

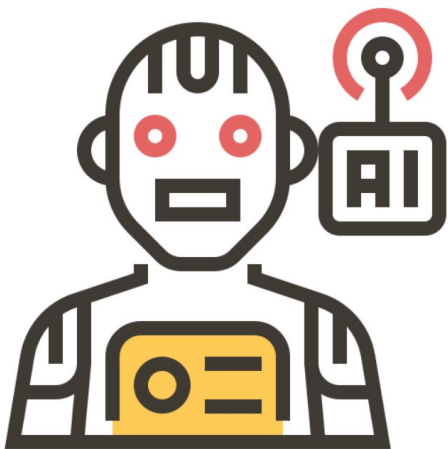
# Building Generalizable and Inclusive Language Processing Models Across Tasks and User Groups

**Da Yin**



Department of Computer Science  
UCLA





Generalizable Across Tasks

## Application Tasks

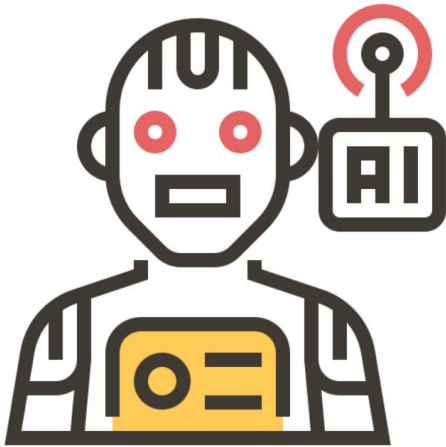
Open-Domain QA

Commonsense Reasoning

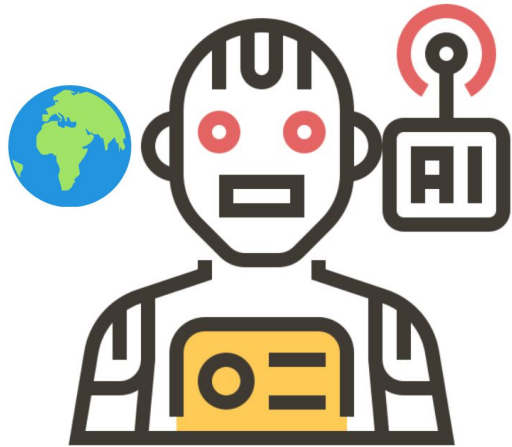
Dialogue Generation

...

## User Groups



Western Countries



## User Groups

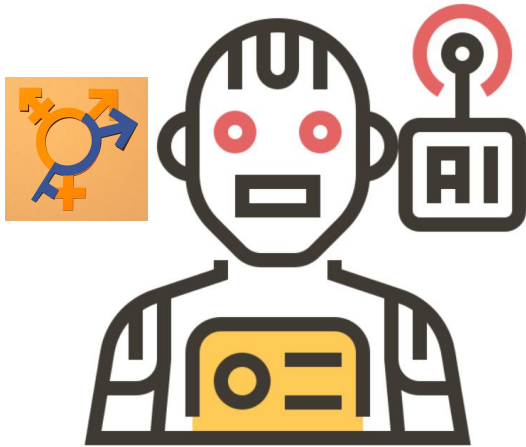
Western Countries

China

India

...

Inclusive Across Regions



Inclusive Across Genders

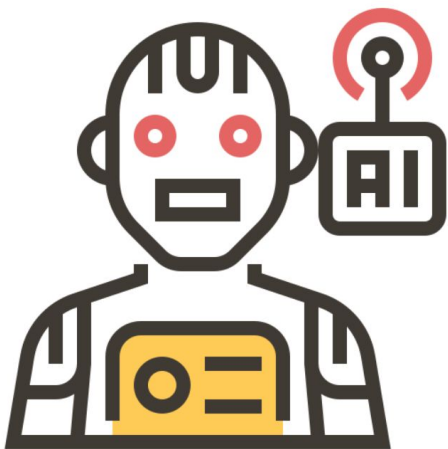
## User Groups

Male

Female

Non-Binary Gender

...



## User Groups

### Western Countries

Open-Domain QA

Commonsense Reasoning

Dialogue Generation

...

China

India

...

**NOT** Geographically Inclusive

# Research Interest

**Building Generalizable and Inclusive Language Processing  
Models Across **Tasks** and **User Groups****

# Research Interest

## Building Generalizable and Inclusive Language Processing Models Across **Tasks** and **User Groups**

Inclusive Across User Groups



GD-VCR, EMNLP 2021, Oral

GCI, NAACL 2021, Oral

GeoMLAMA, EMNLP 2022, Oral

ENTIGEN, EMNLP 2022, Oral

GIVL, Submitted to CVPR 2023





**Knowledge is actually geographically diverse!**

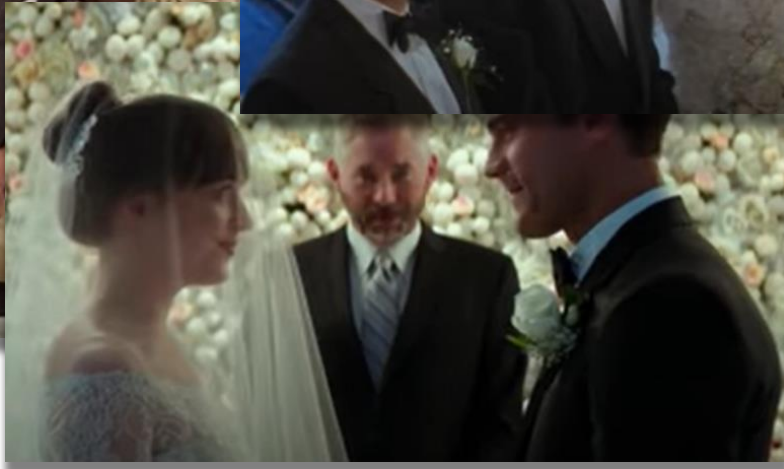


Knowledge is actually

geo-diverse

!





White dress



Black suits



White flowers



...



White dress



Black suits



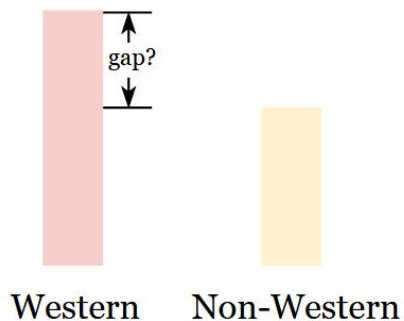
White flowers



...

# Evaluation Methods

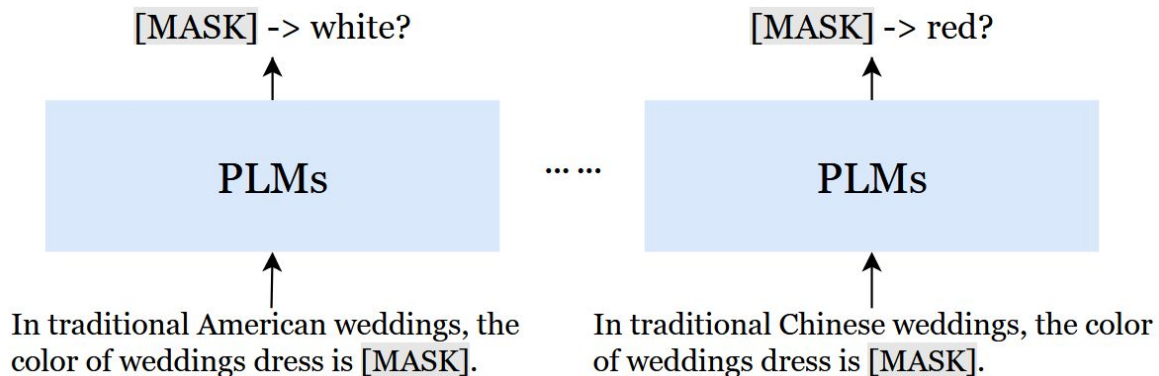
## Downstream Task Performance



GD-VCR, EMNLP 2021, Oral

10-20% gap on West and non-West Images

## Knowledge Probing

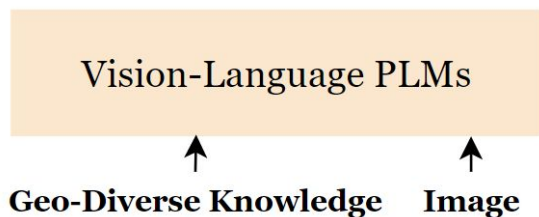


GeoMLAMA, EMNLP 2022, Oral

Only 30-40% probing accuracy, worse on non-West commonsense

# Improving Methods

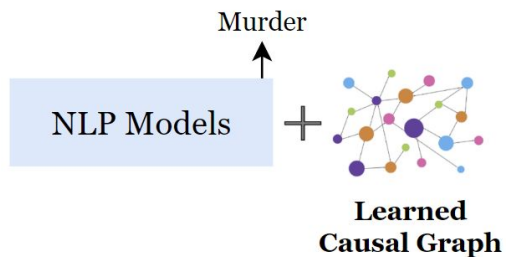
## Pre-Training



GIVL, Submitted to CVPR 2023

SOTA and more balanced performance on geo-diverse V&L tasks with less data;  
Competitive on common V&L tasks

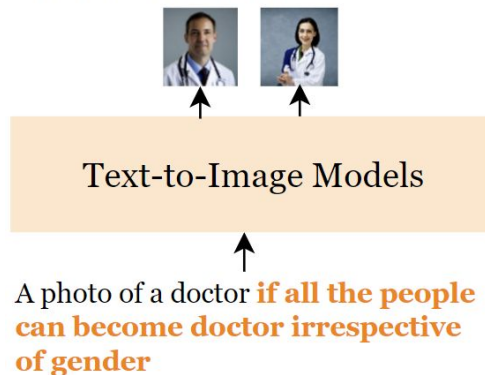
## Causal Inference



GCI, NAACL 2021, Oral

Better overall performance; More balanced performance on male and female cases

## Ethical Intervention



ENTIGEN, EMNLP 2022, Oral

More diverse outputs without losing quality

# Research Interest

## Building Generalizable and Inclusive Language Processing Models Across **Tasks** and **User Groups**

### Generalizable Across Tasks

Automatic and Dynamical Instruction Tuning, Ongoing

### Inclusive Across User Groups

GD-VCR, [EMNLP 2021, Oral](#)

ENTIGEN, [EMNLP 2022, Oral](#)

GeoMLAMA, [EMNLP 2022, Oral](#)

GCI, [NAACL 2021, Oral](#)

GIVL, Submitted to CVPR 2023

Low-Resource MT, Ongoing