

karlstratos@gmail.com

karlstratos.com

KARL STRATOS

Employment

Rutgers University, New Brunswick, NJ (September 2019—present)

Assistant Professor

Bloomberg L.P., New York, NY (June 2020—December 2020)

Consultant

Toyota Technological Institute at Chicago, Chicago, IL (June 2017—August 2019)

Research Assistant Professor

Bloomberg L.P., New York, NY (August 2016—May 2017)

Senior Research Scientist

Columbia University, New York, NY (January 2017—May 2017)

Adjunct Assistant Professor

Education

Columbia University, New York, NY (September 2011—June 2016)

Ph.D. in Computer Science (Advisor: [Michael Collins](#))

Thesis: [Spectral Methods for Natural Language Processing](#)

University of Rochester, Rochester, NY (August 2008—May 2011)

Bachelor of Science in Computer Science

Bachelor of Arts in Mathematics

Minor in History

Research Areas

Natural Language Processing, Machine Learning (code at [GitHub](#))

Awards

[ACL Test-of-Time Paper Award](#), 2022

[ICLR Spotlight Paper](#), 2022

Funding

[Google Faculty Research Award](#), 2020

[Bloomberg Data Science Research Grant](#), 2018

[NVIDIA GPU Grant](#), 2017

Publications

Wenzheng Zhang, Wenyue Hua, and Karl Stratos. [EntQA: Entity Linking as Question Answering](#). In Proceedings of ICLR 2022 (**Spotlight**)

Zewei Chu, Karl Stratos, and Kevin Gimpel. [NatCat: Weakly Supervised Text Classification with Naturally Annotated Resources](#). In Proceedings of AKBC 2021

Sam Wiseman, Arturs Backurs, and Karl Stratos. [Data-to-text Generation by Splicing Together Nearest Neighbors](#). In Proceedings of EMNLP 2021

- Ozan Irsoy, Adrian Benton, and Karl Stratos. [Corrected CBOW Performs as well as Skip-gram](#). In Proceedings of EMNLP 2021, Workshop on Insights from Negative Results in NLP
- Zewei Chu, Karl Stratos, and Kevin Gimpel. [Unsupervised Label Refinement Improves Dataless Text Classification](#). In Findings of ACL 2021
- Wenzheng Zhang and Karl Stratos. [Understanding Hard Negatives in Noise Contrastive Estimation](#). In Proceedings of NAACL 2021
- Rajarshi Bhowmik, Karl Stratos, and Gerard de Melo. [Fast and Effective Biomedical Entity Linking Using a Dual Encoder](#). In Proceedings of EACL 2021, Workshop on Health Text Mining and Information Analysis
- Mingda Chen*, Zewei Chu*, Karl Stratos, and Kevin Gimpel. [Mining Knowledge for Natural Language Inference from Wikipedia Categories](#). In Findings of EMNLP 2020
- Karl Stratos and Sam Wiseman. [Learning Discrete Structured Representations by Adversarially Maximizing Mutual Information](#). In Proceedings of ICML 2020
- Shuning Jin, Sam Wiseman, Karl Stratos, and Karen Livescu. [Discrete Latent Variable Representations for Low-Resource Text Classification](#). In Proceedings of ACL 2020
- David McAllester and Karl Stratos. [Formal Limitations on the Measurement of Mutual Information](#). In Proceedings of AISTATS 2020
- Mingda Chen*, Zewei Chu*, Yang Chen, Karl Stratos, and Kevin Gimpel. [EntEval: A Holistic Evaluation Benchmark for Entity Representations](#). In Proceedings of EMNLP 2019
- Sam Wiseman and Karl Stratos. [Label-Agnostic Sequence Labeling by Copying Nearest Neighbors](#). In Proceedings of ACL 2019
- Karl Stratos. [Mutual Information Maximization for Simple and Accurate Part-Of-Speech Induction](#). In Proceedings of NAACL 2019
- Daniel Edmiston and Karl Stratos. [Compositional Morpheme Embeddings with Affixes as Functions and Stems as Arguments](#). In Proceedings of ACL 2018, Workshop on Relevance of Linguistic Structure in Neural NLP
- Karl Stratos. [A Sub-Character Architecture for Korean Language Processing](#). In Proceedings of EMNLP 2017
- Karl Stratos. [Entity Identification as Multitasking](#). In Proceedings of EMNLP 2017, Workshop on Structured Prediction for NLP
- Karl Stratos. [Reconstruction of Word Embeddings from Sub-Word Parameters](#). In Proceedings of EMNLP 2017, Workshop on Subword and Character Level Models in NLP
- Heejin Choi and Karl Stratos. [Handling Large Structural Costs in Neural Networks with Slack Rescaling](#). In Proceedings of ICML 2017, Workshop on Deep Structured Prediction

- Young-Bum Kim, Karl Stratos, and Dongchan Kim. [Adversarial Adaptation of Synthetic or Stale Data](#). In Proceedings of ACL 2017
- Young-Bum Kim, Karl Stratos, and Dongchan Kim. [Domain Attention with an Ensemble of Experts](#). In Proceedings of ACL 2017
- Young-Bum Kim, Karl Stratos, and Ruhi Sarikaya. [Frustratingly Easy Neural Domain Adaptation](#). In Proceedings of COLING 2016
- Young-Bum Kim, Karl Stratos, and Ruhi Sarikaya. [Domainless Adaptation by Constrained Decoding on a Schema Lattice](#). In Proceedings of COLING 2016
- Karl Stratos, Michael Collins, and Daniel Hsu. [Unsupervised Part-Of-Speech Tagging with Anchor Hidden Markov Models](#). In TACL 2016
- Young-Bum Kim, Karl Stratos, and Ruhi Sarikaya. [Scalable Semi-Supervised Query Classification Using Matrix Sketching](#). In Proceedings of ACL 2016
- Karl Stratos, Michael Collins, and Daniel Hsu. [Model-based Word Embeddings from Decompositions of Count Matrices](#). In Proceedings of ACL 2015
- Young-Bum Kim, Karl Stratos, Ruhi Sarikaya, and Minwoo Jeong. [New Transfer Learning Techniques for Disparate Label Sets](#). In Proceedings of ACL 2015
- Young-Bum Kim, Karl Stratos, and Ruhi Sarikaya. [Pre-training of Hidden-Unit CRFs](#). In Proceedings of ACL 2015
- Young-Bum Kim, Karl Stratos, Xiaohu Liu, and Ruhi Sarikaya. [Compact Lexicon Selection with Spectral Methods](#). In Proceedings of ACL 2015
- Karl Stratos and Michael Collins. [Simple Semi-Supervised POS Tagging](#). In Proceedings of NAACL 2015 Workshop on Vector Space Modeling for NLP
- Young-Bum Kim, Minwoo Jeong, Karl Stratos, and Ruhi Sarikaya. [Weakly Supervised Slot Tagging with Partially Labeled Sequences from Web Search Click Logs](#). In Proceedings of NAACL 2015
- YoungHoon Jung, Karl Stratos, and Luca P. Carloni. [LN-Annotate: An Alternative Approach to Information Extraction from Emails using Locally-Customized Named-Entity Recognition](#). In Proceedings of WWW 2015
- Karl Stratos, Do-kyum Kim, Michael Collins, and Daniel Hsu. [A Spectral Algorithm for Learning Class-Based n-gram Models of Natural Language](#). In Proceedings of UAI 2014
- Shay B. Cohen, Karl Stratos, Michael Collins, Dean Foster and Lyle Ungar. [Spectral Learning of Latent-Variable PCFGs: Algorithms and Sample Complexity](#). In JMLR 2014
- Karl Stratos, Alexander M. Rush, Shay B. Cohen, and Michael Collins. [Spectral Learning of Refinement HMMs](#). In Proceedings of CoNLL 2013
- Shay B. Cohen, Karl Stratos, Michael Collins, Dean Foster and Lyle Ungar. [Experiments with Spectral Learning of Latent-Variable PCFGs](#). In Proceedings of NAACL 2013

Shay B. Cohen, Karl Stratos, Michael Collins, Dean Foster and Lyle Ungar. [Spectral Learning of Latent-Variable PCFGs](#). In Proceedings of ACL 2012

Margaret Mitchell, Xufeng Han, Jesse Dodge, Alyssa Mensch, Amit Goyal, Alex Berg, Kota Yamaguchi, Tamara Berg, Karl Stratos, and Hal Daumé III. [Midge: Generating image descriptions from computer vision detections](#). In Proceedings of EACL 2012 (**ACL Test-of-Time Paper Award, 2022**)

Alexander C. Berg, Tamara L. Berg, Hal Daumé III, Jesse Dodge, Amit Goyal, Xufeng Han, Alyssa Mensch, Margaret Mitchell, Aneesh Sood, Karl Stratos, and Kota Yamaguchi. [Understanding and Predicting Importance in Images](#). In Proceedings of CVPR 2012

Karl Stratos, Lenhart K. Schubert, and Jonathan Gordon. [Episodic Logic: Natural Logic + Reasoning](#). In Proceedings of KEOD 2011

Lenhart Schubert, Jonathan Gordon, Karl Stratos, and Adina Rubinoff. [Towards Adequate Knowledge and Natural Inference](#). In Proceedings of AAAI Fall Symposium on Advances in Cognitive Systems 2011

Teaching

2021 Instructor, [Machine Learning](#), Rutgers University, Fall

2021 Instructor, [Natural Language Processing](#), Rutgers University, Spring

2020 Instructor, [Machine Learning](#), Rutgers University, Fall

2020 Instructor, [Natural Language Processing](#), Rutgers University, Spring

2018 Instructor, [Spectral Techniques for Machine Learning](#), Toyota Technological Institute at Chicago, Fall

2018 Co-Instructor (with Suriya Gunasekar), [Introduction to Machine Learning Summer School](#), Toyota Technological Institute at Chicago, Summer

2017 Instructor, [Natural Language Processing](#), Columbia University, Spring

2013 Teaching Assistant, [Advanced Machine Learning](#), Daniel Hsu, Columbia University, Fall

2012 Teaching Assistant, [Machine Learning for Natural Language Processing](#), Michael Collins, Columbia University, Spring

Professional Activities

Senior Service:

Area Chair, AAAI, 2023

Area Chair (Syntax and Parsing), EMNLP, 2022

Action Editor, ACL Rolling Review, 2021—present

Area Chair (Machine Learning), EMNLP, 2021

Area Chair (Machine Learning), ACL-IJCNLP, 2021

Area Chair (Machine Learning), NAACL, 2021

Area Chair (Machine Learning), EACL, 2020

Editorial Board Reviewer, JMLR, 2020—present

Senior Program Committee, AAAI, 2020, 2021, 2022
Standing Reviewer, TACL, 2019—2023
Area Chair (Machine Learning), EMNLP, 2018 (before AC vs SAC distinction)

Conference/Journal Reviewing:

AAAI: 2018, 2019
ACL: 2014, 2015, 2016, 2017, 2018, 2019, 2020
AISTATS: 2018, 2019, 2020, 2021, 2022
COLING: 2017, 2018
CoNLL: 2018, 2019, 2020, 2021, 2022
EACL: 2017
EMNLP: 2015, 2016, 2017, 2019, 2020
ICLR: 2018, 2019, 2020, 2021, 2022
ICML: 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022
IJCNLP: 2017
JMLR: 2018
LREC: 2018, 2020
NAACL: 2015, 2016, 2018
NeurIPS: 2014, 2015, 2017, 2018, 2019, 2021, 2022

Panelist, National Science Foundation, SBIR/STTR Phase I: NLP (July, 2021); AI (November, 2021)

Ad Hoc Reviewer, National Science Foundation, CISE:Core:RI:Small (2019), SBIR/STTR Phase I: Language-Based AI (2020)

Program Committee, *Midwest Machine Learning Symposium*, 2019

Organizing Committee, *Midwest Speech and Language Days*, 2019

Co-Organizer, [Introduction to Machine Learning Summer School](#), Toyota Technological Institute at Chicago, Summer

Program Committee, *ACL Workshop on Relevance of Linguistic Structure in Neural Architectures for NLP*, 2018

Program Committee, *ACL Workshop on Representation Learning for NLP*, {2016, 2017, 2018, 2019, 2020}

Program Committee, *ACL Workshop on Evaluating Vector Space Representations for NLP*, {2016, 2017}

Program Committee, *NAACL Workshop on Vector Space Modeling for NLP*, 2015

NAACL 2013 Tutorial, *Spectral Learning Algorithms for Natural Language Processing* with Shay Cohen, Michael Collins, Dean Foster, and Lyle Ungar

Program Committee, *Mid-Atlantic Student Colloquium on Speech, Language and Learning*, 2012

Presentations

Invited Speaker, *NLP and Text-as-Data Speaker Series, New York University*, 2021

Invited Speaker, *Summer AI Seminar Series, POSTECH*, 2020

Invited Speaker, *Data Science Seminar, Seoul National University*, 2020

Invited Speaker, *Data Science Speaker Series, Bloomberg L.P.*, 2020

Invited Speaker, *Artificial Intelligence & Machine Learning Group, Princeton University*, 2019

Invited Speaker, *3rd International Workshop on Symbolic-Neural Learning*, 2019

Invited Speaker, *Machine Learning Seminar Series, University of Chicago*, 2019

Invited Speaker, *13th Joint CS Seminar at Toyota Technological Institute*, 2017

Invited Speaker, *CSE Seminar Series, University of Notre Dame*, 2017

Invited Speaker, *COLING Young Researchers Symposium on Natural Language Processing*, 2016

Academic Internships

2014 Research Intern, *Google Inc.*, June—August. Worked with Slav Petrov and Emily Pitler.

2013 Research Intern, *Microsoft*, June—August. Worked with Sham Kakade and T. J. Hazen.

2011 Summer Workshop, *Johns Hopkins University Center for Language and Speech Processing*, July—August. Worked with Alexander and Tamara Berg.

2010 Research Intern, *University of Rochester*, June—August. Worked with Lenhart Schubert.