

Melanie Subbiah

New York, NY, 415-350-4761, melanie.s.subbiah@gmail.com

I enjoy creating AI systems that are capable of creative behavior and problem solving. I value ethical and sustainable approaches to AI research.

EDUCATION

Doctor of Philosophy in Computer Science Sep. 2020 - present
Master of Science in Computer Science Sep. 2020 - May 2022

Columbia University, New York, NY

Advisor: Professor Kathleen McKeown

Honors:

Amazon CAIT PhD Fellowship 2023
NSF GRFP Honorable Mention 2021 (Comp/IS/Eng - Natural Language Processing)
Best Paper Award at NeurIPS 2020 (*Language Models are Few-Shot Learners*)
Presidential and SEAS Fellowship 2020

Bachelor of Arts in Computer Science Sep. 2013 - Jun. 2017

Williams College, Williamstown, MA

Thesis: *Using Text Abstraction and LSTM Language Models for Domain-Independent Narrative Generation*

Advisor: Professor Andrea Danyluk

Honors:

Commencement Phi Beta Kappa Student Speaker
Highest Honors and Best Thesis Presentation award in Computer Science
Magna Cum Laude, Phi Beta Kappa, Sigma Xi
Honorable Mention award in short story writing

EXPERIENCE

Meta, New York, NY Jun. 2022 - Aug. 2022

Research Intern, AI for Augmented Reality Input & Interaction

- Developed deep learning strategies to interpret EMG data for human-computer neural interfaces.
- Conducted an independent research project for the team (formerly CTRL Labs).

OpenAI, San Francisco, CA Aug. 2019 - Aug. 2020

Member of Technical Staff, Language

- Designed and built the evaluation suite for GPT-3 and engineered the model to perform efficiently across many tasks, co-first-author on the GPT-3 paper (*Language Models are Few-Shot Learners*).
- Consulted with customers following the OpenAI API launch, which provided few-shot access to pretrained language models.

Apple, Cupertino, CA

Machine Learning Engineer, AI Research

Sep. 2017 - May 2019

- Directed research in autonomous systems and generative language models.
- Led projects on data center HVAC efficiency, reward function design, domain randomization for sim-to-real transfer in computer vision, and effective QA methods for crowdsourced annotated data.

RESEARCH

G Wang, L Chillrud, K Harwood, A Ananthram, **M Subbiah**, K McKeown. *Check-COVID: Fact-Checking COVID-19 News Claims with Scientific Evidence*. Findings of Association for Computational Linguistics (ACL); Jul 2023; Toronto, CA.

A Storek, **M Subbiah**, K McKeown. *Unsupervised Selective Rationalization with Noise Injection*.

Association for Computational Linguistics (ACL); Jul 2023; Toronto, CA.

M Subbiah*, A Bhattacharjee*, Y Hua, TS Kumarage, H Liu, K McKeown. *Detecting Harmful Agendas in News Articles*. WASSA Workshop at ACL; Jul 2023; Toronto, CA.

S Levy, E Allaway, **M Subbiah**, L Chilton, D Patton, K McKeown, WY Wang. *SafeText: A Benchmark for Exploring Physical Safety in Language Models*. Empirical Methods in Natural Language Processing (EMNLP); Dec 2022; Abu Dhabi, UAE.

A Mei, A Kabir, S Levy, **M Subbiah**, E Allaway, JN Judge, D Patton, B Bimber, K McKeown, WY Wang. *Mitigating Covertly Unsafe Text within Natural Language Systems*. Findings of Empirical Methods in Natural Language Processing (EMNLP); Dec 2022; Abu Dhabi, UAE.

M Subbiah, K McKeown. *Understanding Identity Signalling in Persuasive Online Text*. Presented at International Workshop on Social Sensing at ICWSM; Jun 2021; Online.

TB Brown*, B Mann*, N Ryder*, **M Subbiah***, et al. *Language Models are Few-Shot Learners*. Neural Information Processing Systems (NeurIPS, **Best Paper Award**); Dec 2020; Montreal, CA.

M Subbiah, J lesser, N Apostoloff. *Augmenting Training Data with Simulated Images*. Presented at Women in Machine Learning Workshop at NeurIPS; Dec 2018; Montreal, CA.

M Maher, **M Subbiah**, N Apostoloff. *Cascaded Dataset QA*. Women in Machine Learning Workshop at NeurIPS; Dec 2018; Montreal, CA.

UC Nygaard, Z Li, T Palys, B Jackson, **M Subbiah**, M Malipatlolla, V Sampath, H Maecker, MR Karagas, KC Nadeau. *Cord blood T cell subpopulations and associations with maternal cadmium and arsenic exposures*. PLoS One. 2017 Jun 29;12(6):e0179606. doi: 10.1371/journal.pone.0179606. PMID: 28662050; PMCID: PMC5491028.

Red Teaming for GPT-4. 2023.

Co-organizer of ICLR Workshop on Enormous Language Models. 2021.

SPEAKING

Pearson Publishing's AI Catalyst Conference: NLP with ChatGPT. 2023.

SuperDataScience podcast: GPT-3 for Natural Language Processing. 2022.

Wired 5 Levels of Difficulty video: Machine Learning. YouTube. 2021.

Invited seminar talks on GPT-3 at Stanford University and New York University. 2020.

Invited Computer Science colloquium talk at Williams College. 2020.

Williams College Commencement Phi Beta Kappa speaker. 2017.

TEACHING

Full instructor for *Discrete Mathematics*. Columbia University. Summer Session A 2023.

Guest lecture for *Computational Journalism*. Columbia University. Spring 2023.

TA (and a lecture) for *Natural Language Generation and Summarization*. Columbia University. Fall 2022.

Tutor for *Introductory Computer Science* and *Discrete Math*. Columbia University Athletics. 2021.

TA for *Introductory Computer Science* and *Data Structures*. Williams College. 2014-2017.

MENTORING

Williams College CS Undergraduate Buddy	2022
OpenAI Scholars Program Development	2020

OpenAI Scholars Program Mentor	2020
Institute of International Education TechWomen Mentor	2019
Williams College Underrepresented Identities in CS Leader	2016-2017