

# Xiaochen Zhu

✉ xz479@cam.ac.uk • 📞 +44 (0) 7459 850946 • 🌐 Website  
🎓 Google Scholar • 🐙 GitHub

---

## PROFILE

PhD candidate in Computer Science at the University of Cambridge specializing in Large Language Models, multi-agent systems, reinforcement learning, and controllable generation. Published multiple first-author papers at top-tier AI/HCI/NLP venues (ACL/EMNLP/ICWSM) on LLM, multi-agent systems, and diffusion language models. Hands-on experience building production-grade agentic workflows, evaluation harnesses, and retrieval-augmented pipelines, bridging advanced academic research with industry-focused AI solutions (e.g., Point72, startup ventures).

---

## EDUCATION

### University of Cambridge

*PhD in Computer Science*

Cambridge, UK

Sep. 2023 – Nov. 2026 Expected

Natural Language and Information Processing Group, supervised by [Prof. Andreas Vlachos](#). Research areas: Large Language Models (LLM), Multi-Agent Systems (MAS), Controllable Generation, LLM Agents, Diffusion for Language Modeling, Multi-party Dialogue, Computational Social Science (Deliberation for collective decision).

### University of Cambridge

*MPhil in Advanced Computer Science*

Cambridge, UK

Sep. 2022 – Jun. 2023

Distinction. Research project: *Modeling Human Disagreement in Natural Language Inference* (supervised by [Prof. Andreas Vlachos](#)). Investigated soft-prompt tuning and contrastive learning to represent label distributions and generate aligning justifications.

### University of Edinburgh

*BSc in Artificial Intelligence and Computer Science*

Edinburgh, UK

Sep. 2018 – May. 2022

First Class with Honours. Research project: *Spectral Learning on Latent Weighted Logic Programs* (supervised by [Prof. Shay Cohen](#)). Extended spectral learning to single-parent weighted logic programs with latent states to avoid local optima in Expectation Maximization.

---

## PROFESSIONAL EXPERIENCE

### Point72 Asset Management

*AI/ML Summer Intern, Long/Short Equity*

London, UK

Jun. 2025 – Sep. 2025

- Engineered state-of-the-art multi-agent LLM systems for automated company research, financial analysis, and equity ranking workflows.
- Modeled retail sales and company performance metrics by integrating macroeconomic indicators, credit card transactions, and e-commerce alternative datasets.

---

## RESEARCH EXPERIENCE

### University of Cambridge

*PhD Researcher, Natural Language and Information Processing Group*

Cambridge, UK

Sep. 2023 – Present.

Supervised by [Prof. Andreas Vlachos](#).

- Conducted research in LLM post-training, prompting, and alignment, focusing on uncertainty expression, conformity, multi-agent debate, and controllable diffusion models.
- Co-developed **Delibot** ([delibot.xyz](#)), a dialogue agent platform to facilitate multi-party online deliberation, stimulate diverse reasoning, and improve group decision-making.
- Built high-throughput LLM evaluation harnesses for QA, summarization, long-form text generation, and multi-turn collaborative dialogues.
- Developed custom web annotation interfaces and human-in-the-loop evaluation pipelines for studying user interaction with dialogue systems.

### Heinrich-Heine-Universität Düsseldorf

*Research Intern, Dialogue Systems Group*

Düsseldorf, Germany

Jun. 2021 – Oct. 2021

Supervised by [Prof. Milica Gašić](#).

- Co-developed and evaluated core modules for **ConvLab-3**, a flexible, multi-domain dialogue system toolkit, improving system reproducibility and evaluation pipelines.
- Designed web interfaces and deployed large-scale human-in-the-loop evaluation environments on Amazon Mechanical Turk.

---

## PUBLICATIONS

- **DeliChess: A Multi-party Dialogue Dataset for Deliberation in Chess Puzzle Solving**  
Xiaochen Zhu, Georgi Karadzhov, Tom Stafford, Andreas Vlachos.  
*2606.04987 ArXiv*
- **Demystifying Multi-Agent Debate: The Role of Confidence and Diversity**  
Xiaochen Zhu, Caiqi Zhang, Yizhou Chi, Tom Stafford, Nigel Collier, Andreas Vlachos.  
*Findings of the 64th Annual Meeting of the Association for Computational Linguistics, ACL 2026.*
- **Reinforcement Learning for Better Verbalized Confidence in Long-Form Generation**  
Caiqi Zhang, Xiaochen Zhu, Chengzu Li, Nigel Collier, Andreas Vlachos.  
*Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics, ACL 2026 (Main Conference).*
- **Segment-Level Diffusion: A Framework for Controllable Long-Form Generation with Diffusion Language Models**  
Xiaochen Zhu, Georgi Karadzhov, Chenxi Whitehouse, Andreas Vlachos.  
*Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics, ACL 2025 (Main Conference).*
- **Conformity in Large Language Models**  
Xiaochen Zhu, Caiqi Zhang, Tom Stafford, Nigel Collier, Andreas Vlachos.  
*Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics, ACL 2025 (Main Conference).*
- **Confidence Estimation for LLMs in Multi-turn Interactions**  
Caiqi Zhang, Ruihan Yang, Xiaochen Zhu, Chengzu Li, Tiancheng Hu, Yijiang River Dong, Deqing Yang, Nigel Collier.  
*Findings of the 64th Annual Meeting of the Association for Computational Linguistics, ACL 2026.*
- **Collaborative Evaluation of Deepfake Text with Deliberation-Enhancing Dialogue Systems**  
Jooyoung Lee, Xiaochen Zhu, Georgi Karadzhov, Tom Stafford, Andreas Vlachos, Dongwon Lee.  
*International AAI Conference on Web and Social Media, ICWSM 2026 (Main Conference).*
- **ConvLab-3: A Flexible Dialogue System Toolkit**  
Qi Zhu, Christian Geishauer, Hsien-chin Lin, Carel van Niekerk, Baolin Peng, Zheng Zhang, Shutong Feng, Michael Heck, Nurul Lubis, Dazhen Wan, Xiaochen Zhu, Jianfeng Gao, Milica Gasic, Minlie Huang.  
*Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, EMNLP 2023 (System Demonstrations).*

## ADDITIONAL EXPERIENCE

---

**Precedent AI – Winner of Law x LLM Hackathon** Cambridge, UK  
*King's College Entrepreneurship Lab* Jun. 2024

- Engineered **Precedent AI**, a source-tracked RAG based AI paralegal system that streamlines case reviews, reducing a 13-hour research workflow to under 30 minutes (a 26x efficiency gain).
- Won **First Prize** and the **Amazon AWS Challenge** out of 150 global participants, securing a £10,000 cash award and potential venture investments.

**Best Research Pitch Award** Cambridge, UK  
*Cambridge Centre for Human-Inspired AI (CHIA)* Jun. 2024

- Won best pitch for research on *Latent Diffusion-Based Multi-party Dialogue Simulation* presented to a panel of AI researchers and industry partners at the annual CHIA Conference.

**IRIS – Winner of Entrepreneurship Competition** Shenzhen, China  
*Asia-Pacific Innovation Academy (APIA)* Jul. 2018 – Aug. 2018

- Led the development of **IRIS**, an AI-driven online family doctor symptoms pre-screening and medical consultation portal designed to reduce hospital overcrowding in China.
- Built symptom-checking and automated triage workflows, securing the **Top Team Award** among startups selected across the Asia-Pacific region with guaranteed awards of 10,000RMB and possible investment up to 500,000RMB.

## SKILLS

---

**AI / ML:** Large Language Models, Multi-Agent Systems, Reinforcement Learning, Confidence Calibration, Uncertainty Estimation, Controllable Generation, Diffusion Language Models, Retrieval Augmented Generation (RAG), Evaluation Harnesses

**Frameworks:** PyTorch, HuggingFace Transformers, PyTorch Lightning, vLLM, Weights & Biases, Streamlit, MongoDB, React, Node.js

**Systems & Infra:** GPU Inference, Distributed Training, Experiment Infrastructure, Human Evaluation Pipelines, Data Annotation Workflows, AWS, Linux, Slurm

**Programming:** Python, Julia, JavaScript, Java, C, Haskell, SQL,  $\LaTeX$

**Languages:** Chinese (Native), English (Fluent), German (Intermediate)