# Ivory Yang

Hanover, NH · 347-251-0374 · ivory.yang.gr@dartmouth.edu
linkedin.com/in/ivoryayang/ · https://ivoryayang.github.io/

## Education

**Dartmouth College,** Hanover, NH **Sep 2023 - June 2028 (expected)**
*Ph.D., Computer Science*
*Masters of Science, Computer Science (Awarded May 2025)*
**GPA 4.0/4.0**
Relevant Coursework: Machine Learning, Deep Learning, Artificial Intelligence
Awards: Neukom Research Award, Thomas D. Sayles Research Award, Alumni Research Award, Irving Research Grant

**University of Michigan**, Ann Arbor, MI **Sep 2016 - May 2020**
*Bachelors of Business Administration, Bachelors of Science (Cognitive Science)* **GPA 3.82/4.0**
Honors/Awards: UM Pan-Asia Scholar, James B. Angell Scholar, Global Experience Scholar, University Honors
Activities: Michigan Stocks and Bonds Organization, Equestrian Team, Alpha Omicron Pi Sorority, HEC Paris

**Graduate Coursework**, San Francisco, CA **Dec 2022**
Stanford University - Computer Organization & Systems (CS107) **GPA 4.0/4.0**
Harvard University - Introduction to CS (CS50), Data Structures & Algorithms (CS124)
UC San Diego - Discrete Math (CSE-41243), Linear Algebra (CSE-40023), Intermediate Programming with Objects (CSE-40477)

## Research Interests

As a researcher, I am passionate about the application of NLP and ML to address socially impactful challenges. My current research areas of focus are:

- Computational tools for the Nüshu language; I created the world's first digitized dataset and LLM-based learning framework
- Multilingual capabilities of LLMs; culturally aware adaptation for low-resource and endangered languages
- Resilience of large AI system safety features against persuasion and behavioral-based steering

## Publications

**Recontextualizing Revitalization: A Mixed Media Approach to Reviving the Nüshu Language**
**Ivory Yang**, Xiaobo Guo, Yuxin Wang, Hefan Zhang, Yaning Jia, William Dinauer, Soroush Vosoughi
In *The 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP 2025)* [Main]

**Visibility as Survival: Generalizing NLP for Native Alaskan Language Identification**
**Ivory Yang**, Chunhui Zhang, Yuxin Wang, Zhongyu Ouyang, Soroush Vosoughi
In *The 63rd Annual Meeting of the Association for Computational Linguistics (ACL 2025)* [Findings]

**Is it Navajo? Accurate Language Detection in Endangered Athabaskan Languages**
**Ivory Yang**, Weicheng Ma, Chunhui Zhang, Soroush Vosoughi
In *The 2025 Annual Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics (NAACL 2025)* [Main, Oral Presentation, Top 2.88%]

**NüshuRescue: Revitalization of the Endangered Nüshu Language with AI**
**Ivory Yang**, Weicheng Ma, Soroush Vosoughi
In *The 31st International Conference on Computational Linguistics (COLING 2025)* [Main, Oral Presentation]

**MentalManip: A Dataset for Fine-grained Analysis of Mental Manipulation in Conversations**
Yuxin Wang, **Ivory Yang**, Saeed Hassanpour, Soroush Vosoughi
In *The 62nd Annual Meeting of the Association for Computational Linguistics (ACL 2024)* [Main, Oral Presentation, Top 3.10%]

**Communication is All You Need: Persuasion Dataset Construction via Multi-LLM Communication**
Weicheng Ma, Hefan Zhang, **Ivory Yang**, Shiyu Ji, Joice Chen, Farnoosh Hashemi, Shubham Mohole, Ethan Gearey, Michael Macy, Saeed Hassanpour, Soroush Vosoughi
In *The 2025 Annual Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics (NAACL 2025)* [Main, Oral Presentation, Top 2.88%]

**Enhancing LLM-Based Persuasion Simulations with Cultural and Speaker-Specific Information**

Weicheng Ma, Hefan Zhang, Shiyu Ji, Farnoosh Hashemi, Qichao Wang, **Ivory Yang**, Joice Chen, Juanwen Pan, Michael Macy, Saeed Hassanpour, Soroush Vosoughi

In *The 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP 2025)* [Findings]

## Research & Work Experience

**Algoverse**, Palo Alto, CA **Feb 2025-April 2025**

*Machine Learning Researcher*

- Led seven interdisciplinary student teams conducting applied research in machine learning and NLP, with six papers completed and under review
- Designed and implemented frameworks to increase linguistic data accessibility and contribute to the preservation of underrepresented languages

**Minds, Machines and Society Lab**, Hanover, NH **Jan 2024-Present**

*Machine Learning Research Assistant*

- Supervised by Professor Soroush Vosoughi as part of the Minds, Machines and Society Group, conducting research in the field of machine learning so as to develop computational tools that offer new perspectives on social systems and issues
- Engaged in natural language processing (NLP) and machine learning research, specifically exploring large language models (LLMs) to detect manipulation tactics in speech, so as to harness findings to develop automatic systems to properly handle and mitigate verbal mental manipulation
- Currently exploring LLM generalization capabilities for revitalization of endangered and low-resource languages, with a focus on Nüshu and Native American languages under the Athabaskan family (Navajo, Apache etc.), so as to enhance linguistic data accessibility and preservation.

**Graph Mind Lab**, Hanover, NH **Sep 2024 - June 2025**

*Machine Learning Research Assistant*

- Worked with Professor Yujun Yan on incorporating graph representations into LLM role-play debates to analyze the structure and patterns of persuasive interactions, so as to unveil novel insights into discourse structure and strategies for social AI applications
- Currently developing a heterogeneous graph explainer to enhance the transparency and interpretability of role-play debates, leveraging graph neural networks to model and explain persuasive dynamics in LLM interactions

**Supervised Program for Alignment Research (SPAR),** Berkeley, CA **Mar 2024 - June 2024**

*Machine Learning Research Intern*

- Worked on alignment research with a focus on AI safety and mechanistic interpretability, contributing to the understanding of activation steering vectors and further development of LLM defense mechanisms
- Conducted technical experiments such as testing of refusal dataset with Contrastive Activation Addition (CAA) using LLaMa-2 models, so as to determine the optimal layer for inserting steering vectors to improve model defense performance

## Featured Press & Publicity

**Language Preservation Efforts Get an AI Boost**, [Dartmouth News, 2025]

**Online AI Translators Could Easily Learn Navajo and Related Languages, Study Suggests** [phys.Org, 2025]

**AI Model Could Add Navajo And Related Languages To Online Translators** [ScienceBlog, 2025]

## Honors & Awards

**Irving Institute Research Grant ($5700)**, Dartmouth Arthur L. Irving Institute for Energy & Society

**Thomas D. Sayles Research Award ($3500)**, Dartmouth Ethics Institute

**Neukom Outstanding Research Award ($1000)**, Dartmouth Neukom Institute

**Alumni Research Award ($1000)**, Dartmouth College

**Travel Grant ($2000)**, Widening Natural Language Processing (WiNLP) at EMNLP 2024

**Travel Grant ($1000, $1300, $2500)**, Dartmouth Women in CS (WiCS)

**Guarini Travel Grant ($1000)**, Dartmouth College

**Citation in COSC 189.23 (Top 2% of class)**, Dartmouth College

**Citation in COSC 276 (Top 2% of class)**, Dartmouth College

**Pan-Asia Scholar ($4000)**, University of Michigan
**Global Experience Scholar ($1500)**, University of Michigan
**James B. Angell Scholar**, University of Michigan
**University Honors**, University of Michigan

## Invited Talks & Service

**MIT,** 24.S90 Demystifying Large Language Models (Special feature)
**UC San Diego,** Swartz Center for Computational Neuroscience (Invited talk)
**Language Technologies for All 2025 (LT4All)**, UNESCO (Invited talk)
**Amherst College** (Guest lecture)
**Peer Review Service:** NAACL SRW 2025, ACL 2025, ICWSM 2025, EMNLP 2025
**Program Committee Service**: AmericasNLP 2026

## Teaching experience

**Artificial Intelligence (COSC 276)**, Dartmouth College **Sep 2025 - Dec 2025**
*Teaching Assistant, Instructor: Professor Soroush Vosoughi*

**Artificial Intelligence (COSC 276)**, Dartmouth College **Sep 2024 - Dec 2024**
*Teaching Assistant, Instructor: Professor Soroush Vosoughi*

**Ross Integrative Semester (RIS)**, University of Michigan **Sep 2019 - Dec 2019**
*Teaching Assistant*

## Skills & Interests

Programming Languages/Tech: C, C++, Python
Languages: Mandarin (Fluent), French (Conversational), Korean (Conversational)
Lived in six countries, took a gap year before college to backpack across Asia