# Yuxian **Gu**

*Room 4-503, FIT Building, Tsinghua University, Beijing, China, 100084*

(+86) 15313594019  |  guyx21@mails.tsinghua.edu.cn  |  t1101675

## Education

**Tsinghua University**                                                                    *Beijing, China*

PH.D. STUDENT IN COMPUTER SCIENCE AND TECHNOLOGY                               *Sept. 2021 - Present*

- Member of **Tsinghua Conversational AI Group (CoAI)**
- Research Interest: Efficient Methods for Language Models
- Advisor: Prof. Minlie Huang

**Tsinghua University**                                                                    *Beijing, China*

B. ENG. IN COMPUTER SCIENCE AND TECHNOLOGY                                      *Sept. 2017 - June 2021*

## Selected Publications

[1] **Data Selection via Optimal Learning for Language Models**

**Yuxian Gu**, Li Dong, Hongning Wang, Yaru Hao, Qingxiu Dong, Minlie Huang, Furu Wei. *International Conference on Learning Representations (ICLR)*, 2025. (**Oral**)

[2] **MiniPLM: Knowledge Distillation for Pre-training Language Models**

**Yuxian Gu**, Hao Zhou, Fandong Meng, Jie Zhou, Minlie Huang. *International Conference on Learning Representations (ICLR)*, 2025.

[3] **Instruction Pre-Training: Language Models are Supervised Multitask Learners**

Daixuan Cheng, **Yuxian Gu**, Shaohan Huang, Junyu Bi, Minlie Huang, Furu Wei. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP 2024)*.

[4] **MiniLLM: Knowledge Distillation of Large Language Models**

**Yuxian Gu**, Li Dong, Furu Wei, Minlie Huang. In *Proceedings of the Twelfth International Conference on Learning Representations (ICLR)*, 2024.

[5] **Pre-Training to Learn in Context**

**Yuxian Gu**, Li Dong, Furu Wei, Minlie Huang. In *Proceedings of the 61th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2023. (**Oral**)

[6] **Learning Instructions with Unlabeled Data for Zero-Shot Cross-Task Generalization**

**Yuxian Gu**, Pei Ke, Xiaoyan Zhu, Minlie Huang. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2022. (**Oral**)

[7] **PPT: Pre-Trained Prompt Tuning for Few-Shot Learning**

**Yuxian Gu***, Xu Han*, Zhiyuan Liu, Minlie Huang. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2022.

[8] **CPM-2: Large-Scale Cost-Effective Pre-Trained Language Models**

Zhengyan Zhang*, **Yuxian Gu***, Xu Han*, Shengqi Chen*, Chaojun Xiao*, Zhenbo Sun, Yuan Yao, Fanchao Qi, Jian Guan, Pei Ke, Yanzheng Cai, Guoyang Zeng, Zhixing Tan, Zhiyuan Liu, Minlie Huang, Wentao Han, Yang Liu, Xiaoyan Zhu, Maosong Sun. *AI Open*, 2022.

[9] **Pre-Trained Models: Past, Present and Future**

Xu Han*, Zhengyan Zhang*, Ning Ding*, **Yuxian Gu***, Xiao Liu*, Yuqi Huo*, Jiezhong Qiu, Yuan Yao, Ao Zhang, Liang Zhang, Wentao Han, Minlie Huang, Qin Jin, Yanyan Lan, Yang Liu, Zhiyuan Liu, Zhiwu Lu, Xipeng Qiu, Ruihua Song, Jie Tang, Ji-Rong Wen, Jinhui Yuan, Wayne Xin Zhao, Jun Zhu. *AI Open*, 2022.

[10] **Train No Evil: Selective Masking for Task-Guided Pre-Training**

**Yuxian Gu**, Zhengyan Zhang, Xiaozhi Wang, Zhiyuan Liu, Maosong Sun. In *Proceedings of the Empirical Methods in Natural Language Processing (EMNLP)*, 2020. (short)

## Experience

**Microsoft Research Asia**                                                                *Beijing, China*

RESEARCH INTERN                                                                          *Oct. 2022 - Mar. 2024*

- Mentored by Dr. Li Dong.

**Tsinghua NLP Lab**                                                                       *Beijing, China*

STUDENT RESEARCH ASSISTANT                                                          *Sep. 2018 - Feb. 2020*

- Advised by Prof. Zhiyuan Liu.

# Honors & Awards

| | | |
|---|---|---|
| 2025 | **Apple Scholars in AIML PhD fellowship**<br>Recognizes the contributions of researchers in computer science and engineering at the graduate and postgraduate levels. | *Apple Inc.* |
| 2022-2024 | **Comprehensive Scholarship**<br>Awarded to Ph.D. students with excellent achievements by Tsinghua University. | *Tsinghua University* |
| 2021 | **Excellent Graduate, Tsinghua University (top 5%)**<br>Awarded to graduates with excellent achievements by Tsinghua University. | *Tsinghua University* |
| 2021 | **Outstanding Graduate, Dept. CST, Tsinghua University**<br>Awarded to graduates with excellent achievements by Department of Computer Science and Technology, Tsinghua University. | *Tsinghua University* |
| 2021 | **Outstanding Undergraduate Dissertation, Tsinghua University (top 5%)**<br>Awarded to outstanding undergraduate dissertations by Tsinghua University. | *Tsinghua University* |
| 2020 | **Comprehensive Scholarship**<br>Awarded to students with excellent achievements by Department of Computer Science and Technology, Tsinghua University. | *Tsinghua University* |
| 2019 | **Silver Prize, 2019 Asia Student Supercomputer Challenge**<br>World-wide contest on super-computing. | *Dalian, China* |
| 2018 | **Comprehensive Scholarship**<br>Awarded to students with excellent achievements by Department of Computer Science and Technology, Tsinghua University. | *Tsinghua University* |
| 2017 | **Second Grade Scholarship for Freshman**<br>Awarded to students with excellent overall achievements in high school, Tsinghua University. | *Tsinghua University* |
| 2016 | **1st Prize, Chinese Physics Olympiad in Provinces**<br>Nation-wide contest on physics. | *Beijing, China* |
| 2016 | **1st Prize, Pan-Pearl River Delta and Chinese Elite Schools Physics Olympiad**<br>Nation-wide contest on physics. | *Shenzhen, China* |

# Projects

### CPM-2: Large-scale Cost-effective Pre-trained Language Models

Jan. 2021 - June. 2021

- (By the time of release) The largest open-source Chinese pre-trained language model (11B parameters).
- Encoder-decoder architecture. Based on Megatron-LM and DeepSpeed. Language understanding and generation.
- Open Source: `https://github.com/TsinghuaAI/CPM-2-Pretrain`

### EVA: Large-scale Open-domain Chinese Dialogue System

Sep. 2021 - Jan. 2022

- (By the time of release) The largest open-source Chinese dialogue model with pre-training (2.8B parameters).
- Encoder-decoder architecture. Based on DeepSpeed. Open-domain chit-chat.
- Open Source: `https://github.com/thu-coai/EVA`

# Teaching Assistantships

| | |
|---|---|
| **Object Oriented Programming** | *Tsinghua University* |
| Instructor: Prof. Minlie Huang | *Spring 2022, 2023,2024* |
| **Artificial Neural Networks** | *Tsinghua University* |
| Instructor: Prof. Minlie Huang | *Fall 2021, 2022, 2023* |