Michael J. Ryan

Palo Alto, CA michaeljryan@stanford.edu Google Scholar michryan.com michael-ryan-207
XenonMolecule michaelryan207

Education

Stanford University *Sept 2023 – Jun 2025**
M.Sc. in Computer Science

- GPA: 4.134/4.0
- Advised by: Dr. Diyi Yang
- **Research:** Natural Language Processing (NLP), Large Language Models (LLMs), Language Model Programs, Post-Training, LLM Personalization
- *Expected Graduation June 2025

Georgia Institute of Technology *Sept 2019 – May 2023*
B.Sc. in Computer Science

- GPA: 3.96/4.0
- Advised by: Dr. Wei Xu
- **Research:** Natural Language Processing (NLP), Text Simplification, Fairness
- Thesis title: *A Survey of Non-English Parallel Corpora for Text Simplification*

Awards and Achievements

2024	**Best Social Impact Paper Award**, ACL 2024
	Outstanding Project Award, CS224N Natural Language Processing with Deep Learning
	Best Project Award, CS330 Deep Multi-task and Meta Learning
	Top 5% CA in Stanford CS Department (awarded twice), CS221 Artificial Intelligence Principles and Techniques
2023	**Outstanding Paper Honorable Mention**, ACL 2023
	Guaranteed Course Assistanceship Funding, Stanford University
2022	**Distinction in Research**, Georgia Tech Honors Program
	Outstanding Undergraduate TA for Interactive Computing, Georgia Tech Center for Teaching and Learning
2019-23	**Dean's List**, Georgia Tech

Academic Research Experience

DSPy Optimizers Team *Stanford, CA*
Dr. Chris Potts, Omar Khattab, Stanford University *Dec 2023 – Present*

Social and Language Technologies (SALT) Lab *Stanford, CA*
Dr. Diyi Yang, Stanford University *Sept 2023 – Present*

NLP X Lab *Atlanta, GA*
Dr. Wei Xu, Georgia Institute of Technology *Jan 2021 – Sept 2023*

Industry Research Experience

Research Intern, Snowflake *San Mateo, CA*
Dr. Daniel Campos, Danmei Xu *Jun 2024 – Sept 2024*

Industry Experience

Software Engineering Intern, Microsoft *Redmond, WA*
Windows Servicing and Delivery: Operating System Security Team *May 2022 – Aug 2022*

- Designed and programmed a static analysis tool in C++ for identifying security vulnerabilities throughout Windows OS.

Software Engineering Intern, Microsoft *Virtual*
Windows Servicing and Delivery: Toolkit Team *May 2021 – Aug 2021*

- Updated tooling for porting Windows Updates across versions to run as serverless Azure functions.

Software Engineering Intern, Uber *Virtual*
New Modalities (NeMo) Team *May 2021 – Aug 2021*

- Implemented end-to-end testing service in GoLang for bike, scooter, and moped rentals using virtual vehicles.

Publications and Pre-Prints

ARXIV 2024 Distilling an End-to-End Voice Assistant Without Instruction Training Data.
W Held, M Li, **MJ Ryan**, W Shi, Y Zhang, D Yang

EMNLP 2024 Optimizing Instructions and Demonstrations for Multi-Stage Language Model Programs.
K Opsahl-Ong*, **MJ Ryan***, J Purtell, D Broman, C Potts, M Zaharia, O Khattab
*Equal Contribution

EMNLP 2024 Towards Massively Multi-Domain Multilingual Readability Assessment.
T Naous, **MJ Ryan**, A Lavrouk, M Chandra, W Xu

ACL 2024 Unintended Impacts of LLM Alignment on Global Representation.
MJ Ryan, W Held, D Yang

ACL 2024 Having Beer After Prayer? Measuring Cultural Bias in Large Language Models.
T Naous, **MJ Ryan**, A Ritter, W Xu
Best Social Impact Paper Award

ACL 2023 Revisiting non-English Text Simplification: A Unified Multilingual Benchmark.
MJ Ryan, T Naous, W Xu
Outstanding Paper Honorable Mention

MIT IEEE URTC 2018 Cloud Computed Machine Learning Based Real-Time Litter Detection Using Micro-UAV Surveillance.
A Chung, DY Kim, E Kwok, **M Ryan**, E Tan, R Gamadia

Talks and Presentations

Title:	**DSPy: Prompt Optimization for LM Programs**
Nov 2024	EMNLP 2024
Sep 2024	Transformers at Work 2024
Sep 2024	Bay Area AI
Sep 2024	Snowflake
Jul 2024	UC Berkeley LLM Meetup
Title:	**The Unintended Consequences of Preference Tuning LLMs**
Sep 2024	The Digitalist Papers: Artificial Intelligence and Democracy in America (Stanford University)
Feb 2024	Carnegie Mellon University - Qatar
Jan 2024	Snowflake
Title:	**A Survey of Non-English Parallel Corpora for Text Simplification**
Apr 2023	Georgia Tech Undergraduate Research Symposium

Press Coverage

Nov 2024	Teaser: Unintended Impacts of Alignment on Global Representation Stanford AI Lab Blog
Aug 2024	The Challenge of Aligning AI ChatBots Stanford HAI Blog
Jul 2024	An AI walks into a bar... Can artificial intelligence be genuinely funny? BBC
Jun 2024	MIPRO: A Novel Optimizer that Outperforms Baselines on Five of Six Diverse Language Model LM Programs Using a Best-in-Class Open-Source Model (Llama-3-8B) by 12.9% accuracy MarkTechPost

Mar 2024 LLMs exhibit significant Western cultural bias, study finds VentureBeat

Open Source Software/Data

DSPy MIPROv2 Optimizer StanfordNLP/DSPy
The MIPROv2 optimizer for DSPy improves all the prompts in a multistage LM program by proposing several instruction rewrites and demonstrations and using a bayesian optimization to find the optimal combination.

AskReddit Countries Dataset SALT-NLP/unintended-impacts-of-alignment
The AskReddit Countries Dataset contains 554 question/answer templates collected from r/AskReddit for asking questions *about* different countries. Example: "Which country has the best food?" Such templated questions can be used to measure LLM and Reward Model opinions about countries.

MultiSim Benchmark XenonMolecule/MultiSim
The MultiSim benchmark is a growing collection of text simplification datasets targeted at sentence simplification in several languages. Currently, the benchmark spans 27 resources in 12 languages.

CAMeL Dataset TarekNaous/camel
CAMeL (Culture Appropriateness Measure set for LLMs) is a collection of prompts and culturally relevant entities that measure LLMs' biases toward particular cultures. Specifically, the original version of CAMeL measures Arabic vs. Western bias.

Teaching Experience

Sp 2024 **CS221: Artificial Intelligence Principles and Techniques (Head CA)**
Dr. Nima Anari, Dr. Moses Charikar, Dr. Sanmi Koyejo, Stanford University

Wi 2024 **CS124: From Languages to Information**
Dr. Dan Jurafsky, Stanford University

Fa 2023 **CS221: Artificial Intelligence Principles and Techniques**
Dr. Percy Liang, Dr. Dorsa Sadigh, Stanford University

2021-23 **CS3600: Introduction to Artificial Intelligence (Head TA)**
Dr. Mark Riedl, Dr. James Rehg, Georgia Institute of Technology

Service

2024 **TSAR 2024 @ EMNLP** Reviewer

2020-22 **GT Honors Program Application Review Committee**
Bits of Good Web Development for Atlanta Non-profit Organizations