Chicago Colloquium on Digital Humanities and Computer Science*, 2011, 7 pages.
5. **Jordan Boyd-Graber**. **Linguistic Resource Creation in a Web 2.0 World**. *NSF Workshop on Collaborative Annotation*, 2011, 7 pages.
6. Yuening Hu, **Jordan Boyd-Graber**, and Brianna Satinoff. **Interactive Topic Modeling**. *Association for Computational Linguistics*, 2011, 10 pages (25% Acceptance Rate).
7. Clay Templeton, Kenneth R. Fleischmann, and **Jordan Boyd-Graber**. **Comparing Values and Sentiment Using Mechanical Turk**. *iConference*, 2011, 2 pages.

8. Sonya S. Nikolova, **Jordan Boyd-Graber**, and Christiane Fellbaum. **Collecting Semantic Similarity Ratings to Connect Concepts in Assistive Communication Tools**. *Modeling, Learning and Processing of Text Technological Data Structures*, 2011, 11 pages.
9. Pranav Anand, Joseph King, **Jordan Boyd-Graber**, Earl Wagner, Craig Martell, Douglas W. Oard, and Philip Resnik. **Believe Me: We Can Do This!**. *The AAAI 2011 workshop on Computational Models of Natural Argument*, 2011, 5 pages.

**2010**

1. Eric Hardisty, **Jordan Boyd-Graber**, and Philip Resnik. **Modeling Perspective using Adaptor Grammars**. *Empirical Methods in Natural Language Processing*, 2010, 10 pages (25% Acceptance Rate).
2. Nitin Madnani, **Jordan Boyd-Graber**, and Philip Resnik. **Measuring Transitivity Using Untrained Annotators**. *Creating Speech and Language Data With Amazon's Mechanical Turk*, 2010, 6 pages.
3. **Jordan Boyd-Graber**. **Linguistic Extensions of Topic Models**. 2010, 142 pages.
4. **Jordan Boyd-Graber** and Philip Resnik. **Holistic Sentiment Analysis Across Languages: Multilingual Supervised Latent Dirichlet Allocation**. *Empirical Methods in Natural Language Processing*, 2010, 11 pages (25% Acceptance Rate).

**2009**

1. Sonya S. Nikolova, **Jordan Boyd-Graber**, and Perry Cook. **The Design of ViVA: A Mixed-initiative Visual Vocabulary for Aphasia**. *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*, 2009, 6 pages.
2. Xiaojuan Ma, **Jordan Boyd-Graber**, Sonya S. Nikolova, and Perry Cook. **Speaking Through Pictures: Images vs. Icons**. *ACM Conference on Computers and Accessibility*, 2009, 8 pages (31% Acceptance Rate).
3. Jonathan Chang, **Jordan Boyd-Graber**, Chong Wang, Sean Gerrish, and David M. Blei. **Reading Tea Leaves: How Humans Interpret Topic Models**. *Neural Information Processing Systems*, 2009, 9 pages (24% Acceptance Rate).
4. **Jordan Boyd-Graber** and David M. Blei. **Multilingual Topic Models for Unaligned Text**. *Uncertainty in Artificial Intelligence*, 2009, 8 pages (31% Acceptance Rate).
5. Jonathan Chang, **Jordan Boyd-Graber**, and David M. Blei. **Connections between the Lines: Augmenting Social Networks with Text**. *Knowledge Discovery and Data Mining*, 2009, 9 pages (9% Acceptance Rate).
6. Sonya S. Nikolova, **Jordan Boyd-Graber**, Christiane Fellbaum, and Perry Cook. **Better Vocabularies for Assistive Communication Aids: Connecting Terms using Semantic Networks and Untrained Annotators**. *ACM Conference on Computers and Accessibility*, 2009, 8 pages (31% Acceptance Rate).

**2008**

1. **Jordan Boyd-Graber** and David M. Blei. **Syntactic Topic Models**. *Neural Information Processing Systems*, 2008, 8 pages (25% Acceptance Rate).
2. **Jordan Boyd-Graber** and David M. Blei. **Multilingual Topic Models**. *NIPS Workshop on Unsupervised Latent Variable Models*, 2008.
3. Jonathan Chang, **Jordan Boyd-Graber**, and David M. Blei. **Discovering social networks from free text**. *3rd Annual Machine Learning Symposium*, 2008.

**2007**

1. **Jordan Boyd-Graber** and David M. Blei. **PUTOP: Turning Predominant Senses into a Topic Model for WSD**. *4th International Workshop on Semantic Evaluations*, 2007, 5 pages.
2. **Jordan Boyd-Graber**, David M. Blei, and Xiaojin Zhu. **A Topic Model for Word Sense Disambiguation**. *Empirical Methods in Natural Language Processing*, 2007, 10 pages (27% Acceptance Rate).

**2006**

1. **Jordan Boyd-Graber**, Sonya S. Nikolova, Karyn A. Moffatt, Kenrick C. Kin, Joshua Y. Lee, Lester W. Mackey, Marilyn M. Tremaine, and Maria M. Klawe. **Participatory design with proxies: Developing a desktop-PDA system to support people with aphasia**. *Computer-Human Interaction*, 2006, 10 pages (23% Acceptance Rate).
2. **Jordan Boyd-Graber**, Christiane Fellbaum, Daniel Osherson, and Robert Schapire. **Adding Dense, Weighted, Connections to WordNet**. *Proceedings of the Global WordNet Conference*, 2006, 10 pages.

**2003**

1. Alexander Geyken and **Jordan Boyd-Graber**. **Automatic classification of multi-word expressions in print dictionaries**. *Linguisticae Investigationes*, 2003, 16 pages.

## Contracts and Grants

### Active Funding

**Multilingual Interactive Topic Modeling** *8/2015–7/2019 (DARPA LORELEI)*
**Investigators:** Jordan Boyd-Graber (PI)
**Award:** $426,654 (Share: $325,000)

**Scaling Insight into Science: Assessing the value and effectiveness of machine assisted classification within a statistical system** *8/2014–7/2016 (NSF)*
**Investigators:** Jordan Boyd-Graber (PI)
**Award:** $195,000 (Share: $195,000)

**Closing the User-Model Loop for Understanding Topics in Large Document Collections**[1] *8/2014–7/2018 (NSF)*
**Investigators:** Jordan Boyd-Graber (PI) and Leah Findlater (CO-PI)
**Award:** $650,000 (Share: $325,000)

**Bayesian Thinking on Your Feet—Embedding Generative Models in Reinforcement Learning for Sequentially Revealed Data** *8/2013–7/2016 (NSF)*
**Investigators:** Jordan Boyd-Graber (PI) and Hal Daumé III (CO-PI)
**Award:** $500,000 (Share: $250,000)

### Completed Funding

**Sentiment Analysis in Social Media: Political Spin and Cultural Biases** *8/2013–8/2014 (CASL)*
**Investigators:** Philip Resnik PI and Jordan Boyd-Graber CO-PI
**Award:** $100,000 (Share: $50,000)

**Cross-Language Bayesian Models for Web-Scale Text Analysis** *9/2009–8/2014 (NSF)*
**Investigators:** Jimmy Lin (PI), Philip Resnik (CO-PI), Jordan Boyd-Graber[2] (CO-PI)
**Award:** $350,000 (Share: $175,000)

**Language Evidence for Social Goals** *8/2009–10/2012 (IARPA)*
**Investigators:** Philip Resnik (PI), Pranav Anand (CO-PI), Jordan Boyd-Graber (CO-PI), Deborah Cai (CO-PI), Craig Martell (CO-PI), Doug Oard (CO-PI), Marilyn Walker (CO-PI)
**Award:** $1,454,439 (Share: $100,000)

**Center for Language and Cultural Analysis** *9/2009–8/2012 (ARL)*
**Investigators:** Amy Weinberg (PI), Jordan Boyd-Graber (CO-PI), Michele Gelfand (CO-PI), Philip Resnik (CO-PI, later PI)
**Award:** $ 735,050 (Share: $100,000)

**Advanced Open Source Exploitation Models** *4/2011–12/2011 (Lockheed Martin)*
**Investigators:** Philip Resnik (PI), Jordan Boyd-Graber (CO-PI)
**Award:** $60,000 (Share: $30,000)

**Social Media Scanning** *5/2011–12/2011 (Optimal Solutions Group)*
**Investigators:** Philip Resnik (PI), Jordan Boyd-Graber (CO-PI)
**Award:** $29,849 (Share: $14,925)

## Teaching, Mentoring and Advising

### Courses Taught

**CSCI 5622: Machine Learning** *Colorado, Fall 2015*
104 students

**CSCI 5622: Machine Learning** *Colorado, Spring 2015*
58 students

**CSCI/LING 5832: Natural Language Processing** *Colorado, Fall 2014*
32 students

**INST 737: Digging into Data** UMD, *Spring 2014*
29 students

[1] After I moved to Colorado, Leah Findlater assumed the role of PI to enable a new subcontract to Colorado; the original grant as awarded is provided here.

[2] I wrote this grant while a postdoc working with Philip Resnik; the vast majority of the text and the entirety of the scientific ideas were my own. However, I could not serve as PI while still a postdoc. I was added to the grant as PI after it was awarded and served as sole research advisor to the students funded by the grand while the other PIs were both on sabbatical.

**CMSC/LING 723 / INST 735: Computational Linguistics I** UMD, *Fall 2013*
45 students

**LING 848B / CMSC 828B: Bayesian Nonparametrics** UMD, *Spring 2013*
15 students

**INST 737: Digging into Data** UMD, *Spring 2013*
30 students

**LBSC 690: Introduction to Information Technology** UMD, *Fall 2012*
30 students

**INST728C / CMSC 773 / LING 773: Computational Linguistics II** UMD, *Spring 2012*
11 Students

**LBSC 690: Introduction to Information Technology** UMD, *Fall 2011*
30 students

**INFM 718G: Web Scale Information Processing Applications** UMD, *Spring 2011*
12 students

**LBSC 690: Introduction to Information Technology** UMD, *Fall 2010*
30 students

**COS/LIN 280: Computational Linguistics** *Princeton, Fall 2008*
40 students

## Course or Curriculum Development

- New offering of *CSCI 5622: Machine Learning* (Spring / Fall 2015) as a flipped classroom
- Significant revisions to *LBSC 690: Information Technology* (Fall 2012)
- Chair of committee developing new undergraduate Information Science program at Universities at Shady Grove for University of Maryland (2011-2013)
- Developed new course *INST 737: Digging into Data* (Spring 2013), and recorded lectures for "flipped" classroom in 2014.
- Redesigned both elements of Computational Linguistics I-II sequence (2012 and 2013), and created a "flipped" classroom in 2013 for Computational Linguistics I.

## Guest lectures

- 2010, CMSC 726: Topic Models
- 2011, LING 773: Topic Models
- 2012, CMSC 421: Topic Models
- 2012, CMSC 726: Topic Models
- 2015, CSCI 5832: Topic Models

## Advising: Research Direction (Undergraduate)

1. Stephanie Hwa: Vector word representations for named entities in question answering
2. Danny Bouwman: Crowdsourced coreference annotation

## Advising: Research Direction (Masters)

**Chair or Co-Chair**

1. Brianna Satinoff (UMD, CMSC): Incremental Models for Text Classification [First position: Wellpoint]
2. Alison Smith (UMD, CMSC): Evaluating Interfaces for Interactive Topic Modeling [First position: Decisive Analytics]

**On Committee**

1. Jordan Hoskins (German)
2. Bradley Skaggs (UMD, CMSC) [First position: US Government]

## Doctoral Students

**Chair or Co-chair**

1. Pedro Rodriguez (CSCI): Distributed Machine Learning
2. Shudong Hao (CSCI): Interactive Multilingual Topic Modeling
3. Fenfei Guo (CSCI): Interactive Embedding Learning
4. Forough Poursabzi-Sangdeh (CSCI): Active Labeling with Topic Models
5. Alvin Grissom II (CSCI): Reinforcement Learning for Feature-wise Language Tasks
6. Kim Glasgow[3] (UMD: iSchool): Social Action in Social Media

[3] now advised by Yla Tausczik

7. Thang Nguyen (UMD: CMSC): Evaluation of Topic Models
8. He He (UMD: CMSC): Algorithms that Trade-Off Speed and Accuracy
9. Mohit Iyyer (UMD: CMSC): Deep Learning for Question Answering
10. Viet-An Nguyen (UMD: CMSC, 2015): Detecting Influence in Text [First position: Facebook Data Science]
11. Ke Zhai (UMD: CMSC, 2014): Large Scale Bayesian Inference [First position: Yahoo! Research]
12. Yuening Hu (UMD: CMSC, 2014): Interactive Topic Modeling [First position: Yahoo! Research]

**On committee**

1. Ben London (UMD: CMSC)
2. Irene Eleta (UMD: INFO, 2014)
3. Kevin Dayaratna (UMD: STAT, 2014)
4. Jiarong Jiang (UMD: CMSC, 2014)
5. Jagadeesh Jagarlamudi (UMD: CMSC, 2013) [First position: IBM Research]
6. Amit Goyal (UMD: CMSC, 2013) [First position: Yahoo!]
7. Piyush Rai (Utah: CMSC, 2012) [First position: University of Texas]
8. Arvind Agarwal (UMD: CMSC, 2012) [First position: Xerox Research]
9. Elena Zheleva (UMD: CMSC, 2011) [First position: Living Social]
10. Asad Sayeed (UMD: CMSC, 2011) [First position: Universität des Sarlandes]

### Invited Talks

1. **Interactive Topic Modeling and The US Tea Party**: New Frontiers of Automated Content Analysis in the Social Sciences, 2015 (Zurich, CH)
2. **Thinking on your Feet: Reinforcement Learning for Incremental Language Tasks**: Colorado School of Mines, 2014 (Golden, CO); Harvey Mudd College, 2014 (Claremont, CA); California Institute of Technology, 2014 (Pasadena, CA); Front Range NLP (Boulder, CO); EECS Colloquium, Colorado School of Mines, 2014 (Golden, CO); Brigham Young University, 2014 (Provo, UT); Peking University, 2014 (Beijing, PRC); Darmstadt University, 2014 (Darmstadt, Germany); Hong Kong University of Science and Technology, 2014 (Hong Kong); Cornell University, 2015 (Ithaca, NY)
3. **Big Data Analysis with Topic Models: Human Interaction, Streaming Computation, and Social Science Applications**: University of Colorado Boulder Computer Science Colloquium, 2013 (Boulder, CO); Yandex Machine Learning Conference, 2013 (**Invited Keynote**, Moscow, Russia); DC NLP Meetup, 2014 (Washington, DC); Yahoo! Labs, 2014 (New York, NY); Nortwestern University (Evanston, IL); Renmin University, 2014 (Beijing, PRC); Tsinghua University, 2014 (Beijing, PRC)
4. **Incorporating Human Knowledge and Insights into Probabilistic Models of Text**: Brigham Young University Department of Computer Science Colloquium, 2012 (Provo, UT)
5. **Besting the Quiz Master: Crowdsourcing Incremental Classification Games**: Rutgers University, 2012 (New Brunswick, NJ); Brigham Young University, 2012 (Provo, UT)
6. **Making Topic Models More Human(e)**: Colorado University, 2012 (Boulder, CO); University of Maryland Institute for Technology and Humanities, 2012 (College Park, MD)
7. **When Topic Models Go Bad: Diagnosing and Improving Models for Exploring Large Corpora**: Johns Hopkins University, 2011 (Baltimore, MD); Rutgers University, 2011 (New Brunswick, NJ)
8. **Inference and Validation of Probabilistic Models of Language in the Cloud**: UMD Winter Storm, 2011 (College Park, MD)
9. **Interactive Topic Models**: Harvard University's New Directions in Text Analysis Symposium, 2011 (Cambridge, MA); Princeton University, 2011 (Princeton, NJ); Maryland Institute for Technology and the Humanities: Topic Modeling and the Humanities Workshop, 2012 (College Park, MD)
10. **Putting Words Together: Crowdsourcing Data Collection for Lexical Similarity and Topical Coherence**: University of Massachusetts, 2010 (Amherst, Massachusetts)
11. **Topic Models, Mechanical Turk, and WordNet**: Harvard University, 2010 (Cambridge, MA)
12. **Topic Models and Hierarchical Models**: Johns Hopkins Summer Workshop for SMT, 2010 (Baltimore, MD)
13. **Linguistic Extensions to Topic Models**: University of Massachusetts, 2009 (Amherst, Massachusetts); Center for Communications Research, 2009 (Princeton, NJ); Center of Excellence, 2009 (Baltimore, MD); Columbia University, 2009 (New York, NY)

## Professional Service

**Conference Reviewing**

- Program Committee *2014 Workshop on Language Technologies and Computational Social Science*
- Program Committee *Neural Information Processing Systems*: 2014, 2013, 2012, 2011, 2010, 2009

- Program Committee *Association for Computational Linguistics*: 2014, 2012, 2011, 2010
- Program Committee *Empirical Methods in Natural Language Processing*: 2014, 2013, 2012, 2011, 2008
- Program Committee *International Conference of Machine Learning*: 2014, 2013, 2012, 2011, 2010, 2009
- Program Committee *International Conference on the Web and Social Media*: 2014
- Program Committee *World Wide Web Conference*: 2014
- Program Committee *AISTATS*: 2012, 2011
- Program Committee *North American Association for Computational Linguistics*: 2012
- Program Committee *NIPS 2010 Workshop on Computational Social Science and the Wisdom of Crowds*
- Program Committee *NAACL 2010 Workshop on Creating Speech and Text Language Data With Amazon's Mechanical Turk*
- Reviewer for *COLING 2010*
- Program Committee *Global WordNet Association Conference*: 2010, 2008, 2006
- Assistant Reviewer for *UAI 2007*
- Reviewer, Works in Progress *2006 SIGCHI*

**Reviewing and Editing for Journals**

- Reviewer for *Machine Learning Journal*: 2014
- Reviewer for *Transactions of the Association of Computational Linguistics*: 2011,2012,2012,2013,2013,2013,2014
- Reviewer for *IEEE Transactions on Pattern Analysis and Machine Intelligence*: 2013
- Reviewer for *Computational Linguistics*: 2013
- Reviewer for *Scientometrics*: 2012
- Reviewer for *Information Visualization*: 2012
- Reviewer for *Transactions on Knowledge Discovery from Data*: 2011
- Reviewer for *Annals of Applied Statistics*: 2011
- Reviewer for *Journal of Machine Learning Research*: 2011,2012
- Reviewer for *Elsivier Computer Speech and Language*: 2007

**Reviewing activities for agencies**

- NSF IIS Review Panel (2015)
- NSF External Review (2014)
- NSF IIS Review Panel (2012)
- NSF BIGDATA Review Panel (2012)

**Other unpaid services to local, state, and federal agencies**

- Consultant for *Interactive Topic Modeling*, National Institute of Food and Agriculture (2013)
- Consultant for *Interactive Topic Modeling*, National Institutes for Health (2010)
- Collaboration on *Dirichlet Process Protein Clustering*, National Institues for Health (2012)
- Collaboration on *Nonparametric Beta Coalescent Clustering*, US Department of Agriculture (2013)

**Leadership in Professional Organizations**

- **Area Chair** for Machine Learning, *EMNLP 2015*
- **Area Chair**, *ICML 2015*
- **Area Chair** for Document Classification and Topic Clustering, *NAACL 2015*
- **Co-organizer** for *ACL 2014 Student Research Workshop*
- **Co-organizer** for *NIPS 2013 Workshop on Topic Models*
- Computational Committee *North American Computational Linguistics Olympiad* 2012-2014
- **Area Chair** for Document Classification and Topic Clustering, *NAACL 2012*
- **Co-organizer** for *NIPS 2009 Workshop on Topic Model Applications: Text and Beyond*

## Paid consultancies

- Consultant: Norwegian Research Council, 2014
- Consultant: Barquin International, 2013-2014
- Consultant: New Brand Analytics, 2012-2014

## Campus Service

**Departmental Service** *Computer Science, Colorado*

- Member: Graduate Committee (2014-)
- Member: Search Committee (Machine Learning, 2014-2015)

**Departmental Service** *Institute for Advanced Computer Studies*, UMD

- Member: Appointments and Promotion (2012-2013)
- Coordinator: Computational Linguistics and Information Processing Lab Colloquium (2010-2012)
- Data Czar: Computational Linguistics and Information Processling Lab (2011-2014)

**College Service** *College of information Studies*, UMD

- Chair: College of Information Studies Undergraduate Education Committee (2011-2013)
- Member: College of Information Studies Undergraduate Education Committee (2011-2014)
- Secretary: College of Information Studies Assembly (2011-2012)
- Member: College of Information Studies Programs, Courses, Curriculum Committee (2011-2013)
- Member: College of Information Studies Research Committee (2010-2013)

**College Service** *Engineering and Applied Science, Colorado*

- Yellowshirt Interviewer, 2015

**University Service** UMD

- Faculty Advisor: Maryland Academic Quiz Team (2010-2014)

## Press Coverage

- STUDY: Changes in language, tone could reveal impending betrayal (Diamondback)
- Beware Polite People (CNN)
- What You Should Know About Really Polite People (Huffington Post)
- Here's a Good Reason to Be Wary of Overly Polite People (NY Mag, Science of US)
- Detecting a Coming Betrayal (Pacific Standard)
- A few key signs betray betrayal (Science News)
- Language analysis predicts a coming betrayal (Cornell Chronicle)
- Language analysis predicts a coming betrayal (Phys.org)
- Interview with collaborator Eric Ringger (BYU Radio)
- Writeup of our quiz bowl playing robot (CU CS Newsletter)