HE ZHU

1088 Xueyuan Avenue, Nanshan◇ Shenzhen, China
(+86) · 189 · 0301 · 0858 ◇ zhuhe@stu.pku.edu.cn

EDUCATION

Peking University Sep 2024 - Present
M.S., major in Smart City and Big Data Beijing, China
Prof. Wenjia Zhang

Southern University of Science and Technology Sep 2020 - Jul 2024
B.E., major in Computer Science Shenzhen, China
GPA: 90.2/100, Rankings: 10%. Prof. Xuan Song

PROJECT

PlanGPT Series: Large Model for Urban Planning and Design Sep 2023 – Present
Peking University, Beijing, China *First Author and Technical Lead*

- Led the PlanGPT project, the first systematic exploration of LLMs in urban planning, including: (1) ***PlanGPT***: language model for planning scheme generation, (2) ***PlanGPT-VL***: first vision-language model for planning maps, (3) ***PlanGPT-R1***: reasoning-enhanced model for document understanding, (4) ***UP-Bench***: first benchmark dataset for urban planning LLM tasks, (5) applied systems (Urban-Thinking, **Coplanner**, AuditPlanner) deployed in top planning institutes.
[GitHub: https://plangpt.github.io/]

SELECTED PUBLICATIONS

PlanGPT: Enhancing Urban Planning with Tailored Language Model ACL Industry, 2025 (Oral)
He Zhu, *Wenjia Zhang, Guanhua Chen*

- The first LLM specifically developed for urban planning and design, capable of generating high-quality urban planning schemes and design suggestions.
[Available at: https://arxiv.org/abs/2402.19273]

Augmenting High-Quality Instruction Data with Open-Sourced LLMs Only ACL, 2025 (Findings)
He Zhu, *Yifan Ding, Yicheng Tao, Zhiwen Ruan, Yixia Li, Wenjia Zhang, Yun Chen, Guanhua Chen*

- Proposed FANNO, a framework synthesizing instruction data using open-source LLMs and unlabeled documents. Introduced a tagging-based prompt method and a UCB-based approach to expand instructions. Addressed the high cost of manual annotation by enabling high-quality data generation without pre-crafted seed data.
[Available at: https://aclanthology.org/2025.findings-acl.906/]

Tag-Instruct: A Scalable Framework for Controlled Instruction-Tuning Data Synthesis ACL, 2025 (Findings)
He Zhu, *Zhiwen Ruan, Junyou Su, Xingwei He, Yun Chen, Wenjia Zhang, Guanhua Chen*

- Proposed Tag-Instruct, a VAE-inspired framework that extracts latent tags from instructions for controlled data generation. Introduced augmentation mechanisms for difficulty and variety control. Achieved SOTA results across multiple domains with minimal human effort.
[Available at: https://aclanthology.org/2025.findings-acl.911/]

PlanGPT-VL: Vision-Language Model for Urban Planning Maps EMNLP Industry, 2025
He Zhu*, *Junyou Su**, Minxin Chen**, Yun Chen, Guanhua Chen, Wenjia Zhang*

- Developed PlanGPT-VL, a 7B model tailored for planning maps with custom data, hallucination control, and hybrid training. Achieves strong results on PlanBench-V, surpassing larger general VLMs.
[Available at: https://arxiv.org/pdf/2505.14481]

Personalized Individual Trajectory Prediction via Meta-Learning SIGSPATIAL, 2022
***He Zhu**, Liyu Zhang, Zipei Fan* *Seattle, USA*

- Developed a pre-trained trajectory prediction model using meta-learning to address inhomogeneous distribution challenges with a grid-based classification approach.
[Available at: https://dl.acm.org/doi/abs/10.1145/3557915.3565536]

WORKING PAPERS

Anchored Supervised Fine-Tuning: A Principled Framework for Post-Training LLMs ICLR, 2026 (submit)
He Zhu*, Junyou Su*, Peng Lai*, Wenjia Zhang, Linyi Yang, Guanhua Chen*

AlignDiff: Exploiting Model-Intrinsic Information for Better Data Selection ICLR, 2026 (submit)
Peng Lai, **He Zhu***, Zhiwen Ruan, Dongdong Zhang, Yun Chen, Peng Li, Furu Wei, Yang Liu, Guanhua Chen*

Towards Fair and Comprehensive Evaluation of Routers in Collaborative LLM Systems ICLR, 2026 (submit)
Wanxing Wu, **He Zhu***, Yixia Li*, Yun Chen, Guanhua Chen*

Measuring Distribution Alignment through Loss Landscape Geometry CVPR 2026 (To be sumbit)
***He Zhu**, Ren Ma, Junyou Su, Xingwei He, Yun Chen, Wenjia Zhang, Guanhua Chen*

PROFESSIONAL EXPERIENCE SERVICE

Shanghai AI Laboratory May 2025 - Oct 2025
Researcher, Foundation Language Models *Shanghai, China*

SenseTime June 2024 - Sep 2024
Researcher, Foundation Language Models *Shenzhen, China*

LocationMind June 2023 - Dec 2023
Project Assistant, Autonomous Driving Algorithm Engineer *Tokyo, Japan*

RESEARCH EXPERIENCE

SUSTech-NLP Group, SUSTech Nov 2023 - Present
Research Assistant, Supervisor: Prof. Guanhua Chen *Shenzhen, China*

Focus on instruction data synthesis and distillation, leveraging LLMs for high-quality data generation.

Center for Spatial Information Science (CSIS), UTokyo Feb 2022 - May 2022, Jul 2023 - Sep 2023
Research Assistant, Supervisor: SHIBASAKI Ryosuke, Prof. Zipei Fan *Tokyo, Japan*

Focus on end-to-end trajectory prediction and group interaction modeling for traffic agents.

School of Computing Summer Workshop, NUS May 2022 - Jul 2022
Visiting Student *Singapore*

Focus on human behavior prediction, trajectory tracking, and IoT.

SELECTED AWARDS

Outstanding Graduate, Department of Computer Science, SUSTech (top 5%) 2024
Outstanding Graduate, SUSTech (top 5%) 2024
Outstanding Thesis, SUSTech (top 10%) 2024
Annual Outstanding Student, SUSTech (top 10%) 2021, 2022, 2023