

David Mueller

📍 New York, New York 📩 davidmlr@amazon.com 🌐 damueller.com 💬 davidandym 🎙️ davidandym

Education

| | |
|--|---|
| Johns Hopkins University Ph.D. in Computer Science | Sept 2018 – Feb 2026 <i>Baltimore MD</i> |
| Johns Hopkins University M.Sc. in Computer Science | Sept 2018 – May 2020 <i>Baltimore MD</i> |
| University of Texas B.Sc. in Computer Science | Sept 2012 – May 2016 <i>Austin, TX</i> |

Industry Research

| | |
|--|---|
| Applied Scientist II Amazon | Feb 2026 – Present <i>New York, NY</i> |
| ◦ Member of Stores Foundational AI team, working on Amazon Rufus ◦ Current Focus: Long context compression for LLMs with embeddings | |
| Applied Scientist Amazon | Jan 2025 – Feb 2026 <i>New York, NY</i> |
| ◦ Member of Stores Foundational AI team, working on Amazon Rufus ◦ Focus: Improving and evaluating data for reasoning in large language model pre-training. | |
| Machine Learning Research Intern Netflix | June 2023 – Sept 2023 <i>Los Gatos, CA</i> |
| ◦ Member of the Promotional Media Team ◦ Focus: Multi-modal and multi-task optimization for scene ranking and retrieval | |

Academic Research

| | |
|--|--|
| Research Assistant (Ph.D.) Johns Hopkins University - <i>Advisors: Mark Dredze & Nicholas Andrews</i> | Sept 2018 – Dec 2024 <i>Baltimore, MD</i> |
| ◦ Center for Language and Speech Processing (CLSP) ◦ PhD Thesis: Understanding multi-task optimization and transfer in deep neural networks | |
| Undergraduate Research Assistant University of Texas - <i>Advisor: Greg Durrett</i> | June 2017 – June 2018 <i>Austin, TX</i> |
| ◦ Focus: Efficient models for entity linking in noisy contexts | |

Peer-Reviewed Publications

- [1] [David Mueller](#), Mark Dredze, Nicholas Andrews, Can Optimization Trajectories Explain Multi-Task Transfer?, *Transactions on Machine Learning Research* (TMLR 2024)
- [2] Suzanna Sia, [David Mueller](#), Kevin Duh, Where does In-Context Learning Happen in Large Language Models?, *Proceedings of the 38th Conference on Neural Information Processing Systems* (NeurIPS 2024)
- [3] [David Mueller](#), Mark Dredze, Nicholas Andrews, Multi-Task Transfer Matters During Instruction-Tuning, *Findings of the 62nd Annual Meeting of the Association for Computational Linguistics* (ACL 2024)
- [4] [David Mueller](#), Mark Dredze, Nicholas Andrews, The Importance of Temperature in Multi-Task Learning, *Optimization for Machine Learning Workshop @ NeurIPS 2022* (OPT 2022)
- [5] [David Mueller](#), Nicholas Andrews, Mark Dredze, Do Text-to-Text Multi-Task Learners Suffer from Task-Conflict?, *Findings of the 2022 Conference on Empirical Methods in Natural Language Processing* (EMNLP 2022)

- [6] Steven Reich, [David Mueller](#), Nicholas Andrews, Ensemble Distillation for Structured Prediction, *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing* (EMNLP 2020)
- [7] [David Mueller](#), Nicholas Andrews, Mark Dredze, Sources of Transfer in Multilingual Named Entity Recognition, *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (ACL 2020)
- [8] [David Mueller](#), Greg Durrett, Effective Use of Context in Noisy Entity Linking, *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing* (EMNLP 2018)

Pre-Prints and Ongoing Work

- [9] Sophia Hager, [David Mueller](#), Kevin Duh, Nicholas Andrews, Uncertainty Distillation: Teach Language Models to Express Semantic Confidences, *Pre-print, Under submission to ICML* (2025).
- [10] Kaustubh Dhole, [David Mueller](#), Neeraj Varshney, Sanket Laogenkoar, Colin Lockard, Nasser Zalmout, Xian Li, Unlocking Math Reasoning at Scale: Choosing the Most “Influential” Datasets through Small Scale Experiments, *Pre-print, Under submission to ACL* (2025).

Invited Talks

- Can Optimization Trajectories Explain Multi-Task Transfer?** Sept 2024
Amazon Search (Stores Foundational AI Group) Remote
- Where Does Negative Transfer Come From? On the Implicit Bias of SGD in Multi-Task Learning** Sept 2023
Johns Hopkins University *Baltimore, MD*
o Video Recording 

Teaching Experience

- Teaching Assistant (TA) for Machine Learning (CS 601.475)** Fall 2019
Johns Hopkins University (*with Prof. Mark Dredze*)
o Responsible for lectures, office hours, homework & exam writing, covering graduate-level machine learning topics and algorithms.

Community Service

- Graduate Admissions Committee** 2020 - 2024
Center for Language and Speech Processing (CLSP)
- Graduate Admissions Mentor** 2024
CLSP Application Support Program

Reviewer Service

- Neural Information Processing Systems (*NeurIPS*) 2022 - 2025
- International Conference on Machine Learning (*ICML*) 2020 - 2025
- International Conference on Learning Representations (*ICLR*) 2020 - 2025
- Association for Computational Linguistics Rolling Review (*ARR*) 2022 - 2025
- International Conference on Artificial Intelligence and Statistics (*AISTATS*) 2022 - 2025
- Conference on Empirical Methods in Natural Language Processing (*EMNLP*) 2019 - 2022
- Annual Meeting of the Association for Computational Linguistics (*ACL*) 2019 - 2022
- Conference on Language Modeling (*COLM*) 2024
- Workshop on Optimization for Machine Learning (*OPT @ NeurIPS*) 2023

Prior Industry Experience

Full-Stack Software Developer
Cars.com

May 2016 – Sept 2018
Austin, TX

- Responsible for back-end database design and multithread data-processing using graph databases and Go.