HAL DAUMÉ III — CURRICULUM VITAE

(Last Updated: 23 March 2018)

Contact Information

University of Maryland
Computer Science Department
A.V. Williams #3227
College Park, MD 20742

phone: (301) 405-1073
fax: (301) 314-9658
email: me@hal3.name
web: http://hal3.name/
blog: http://nlpers.blogspot.com/

Employment History

Associate Professor July 2013 – Present
Computer Science Department, University of Maryland – College Park
and Language Science Center *and* Institute for Advanced Computer Studies

Assistant Professor July 2010 – June 2013
Computer Science Department, University of Maryland – College Park
and University of Maryland Institute for Advanced Computer Studies *and* Department of Linguistics

Assistant Professor July 2006 – June 2010
School of Computing, University of Utah

Research Intern Summer 2003
Machine Learning and Applied Statistics Group, Microsoft Research

Education

University of Southern California August 2006
Ph.D., Computer Science (advisor: Daniel Marcu)
Honors: Outstanding Research Assistant Award, Dean's Doctoral Merit Fellowship

University of Southern California May 2003
M.S., Computer Science

Carnegie Mellon University August 1998 – May 2001
B.S., Mathematical Sciences (advisor: Rami Grossberg)
Honors: Research Honors

Awards

- Best paper award NAACL 2016 for "Feuding Families and Former Friends: Unsupervised Learning for Dynamic Fictional Relationships"
- Best demonstration award NIPS 2015 for "Interactive Incremental Question Answering"
- Best paper award CEAS 2011 for "Using Classifier Cascades for Scalable E-Mail Classification"
- Best paper award ECML 2009 for "A geometric view of conjugate priors"

Research Interests

Natural Language Processing: semantics, language grounding, typology, translation, discourse
Machine Learning: structured prediction, domain adaptation, sequential decision making

Current Advisees

Ph.D. Students:
- Sudha Rao (expected graduation $\sim$ 2018)
- Amr Sharaf (expected graduation $\sim$ 2019)
- Kianté Brantley (expected graduation $\sim$ 2020)
- Khanh Ngyuen (expected graduation $\sim$ 2020)

Past Advisees

Postdoctoral Researchers:
- Hadi Amiri (Ph.D. National University of Singapore; Postdoc $2013 - 2015$; now Research Scientist at Harvard Medical)
- Dan Goldwasser (Ph.D. UIUC; Postdoc $2012 - 2014$; now Assistant Professor at Purdue University)
- Taesun Moon (Ph.D. UT Austin; Postdoc $2011 - 2013$; now Research Scientist at IBM Research)

Ph.D. Students:
- Piyush Rai (Ph.D. 2012; now Assistant Professor at IIT Kanpur)
- Arvind Agarwal Ph.D. 2012; now Resarch Scientist at Xerox Labs)
- Amit Goyal Ph.D. 2013; now Resarch Scientist at Yahoo!)
- Abhishek Kumar (Ph.D. 2013; now Resarch Scientist at IBM Research)
- Jagadeesh Jagaralamudi (Ph.D. 2013; now Resarch Scientist at Google)
- Jiarong Jiang (Ph.D. $\sim$ 2013; now Scientist at TwoSigma)
- He He (Ph.D. 2016; now Postdoc at Stanford)

  won the Larry S. Davis Doctoral Dissertation award 2016

  won the CMNS Board of Visitors Outstanding Graduate Student Award in 2016
- Snigdha Chaturvedi (Ph.D. 2016, now Assistant Professor at UCSC)
- Mohit Iyyer (Ph.D., 2017, now Assistant Professor at UMass Amherst)

Undergraduate and M.S. Students:
- Scott Alfeld (B.S. 2010; now Ph.D. student at UW Madison)
- Adam Teichert (M.S. 2010; now Ph.D. student at JHU)
- J. Ozzie Fallick (CS honors student $2011 - 2014$)
- Anna Gale (CS honors student $2012 - 2013$)
- Daniel Rabinovich (CS honors student $2011 - 2014$)
- Cynthia Gan (CS honors student $2014 - 2016$)
- Meir Friedenberg (CS honors student $2015 - 2017$)
- Harita Kannan (CS honors student $2016 - 2017$)
- Ephraim Rothschild (CS honors student $2016 - 2017$)

Publications

Natural language processing papers are marked with a $\circ$ and machine learning papers are marked with a $\star$. Top-tier NLP/ML conferences: ACL (8 papers), NIPS (10 papers), EMNLP (12 papers) and ICML (9 papers). In both natural language processing and machine learning, conferences are more competitive, more highly regarded, and more highly cited than journals.

Citation counts from Google scholar for papers with at least 10 citations. My six most cited papers are: [110] (1110 citations), [7] (649 citations), [85] (366 citations), [63] (332 citations), [6] (332 citations), [115] (237 citations).

Ph.D. Thesis

NLP ML

○ ⋆ [1] **Hal Daumé III**. *Practical Structured Learning Techniques for Natural Language Processing*. 2006. Ph.D. Thesis. (108 citations)

Journal Articles

NLP ML
○   [2] Amit Goyal, Ellen Riloff and **Hal Daumé III**. *A Computational Model for Plot Units*. 2013. Computational Intelligence Journal.

○   [3] Ann Irvine, John Morgan, Marine Carpuat, **Hal Daumé III** and Dragos Munteanu. *Measuring Machine Translation Errors in New Domains*. 2013. Transactions of the Association for Computational Linguistics (TACL). (18 citations)

⋆ [4] Arvind Agarwal and **Hal Daumé III**. *A geometric view of conjugate priors*. 2010. Machine Learning Journal (MLJ). (17 citations)

⋆ [5] Pu Liu, Qiang Shi, **Hal Daumé III** and Gregory Voth. *A Bayesian Statistics Approach to Multiscale Coarse Graining*. 2009. Journal of Chemical Physics (J.ChPhys). (33 citations)

○ ⋆ [6] **Hal Daumé III**, John Langford and Daniel Marcu. *Search-based Structured Prediction*. 2009. Machine Learning Journal (MLJ). (332 citations)

○ ⋆ [7] **Hal Daumé III** and Daniel Marcu. *Domain Adaptation for Statistical Classifiers*. 2006. Journal of Artificial Intelligence Research (JAIR). (649 citations)

○   [8] **Hal Daumé III** and Daniel Marcu. *Induction of Word and Phrase Alignments for Automatic Document Summarization*. 2005. Computational Linguistics (CL). (39 citations)

⋆ [9] **Hal Daumé III** and Daniel Marcu. *A Bayesian Model for Supervised Clustering with the Dirichlet Process Prior*. 2005. Journal of Machine Learning Research (JMLR).

Conference Papers

NLP ML
[10] Hoang M Le, Nan Jiang, Alekh Agarwal, Miroslav Dudk, Yisong Yue and **Hal Daumé III**. *Hierarchical Imitation and Reinforcement Learning*. 2018. arxiv.

[11] **Hal Daumé III**, John Langford and Amr Sharaf. *Residual Loss Prediction: Reinforcement Learning with no Incremental Feedback*. 2018. ICLR.

[12] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, **Hal Daumé III** and Kate Crawford. *Datasheets for Datasets*. 2018. arxiv.

[13] Amr Sharaf, Shi Feng, Khanh Nguyen, Kiant Brantley and **Hal Daumé III**. *The UMD Neural Machine Translation Systems [at WMT17 Bandit Learning Task]*. 2017. WMT.

[14] Khanh Nguyen, **Hal Daumé III** and Jordan Boyd-Graber. *Reinforcement Learning for Bandit Neural Machine Translation with Simulated Human Feedback*. 2017. Conference on Empirical Methods in Natural Language Processing (EMNLP).

⋆ [15] Akshay Krishnamurthy, Alekh Agarwal, Tzu-Kuo Huang, **Hal Daumé III** and John Langford. *Active Learning for Cost-Sensitive Classification*. 2017. International Conference on Machine Learning (ICML).

[16] Sudha Rao, Daniel Marcu, Kevin Knight and **Hal Daumé III**. *Biomedical event extraction using abstract meaning representation*. 2017. BioNLP.

[17] Emily Bender, **Hal Daumé III**, Allyson Ettinger and Sudha Rao. *Proceedings of the First Workshop on Building Linguistically Generalizable NLP Systems*. 2017. Conference of the Association for Computational Linguistics (ACL).

[18] Mohit Iyyer, Varun Manjunatha, Anupam Guha, Yogarshi Vyas, Jordan Boyd-Graber, **Hal Daumé III** and Larry Davis. *The amazing mysteries of the gutter: Drawing inferences between panels in comic book narratives*. 2017. Computer Vision and Pattern Recognition (CVPR). (10 citations)

○ [19] Snigdha Chaturvedi, Mohit Iyyer and **Hal Daumé III**. *Unsupervised Learning of Evolving Relationships Between Literary Characters*. 2017. National Conference on Artificial Intelligence (AAAI).

⋆ [20] **Hal Daumé III**, Nikos Karampatziakis, John Langford and Paul Mineiro. *Logarithmic time one-against-some*. 2017. International Conference on Machine Learning (ICML). (11 citations)

⋆ [21] Kai-Wei Chang, He He, Stéphane Ross, **Hal Daumé III** and John Langford. *A Credit Assignment Compiler for Joint Prediction*. 2016. Advances in Neural Information Processing Systems (NIPS).

○ [22] Snigdha Chaturvedi, Shashank Srivastava, **Hal Daumé III** and Chris Dyer. *Modeling Evolving Relationships Between Characters in Literary Novels*. 2016. National Conference on Artificial Intelligence (AAAI).

[23] Mohit Iyyer, Varun Manjunatha, Anupam Guha, Yogarshi Vyas, Jordan Boyd-Graber, **Hal Daumé III** and Larry Davis. *The Amazing Mysteries of the Gutter: Drawing Inferences Between Panels in Comic Book Narratives*. 2016. arXiv Preprint. (10 citations)

○ [24] He He, Jordan Boyd-Graber and **Hal Daumé III**. *Interpretese vs. Translationese: The Uniqueness of Human Strategies in Simultaneous Interpretation*. 2016. Conference of the North American Chapter of the Association for Computational Linguistics (NAACL).

○ [25] Hadi Amiri and **Hal Daumé III**. *Short Text Representation for Detecting Churn in Microblogs*. 2016. National Conference on Artificial Intelligence (AAAI).

○ [26] Mohit Iyyer, Anupam Guha, Snigdha Chaturvedi, Jordan Boyd-Graber and **Hal Daumé III**. *Feuding Families and Former Friends: Unsupervised Learning for Dynamic Fictional Relationships*. 2016. Conference of the North American Chapter of the Association for Computational Linguistics (NAACL). (24 citations)

○ [27] Fabienne Braune, Alexander Fraser, **Hal Daumé III** and Aleš Tamchyna. *A Framework for Discriminative Rule Selection in Hierarchical Moses*. 2016. WMT.

[28] Vicente Ordóñez, Xufeng Han, Polina Kuznetsova, Girish Kulkarni, Margaret Mitchell, Kota Yamaguchi, Karl Stratos, Amit Goyal, Jesse Dodge, Alyssa Mensch, **Hal Daumé III**, Alexander C. Berg, Yejin Choi and Tamara L. Berg. *Large scale retrieval and generation of image descriptions*. 2016. IJCV.

⋆ [29] **Hal Daumé III**, Nikos Karampatziakis, John Langford and Paul Mineiro. *Logarithmic Time One-Against-Some*. 2016. arXiv preprint. (11 citations)

○ [30] Hadi Amiri, Philip Resnik, Jordan Boyd-Graber and **Hal Daumé III**. *Learning Text Pair Similarity with Context-sensitive Autoencoders*. 2016. Conference of the Association for Computational Linguistics (ACL).

○ [31] Kathy McKeown, **Hal Daumé III**, Snigdha Chaturvedi, John Paparrizos, Kapil Thadani, Pablo Barrio, Or Biran, Suvarna Bothe, Michael Collins, Kenneth R Fleischmann, Luis Gravano, Rahul Jha, Ben King, Kevin McInerney, Taesun Moon, Arvind Neelakantan, Diarmuid O'Seaghdha, Dragomir Radev, Clay Templeton and Simone Teufel. *Predicting the impact of scientific concepts using fulltext features*. 2016. JAIST. (14 citations)

⋆ [32] He He, Jordan Boyd-Graber, Kevin Kwok and **Hal Daumé III**. *Opponent Modeling in Deep Reinforcement Learning*. 2016. International Conference on Machine Learning (ICML). (20 citations)

○ [33] Snigdha Chaturvedi, Dan Goldwasser and **Hal Daumé III**. *Ask, and Shall You Receive? Understanding Desire Fulfillment in Natural Language Text*. 2016. National Conference on Artificial Intelligence (AAAI).

[34] Dong Xu, Rama Chellappa, Trevor Darrell and **Hal Daumé III**. *Guest Editors Introduction to the Special Issue on Domain Adaptation for Vision Applications*. 2015. IJCV.

[35] Sudha Rao, Allyson Ettinger, **Hal Daumé III** and Philip Resnik. *Dialogue focus tracking for zero pronoun resolution*. 2015. Conference of the North American Chapter of the Association for Computational Linguistics (NAACL).

○ ⋆ [36] Mohit Iyyer, Varun Manjunatha, Jordan Boyd-Graber and **Hal Daumé III**. *Deep unordered composition rivals syntactic methods for text classification*. 2015. Conference of the Association for Computational Linguistics (ACL). (150 citations)

⋆ [37] **Hal Daumé III**, Samir Khuller, Manish Purohit and Gregory Sanders. *On Correcting inputs: Inverse Optimization for Online Structured Prediction*. 2015. FSTTCS.

⋆ [38] Kai-Wei Chang, Akshay Krishnamurthy, Alekh Agarwal, **Hal Daumé III** and John Langford. *Learning to search better than your teacher*. 2015. International Conference on Machine Learning (ICML).

⋆ [39] Alina Beygelzimer, **Hal Daumé III**, John Langford and Paul Mineiro. *Learning Reductions that Really Work*. 2015. IEEE Proceedings.

○ [40] Naho Orita, Eliana Vornov, Naomi H Feldman and **Hal Daumé III**. *Why discourse affects speakers choice of referring expressions*. 2015. Conference of the Association for Computational Linguistics (ACL).

○ [41] He He, Alvin Grissom II, Jordan Boyd-Graber and **Hal Daumé III**. *Syntax-based Rewriting for Simultaneous Machine Translation*. 2015. Conference on Empirical Methods in Natural Language Processing (EMNLP).

⋆ [42] Kai-Wei Chang, **Hal Daumé III**, John Langford and Stéphane Ross. *Efficient programmable learning to search*. 2014. arXiv.

○ [43] Snigdha Chaturvedi, Dan Goldwasser and **Hal Daumé III**. *Predicting Instructor Intervention in MOOC Forums*. 2014. Conference of the Association for Computational Linguistics (ACL). (34 citations)

○ [44] Dan Goldwasser and **Hal Daumé III**. *"I Object!" Modeling Latent Pragmatic Effects in Courtroom Dialogues*. 2014. Conference of the European Association for Computational Linguistics (EACL).

⋆ [45] Arti Ramesh, Dan Goldwasser, Bert Huang, **Hal Daumé III** and Lise Getoor. *Learning Latent Engagement Patterns of Students in Online Courses*. 2014. National Conference on Artificial Intelligence (AAAI). (50 citations)

⋆ [46] Arti Ramesh, Dan Goldwasser, Bert Huang, **Hal Daumé III** and Lise Getoor. *Uncovering Hidden Engagement Patterns for Predicting Learner Performance in MOOCs*. 2014. Learning at Scale. (25 citations)

○ [47] Aleš Tamchyna, Fabienne Braune, Alexander Fraser, Marine Carpuat, **Hal Daumé III** and Chris Quirk. *Integrating a Discriminative Classifier into Phrase-based and Hierarchical Decoding*. 2014. The Prague Bulletin of Mathematical Linguistics.

○ [48] Alvin Grissom II, Jordan Boyd-Graber, He He, John Morgan and **Hal Daumé III**. *Don't Until the Final Verb Wait: Reinforcement Learning for Simultaneous Machine Translation*. 2014. Conference on Empirical Methods in Natural Language Processing (EMNLP). (18 citations)

○ ⋆ [49] Mohit Iyyer, Jordan Boyd-Graber, Leonardo Claudino, Richard Socher and **Hal Daumé III**. *A Neural Network for Factoid Question Answering over Paragraphs*. 2014. Conference on Empirical Methods in Natural Language Processing (EMNLP). (181 citations)

○ ⋆ [50] He He, **Hal Daumé III** and Jason Eisner. *Dynamic Feature Selection for Dependency Parsing*. 2013. Conference on Empirical Methods in Natural Language Processing (EMNLP). (34 citations)

⋆ [51] Yuancheng Luo, Dmitry N. Zotkin, **Hal Daumé III** and Ramani Duraiswami. *Kernel Regression for Head-Related Transfer Function Interpolation and Spectral Extrema Extraction*. 2013. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).

○ [52] Ann Irvine, Chris Quirk and **Hal Daumé III**. *Monolingual Marginal Matching for Translation Model Adaptation*. 2013. Conference on Empirical Methods in Natural Language Processing (EMNLP). (13 citations)

⋆ [53] Yuening Hu, Jordan Boyd-Graber, **Hal Daumé III** and Z. Irene Ying. *Binary to Bushy: Bayesian Hierarchical Clustering with the Beta Coalescent*. 2013. Advances in Neural Information Processing Systems (NIPS).

○ [54] Jeffrey Ferraro, **Hal Daumé III**, Scott DuVall, Wendy Chapman, Henk Harkema and Peter Haug. *Improving performance of natural language processing part-of-speech tagging on clinical narratives through domain adaptation*. 2013. Jornal of the American Medical Informatics Association. (31 citations)

⋆ [55] Mohammad Rastegari, Jonghyun Choi, Shobeir Fakhraei, **Hal Daumé III** and Larry S. Davis. *Predictable Dual-View Hashing*. 2013. International Conference on Machine Learning (ICML).

○ [56] Junhui Li, Philip Resnik and **Hal Daumé III**. *Modeling Syntactic and Semantic Structures in Hierarchical Phrase-based Translation*. 2013. Conference of the North American Chapter of the Association for Computational Linguistics (NAACL). (20 citations)

○ ⋆ [57] Snigdha Chaturvedi, **Hal Daumé III** and Taesun Moon. *Discriminatively Enhanced Topic Models*. 2013. International Conference on Data Mining (ICDM).

○ [58] Marine Carpuat, **Hal Daumé III**, Katharine Henry, Ann Irvine, Jagadeesh Jagarlamudi and Rachel Rudinger. *SenseSpotting: Never let your parallel data tie you to an old domain*. 2013. Conference of the Association for Computational Linguistics (ACL).

○ [59] Margaret Mitchell, Jesse Dodge, Amit Goyal, Kota Yamaguchi, Karl Stratos, Xufeng Han, Alyssa Mensch, Alexander C. Berg, Tamara L. Berg and **Hal Daumé III**. *Midge: Generating Image Descriptions From Computer Vision Detections*. 2012. European Chapter of the Association for Computational Linguistics (EACL).

⋆ [60] Alexandre Passos, Piyush Rai, Jacques Wainer and **Hal Daumé III**. *Flexible Modeling of Latent Task Structures in Multitask Learning*. 2012. International Conference on Machine Learning (ICML). (33 citations)

[61] He He, **Hal Daumé III** and Jason Eisner. *Imitation Learning by Coaching*. 2012. Neural Information Processing Systems (NIPS). (49 citations)

○ [62] Ching Lik Teo, Yezhou Yang, **Hal Daumé III**, Cornelia Fermüller and Yiannis Aloimonos. *Towards a Watson That Sees: Language-Guided Action Recognition for Robots*. 2012. ICRA. (22 citations)

[63] Abhishek Sharma, Abhishek Kumar, **Hal Daumé III** and David Jacobs. *Generalized Multiview Analysis: A Discriminative latent space*. 2012. Computer Vision and Pattern Recognition (CVPR). (344 citations)

⋆ [64] **Hal Daumé III**, Jeff M. Phillips, Avishek Saha and Suresh Venkatasubramanian. *Efficient Protocols for Distributed Classification and Optimization*. 2012. ALT. (29 citations)

⋆ [65] Abhishek Kumar and **Hal Daumé III**. *Learning Task Grouping and Overlap in Multi-task Learning*. 2012. International Conference on Machine Learning (ICML). (224 citations)

○ [66] Amit Goyal, **Hal Daumé III** and Graham Cormode. *Sketch Algorithms for Estimating Point Queries in NLP*. 2012. Empirical Methods in Natural Language Processing (EMNLP). (23 citations)

⋆ [67] Piyush Rai, Abhishek Kumar and **Hal Daumé III**. *Simultaneously Leveraging Output and Task Structures for Multiple-Output Regression*. 2012. Advances in Neural Information Processing Systems (NIPS). (29 citations)

○ [68] Karl Stratos, Aneesh Sood, Alyssa Mensch, Xufeng Han, Margaret Mitchell, Kota Yamaguchi, Jesse Dodge, Amit Goyal, **Hal Daumé III**, Alexander C. Berg and Tamara L. Berg. *Understanding and Predicting Importance in Images*. 2012. Computer Vision and Pattern Recognition (CVPR). (94 citations)

[69] Jagadeesh Jagarlamudi and **Hal Daumé III**. *Regularized Interlingual Projections: Evaluation on Multilingual Transliteration*. 2012. 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. (10 citations)

○ ⋆ [70] Jagadeesh Jagarlamudi and **Hal Daumé III**. *Low-dimensional Discriminative Reranking*. 2012. Conference on North American Chapter of the Association for Computational Linguistics.

○ [71] Amit Goyal, **Hal Daumé III** and Raul Guerra. *Fast Large-Scale Approximate Graph Construction for NLP*. 2012. Empirical Methods in Natural Language Processing (EMNLP). (12 citations)

○ ⋆ [72] Jordan Boyd-Graber, Brianna Satinoff, He He and **Hal Daumé III**. *Besting the quiz master: crowdsourcing incremental classification games*. 2012. Conference on Empirical Methods in Natural Language Processing (EMNLP). (27 citations)

○ [73] Jagadeesh Jagarlamudi, **Hal Daumé III** and Raghavendra Udupa. *Incorporating Lexical Priors into Topic Models*. 2012. Conference on European Chapter of the Association for Computational Linguistics (EACL). (92 citations)

○ ⋆ [74] Jiarong Jiang, Adam Teichert, **Hal Daumé III** and Jason Eisner. *Learned Prioritization for Trading Off Accuracy and Speed*. 2012. Advances in Neural Information Processing Systems (NIPS). (25 citations)

○ [75] Jesse Dodge, Amit Goyal, Xufeng Han, Alyssa Mensch, Margaret Mitchell, Karl Stratos, Kota Yamaguchi, Yejin Choi, **Hal Daumé III**, Alexander C. Berg and Tamara L. Berg. *Detecting Visual Text*. 2012. North American Chapter of the Association for Computational Linguistics (NAACL).

⋆ [76] Abhishek Kumar, Alexandru Niculescu-Mizil, Koray Kavukcuoglu and **Hal Daumé III**. *A Binary Classification Framework for Two-Stage Multiple Kernel Learning*. 2012. International Conference on Machine Learning (ICML). (35 citations)

⋆ [77] Arvind Agarwal and **Hal Daumé III**. *Generative Kernels for Exponential Families*. 2011. Conference on Artificial Intelligence and Statistics (AI-Stats). (13 citations)

⋆ [78] Abhishek Kumar and **Hal Daumé III**. *A Co-training Approach for Multiview Spectral Clustering*. 2011. International Conference on Machine Learning (ICML). (302 citations)

⋆ [79] Jiarong Jiang, Piyush Rai and **Hal Daumé III**. *Message-Passing for Approximate MAP Inference with Latent Variables*. 2011. Conference on Neural Information Processing Systems (NIPS). (19 citations)

○ [80] Jagadeesh Jagarlamudi, **Hal Daumé III** and Raghavendra Udupa. *From Bilingual Dictionaries to Interlingual Document Representations*. 2011. Association for Computational Linguistics (ACL). (12 citations)

⋆ [81] Jay Pujara, **Hal Daumé III** and Lise Getoor. *Using Classifier Cascades for Scalable E-Mail Classification*. 2011. CEAS. (19 citations)

⋆ [82] Anusua Trivedi, Piyush Rai, **Hal Daumé III** and Scott L. DuVall. *Leveraging Social Bookmarks from Partially Tagged Corpus for Improved Webpage Clustering*. 2011. ACM Transactions on Intelligent Systems and Technology. (10 citations)

○ [83] Yezhou Yang, Ching Lik Teo, **Hal Daumé III** and Yiannis Aloimonos. *Corpus-Guided Sentence Generation of Natural Images*. 2011. Empirical Methods in Natural Language Processing (EMNLP). (203 citations)

○ [84] Jagadeesh Jagarlamudi, Raghavendra Udupa, **Hal Daumé III** and Abhijit Bhole. *Improving Bilingual Projections via Sparse Covariance Matrices*. 2011. Empirical Methods in Natural Language Processing (EMNLP). (12 citations)

⋆ [85] Abhishek Kumar, Piyush Rai and **Hal Daumé III**. *Co-regularized Multi-view Spectral Clustering*. 2011. Conference on Neural Information Processing Systems (NIPS). (366 citations)

○ [86] Amit Goyal and **Hal Daumé III**. *Lossy Conservative Update (LCU) sketch: Succinct approximate count storage*. 2011. Conference on Artificial Intelligence (AAAI). (10 citations)

[87] Avishek Saha, Piyush Rai, **Hal Daumé III**, Suresh Venkatasubramanian and Scott L. Du-Vall. *Active Supervised Domain Adaptation*. 2011. European Conference on Machine Learning (ECML). (34 citations)

⋆ [88] Piyush Rai and **Hal Daumé III**. *Beam Search based MAP Estimates for the Indian Buffet Process*. 2011. International Conference on Machine Learning (ICML).

○ [89] **Hal Daumé III** and Jagadeesh Jagarlamudi. *Domain Adaptation for Machine Translation by Mining Unseen Words*. 2011. Association for Computational Linguistics. (104 citations)

○ [90] Amit Goyal and **Hal Daumé III**. *Approximate Scalable Bounded Space Sketch for Large Data NLP*. 2011. Empirical Methods in Natural Language Processing (EMNLP). (20 citations)

⋆ [91] Avishek Saha, Piyush Rai, **Hal Daumé III** and Suresh Venkatasubramanian. *Online Learning of Multiple Tasks and Their Relationships*. 2011. Conference on Artificial Intelligence and Statistics (AI-Stats). (72 citations)

⋆ [92] Arvind Agarwal, Samuel Gerber and **Hal Daumé III**. *Learning Multiple Tasks using Manifold Regularization*. 2010. Conference on Neural Information Processing Systems (NIPS). (71 citations)

○ ⋆ [93] Jagadeesh Jagarlamudi and **Hal Daumé III**. *Extracting Multilingual Topics from Unaligned Corpora*. 2010. European Conference on Information Retrieval (ECIR). (67 citations)

○ [94] Amit Goyal, Ellen Riloff and **Hal Daumé III**. *Automatically Producing Plot Unit Representations for Narrative Text*. 2010. Empirical Methods in Natural Language Processing (EMNLP). (56 citations)

⋆ [95] Piyush Rai and **Hal Daumé III**. *Infinite Predictor Subspace Models for Multitask Learning*. 2010. Conference on Artificial Intelligence and Statistics (AI-Stats). (59 citations)

○ ⋆ [96] Jagadeesh Jagarlamudi, Seth Juarez and **Hal Daumé III**. *Kernelized Sorting for Natural Language Processing*. 2010. Conference on Artificial Intelligence (AAAI). (13 citations)

⋆ [97] Abhishek Kumar, Avishek Saha and **Hal Daumé III**. *A Co-regularization Based Semi-supervised Domain Adaptation*. 2010. Conference on Neural Information Processing Systems (NIPS). (108 citations)

⋆ [98] Arvind Agarwal and **Hal Daumé III**. *Exponential Family Hybrid Semi-Supervised Learning*. 2009. International Joint Conference on Artificial Intelligence (IJCAI). (12 citations)

○ ⋆ [99] **Hal Daumé III**. *Unsupervised Search-based Structured Prediction*. 2009. International Conference on Machine Learning (ICML). (36 citations)

⋆[100] **Hal Daumé III**. *Bayesian Multitask Learning with Latent Hierarchies*. 2009. Conference on Uncertainty in Artificial Intelligence (UAI). (93 citations)

○ [101] **Hal Daumé III**. *Non-Parametric Bayesian Model Areal Linguistics*. 2009. North American Chapter of the Association for Computational Linguistics (NAACL). (40 citations)

⋆[102] Piyush Rai, **Hal Daumé III** and Suresh Venkatasubramanian. *Streamed Learning: One-Pass SVMs*. 2009. International Joint Conference on Artificial Intelligence (IJCAI). (25 citations)

○ ⋆[103] **Hal Daumé III**. *Markov Random Topic Fields*. 2009. Association for Computational Linguistics (ACL). (23 citations)

○ [104] Amit Goyal, **Hal Daumé III** and Suresh Venkatasubramanian. *Streaming for Large Scale NLP: Language Modeling*. 2009. North American Chapter of the Association for Computational Linguistics (NAACL). (37 citations)

⋆[105] Piyush Rai and **Hal Daumé III**. *Multi-Label Prediction via Sparse Infinite CCA*. 2009. Conference on Neural Information Processing Systems (NIPS). (66 citations)

⋆[106] Piyush Rai and **Hal Daumé III**. *The Infinite Hierarchical Factor Regression Model*. 2008. Conference on Neural Information Processing Systems (NIPS). (63 citations)

○ [107] Ulf Hermjakob, Kevin Knight and **Hal Daumé III**. *Name Translation in Statistical Machine Translation: Learning When to Transliterate*. 2008. Conference of the Association for Computational Linguistics (ACL). (97 citations)

○ ⋆[108] Percy Liang, **Hal Daumé III** and Dan Klein. *Structure Compilation: Trading Structure for Features*. 2008. International Conference on Machine Learning (ICML).

○ ⋆[109] **Hal Daumé III**. *Cross-Task Knowledge-Constrained Self Training*. 2008. Empirical Methods in Natural Language Processing (EMNLP). (19 citations)

○ ⋆[110] **Hal Daumé III**. *Frustratingly Easy Domain Adaptation*. 2007. Conference of the Association for Computational Linguistics (ACL). (1110 citations)

○ [111] **Hal Daumé III** and Lyle Campbell. *A Bayesian Model for Discovering Typological Implications*. 2007. Conference of the Association for Computational Linguistics (ACL). (34 citations)

⋆[112] **Hal Daumé III**. *Fast search for Dirichlet process mixture models*. 2007. Eleventh International Conference on Artificial Intelligence and Statistics (AIStats). (51 citations)

⋆[113] Yee Whye Teh, **Hal Daumé III** and Daniel Roy. *Bayesian Agglomerative Clustering with Coalescents*. 2007. Conference on Neural Information Processing Systems (NIPS). (93 citations)

○ [114] **Hal Daumé III** and Daniel Marcu. *Bayesian Query-Focused Summarization*. 2006. Conference of the Association for Computational Linguistics (ACL). (203 citations)

⋆[115] **Hal Daumé III** and Daniel Marcu. *Learning as Search Optimization: Approximate Large Margin Methods for Structured Prediction*. 2005. International Conference on Machine Learning (ICML). (237 citations)

○ [116] **Hal Daumé III** and Daniel Marcu. *A Large-Scale Exploration of Effective Global Features for a Joint Entity Detection and Tracking Model*. 2005. Joint Conference on Human Language Technology and Empirical Methods in Natural Language Processing (HLT/EMNLP). (84 citations)

○ [117] **Hal Daumé III** and Daniel Marcu. *NP Bracketing by Maximum Entropy Tagging and SVM Reranking*. 2004. Empirical Methods in Natural Language Processing. (16 citations)

○ [118] **Hal Daumé III** and Eric Brill. *Web Search Intent Induction via Automatic Query Reformulation*. 2004. North American Chapter of the Association for Computational Linguistics (NAACL). (30 citations)

○ [119] **Hal Daumé III** and Daniel Marcu. *A Phrase-Based HMM Approach to Document/Abstract Alignment*. 2004. Empirical Methods in Natural Language Processing (EMNLP). (33 citations)

○ [120] **Hal Daumé III**, Kevin Knight, Irene Langkilde-Geary, Daniel Marcu and Kenji Yamada. *The Importance of Lexicalized Syntax Models for Natural Language Generation Tasks*. 2002. 2002 International Conference on Natural Language Generation (INLG).

○ [121] **Hal Daumé III** and Daniel Marcu. *A Noisy-Channel Model for Document Compression*. 2002. 40th Annual Meeting of the Association for Computational Linguistics (ACL). (85 citations)

○ [122] Eric Nyberg and **Hal Daumé III**. *Integrated Information Management: An Interactive, Extensible Architecture for Information Retrieval*. 2001. 2001 Human Language Technology Conference (HLT).

Reviews, Commentaries and Invited Papers

NLP ML
[123] Roger Levy and **Hal Daumé III**. *Computational methods are invaluable for typology, but the models must match the questions: Commentary on Dunn et al. (2011)*. 2011. Unpublished.

○ ⋆[124] **Hal Daumé III**. *Semi-supervised or Semi-unsupervised?*. 2009. Unpublished. (10 citations)

[125] **Hal Daumé III**. *Book Review: Automatic Summarization (I. Mani)*. 2004. Unpublished.

Workshop Papers

NLP ML
[126] Amr Sharaf and **Hal Daumé III**. *Structured prediction via learning to search under bandit feedback*. 2017. Workshop on Structured Prediction for NLP.

○ [127] Arti Ramesh, Dan Goldwasser, Bert Huang, **Hal Daumé III** and Lise Getoor. *Modeling Learner Engagement in MOOCs using Probabilistic Soft Logic*. 2013. NIPS Workshop on Data Driven Education. (77 citations)

⋆[128] Snigdha Chaturvedi, **Hal Daumé III**, Taesun Moon and Shashank Srivastava. *A Topical Graph Kernel for Link Prediction in Labeled Graphs*. 2013. ICML workshop on Mining and Learning with Graphs (MLG).

○ ⋆[129] Dan Goldwasser and **Hal Daumé III**. *Predicting Dialogue Outcomes over Structured Latent Representations*. 2013. NIPS Workshop on Output Representation Learning.

⋆[130] Jiarong Jiang, Taesun Moon, **Hal Daumé III** and Jason Eisner. *Prioritized Asynchronous Belief Propagation*. 2013. ICML Workshop on Inferning.

⋆[131] **Hal Daumé III**, Jeff Phillips, Avishek Saha and Suresh Venkatasubramanian. *Protocols for Learning Classifiers on Distributed Data*. 2012. Workshop on Artificial Intelligence and Statistics (AI-Stats). (24 citations)

⋆[132] He He, **Hal Daumé III** and Jason Eisner. *Cost-sensitive Dynamic Feature Selection*. 2012. ICML 2012 Workshop on Interactions between Inference and Learning (Inferning). (26 citations)

○ [133] Ching L. Teo, Yezhou Yang, **Hal Daumé III**, Cornelia Fermüller and Yiannis Aloimonos. *A Corpus-Guided Framework for Robotic Visual Perception*. 2011. AAAI Workshop on Language-Action Tools for Cognitive Artificial Agents.

○ [134] Amit Goyal and **Hal Daumé III**. *Generating Semantic Orientation Lexicon using Large Data and Thesaurus*. 2011. ACL Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (WASSA).

⋆[135]   Amit Goyal, Piyush Rai and **Hal Daumé III**. *Multiple Hash Functions for Learning*. 2011. NIPS Big Learning Workshop.

○   ⋆[136]   Jason Eisner and **Hal Daumé III**. *Speed-Accuracy Tradeoffs in Nondeterministic Inference Algorithms*. 2011. COST: NIPS 2011 Workshop on Computational Trade-offs in Statistical Learning.

○   [137]   Amit Goyal, Jagadeesh Jagarlamudi, **Hal Daumé III** and Suresh Venkatasubramanian. *Sketch Techniques for Scaling Distributional Similarity to the Web*. 2010. GEometrical Models of Natural Language Semantics Workshop (GEMS) at ACL. (15 citations)

⋆[138]   Piyush Rai, Avishek Saha, **Hal Daumé III** and Suresh Venkatasubramanian. *Domain Adaptation meets Active Learning*. 2010. HLT/NAACL Workshop on Active Learning for NLP (ALNLP). (69 citations)

○   ⋆[139]   **Hal Daumé III**, Abhishek Kumar and Avishek Saha. *Frustratingly Easy Semi-Supervised Domain Adaptation*. 2010. Workshop on Domain Adaptation for NLP. (118 citations)

⋆[140]   Avishek Saha, Piyush Rai, **Hal Daumé III** and Suresh Venkatasubramanian. *Active Online Multitask Learning*. 2010. ICML 2010 Workshop on Budgeted Learning (Budget).

○   [141]   Amit Goyal, Ellen Riloff, **Hal Daumé III** and Nathan Gilbert. *Toward Plot Units: Automatic Affect State Analysis*. 2010. HLT/NAACL Workshop on Computational Approaches to Analysis and Generation of Emotion in Text (CAET). (14 citations)

⋆[142]   Piyush Rai, Anusua Trivedi, **Hal Daumé III** and Scott L. DuVall. *Multiview Clustering with Incomplete Views*. 2010. NIPS Workshop on Machine Learning for Social Computing. (20 citations)

⋆[143]   Piyush Rai and **Hal Daumé III**. *Multitask Learning via Mixture of Linear Subspaces*. 2010. NIPS Workshop on Transfer Learning by Learning Rich Generative Models.

⋆[144]   Abhishek Kumar, Piyush Rai and **Hal Daumé III**. *Co-regularized Spectral Clustering with Multiple Kernels*. 2010. NIPS Workshop on New Directions in Multiple Kernel Learning. (12 citations)

⋆[145]   Anusua Trivedi, Piyush Rai, Scott L. DuVall and **Hal Daumé III**. *Exploiting Tag and Word Correlations for Improved Webpage Clustering*. 2010. CIKM Workshop on Search and Mining User-generated Contents (SMUC). (13 citations)

○   [146]   Amit Goyal, Jagadeesh Jagarlamudi, **Hal Daumé III** and Suresh Venkatasubramanian. *Sketching Techniques for Large Scale NLP*. 2010. HLT/NAACL Workshop on the Web as a Corpus (WAC). (17 citations)

⋆[147]   Piyush Rai and **Hal Daumé III**. *Fast Search for Infinite Latent Feature Models*. 2009. NIPS Workshop on Non-parametric Bayes (NP-Bayes).

⋆[148]   Piyush Rai and **Hal Daumé III**. *Multitask Learning using Nonparametrically Learned Predictor Subspaces*. 2009. NIPS Workshop on Learning from Multiple Sources.

○   [149]   Adam R. Teichert and **Hal Daumé III**. *Unsupervised Part of Speech Tagging Without a Lexicon*. 2009. NIPS Workshop on Grammar Induction, Representation of Language and Language Learning (GIRLLL).

⋆[150]   Amrish Kapoor, Piyush Rai and **Hal Daumé III**. *Factor Regression Combining Heterogeneous Sources of Information*. 2009. NIPS Workshop on Learning From Multiple Sources with Applications to Robotics (LMS).

⋆[151] Devyani Ghosh, John Carter and **Hal Daumé III**. *Perceptron-based Coherence Predictors*. 2008. 2nd Workshop on Chip Multiprocessor Memory Systems and Interconnects (ICSA).

○ [152] **Hal Daumé III** and Daniel Marcu. *Bayesian Summarization at DUC and a Suggestion for Extrinsic Evaluation*. 2005. Document Understanding Conference (DUC). (17 citations)

○ [153] **Hal Daumé III** and Daniel Marcu. *Bayesian Multi-Document Summarization at MSE*. 2005. Workshop on Multilingual Summarization Evaluation (MSE). (35 citations)

⋆[154] **Hal Daumé III**, John Langford and Daniel Marcu. *Search-Based Structured Prediction as Classification*. 2005. NIPS Workshop on Advances in Structured Learning for Text and Speech Processing (ASLTSP).

○ ⋆[155] **Hal Daumé III** and Daniel Marcu. *A Tree-Position Kernel for Document Compression*. 2004. Fourth Document Understanding Conference (DUC). (16 citations)

○ [156] **Hal Daumé III** and Daniel Marcu. *Generic Sentence Fusion is an Ill-Defined Summarization Task*. 2004. Text Summarization Branches Out Workshop at ACL (TextSum). (24 citations)

⋆[157] **Hal Daumé III** and Daniel Marcu. *Supervised clustering with the Dirichlet process*. 2004. NIPS Workshop on Learning With Structured Outputs (LwSO).

○ [158] **Hal Daumé III**, Abdesammad Echihabi, Daniel Marcu, Dragos Stefan Munteanu and Radu Soricut. *GLEANS: A Generator of Logical Extracts and Abstracts for Nice Summaries*. 2002. Second Document Understanding Conference (DUC). (23 citations)

Unpublished

NLP ML

○ ⋆[159] **Hal Daumé III**, John Langford and Daniel Marcu. *Searn in Practice*. 2006. Unpublished.

⋆[160] **Hal Daumé III**. *From Zero to Reproducing Kernel Hilbert Spaces in Twelve Pages or Less*. 2004. Unpublished. (20 citations)

⋆[161] **Hal Daumé III**. *Carefully Approximated Bayes Factors for Feature Selection in MaxEnt Models*. 2004. Unpublished.

⋆[162] **Hal Daumé III**. *Notes on CG and LM-BFGS Optimization of Logistic Regression*. 2004. Unpublished. (148 citations)

[163] **Hal Daumé III**. *Yet Another Haskell Tutorial*. 2002. Unpublished. (32 citations)

○ [164] **Hal Daumé III**. *A Phrase-Based HMM*. 2002. Unpublished.

[165] **Hal Daumé III**. *Asymmetry of Coordination*. 2001. Unpublished.

Publicly Available Software

VW: Significant contributer to vowpal wabbit online training system, originally developed by John Langford: contributions include multiclass classification, structured learning, etc.

BRAQUE: Developer (with Percy Liang) of http://braque.cc: the news-for-researchers web site.

MEGAM: Optimization software for maximum entropy models; implementation based on conjugate gradient and limited-memory BFGS methods. (Downloaded over 20,000 times.)

http://hal3.name/megam/

HBC: Software that compiles statistical models into executable C code. (Downloaded over 6,000 times.)

http://hal3.name/HBC/.

TAGCHUNK: Software for performing joint inference over part of speech tags and syntactic chunks. (Downloaded over 400 times.)

http://hal3.name/TagChunk/

SVMSEQUEL: Optimization software for support vector machines; implementation based on the kernel minover method. Supports linear, polynomial and RBF kernels, as well as string and tree kernels. (Downloaded over 900 times.)

http://hal3.name/SVMsequel/

FASTDT: Software for learning decision trees, supporting boosting and bagging. (Downloaded over 750 times.)

http://hal3.name/FastDT/

. . . many more, see http://hal3.name/software.html.

Teaching Experience

CMSC 421, Artificial Intelligence UMD, CS
Fall 2016 (enrollment: 90)
Spring 2012 (enrollment: 40)

CMSC 422, Undergraduate Machine Learning UMD, CS
Fall 2015 (enrollment: 90)
Spring 2013 (enrollment: 44)

CMSC 723, Computational Linguistics I UMD, CS
Fall 2012 (enrollment: 51)
Fall 2010 (enrollment: 43)

CMSC 726, Machine Learning UMD, CS
Spring 2016 (enrollment: 45)
Fall 2011 (enrollment: 47)

CMSC 828, Reinforcement Learning UMD, CS
Fall 2016 (enrollment: 35)

CMSC 828, Advanced Topics in Natural Language Processing UMD, CS
Fall 2011 (enrollment: 10)

CS 5300/6300, Artificial Intelligence U Utah, CS
Spring 2009, Spring 2010 (enrollments: 78, 74)
Received Dean's letter for Excellence in Teaching

CS 5350/6350, Machine Learning U Utah, CS
Spring 2007, Spring 2008, Fall 2008, Fall 2009 (enrollments: 28, 32, 36, 46)

CS 5964/6964, Applications of Natural Language Processing U Utah, CS
Fall 2007 (enrollment: 28)

CS 7941, Topics in Machine Learning U Utah, CS
Every semester, 2007–current (enrollments: about 15)
Received Dean's letter for Excellence in Teaching

Various Teaching as a Student   USC CS and CMU Math
USC CSCI 544, Natural Language Processing (Teaching Assistant) (Spring 2005), Outstanding TA Award
USC CSCI 544, Natural Language Processing (Invited Lecturer) (Spring 2004)
USC CSCI 599, Topics in Statistical Learning (Invited Lecturer) (Spring 2004)
USC CSCI 562, Empirical Methods in NLP (Invited Lecturer) (Fall 2003)
CMU MATH 229, Set Theory (Grader) (Spring 2000)
CMU MATH 115/116, Calculus I & II (Teaching Assistant) (Fall 1999)

Invited Talks

Learning language through interaction   December 2016
Georgetown, Amazon, USC, GATech, NYU, ...

Bias in AI   November 2016
UMD MCWIC Diversity Summit

Imitation learning and recurrent neural networks mashup   December 2015
CIFAR NCAP Workshop

Algorithms that learn to think on their feet   October 2015
UC Santa Cruz, Computer Science Seminar

Interpretese vs Translationese   August 2015
Shonan Workshop

Algorithms that learn to think on their feet   May 2015
Invited Talk at ICLR

A picture is worth 13.6 words (on average)   February 2015
MSR NY Tea Talk

Domain adaptation: the problem of new labels   December 2014
NIPS workshop on Transfer and Multitask Learning

Algorithms that learn to think of their feet   October 2014
Columbia University Data Science Institute

Understanding and adapting statistical models: an exploration in language May 2014
University of Tornoto, Machine Learning Colloquium

Predicting linguistic structures accurately and efficiently   November 2013
Université de Montreal, CS Colloquium

A picture is worth 13.6 words on average   November 2013
University de Montreal, LISA Lab Talk Series

Better! Faster! Stronger! Learning to balance accuracy and efficiency when predicting linguistic structures   September 2013
National Research Council of Canada, NLP Colloquium

Imperative Learning   June 2013
Prediction with Sequential Models workshop at ICML (joint invited talk with John Langford)

Better! Faster! Stronger! Learning to balance accuracy and efficiency when predicting linguistic structures   June 2013
Inferning workshop at ICML

Dynamic Feature Selection   June 2013
Budgeting workshop at ICML

Better! Faster! Stronger! Learning to balance accuracy and efficiency when predicting linguistic structures   January 2013
Microsoft Research, New York

Complex Predictions Need not be Slow   September 2012
University of Southern California, CS Colloquium

Transfer learning in language   September 2012
Machine Learning for Spoken Language Processing workshop at ICASSP

NLP for smart people who know nothing about NLP   August 2012
Course at the Machine Learning Summer School, UC Santa Cruz

Complex predictions need not be slow   May 2012
University of Toronto, Statistics Colloquium

Structured prediction in lingustic domains   March 2012
University of Utah, CS Colloquium

Complex prediction need not be slow   March 2012
Tokyo Institute of Technology, CS Colloquium Series

Transfer learning in Language   March 2012
International Workshop on Statistical Machine Learning, Tokyo

Moving between Tasks and Domains   November 2011
Temple University, AI Seminar Series

Structured Prediction Need Not Be Slow   June 2011
TextGraphs Workshop at NAACL

Learning Structured Prediction by Deomonstration   June 2011
Laboratory for Telecommunication Science

Structure and Knowledge in Natural Language Processing   June 2011
University of Trento, Italy

Structure and Knowledge in Natural Language Processing   April 2011
Microsoft Research Tutorial, Redmond, WA

Structure and Knowledge in Natural Language Processing   March 2011
Carnegie Mellon University, AI Seminar Series

Transfer Learning in Language: We've Got a Long Way to Go...   December, 2010
Workshop on Transfer Learning via Rich Generative Models at NIPS

Linguists get the abstraction, machines get the details   July, 2010
Linguistics Meets NLP Workshop at ACL

Structure and Knowledge in Natural Language Processing   April, 2010
University of Illinois Urbana/Champaigne, Computer Science

Structure and Knowledge in Natural Language Processing   November, 2009
University of Maryland, Computer Science

Statistics, Typology and Language Processing   October, 2009
University of Utah, Department of Linguistics

{ϵ-,Semi-,Un-}supervised Search-based Structured Prediction   May, 2009
Yahoo! Research, New York

Non-parametric Models with Latent Hierarchies   April, 2009
Columbia University, Department of Statistics

Domain Adaptation in Natural Language   March, 2009
Johns Hopkins University, Center for Language and Speech Processing

Domain Adaptation in Natural Language   November, 2008
University of Texas at Austic, Computer Science

Summer School in Computational- and Psycho-Linguistics   September 2008
University of Edinburgh

Automatic Document Summarization   January 2007
Information Sciences Institute

Search-based Structured Prediction   March 2007
University of Maryland, College Park

Bayesian Techniques for Natural Language Processing   July 2007
Microsoft Research Asia, Beijing, China

Tutorials

Advances in Structured Prediction   June 2015
ICML Tutorial (with John Lanford)

Hands-on Learning to Search for Joint Prediction   May 2015
NAACL Tutorial (with John Lanford)

From Structured Prediction to Inverse Reinforcement Learning   August 2011
Tutorial at the Conference of the Association for the Advancement of Artificial Intelligence (AAAI)

Beyond Structured Prediction: Inverse Reinforcement Learning   July 2011
Tutorial at the Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)

Structured Prediction 101   May 2011
University of Trento, Italy

Domain Adaptation   April 2011
Microsoft Research Tutorial, Redmond, WA

From Structured Prediction to Inverse Reinforcement Learning   July 2010
Tutorial at the Conference of the Association for Computational Linguistics (ACL)

Domain Adaptation   June 2010
Tutorial at the International Conference in Machine Learning (ICML)
Co-presented with John Blitzer (Google)

Bayesian Techniques for HLT Reseachers   June 2006
Tutorial at the Human Language Technology Conference (HLT/NAACL)

Statistical Learning Theory and SVMs for AI Researchers   Feburary 2006
Tutorial at the ISI Intelligent Systems Division Retreat

NLP for Smart People Who Know Nothing about NLP   January 2006
Toyota Technological Institute at Chicago

Bayesian Methods for Natural Language Processing   March 2005
Tutorial at the Information Sciences Institute, AI seminar

Funding

Collaborative Research: Adaptive Heads-up Displays for Simultaneous Interpretation 2017–2019

NSF RI: EAGER: $150k to UMD; PI: Graham Neubig (CMU), Co-PIs: me, Jordan Boyd-Graber and Leah Findlater (UW)

System for Cross Language Information Processing, Translation and Summarization 2017–2021

IARPA Grant, $14m over four years; PI Kathy McKeown (Columbia), UMD PI Doug Oard.

Linguistic Semantics and Discourse from Leaky Distant Supervision 2016–2020

NSF-IIS Small Grant, $500k over three years. Sole PI.

Neural Machine Translation from Weak User Feedback 2016

Amazon Research Award, $65k; Sole PI.

Discrete Algorithms in NLP 2014

NSF-IIS EAGER Grant, $75k for one year. Co-PI with Samir Khuller (UMD).

Learning Integer Linear Programming for Natural Language Processing 2014

Google Research Grant, $65k; Sole PI.

NRT-DESE: Flexibility in Language Processes and Technology: Human- and Global-Scale 2013–2019

NSF NRT-DESE Grant, $3.0m over five years. PI: Colin Phillips (UMD).

Bayesian Thinking on Your Feet—Embedding Generative Models in Reinforcement Learning for Sequentially Revealed Data 2013–2016

NSF-IIS Small Grant, $550k over three years. PI: Jordan Boyd-Graber (UC Boulder).

Developing the Next Generation of Community Financial CyberInfrastructure for Monitoring and Modeling Financial Eco-Systems and for Managing Systemic Risk 2013–2014

NSF CI-PI Infrastructure Grant, $100k for one year. PI: Louiqa Raschid (UMD).

Domain Adaptation and Translation 2012–2013

DARPA CSSG Grant, $400k over two years. Sole PI.

Learning the Relationship between the Anatomy and Spatial Hearing 2011–2012

NSF-IIS Small Grant, $164k for one year. PI: Ramani Duraiswami (UMD).

Distributed Empirical Language Processing for Human Interaction 2011–2016

DARPA "BOLT" Grant, $3.3m for UMD over five years. PI: Doug Oard (UMD).

Discovering and explaining technical emergence through analysis of the language and structure of scientific publications 2011–2016

IARPA "FUSE" Grant, $2.8m for UMD over five years. PI: Kathy McKeown (Columbia University).

Distributed Domain Adaptation for Millions of Domains 2011

DARPA CSSG Grant, $92k over one year. Sole PI.

Learned Dynamic Prioritization 2010–2013

NSF-IIS Medium Grant, $496k for UMD over three years. Co-PI with Jason Eisner (JHU)

Statistical Linguistic Typology 2009–2012

NSF-IIS Small Grant, $550k over three years. Sole PI.

Graph Algorithms for Large-Data NLP 2009–2010

Google Research Grant, $55k; Co-PI with Suresh Venkatasubramanian.

Language Understanding for Research Papers 2009

U Utah URC Faculty Research & Creative Grant Proposal, $6k. Sole PI.

Advanced Learning Technology for Translation 2009
U Utah Seed Grant, $35k$. Sole PI.

Computational Thinking Olympiad: Brainstorming Workshop 2008
NSF-IIS grant, $20k$ over two years. Sole PI.

Cross-task Learning for Natural Language Processing *(REU)* 2008
NSF-IIS Research Experience for an Undergraduate Supplement, $6k$ over one year. Sole PI.

Cross-task Learning for Natural Language Processing 2007–2010
NSF-IIS Small Grant, $404k$ over three years. Sole PI.

External Service

Diversity and Inclusion Co-Chair, NIPS 2018

Coordinating Committee, ACL 2017

Executive Board Chair, NAACL, 2014–2016

Sponsorship Chair, ICML 2010–2011

Publicity Chair, COLING 2010

Publications Chair, ICML 2009–2010

Computational Thinking Olympiad
Co-program chair of an NSF-sponsored Olympiad in *computational thinking*

Executive Board Member, NAACL
Executive Board Member (elected) of the North American Association for Computational Linguistics

Publicity Chair, ACL 2008

Research blog: Natural language processing blog
Since 2006, `http://nlpers.blogspot.com`

Workshop Organization

Co-organizer, Building Linguistically Generalizable NLP Systems

    Workshop at the European Association for Computational Linguitics, 2017

Co-organizer, Let's Discuss: Learning Methods for Dialogue

    Workshop at the Conference on Neural Information Processing Systems (NIPS) 2016

Co-organizer, Workshop on Human-Computer Question Answering

    Workshop at the Human Language Technology Conference (HLT/NAACL), 2016

Co-organizer, Simultaneous Machine Interpretation

    Shonan Workshop, Japan, 2015

Co-organizer, Discrete Algorithms for Machine Learning

    Shonan Workshop, Japan, 2013

Co-organizer, Combining Strategies for Reducing the Label Cost

Workshop at the International Conference on Machine Learning (ICML), 2011

Co-organizer, Prior knowledge in text and language processing

Workshop at the International Conference on Machine Learning (ICML), 2008

Co-organizer, Computationally Hard Problems in Speech & Language

Workshop at the Human Language Technology Conference (HLT/NAACL), 2006

Co-organizer, Bayesian Methods in Natural Language Processing

Workshop at the Conference on Neural Information Processing Systems (NIPS) 2005

Journal Editorial Roles

Editorial Boards: Transactions of the ACL (2013–2017), Computational Linguitics (2011–2013), Machine Learning Journal (2008–2016), IEEE Transactions on Speech and Language Processing (2008–2012)

Ad-hoc Reviews: Journal of Machine Learning Research, Journal of Computational and Graphical Statistics, Journal of Natural Language Engineering, IEEE's Transactions on Knowledge and Data Engineering, ACM Transactions on Speech and Language Processing, ACM Transactions on Pattern Analysis and Machine Intelligence, Journal of Artificial Intelligence Research, IBM Systems Journal

Conference Chair Roles

Conference Program Chair: North American Chapter of the ACL (2013)

Conference Area Chair: many times at ICML, IJCAI, NIPS, NAACL, ACL, EMNLP, AIStats.

Conference PCs: many times for the above conferences and more

Internal Service

UMD CS Diversity Committee	2014–2017
UMD CS Department Council	2015–2016
UMD CS Artificial Intelligence Field Committee Chair	2013–2016
UMD Computer Science Hiring Committee	2014
Director of Computational Linguistics & Information Processing Lab	2013–2017
Executive Board, Language Science Center	2013–2017
UMIACS APT Committee Member	2012
UMD Computer Science Hiring Committee	2012
Artificial Intelligence Brown Bag Lunch	2012
Co-director of Computational Linguistics & Information Processing Lab	2011–2012
UMD Admissions Committee Member	2010–2010

Utah Curriculum Committee Chair 2008–2010
Overseen significant advances in math/science requirements, formation of undergraduate tracks, and development of four new undergraduate courses

Utah Masters in Statistics Program, SoC Representative 2008–2010
Representative from the School of Computing to the Masters in Statistics Program

Utah Committee Memberships
Admissions (2007, 2008); Curriculum (2007, 2008); College Council (2007, 2008)

Organizer, Natural Language Seminar Series   2001–2006
Weekly seminar at the Information Sciences Institute