

# Dmitriy Dligach, Ph.D.

Boston Children's Hospital  
Harvard Medical School  
300 Longwood Ave  
Boston, MA 02115

Phone: (617) 919-3596  
Fax: (617) 730-0921  
Email: [dmitriy.dligach@childrens.harvard.edu](mailto:dmitriy.dligach@childrens.harvard.edu)  
Web: <https://sites.google.com/site/dmitriydligach/>

## Research Interests

Natural language processing, machine learning, data mining, medical informatics.

## Education

Ph.D. Computer Science, University of Colorado, Boulder, 2010.

M.S. Computer Science, State University of New York, Buffalo, 2004.

B.S. Computer Science, Loyola University, Chicago, 1999.

## Professional Experience

Instructor (entry-level faculty appointment), Harvard Medical School, 2014 - present

Research Fellow, Harvard Medical School, 2011 - 2014

Research Fellow, Boston Children's Hospital, 2011 - 2014

Postdoctoral Research Associate, University of Colorado, Boulder, 2010 - 2011

Research Assistant, University of Colorado, Boulder, 2005 - 2010.

Research Scientist, Pronto Inc., Boulder, CO, 2005.

Teaching Assistant, State University of New York, Buffalo, 2003 - 2005.

Senior Software Engineer, Motorola, Inc., Arlington Heights, IL, 2000 - 2003.

Development Analyst, Commonwealth Edison, Chicago, IL, 1999 - 2000.

Network Administrator, Graver, Bokhof, Goodwin & Sullivan, Chicago, IL, 1997 - 1998.

Internship, Trans Union Corporation, Chicago, IL, 1997.

## Publications

### *Journal Articles*

\*Chen Lin, \*Elizabeth W. Karlson, \*Dmitriy Dligach, Monica Ramirez, Timothy A. Miller, Huan Mo, Natalie S. Braggs, Andrew Cagan, Joshua C. Denny, Guergana K. Savova. Automatic Identification of Methotrexate-induced Liver Toxicity in Rheumatoid Arthritis Patients from the Electronic Medical Record. 2014. Accepted for publication in JAMIA. \* Indicates co-first authors

Dmitriy Dligach, Steven Bethard, Lee Becker, Timothy A. Miller, Guergana K. Savova. Discovering Body Site and Severity Modifiers in Clinical Texts. *Journal of American Medical Informatics Association*. 2014. 2014;21:438-447. doi:10.1136/amiajnl-2013-001949.

Chen Lin, Elizabeth W. Karlson, Helena Canhao, Timothy A. Miller, Dmitriy Dligach, Pei Jun Chen, Raul Natanael Guzman Perez, Tianxi Cai, Michael E. Weinblatt, Nancy A. Shadick, Robert M. Plenge, Guergana K. Savova. Automatic Prediction of Rheumatoid Arthritis Disease Activity from the Electronic Medical Records. *PLoS ONE*. 2013. 8(8): e69932. doi:10.1371/journal.pone.0069932.

Daniel Albright, Arrick Lanfranchi, Anwen Fredriksen, William F. Styler, Colin Warner, Jena D. Hwang, Jinho D. Choi, Dmitriy Dligach, Rodney D. Nielsen, James Martin, Wayne Ward, Martha Palmer, Guergana K. Savova. Towards comprehensive syntactic and semantic annotations of the clinical narrative. *Journal of American Medical Informatics Association*. 2013. Sep-Oct;20(5):922-30. doi: 10.1136/amiajnl-2012-001317.

Stephen T. Wu, Vinod C. Kaggal, Dmitriy Dligach, James J. Masanz, Pei Chen, Lee Becker, Wendy W. Chapman, Guergana K. Savova, Hongfang Liu and Christopher G. Chute. A common type system for clinical Natural Language Processing. *Journal of Biomedical Semantics*. January 2013, 4:1.

Dmitriy Dligach and Martha Palmer. Word Sense Disambiguation with Automatically Retrieved Semantic Knowledge. *International Journal of Semantic Computing (IJSC)*. 2:3, pp. 365-380, 2008.

### Conference Proceedings

Chen Lin, Timothy Miller, Alvin Kho, Steven Bethard, Dmitriy Dligach, Sameer Pradhan and Guergana Savova. Descending-Path Convolution Kernel for Syntactic Structures. In *ACL '14: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*. June 22 - 27, 2014, Baltimore, MD.

Dmitriy Dligach and Martha Palmer. Good Seed Makes a Good Crop: Accelerating Active Learning Using Language Modeling. In *ACL '11: Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. June 19 - 24, 2011, Portland, OR.

Susan Windisch Brown, Dmitriy Dligach, and Martha Palmer. VerbNet Class Assignment as a WSD Task. In *IWSC 2011: Proceedings of the 9th International Conference on Computational Semantics*. January 12 - 14, 2011, Oxford, UK.

Dmitriy Dligach and Martha Palmer. Improving Verb Sense Disambiguation with Automatically Retrieved Semantic Knowledge. In *ICSC '08: Proceedings of the Second IEEE International Conference on Semantic Computing*. August 4 - 7, 2008, Santa Clara, CA.

Dmitriy Dligach and Martha Palmer. Novel Semantic Features for Verb Sense Disambiguation. In *ACL '08: Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. June 15 - 20, 2008, Columbus, OH.

Jinying Chen, Dmitriy Dligach, and Martha Palmer. Towards Large-scale High-Performance English Verb Sense Disambiguation by Using Linguistically Motivated Features. In *ICSC '07: Proceedings of the First IEEE International Conference on Semantic Computing*. Sep 17 - 19, 2007, Irvine, CA.

Cecily Jill Duffield, Jena D. Hwang, Susan Windisch Brown, Dmitriy Dligach, Sarah E. Vieweg, Jenny Davis, and Martha Palmer. Criteria for the Manual Clustering of Verb Senses. In *CogSci '07: Proceedings of the 29th Annual Conference of the Cognitive Science Society*. August 1-4, 2007, Nashville, TN.

## Workshop Proceedings

Dmitriy Dligach, Timothy A. Miller, Guergana K. Savova. Active Learning for Phenotyping Tasks. In *Proceedings of the 2013 NLP for Medicine and Biology workshop held in conjunction with RANLP-2013*. September 2013. Hissar, Bulgaria.

Timothy A. Miller, Steven Bethard, Dmitriy Dligach, Sameer Pradhan, Chen Lin, Guergana K. Savova. Discovering Narrative Containers in Clinical Text. In *Proceedings of the 2013 Workshop on Biomedical Natural Language Processing (BioNLP 2013) held in conjunction with NAACL-2013*. August 2013, Sophia, Bulgaria.

Chen Lin, Timothy Miller, Dmitriy Dligach, Robert M. Plenge, Elizabeth W. Karlson, Guergana Savova. Maximal Information Coefficient for Feature Selection for Clinical Document Classification. In *International Conference on Machine Learning (ICML 2012) Workshop on Machine Learning for Clinical Data Analysis*. June 2012, Edinburgh, Scotland.

Chen Lin, Timothy Miller, Dmitry Dligach, Robert Plenge, Elizabeth Karlson, Guergana Savova. Feature Engineering and Selection for Rheumatoid Arthritis Disease Activity Classification Using Electronic Medical Records. In *International Conference on Machine Learning (ICML 2012) Workshop on Machine Learning for Clinical Data Analysis*. June 2012, Edinburgh, Scotland.

Timothy A. Miller, Dmitriy Dligach, and Guergana K. Savova. Active Learning for Coreference Resolution. In *Proceedings of the 2012 Workshop on Biomedical Natural Language Processing (BioNLP 2012) held in conjunction with NAACL-2012*. June 2012, Montreal, Canada.

Stephen T. Wu, Vinod C. Kaggal, Guergana K. Savova, Hongfang Liu, Dmitriy Dligach, Jiaping Zheng, Wendy W. Chapman, Christopher G. Chute. Generality and Reuse in a Common Type System for Clinical Natural Language Processing. In *Managing Interoperability and complexity in Health Systems (MIXHS-2011)*. October 2011, Glasgow, Scotland, U.K.

Dmitriy Dligach and Martha Palmer. Reducing the Need for Double Annotation. In *Proceedings of the Fifth Linguistic Annotation Workshop (LAW V) held in conjunction with ACL-2011*. June, 2011, Portland, OR.

Dmitriy Dligach, Rodney D. Nielsen, and Martha Palmer. To Annotate More or to Annotate More Accurately. In *Proceedings of the Fourth Linguistic Annotation Workshop (LAW IV) held in conjunction with ACL-2010*. July, 2010, Uppsala, Sweden.

Suresh Manandhar, Ioannis P. Klapaftis, Dmitriy Dligach, and Sameer S. Pradhan. SemEval-2010 Task 14: Word Sense Induction & Disambiguation. In *Proceedings of the Fifth International Workshop on Semantic Evaluations (SemEval-2010) held in conjunction with ACL-2010*. July, 2010, Uppsala, Sweden.

Dmitriy Dligach and Martha Palmer. Using Language Modeling to Select Useful Annotation Data. In *Proceedings of the Student Research Workshop and Doctoral Consortium Held in Conjunction with NAACL-HLT 2009*. May 31 - June 5, 2009. Boulder, Colorado.

Sameer Pradhan, Edward Loper, Dmitriy Dligach, and Martha Palmer. SemEval-2007 Task-17: English Lexical Sample, SRL and All Words. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007) held in conjunction with ACL-2007*. June, 2007, Prague, the Czech Republic.

Cecily Jill Duffield, Jena D. Hwang, Susan Windisch Brown, Dmitriy Dligach, Sarah E. Vieweg, Jenny Davis, and Martha Palmer. 2007. Criteria for the Manual Grouping of Verb Senses. In *Proceedings of the Linguistic Annotation Workshop held in conjunction with ACL 2007*. June, 2007, Prague, the Czech Republic.

## Book Chapters

Susan Windisch Brown, Dmitriy Dligach, and Martha Palmer. VerbNet Class Assignment as a WSD Task. 2014. *Computing Meaning Vol. 4*. Springer. pp 203-216.

## Conference Abstracts

Dmitriy Dligach, Steven Bethard, Lee Becker, Timothy A. Miller, Guergana K. Savova. Discovering Body Site and Severity Modifiers in Clinical Texts. In *American Medical Informatics Association Symposium*. Washington DC, November, 2013.

Chen Lin, Elizabeth W. Karlson, Helena Canhao, Timothy A. Miller, Dmitriy Dligach, Pei Jun Chen, Raul Natanael Guzman Perez, Tianxi Cai, Michael E. Weinblatt, Nancy A. Shadick, Robert M. Plenge, Guergana K. Savova. Automatic Prediction of Rheumatoid Arthritis Disease Activity from the Electronic Medical Records. In *American Medical Informatics Association Symposium*. Washington DC, November, 2013.

Timothy Miller, Dmitriy Dligach, Steven Bethard, Sameer Pradhan, Chen Lin, Guergana K. Savova. Discovering Time Expressions in Clinical Text. In *American Medical Informatics Association Symposium*. Washington DC, November, 2013.

## Poster Presentations

Susan Windisch Brown, Cecily Jill Duffield, Jena D. Hwang, Dmitriy Dligach, Sarah E. Vieweg, Jenny Davis, Martha Palmer. The role of functional nouns in manual grouping of verb senses. In *Poster Session of Concept Types and Frames in Language, Cognition, and Science*. August 20-22, 2007, Dusseldorf, Germany.

## Ph.D. Thesis

Dmitriy Dligach. High-performance Word Sense Disambiguation with Less Manual Effort. Ph.D. thesis, University of Colorado at Boulder, August 2010.

## Minor Contributions

William J. Rapaport. Philosophy of Computer Science: An Introductory Course. In *Teaching Philosophy* 28(4): 319-341. 2005.

## Invited Talks

Semantic Analysis of Clinical Narratives. NorthShore University HealthSystem Research Institute. Evanston, IL. June 2014.

Active Learning for Phenotyping Tasks. i2b2 3rd Annual Academic User Group Conference and NLP Workshop. Boston, MA. June 2013.

Discovering Body Site and Severity Modifiers in Rheumatoid Arthritis Data. i2b2 2nd Annual Academic User Group Conference and NLP Workshop. Boston, MA. June 2012.

Discovering Severity and Body Site Modifiers: A Relation Extraction Task. Strategic Health IT Advanced Research Project Annual Meeting. Rochester, MN. June, 2012.

### *Courses Taught*

CSCI 3202: Introduction to Artificial Intelligence. University of Colorado, Boulder, Invited lecture, Fall 2008.

LING 7800: Advanced Computational Linguistics: Lexical Semantics. University of Colorado, Boulder, Invited lecture, Spring 2007.

LING 5200: Computational Corpus Linguistics. University of Colorado, Boulder, Invited lecture, Spring 2006

CSE/PHI 510: Philosophy of Computer Science. State University of New York, Buffalo, Teaching Assistant, Fall 2004.

CSE 116 : Introduction To Computer Science for Majors II. State University of New York, Buffalo, Teaching Assistant, Fall 2003, Spring 2004.

### *Courses Prepared to Teach*

Natural Language Processing / Computational Linguistics

Machine Learning

Artificial Intelligence

Data Mining

### *Honors*

Presidential Fellowship, State University of New York, Buffalo.

Fr. Gerst Memorial Award ("presented to a single graduating senior who has displayed outstanding academic achievement in mathematics or computer science"), Loyola University, Chicago, 1999.

Graduated Magna Cum Laude, Loyola University, Chicago, 1999.

Loyola University scholarship, Loyola University, Chicago, 1996.

### *Service*

RANLP NLP for Medicine and Biology Workshop. Paper Reviewer. September 2013.

American Medical Informatics Association Symposium. Paper Reviewer. May 2013.

International Conference on Computational Semantics (IWCS). Program Committee. March 2013.

International Conference on Computational Semantics (IWCS). Paper Reviewer. March 2013.

i2b2/VA/Cincinnati 2011 Shared Task and Workshop. Paper Reviewer. September 2011.

Computational Linguistics. Paper Reviewer. May 2011.

SemEval 2010. Program Committee. April 2010.

SemEval 2010. Paper Reviewer. April 2010.

SemEval 2010 Tasks 14: Word Sense Induction & Disambiguation Task. Organizing Committee. April 2010.