

Yuge Tu

☎ +86 188 1133 9296 | ✉ tuyuge22@gmail.com | 🌐 GitHub | 📁 Portfolio | 📍 Beijing, China

Integrity, Resilience, Patience, Courage, Innovation

EDUCATION

Beijing Foreign Studies University

Beijing, China

M.A. in Applied Linguistics (direct admission without examination); GPA: 3.65/4.00

Oct. 2019 – Jun. 2023

B.A. in English Language and Literature; GPA: 3.54/4.00

Sep. 2015 – Jun. 2019

Major coursework: Psycholinguistics, Cognitive Linguistics, Second Language Acquisition, Syntax, Semantics, Phonetics, Pragmatics, Stylistics, Discourse Analysis, Language Assessment, Cultural Studies

Online courses: Mathematics for Machine Learning, Deep Learning, Advanced Learning Algorithms, Natural Language Processing, Databases and Data Analysis (see [certificates](#))

PUBLICATIONS

MiniCPM: Unveiling the Potential of Small Language Models with Scalable Training Strategies.

Shengding Hu, [Yuge Tu](#), Xu Han ... Zhiyuan Liu, Maosong Sun. Submitted to *COLM 2024*. [Arxiv](#). [Code](#).

Optimized scaling law based on MUP, conducted extensive model wind tunnel experiments, involving Bayesian parameter searches on hyper-parameters, batch size, learning rate, learning rate scheduler (i.e. WSD), and data strategy.

LEGENT: Open Platform for Embodied Agents.

Zhili Cheng, Zhitong Wang, Jinyi Hu, Shengding Hu, An Liu, [Yuge Tu](#) ... Zhiyuan Liu, Maosong Sun. Submitted to *ACL 2024*. [Arxiv](#). [Code](#).

Replicated Holodeck for LLM-guided scene generation in the interactive 3D environment LEGENT. Also refined prompt for instruction-following and QA data generation.

A Comprehensive Framework for Designing and Evaluating Vocabulary Learning Apps from Multiple Perspectives.

[Yuge Tu](#), Di Zou, & R Zhang. *International Journal of Mobile Learning and Organisation*. 2020. [DOI](#).

Designed the evaluation checklist from aspects like attention; retrieval; cognitive depth; engagement and gamification; scaffolding, self-paced and situated learning; immediacy, connectivity and data tracking; etc.

EXPERIENCE

Tsinghua NLP Lab & ModelBest Inc.

Beijing, China

Research Intern, full-time

Sep. 2023 – Present

Worked on LLM pretraining, predictable scaling and data synthesis mentored by [Shengding Hu](#), and multi-modality and embodied AI mentored by [Jinyi Hu](#). Earlier engaged in modular design of models, and LLM evaluation.

Zhipu AI

Beijing, China

Research Intern, full-time

Sep. 2021 – Oct. 2021

Investigated on techniques of knowledge extraction and abstractive summarization from massive academic literature.

Tsinghua AI TIME

Beijing, China

Ex-Leader of Editorial Group, part-time

Sep. 2019 – Present

Facilitated brainstorming sessions and debates about theories, principles, applications and social ramifications of AI.

AWARDS & ACHIEVEMENTS

The Future Cup of Artificial Intelligence & Robotics Project National Third Prize (Oct. 2021)

National Second-Class Academic Scholarship (Oct. 2019)

Beijing Outstanding Undergraduate Thesis (Jul. 2019)

T. S. Eliot International Summer School Full Scholarship (Aug. 2018 in London)

Excellent in University Student Innovation and Entrepreneurship Competition (Oct. 2017)

National Inspirational Scholarship (Sep. 2016 – Jun. 2018)

National High School Mathematical League Third Prize in Hubei Province (Sep. 2013)

SKILLS

Engineering: PyTorch, BMTrain, HF-Transformers, vLLM, TensorFlow, Keras, Numpy, Pandas, Matplotlib, DeepSpeed, PEFT, Scikit-learn, Docker, NetworkX, Flask, SQL, Django, Scrapy, L^AT_EX

Mathematics: Calculus, Linear Algebra, Probability Theory, Statistics, Discrete Mathematics

Language: Native in Chinese, fluent in English (TEM 8, IELTS 7.0), familiar with French and Japanese.