

John DeNero

CONTACT INFORMATION	781 Soda Hall UC Berkeley Berkeley, CA 94720	<i>Phone:</i> (415) 203-1943 <i>Email:</i> denero@berkeley.edu <i>Web:</i> http://www.denero.org
POSITIONS	Faculty Director for Data Science Undergraduate Studies	2023-present
	Associate Dean for Undergraduate Studies, Division of Computing, Data Science, and Society	2021-2023
	Associate Teaching Professor, UC Berkeley, in Computer Science	2014-present
	Senior Research Scientist, Google Inc., in Google Translate	2010-2014
EDUCATION	University of California, Berkeley Ph.D., Computer Science <i>Advisor:</i> Dan Klein	2005-2010
	Stanford University Master of Arts, Philosophy	2001-2002
	Bachelor of Science, with distinction <i>Major:</i> Mathematical and Computational Science <i>Secondary Major:</i> Symbolic Systems	1998-2002
AWARDS	Distinguished Teaching Award, UC Berkeley	2018
	Giancarlo Teaching Fellowship, College of Engineering	2017
	Jim and Donna Gray Award for Excellence in Undergraduate Teaching, EECS Department	2017
	Diane McEntyre Award for Excellence in Teaching, EECS Department	2016
	Award for Excellence in Promoting Women in Undergraduate Computing, Honorable Mention, NCWIT	2016
	Outstanding Faculty of the Year Award, Tau Beta Pi	2015
PUBLICATIONS	Ani Adhikari, John DeNero, and David Wagner. <i>Computational and Inferential Thinking</i> . Online textbook, http://www.inferentialthinking.com , 2016.	
	John DeNero. <i>Composing Programs</i> . Online textbook, http://www.composingprograms.com , 2013.	
	John DeNero. <i>Phrase Alignment Models for Statistical Machine Translation</i> . Ph.D. Dissertation in Computer Science, UC Berkeley, 2010.	
	Refereed Publications	
	Jessy Lin, Geza Kovacs, Aditya Shastri, Joern Wuebker, and John DeNero. “Automatic Correction of Human Translations,” in <i>Proceedings of the North American Association of Computational Linguistics</i> , 2022. Awards: Best New Resource and Best New Task (tied) and Honorable mention for contribution to the special theme on human-centered NLP.	

Samee Ibrahim, Gaoyue Zhou, and John DeNero, “Putting the Con in Context: Identifying Deceptive Actors in the Game of Mafia,” in *Proceedings of the North American Association of Computational Linguistics*, 2022.

Geza Kovacs, and John DeNero, “Measuring the Effects of Human and Machine Translation on Website Engagement,” in *Association for Machine Translation in the Americas*, 2022. **Best presentation award.**

Thomas Zenkel, Joern Wuebker, and John DeNero, “Automatic Bilingual Markup Transfer,” in *Findings of EMNLP*, 2021.

Nicholas Altieri, Briton Park, Mara Olson, John DeNero, Anobel Y Odisho, and Bin Yu, “Supervised Line Attention for Tumor Attribute Classification from Pathology Reports: Higher Performance with Less Data,” in *Journal of Biomedical Informatics*, 2021.

David Gaddy, Daniel Fried, Nikita Kitaev, Mitchell Stern, Rodolfo Corona, John DeNero, and Dan Klein, “Interactive Assignments for Teaching Structured Neural NLP,” in *Proceedings of the Fifth Workshop on Teaching NLP*, 2021.

Briton Park, Nicholas Altieri, John DeNero, Anobel Y. Odisho, and Bin Yu, “Improving natural language information extraction from cancer pathology reports using transfer learning and zero-shot string similarity,” in *Journal of the American Medical Informatics Association*, 2021.

Ani Adhikari, John DeNero, and Michael I. Jordan, “Interleaving computational and inferential thinking in an undergraduate data science curriculum,” in *Harvard Data Science Review*, 2021.

Kevin Yang, Violet Yao, John DeNero, and Dan Klein, “A Streaming Approach For Efficient Batched Beam Search,” in *Proceedings of Empirical Methods in Natural Language Processing*, 2020.

Thomas Zenkel, Joern Wuebker, and John DeNero, “End-to-end neural word alignment outperforms GIZA++,” in *Proceedings of the Association for Computational Linguistics*, 2020.

Patrick Simianer, Joern Wuebker, and John DeNero, “Measuring Immediate Adaptation Performance for Neural Machine Translation,” in *Proceedings of the North American Association of Computational Linguistics*, 2019.

John D. Co-Reyes, Abhishek Gupta, Suvansh Sanjeev, Nicholas Altieri, John DeNero, Pieter Abbeel, and Sergey Levine. “Guiding Policies with Language via Meta-Learning,” in *Proceedings of the International Conference on Learning Representations*, 2019.

Joern Wuebker, Patrick Simianer, and John DeNero, “Compact Personalized Models for Neural Machine Translation,” in *Proceedings of Empirical Methods in Natural Language Processing*, 2018.

Samee Ibraheem, Nicholas Altieri, and John DeNero. “Learning an Interactive Attention Policy for Neural Machine Translation,” in *MT Summit*, 2017.

Nicholas Altieri, Sherdil Niyaz, Samee Ibraheem, and John DeNero. “Improved Word and Symbol Embedding for Part-of-Speech Tagging,” in the *International Workshop on Symbolic-Neural Learning*, 2017.

John DeNero, Sumukh Sridhara, Manuel Pérez-Quinones, Aatish Nayak, and Ben Leong. “Beyond Autograding: Advances in Student Feedback Platforms,” panel in the *ACM Special Interest Group on Computer Science Education*, 2017.

Joern Wuebker, Spence Green, John DeNero, Sasa Hasan and Minh-Thang Luong. “Models and

- Inference for Prefix-Constrained Machine Translation,” in the *Proceedings of the Association for Computational Linguistics*, 2016.
- Brian Hou, Sumukh Sridhara, and John DeNero. “Fuzz Testing Projects in Massive Courses,” in *Learning @ Scale*, 2016.
- Kristin Stephens-Martinez, An Ju, Colin Schoen, John DeNero, and Armando Fox. “Identifying Student Misunderstandings using Constructed Responses,” in *Learning @ Scale, Work-in-progress track*, 2016.
- Brian Hou, Marvin Zhang, and John DeNero. “Restaurant Recommendations,” in the *Nifty Assignments track of the ACM Special Interest Group on Computer Science Education*, 2016.
- Daniel D. Garcia, Jennifer Campbell, John DeNero, Mary Lou Dorf, and Stuart Reges. “CS10K Teachers by 2017? Try CS1K+ students NOW,” panel in the *ACM Special Interest Group on Computer Science Education*, 2016.
- Aneiss Ghodsi and John DeNero. “An Analysis of the Ability of Statistical Language Models to Capture the Structural Properties of Language,” in the *International Natural Language Generation conference*, 2016.
- Joern Wuebker, Spence Green, and John DeNero. “Hierarchical Incremental Adaptation for Statistical Machine Translation,” in the *Proceedings of Empirical Methods in Natural Language Processing*, 2015.
- Rohan Chitnis and John DeNero. “Variable-Length Word Encodings for Neural Translation Models,” in the *Proceedings of Empirical Methods in Natural Language Processing*, 2015.
- Soumya Basu, Brian Hou, Albert Wu, and John DeNero. “Problems Before Solutions: Automated Problem Clarification at Scale,” in *Learning @ Scale*, 2015.
- John DeNero and Stephen Martinis. “Teaching Composition Quality at Scale,” in the *Proceedings of the ACM Special Interest Group on Computer Science Education*, 2014.
- John DeNero, Tom Magrino, and Eric Tzeng. “Ants Vs SomeBees,” in the *Nifty Assignments track of the ACM Special Interest Group on Computer Science Education*, 2014.
- Karl Pichotta and John DeNero. “Identifying Phrasal Verbs Using Many Bilingual Corpora,” in the *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2013.
- Greg Durrett and John DeNero. “Supervised Learning of Complete Morphological Paradigms,” in the *Proceedings of the North American Association of Computational Linguistics*, 2013.
- John DeNero and Aditi Muralhindaran. “Twitter Trends,” in the *Nifty Assignments track of the ACM Special Interest Group on Computer Science Education*, 2013.
- David Golland, John DeNero, and Jakob Uszkoreit. “A Feature-Rich Constituent Context Model for Grammar Induction,” in the *Proceedings of the Association of Computational Linguistics*, 2012.
- Spence Green and John DeNero. “A Class-Based Agreement Model for Generating Accurately Inflected Translations,” in the *Proceedings of the Association of Computational Linguistics*, 2012.
- Mohit Bansal, John DeNero, and Dekang Lin. “Unsupervised Translation Sense Clustering,” in the *Proceedings of the North American Association of Computational Linguistics*, 2012.
- Robert Moore and John DeNero. “L1 and L2 Regularization for Multiclass Hinge Loss Models,” in

- the *Proceedings of the Symposium on Machine Learning in Speech and Language Processing*, 2011.
- John DeNero and Jakob Uszkoreit. “Inducing Sentence Structure from Parallel Corpora for Reordering,” in the *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2011.
- John DeNero and Klaus Machery. “Model-Based Aligner Combination Using Dual Decomposition,” in the *Proceedings of the Association of Computational Linguistics*, 2011.
- John DeNero and Dan Klein. “Discriminative Modeling of Extraction Sets for Machine Translation,” in the *Proceedings of the Association of Computational Linguistics*, 2010.
- John DeNero, Shankar Kumar, Ciprian Chelba, and Franz Och. “Model Combination for Machine Translation,” in the *Proceedings of the North American Association of Computational Linguistics*, 2010.
- John DeNero and Dan Klein. “Teaching Introductory Artificial Intelligence with Pac-Man,” in the *Proceedings of the Symposium on Educational Advances in Artificial Intelligence*, 2010.
- John DeNero and Dan Klein. “The Pac-Man Projects Software Package for Introductory Artificial Intelligence,” in the *Proceedings of the Symposium on Educational Advances in Artificial Intelligence, Model Assignments Track*, 2010.
- Taylor Berg-Kirkpatrick, Alexandre Bouchard-Côté, John DeNero, and Dan Klein. “Painless Unsupervised Learning with Features,” in the *Proceedings of the North American Association of Computational Linguistics*, 2010.
- John DeNero, David Chiang, and Kevin Knight. “Fast Consensus Decoding over Translation Forests,” in the *Proceedings of the Association of Computational Linguistics*, 2009.
- Adam Pauls, John DeNero, and Dan Klein. “Consensus Training for Consensus Decoding in Machine Translation,” in the *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2009.
- John DeNero, Adam Pauls, and Dan Klein. “Asynchronous Binarization for Synchronous Grammars,” in the *Proceedings of the Association of Computational Linguistics, Short Paper Track*, 2009.
- Aria Haghighi, John Blitzer, John DeNero, and Dan Klein. “Better Word Alignments with Supervised ITG Models,” in the *Proceedings of the Association of Computational Linguistics*, 2009.
- John DeNero, Mohit Bansal, Adam Pauls, and Dan Klein. “Efficient Parsing for Transducer Grammars,” in the *Proceedings of the North American Association of Computational Linguistics*, 2009.
- John DeNero, Alex Bouchard-Côté, and Dan Klein. “Sampling Alignment Structure under a Bayesian Translation Model,” in the *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2008.
- John DeNero and Dan Klein. “The Complexity of Phrase Alignment Models,” in the *Proceedings of the Association of Computational Linguistics, Short Paper Track*, 2008.
- John DeNero and Alexandre Bouchard. “A Hierarchical Dirichlet Process Prior for a Conditional Model of Phrase Alignment,” in the *Workshop on Unsupervised Models in NLP, Neural and Information Processing Systems*, 2008.
- Aria Haghighi, John DeNero, and Dan Klein. “A* Search via Approximate Factoring,” in the

Proceedings of American Association of Artificial Intelligence, Nectar Track, 2007.

John DeNero and Dan Klein. “Tailoring Word Alignments to Syntactic Machine Translation,” in the *Proceedings of the Association of Computational Linguistics*, 2007.

Aria Haghighi, John DeNero, and Dan Klein. “Approximate Factoring for A* Search,” in the *Proceedings of the North American Association of Computational Linguistics*, 2007.

John DeNero, Dan Gillick, James Zhang, and Dan Klein. “Why Generative Phrase Models Underperform Surface Heuristics,” in the *Proceedings of the Workshop on Statistical MT*, 2006.

ACTIVITIES

Organizing Committee: EAAI ’13, EAAI ’14, NACLO/ELCLO ’13, NACLO/ELCLO ’14, EMNLP ’15 (Session chair), NAACL ’16 (Demonstration chair), EMNLP ’16 (Area chair, translation)

Conference Reviewer: EMNLP ’08, NAACL ’09, ACL ’09, EMNLP ’09, MT Summit ’09, NAACL ’10, ACL ’10, COLING ’10, ACL ’11, AAAI ’11, EMNLP ’11, NAACL ’12, ACL ’12, EMNLP ’12, EAAI ’12, NAACL ’13, ACL ’13, EMNLP ’13, EACL ’14, NAACL ’15, ACL ’15, EMNLP ’15, SIGCSE ’17, ACL ’19, ACL ’20

EXTERNAL INVITED TALKS	Invited Talk on Data Science and Computer Science Education, The Blake School <i>Host:</i> Jon Friedell	2022
	Invited Talk on Neural Machine Translation, TAUS Massively Multilingual Conference <i>Host:</i> Jaap van der Meer	2022
	Invited Talk on Data Science Education, Ben-Gurion University <i>Host:</i> Samuel Itzikowitz	2022
	Invited Talk on Interactive Machine Translation, Google <i>Host:</i> Melvin Johnson	2018
	Invited Panel Moderator on Neural Machine Translation, TAUS <i>Host:</i> Jaap van der Meer	2017
	Invited Talk on Computer Science Capacity, Google New York <i>Host:</i> Chris Stephenson	2017
	Invited Talk on Data Science Education, San Jose State University <i>Host:</i> Belle Wei	2017
	Invited Talk on Data Science Education, UC San Diego <i>Host:</i> Ramamohan Paturi	2017
	Invited Panelist on Data Science Education, American Statistical Association (ASA) and Association for Computing Machinery (ACM) <i>Host:</i> Steve Pierson	2017
	Invited Talk on Computer Science Capacity, Google <i>Host:</i> Chris Stephenson	2015
	Invited Research Talk, Google Research <i>Host:</i> Jay Ponte	2010

Invited Research Talk, Yahoo! Research <i>Host:</i> Patrick Pantel	<i>2010</i>
Invited Research Talk, Microsoft Research <i>Host:</i> Robert Moore	<i>2010</i>
Invited Research Talk, Toyota Technological Institute, Chicago <i>Host:</i> Karen Livescu	<i>2010</i>
Invited Research Talk, New York University <i>Host:</i> Richard Cole	<i>2010</i>
Invited Research Talk, Johns Hopkins University <i>Host:</i> Mark Dredze	<i>2010</i>
Invited Research Talk, University of Maryland <i>Host:</i> Jimmy Lin	<i>2010</i>
Invited Research Talk, Yahoo! Research <i>Host:</i> Patrick Pantel	<i>2008</i>
Invited Research Talk, Information Sciences Institute <i>Host:</i> Kevin Knight	<i>2008</i>
Invited Research Talk, International Computer Science Institute <i>Host:</i> Dilek Hakkani-Tür	<i>2008</i>
Invited Research Talk, Microsoft Research Asia <i>Host:</i> Mu Li	<i>2008</i>
Invited Research Talk, SRI International <i>Host:</i> Jing Zheng	<i>2008</i>