

# William B. Held

MACHINE LEARNING (ML) & NATURAL LANGUAGE PROCESSING (NLP)

☎ (+1) 646-961-2627 | ✉ wheld3@gatech.edu | 🏠 www.williamheld.com | 📱 Helw150 | 🌐 williambarrheld

## Education

### Georgia Institute of Technology

Atlanta, GA, United States

3.83 GPA Ph.D. in Machine Learning

Aug. 2021 - PRESENT

- Advised by Prof. Diyi Yang; Research on NLP and ML for Low-Resource Languages, Language Variants, and Dialects
- NSF Graduate Research Fellowship Program Honorable Mention Recipient.

### New York University Abu Dhabi

Abu Dhabi, United Arab Emirates

3.73 GPA B.S. in Computer Science with a concentration in Economics

Aug. 2015 - May 2019

- Research focused on semantic relation extraction and Arabic computational linguistics under Prof. Nizar Habash

## Highlighted Work

### TADA : Task Agnostic Dialect Adapters for English

William Held, Caleb Ziems, Diyi Yang

61st Annual Meeting of the Association for Computational Linguistics

2023

### Can Large Language Models Transform Computational Social Science?

Caleb Ziems, William Held, Omar Shaikh, Jiaao Chen, Zhehao Zhang, Diyi Yang

Journal of Computational Linguistics

2023

### Multi-VALUE: A Framework for Cross-Dialectal English NLP

William Held 🔥, Caleb Ziems 🔥, Jingfeng Yang, Jwala Dhamala, Rahul Gupta, Diyi Yang

🔥 Equal Contribution

61st Annual Meeting of the Association for Computational Linguistics

2023

### DAMP: Doubly Aligned Multilingual Parser for Task-Oriented Parsing

William Held, Chris Hidey, Fei Liu, Eric Zhu, Rahul Goel, Diyi Yang, Rushin Shah

61st Annual Meeting of the Association for Computational Linguistics

2023

## Professional Experience

### Stanford University

Palo Alto, CA, United States

Visiting Student Researcher - NLP Group

August 2023-PRESENT

### Meta AI

Research Scientist Intern - GenAI Research

Menlo Park, CA, United States

May - August 2023

- Worked on methods to improve the Multilinguality of the LLaMA Language Models.

### Google Research

New York, NY, United States

Research Intern - Neural Semantic Parsing

May - August 2022

- Worked on methods to help extend Google Assistant's Multilingual capabilities. Work published in our paper: "DAMP".

### Sunshine/Lumi Labs

Palo Alto, CA, United States

Software Engineer - Machine Learning

June 2019 - June 2021

- Worked with Prof. Dan Jurafsky on coreference resolution. Work published in our paper: "Focus on What Matters".
- 9th engineer hired to build out AI-powered applications focused on contact management

### Foursquare

New York, NY, United States

Machine Learning Engineer Intern - Pilgrim Data Team

May 2018 - Aug 2018

### Knotch

New York, NY, United States

Data Scientist - Knowledge Search Engine

May 2017 - February 2018

### Quorum Analytics

Washington, D.C., United States

Software Engineering Intern - User Accounts and Permissions

January - May 2017

### Sandia National Laboratories

Albuquerque, NM, United States

Research Intern - Computer Science Research Institute

April - August 2015 & 2016

## Service

**Reviewer** ACL 2024 NeurIPS 2022-2024, ICML 2023EMNLP 2023, NAACL 2024, EACL 2023-2024

**Volunteer** ACL 2024 Student Research Workshop

**Organizer** 2021 Georgia Tech Graduate Application Support (GT-GAS)

**Volunteer Coordinator** The 2021 Conference on Empirical Methods in Natural Language Processing

## Languages

**Programming** Scala(Spark), Python(PyTorch, Tensorflow), Javascript(React) **Natural** English, Arabic

# Mentorship

## Michael Ryan

Master's Student

- Worked with Michael on his first project with the SALT lab, resulting in an Arxiv Pre-Print in submission.

Stanford University

August 2023-PRESENT

## Seonghee Lee

Master's Student

Stanford University

August 2023-PRESENT

## Vyoma Raman

Master's Student

Stanford University

August 2023-PRESENT

## Faye Holt

Undergraduate Student

Georgia Tech

August 2022-PRESENT

- Worked with Faye at Georgia Tech, resulting in a her publication featured as an oral presentation at SouthNLP.

## Zedian Xiao

Master's Student

Stanford University

January 2023-May 2023

- Worked with Zedian at Stanford, helping him get his first publication in NLP at the EMNLP 2023 Main Conference.

## Yanchen Liu

Master's Student

Harvard University

January 2023-May 2023

- Worked with Yanchen on his first project with the SALT lab, resulting in a publication at the EMNLP 2023 Main Conference.

# All Publications

## Unintended Impacts of LLM Alignment on Global Representation

Michael Ryan, William Held, Diyi Yang

Pre-Print

2024

## Perceptions of Language Technology Failures from South Asian English Speakers

Faye Holt, William Held, Diyi Yang



Equal Contribution

Oral Presentation (Non-Archival Conference) - South NLP Symposium

2024

## A Material Lens on Coloniality in NLP

William Held, Camille Harris, Michael Best, Diyi Yang

Pre-Print

2023

## Task-Agnostic Low-Rank Adapters for Unseen English Dialects

Zedian Xiao, William Held, Yanchen Liu, Diyi Yang

The 2023 Conference on Empirical Methods in Natural Language Processing

2023

## DADA: Dialect Adaptation via Dynamic Aggregation of Linguistic Rules

Yanchen Liu, William Held, Diyi Yang

The 2023 Conference on Empirical Methods in Natural Language Processing

2023

## TADA : Task Agnostic Dialect Adapters for English

William Held, Caleb Ziems, Diyi Yang

61st Annual Meeting of the Association for Computational Linguistics

2023

- Findings, Short Paper Acceptance Rate: 35.58%

## Can Large Language Models Transform Computational Social Science?

Caleb Ziems, William Held, Omar Shaikh, Jiaao Chen, Zhehao Zhang, Diyi Yang

Journal of Computational Linguistics

2023

## Modeling Cross-Cultural Pragmatic Inference with Codenames Duet

Omar Shaikh, Caleb Ziems, William Held, Aryan Pariani, Fred Morstatter, Diyi Yang

61st Annual Meeting of the Association for Computational Linguistics

2023

- Findings, Long Paper Acceptance Rate: 41.89%

## Multi-VALUE: A Framework for Cross-Dialectal English NLP

William Held, Caleb Ziems, Jingfeng Yang, Jwala Dhamala, Rahul Gupta, Diyi Yang



Equal Contribution

61st Annual Meeting of the Association for Computational Linguistics

2023

- Main Conference, Long Paper Acceptance Rate: 20.75%

## DAMP: Doubly Aligned Multilingual Parser for Task-Oriented Parsing

William Held, Chris Hidey, Fei Liu, Eric Zhu, Rahul Goel, Diyi Yang, Rushin Shah

61st Annual Meeting of the Association for Computational Linguistics

2023

- Main Conference, Long Paper Acceptance Rate: 20.75%

## **On Second Thought, Let's Not Think Step by Step! Bias and Toxicity in Zero-Shot Reasoning**

Omar Shaikh, Hongxin Zhang, **William Held**, Michael Bernstein, Diyi Yang

*61st Annual Meeting of the Association for Computational Linguistics*

2023

- Main Conference, Long Paper Acceptance Rate: 20.75%

## **Shapley Head Pruning: Identifying and Removing Interference in Multilingual Transformers**

**William Held**, Diyi Yang

*17th Conference of the European Chapter of the Association for Computational Linguistics*

2022

- Main Conference, Long Paper Acceptance Rate: 24.7%

## **Focus on what matters: Applying Discourse Coherence Theory to Cross Document Coreference**

**William Held**, Dan Iter, Dan Jurafsky

*Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*

2021

- Selected for Oral Presentation, Long Paper Acceptance Rate: 24.6%

## **The Effectiveness of Simple Hybrid Systems for Hypernym Discovery**

**William Held**, Nizar Habash

*Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*

2019

- Short Paper Acceptance Rate: 18.2%

## **Transitioning Green-Gauss Gradients to the Kokkos Framework**

**William Held**, Andrew Bradley

*Sandia National Labs Center for Computing Research Summer Proceedings*

2016

## **A Testing Framework for a Hybrid Triangular Solver**

**William Held**, Andrew Bradley

*Sandia National Labs Center for Computing Research Summer Proceedings*

2015