

Ryan McDonald

Google Inc.
76 Ninth Avenue
New York, NY 10011
ryanmcd@gmail.com
<http://www.ryanmcd.com>

Date of Birth: June 26th, 1979
Citizenship: Canadian and N. Irish (British/EU)
Visa: H-1 Visa (United States)

Research Interests

- Natural language processing, multilingual syntactic parsing, information extraction, opinion mining, summarization, machine learning for natural language processing, large-scale machine learning

Employment

- **Google Inc.**, New York, NY (2006 – Present)
Staff Research Scientist (promoted May 2010)
Senior Research Scientist (promoted December 2007)
Research Scientist
- **University of Pennsylvania**, Philadelphia, PA (2002 – 2006)
Research Assistant
- **University of Toronto**, Toronto, ON (2001 – 2002)
Research Assistant
- **Canadian Imperial Bank of Commerce**, Toronto, ON (1997-2000)
Database Specialist (Part-time and Summer)

Education

- **University of Pennsylvania**, Philadelphia, PA, USA
Ph.D. in Computer and Information Science, July 2006
Thesis: Discriminative Training and Spanning Tree Algorithms for Dependency Parsing
Advisor: Fernando Pereira
- **University of Pennsylvania**, Philadelphia, PA, USA
M.S.E. in Computer and Information Science, May 2004
- **University of Toronto**, Toronto, ON, Canada
Hon.B.Sc. majoring in Computer Science and Philosophy, May 2002

Professional Services

- Editorial Board of Computational Linguistics (2011-2013)
- Senior Program Committee/Area Chair for the Association for Computational Linguistics (ACL 2011), North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2009), International Conference on Machine Learning (ICML 2008), North American Chapter of the Association for Computational Linguistics (HLT-NAACL 2007)
- Best Paper Award Committee for the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2009)

- Co-organizer of the shared task on multilingual dependency parsing at the Conference on Computational Natural Language Learning (2007)
- Co-organizer of the Workshop on Computationally Hard Problems in Speech and Language Processing at HLT-NAACL (2006)
- Co-organizer of the first North East Student Conference on AI (NESCAI 2006)
- Grant reviewer for the National Science Foundation (NSF 2011), U.S.-Israel Binational Science Foundation (BSF 2011)
- Program Committee/Reviewer for the Association for Computational Linguistics (ACL 2006, 2007, 2008, 2009, 2010, 2012), North American Chapter of the ACL (NAACL 2006, 2008, 2010, 2012), European Chapter of the ACL (EACL 2006, 2009, 2012), Empirical Methods in Natural Language Processing (EMNLP 2007, 2008, 2009, 2010, 2011), Conference on Computational Natural Language Learning (CoNLL 2007, 2009, 2010, 2011), Conference on Computational Linguistics (COLING 2010), Human Language Technologies (HLT 2006, 2008, 2010), International Joint Conference on Natural Language Processing (IJCNLP 2007, 2009, 2011), IEEE International Conference on Semantic Computing (ICSC 2008), International Conference on Weblogs and Social Media (ICWSM 2010), International Conference on Machine Learning (ICML 2009), Neural Information Processing Systems (NIPS 2007, 2008), Conference on Artificial Intelligence and Statistics (AI-STATS 2010), International Joint Conference on Artificial Intelligence (IJCAI 2007, 2011), Computational Linguistics (CL 2007, 2008, 2010, 2011), Journal of Machine Learning Research (JMLR 2006, 2007, 2008), Journal of Artificial Intelligence Research (JAIR 2008, 2011), Journal of Natural Language Engineering (JNLE, 2006, 2008, 2011), International Symposium on Semantic Mining in Biomedicine (2006), IEEE Transactions on Knowledge and Data Engineering (TKDE 2006), BMC Bioinformatics (2006), BioNLP (2009), Workshop on High-level Information Extraction (2008), Workshop on Integer Linear Programming for Natural Language Processing (2009), Workshop on Web Search Result Summarization and Presentation (2009), NIPS Workshop on Learning on Cores, Clusters, and Clouds (2010), Workshop on Automatic Summarization for Different Genres, Media, and Languages (2011), Workshop on Workshop on Monolingual Text-To-Text Generation (2011), TextGraphs-6: Graph-based Methods for Natural Language Processing (2011)
- Member of ACL, SIGNLL, SIGPARSE, SIGDAT

Publications

- Statistics via Google Scholar (September 2011):
 - Total citations \approx 2300
 - Total citations (w/o self citations) \approx 2200
 - Maximum citation = 321
 - Average citation \approx 70
 - h-index = 25.

- Books / Thesis -

- **Dependency Parsing**
S. Kübler, R. McDonald, and J. Nivre
Synthesis Lectures on Human Language Technologies, G. Hirst ed.
Morgan & Claypool, 2009
- **Discriminative Training and Spanning Tree Algorithms for Dependency Parsing**
R. McDonald (under the supervision of Fernando Pereira)
Ph.D. Thesis, University of Pennsylvania, 2006

- Journal Articles -

- **Analyzing and Integrating Dependency Parsers**
R. McDonald and J. Nivre
Computational Linguistics, 37:1, 2011
- **Automated recognition of malignancy mentions in biomedical literature**
Y. Jin, R. McDonald, K. Lerman, M. Mandel, S. Carroll, M. Liberman, F. Pereira, R. S. Winters and P. S. White
BMC Bioinformatics, 7:492, 2006
- **An automated procedure to identify biomedical articles that contain cancer-associated gene variants**
R. McDonald, R.S. Winters, C.K. Ankuda, J.A. Murphy, A.E. Rogers, F. Pereira, M.S. Greenblatt, and P.S. White
Human Mutation, Volume 27 Issue 9, 2006
- **Automatically annotating documents with normalized gene lists**
J. Crim, R. McDonald, and F. Pereira
BMC Bioinformatics, 6(Suppl 1):S13, 2005
- **Identifying gene and protein mentions in text using conditional random fields**
R. McDonald and F. Pereira
BMC Bioinformatics, 6(Suppl 1):S6, 2005
- **An entity tagger for recognizing acquired genomic variations in cancer literature**
R. McDonald, R.S. Winters, M. Mandel, Y. Jin, P.S. White, and F. Pereira
Bioinformatics, Volume 20 Issue 17, 2004

- International Conference Proceedings (Refereed) -

- **Multi-source Transfer of Delexicalized Dependency Parsers**
R. McDonald, S. Petrov and K. Hall
Empirical Methods in Natural Language Processing (EMNLP), 2011
- **Training dependency parsers by jointly optimizing multiple objectives**
K. Hall, R. McDonald, J. Katz-Brown, M. Ringgaard
Empirical Methods in Natural Language Processing (EMNLP), 2011
- **Training a Parser for Machine Translation Reordering**
J. Katz-Brown, S. Petrov, R. McDonald, D. Talbot, F. Och, H. Ichikawa, M. Seno and H. Kazawa
Empirical Methods in Natural Language Processing (EMNLP), 2011
- **Semi-supervised Latent Variable Models for Sentence-level Sentiment Analysis**
O. Täckström and R. McDonald
Association for Computational Linguistics (ACL), 2011
- **Discovering fine-grained sentiment with latent variable structured prediction models**
O. Täckström and R. McDonald
European Conference on Information Retrieval (ECIR), 2011
- **Evaluation of Dependency Parsers on Unbounded Dependencies**
J. Nivre, L. Rimell, R. McDonald, and C. Gómez-Rodríguez
International Conference on Computational Linguistics (COLING), 2010
- **The Viability of Web-Derived Polarity Lexicons**
L. Velikovich, S. Blair-Goldensohn, K. Hannan, and R. McDonald
North American Association for Computational Linguistics (NAACL), 2010

- **Distributed Training Strategies for the Structured Perceptron**
R. McDonald, K. Hall, and G. Mann
North American Association for Computational Linguistics (NAACL), 2010
- **Efficient Large-Scale Distributed Training of Conditional Maximum Entropy Models**
G. Mann, R. McDonald, M. Mohri, N. Silberman, and D. Walker
Neural Information Processing Systems (NIPS), 2009
- **Contrastive Summarization: An Experiment with Consumer Reviews**
K. Lerman and R. McDonald
North American Association for Computational Linguistics (NAACL), 2009
- **Sentiment Summarization: Evaluating and Learning User Preferences**
K. Lerman, S. Blair-Goldensohn and R. McDonald
European Association for Computational Linguistics (EACL), 2009
- **A Joint Model of Text and Aspect Ratings for Sentiment Summarization**
I. Titov and R. McDonald
Association for Computational Linguistics (ACL), 2008
- **Integrating Graph-based and Transition-based Dependency Parsers**
J. Nivre and R. McDonald
Association for Computational Linguistics (ACL), 2008
- **Modeling Online Reviews with Multi-Grain Topic Models**
I. Titov and R. McDonald
International Conference on the World Wide Web (WWW), 2008
- **Characterizing the Errors of Data-Driven Dependency Parsing Models**
R. McDonald and J. Nivre
Empirical Methods in Natural Language Processing (EMNLP-CoNLL), 2007
- **The CoNLL 2007 Shared Task on Dependency Parsing**
J. Nivre, J. Hall, S. Kübler, R. McDonald, J. Nilsson, S. Riedel, and D. Yuret
Empirical Methods in Natural Language Processing (EMNLP-CoNLL), 2007
- **On the Complexity of Non-Projective Data-Driven Dependency Parsing**
R. McDonald and G. Satta
International Conference on Parsing Technologies (IWPT), 2007
- **Structured Models for Fine-to-Coarse Sentiment Analysis**
R. McDonald, K. Hannan, T. Neylon, M. Wells, and J. Reynar
Association for Computational Linguistics (ACL), 2007
- **A Study of Global Inference Algorithms in Multi-Document Summarization**
R. McDonald
European Conference on Information Retrieval (ECIR), 2007
- **Domain adaptation with structural correspondence learning**
J. Blitzer, R. McDonald, and F. Pereira
Empirical Methods in Natural Language Processing (EMNLP), 2006
- **Multilingual dependency analysis with a two-stage discriminative parser**
R. McDonald, K. Lerman, and F. Pereira
Conference on Natural Language Learning (CoNLL), 2006
- **Online learning of approximate dependency parsing algorithms**
R. McDonald and F. Pereira
European Association for Computational Linguistics (EACL), 2006

- **Discriminative sentence compression with soft syntactic evidence**
R. McDonald
European Association for Computational Linguistics (EACL), 2006
- **Non-projective dependency parsing using spanning tree algorithms**
R. McDonald, F. Pereira, K. Ribarov, and J. Hajič
Empirical Methods in Natural Language Processing (HLT-EMNLP), 2005
- **Flexible text segmentation with multilabel structured classification**
R. McDonald, K. Crammer, and F. Pereira
Empirical Methods in Natural Language Processing (HLT-EMNLP), 2005
- **Simple algorithms for complex relation extraction with applications to biomedical IE**
R. McDonald, F. Pereira, S. Kulick, S. Winters, Y. Jin, and P. White
Association for Computational Linguistics (ACL), 2005
- **Online large margin training of dependency parsers**
R. McDonald, K. Crammer, and F. Pereira
Association for Computational Linguistics (ACL), 2005
- **Flexible web document analysis for delivery to narrow-bandwidth devices**
G. Penn, J. Hu, H. Luo and R. McDonald
International Conference on Document Analysis and Recognition (ICDAR), 2001

- International Workshop Proceedings (Refereed) -

- **Learning to Classify the Scope of Negation for Improved Sentiment Analysis**
I. Council, R. McDonald, and L. Velikovich
Workshop on Negation and Speculation in Natural Language Processing (NeSp-NLP), 2010
- **Building a Sentiment Summarizer for Local Service Reviews**
S. Blair-Goldensohn, K. Hannan, R. McDonald, T. Neylon, G. Reis, and J. Reynar
WWW Workshop on NLP Challenges in the Information Explosion Era (NLPiX), 2008
- **Identifying and extracting malignancy types in cancer literature**
Y. Jin, R. McDonald, K. Lerman, M. Mandel, M. Liberman, F. Pereira, R.S. Winters, and P.S. White
Linking Literature, Information and Knowledge for Biology (BioLink), 2005
- **New large margin algorithms for structure prediction**
K. Crammer, R. McDonald, and F. Pereira
NIPS Workshop on Learning With Structured Outputs, 2004
- **Scalable large margin online learning algorithms for structured classification**
R. McDonald, K. Crammer, and F. Pereira
NIPS Workshop on Learning With Structured Outputs, 2004
- **Integrated annotation for biomedical information extraction**
S. Kulick, A. Bies, M. Liberman, M. Mandel, R. McDonald, M. Palmer, A. Schein, L. Ungar, S. Winters, and P. White
Linking Biological Literature, Ontologies and Databases (BioLink) 2004
- **Exploiting sequent structure in membership algorithms for the Lambek Calculus**
R. McDonald
European Summer School in Logic Language and Information (ESSLLI), 2003

- Edited Proceedings -

- **Proceedings of the 2007 CoNLL Shared Task**
J. Nivre, J. Hall, S. Kübler, R. McDonald, J. Nilsson, S. Riedel, and D. Yuret (eds.)
Association for Computational Linguistics, 2007
- **Proceedings of the Workshop on Computationally Hard Problems in Speech and Language Processing**
R. McDonald, C. Sutton, H. Daumé, A. McCallum, J. Bilmes, and F. Pereira (eds.)
Association for Computational Linguistics, 2006

Invited Talks

- Learning to Mine Fine-grained Opinions, Washington University, 2010
- Learning to Mine Fine-grained Opinions, University of Edinburgh, 2010
- Web-scale Multilingual Opinion Mining, New York University, 2009
- Web-scale Multilingual Opinion Mining, University of Uppsala, 2009
- Graph-based Dependency Parsing, University of Colorado, 2009
- Mining Opinions from the Web, Northwestern University, 2009
- Mining Opinions from the Web, University of Pennsylvania, 2009
- Mining Opinions from the Web, University of Virginia, 2008
- Analyzing and Integrating Dependency Parsers, Massachusetts Institute of Technology (MIT), 2008
- Summarizing User Generated Content, University of Tokyo, 2008
- The Complexity of Data-Driven Dependency Parsing, University of Uppsala, 2007
- The Complexity of Data-Driven Dependency Parsing, University of Utah, 2007
- Multilingual Dependency Parsing with Spanning Tree Algorithms, Center for Language and Speech Processing, Johns Hopkins University, Baltimore, MD, 2006
- Multilingual Dependency Parsing with Spanning Tree Algorithms, BBN Technologies, Cambridge, MA, 2006
- Flexible Text Segmentation with Structured Multilabel Classification, Google, New York, NY, 2006
- Discriminative Training for Dependency Parsing, Microsoft Research, Redmond, WA, 2005

Teaching Experience

- **Guest Lecturer, Computer Science, New York University** (2010)
Single lecture entitled *Dependency parsing* in graduate level NLP course.
- **Guest Lecturer, Swedish Graduate School of Language Technology** (2007 & 2009)
Designed and taught a day course entitled *Generalized Linear Classifiers in NLP*.
- **Guest Lecturer, Department of Computer Science, Columbia University** (2008)
Single lecture entitled *Graph-based dependency parsing* in network theory course.
- **Lecturer, European Summer School in Logic Language and Information** (2007)
Designed and co-taught a week long course entitled *Introduction to Data-Driven Dependency Parsing* with Joakim Nivre.

- **Guest Lecturer, Department of Computer Science, New York University (2007)**
Single lecture entitled *Structured Learning* in machine learning course.
- **Guest Lecturer, Computer and Information Science, University of Pennsylvania (2005)**
Gave a series of lectures on probabilistic parsing and discriminative online learning for machine learning course.
- **Lecturer and Teaching Assistant, University of Pennsylvania (2003 – 2004)**
 - Programming Languages and Techniques I (Fall 2003)
- Lecturer for corequisite half credit lab
 - Programming Languages and Techniques II (Spring 2004)
- Lecturer for corequisite half credit lab

Nominated for the *Penn Prize for Excellence in Teaching by Graduate Students* and the *Computer and Information Science Outstanding Teaching Assistant Award*.
- **Teaching Assistant, University of Toronto (January 2001 – May 2002)**
 - Introduction to Computer Science (Spring 2001)
 - Discrete Mathematics for Computer Science (Fall 2001)
 - Introduction to Computer Programming (Spring 2002)

Research Supervision & Thesis Committees

- Oscar Täckström, Summer internship supervisor
Title: *Multilingual clustering and its applications in NLP*
Google, 2011
- Oscar Täckström, Summer internship supervisor
Title: *Sub-document sentiment analysis via document level supervision*
Google, 2010
- Jordan Boyd-Graber, Thesis committee member
Title: *Linguist Features in Topic Models*
Department of Computer Science, Princeton University, 2010
- Jens Nilsson, Faculty opponent
Title: *Transformation and Combination in Data-Driven Dependency Parsing*
Department of Computer Science, Växjö University, 2009
- Kevin Lerman, Summer internship supervisor
Title: *Evaluating and Learning User Preferences in Opinion Summarization*
Google, 2008
- Slav Petrov, Summer internship supervisor
Title: *Anchor Topic Models for Entity Centric Text*
Google, 2008
- Ivan Titov, Summer internship supervisor
Title: *Topic Models for Review Analysis and Summarization*
Google, 2007
- Kevin Lerman, Undergraduate research supervisor
Title: *Dependency Parsing and Biomedical Information Extraction*
University of Pennsylvania, 2004-2006