

# Kaiyu He

917-238-2465 | kaiyu.he@utdallas.edu | Github | Homepage | Google Scholar  
Research Interests: Large Language Models (LLMs); LLM-based Language Agents;

## EDUCATION

<b>University of Texas at Dallas, TX, US</b> <i>Ph.D. in Computer Science, GPA:4.0/4.0</i>	Aug 2023– Jun 2028(expected)
<b>Columbia University in the City of New York, New York, US</b> <i>M.S in Biostatistics, GPA:4.0/4.4</i>	Sep 2021– Jun 2023
<b>Renmin University of China, Beijing, China</b> <i>B.S in Applied Statistics, GPA:3.2/4.0</i>	Sep 2016 – Jun 2020

## PUBLICATIONS

**1, IDEA: Enhancing the Rule Learning Ability of Large Language Model Agent through Induction, Deduction, and Abduction**  
**Kaiyu He**, Mian Zhang, Shuo Yan, Peilin Wu, Zhiyu Zoey Chen  
ACL 2025 Findings. [Paper Link](#)

**2, Evaluating completion rates of COVID-19 contact tracing surveys in New York City**  
**Kaiyu He**, Steffen Foerster, Neil M. Vora, Kathleen Blaney, Chris Keeley, Lisa Hendricks, Jay K. Varma, Theodore Long, Jeffrey Shaman, Sen Pei  
BMC Public Health. [Paper Link](#)

**3, From Reasoning to Learning: A Survey on Hypothesis Discovery and Rule Learning with Large Language Models**  
**Kaiyu He**, Zhiyu Zoey Chen  
[TMLR Survey Certification](#) [Paper Link](#)

**4, LMR-BENCH: Evaluating LLM Agent's Ability on Reproducing Language Modeling Research**  
Shuo Yan, Ruochen Li, Ziming Luo, Zimu Wang, Daoyang Li, Liqiang Jing, **Kaiyu He**, Peilin Wu, George Michalopoulos, Yue Zhang, Ziyang Zhang, Mian Zhang, Zhiyu Chen, Xinya Du  
EMNLP, 2025 [Paper Link](#)

**5, GEAR: A General Evaluation Framework for Abductive Reasoning**  
**Kaiyu He**, Peilin Wu, Mian Zhang, Kun Wan, Wentian Zhao, Xinya Du, Zhiyu Chen  
In submission to ICLR 2026 [Paper & Code](#)

**6, HiPRAG: Hierarchical Process Rewards for Efficient Agentic Retrieval Augmented Generation**  
Peilin Wu, Mian Zhang, Kun Wan, Wentian Zhao, **Kaiyu He**, Xinya Du, Zhiyu Chen  
In submission to ICLR 2026

## POSTERS

**A Decoded EEG Neurofeedback Platform Using Muse2**  
Junsheng Shi, **Kaiyu He**, Ming Li, Adam Brown, Alfredo Spagna, Xiaofu He  
Real-Time Functional Imaging and Neurofeedback meeting (rtFIN 2022, poster and oral presentation) [Poster Link](#)

## RESEARCH EXPERIENCES

<b>Human Language Technology Research Institute at UTD</b> Research Assistant (Advisor: <a href="#">Zhiyu Zoey Chen</a> )	May 2024 – Present Dallas, TX
<b>AI Safety Laboratory at UTD</b> Research Assistant (Advisor: <a href="#">Feng chen</a> )	Sep 2023 – May 2024 Dallas, TX
<b>Columbia NLP group at Columbia University</b> Research Assistant (Advisor: <a href="#">Kathleen McKeown</a> )	Jan 2023 – May 2024 Manhattan, NY
<b>Data Science Institute at Columbia University</b> Research Assistant (Advisor: <a href="#">Xiaofu He</a> )	Apr 2022 – Dec 2022 Manhattan, NY

## OTHER PROJECTS

---

### Latent Space Discovery of DCGAN Model

*Personal Project*

May 2022 - June 2022  
GAN, Knowledge representation

- Train DCGAN models using TensorFlow to generate images.
- 3-D Visualization of latent representations of images to show the latent representation from GAN.
- [Source Code Link](#)

### Abstract text summarization using Sim-CLS

*Personal Project*

Oct 2022 - Dec 2022  
NLP summarization

- SimCLS tuning in more recent models and see how SimCLS boost different types of models.
- Using medical datasets to train a SimCLS model to auto titling medical papers.
- [Blog link](#)
- [Source Code Link](#)

### 2022 Columbia University Innovation Award (VR)

*Teachers College, Columbia University*

Mar 2022 - May 2022  
VR video game

- Built an VR educational game for teaching physics using Unity Engine.
- [Demo Video](#)
- [Source Code](#)

## WORKING EXPERIENCE

---

### 89 Trillion, Art of War: Legions

*Game Numeric Designer*

Aug 2020 – Apr 2021  
Beijing, China

- Worked with the AI team to select the in-game numeric features used for Recommend System.
- Conducted A-B tests to determine the performance of new systems in the game.
- Set numeric values for game content according to the statistic model.
- Design game content. [Game BOSS I designed.](#)

## TECHNICAL SKILLS

---

**Programming Languages:** Python, C#, R, C/C++, SQL, Shell, Latex

**Machine Learning Frameworks:** Tensorflow, PyTorch, Caret(R), SciKit-Learn, Transformers

**Others:** Unity 3D

## CERTIFICATIONS

---

- TensorFlow Developer Certificate