

Chenxi Whitehouse

RESEARCH SCIENTIST · NLP · LLM · META MSL

[✉ chenxi.whitehouse@gmail.com](mailto:chenxi.whitehouse@gmail.com) | [🏡 chenxwh.github.io](https://chenxwh.github.io) | [👤 chenxwh](https://chenxwh.com) | [🔗 chenxwh](https://www.linkedin.com/in/chenxwh/)

Summary

Research Scientist specializing in Natural Language Processing and Large Language Models. Currently at Meta Superintelligence Lab, leading projects with the FAIR Alignment team on Post-training and Reasoning, with a focus on Reinforcement Learning, LLM-as-a-Judge, and Generative Reward Modeling.

Professional Experience

RESEARCH

- 2024-2025 **Research Scientist, Meta Superintelligence Lab**, London, UK
- Working with FAIR Alignment on reasoning and post-training.
 - Focusing on Reinforcement Learning, Thinking LLM-judges, and Generative Reward Models.
 - Worked on production for post-training models to improve Instruction Hierarchy and Instruction Following.
- 2023 **Research Intern, Google DeepMind**, Amsterdam, Netherlands
- Worked on Parameter Efficient Fine-tuning (LoRA) for planning and summarization.
- 2022 **Research Intern, Amazon AGI**, Cambridge, UK
- Worked on Information Extraction and Verification on Alexa Knowledge.
- 2021 **Research Intern, Huawei**, Noah's Ark Lab, London, UK

ENGINEERING

- 2021-2024 **Machine Learning Engineer**, Replicate, Part-time
- 2018-2020 **Software Engineer**, Inpixon, Berlin, Germany and HP Inc, London, UK

Education

University of Cambridge	Cambridge, UK
POSTDOC RESEARCH ASSOCIATE IN NATURAL LANGUAGE PROCESSING	2024
City, University of London	London, UK
PHD IN COMPUTER SCIENCE, NATURAL LANGUAGE PROCESSING	2020 - 2023
• Thesis: <u>Towards Knowledge-Grounded Natural Language Understanding and Generation</u>	
• Research Focus: Knowledge-intensive NLP, Multilingual NLP	
University College London & University of Erlangen-Nuremberg	London, UK & Erlangen, Germany
MSC ELECTRICAL ENGINEERING & SIGNAL PROCESSING	2015 - 2018
Xi'an Jiaotong University	Xi'an, China
BSC INFORMATION & COMPUTER ENGINEERING	2011 - 2015

Selected Publications

1. Shashwat Goel, Rishi Hazra, Yoram Bachrach, Jonas Geiping, and **Chenxi Whitehouse**. “PlanRL: Rubric-Guided Reinforcement Learning for Research Plan Generation”. (2025, *work conducted at Meta MSL*)
2. **Chenxi Whitehouse**, Swarnadeep Saha, Ilia Kulikov, Ping Yu, Xian Li, and Jason Weston. “RLLM: Reinforcement Learning with Reasoning Language Model as Reward Model”. (2025, *work conducted at Meta FAIR*)
3. **Chenxi Whitehouse**, Tianlu Wang, Ping Yu, Xian Li, Jason Weston, Ilia Kulikov, and Swarnadeep Saha. “Incentivizing Thinking in LLM-as-a-Judge via Reinforcement Learning”. (Submitted to ICLR 2026, *work conducted at Meta FAIR*)

4. **Chenxi Whitehouse**, Sebastian Ruder, ..., Paco Guzmán. “MENLO: From Preferences to Proficiency – Evaluating and Modeling Native-like Quality Across 47 Languages via RL”. (Submitted to ICLR 2026, *work conducted at Meta MSL*)
5. Bingchen Zhao, Jenny Zhang, **Chenxi Whitehouse**, Minqi Jiang, ..., Yoram Bachrach. “APRES: An Agentic Paper Revision and Evaluation System”. (Submitted to ICLR 2026, *work conducted at Meta MSL*)
6. Dongqi Liu, **Chenxi Whitehouse**, Xi Yu, Louis Mahon, Rohit Saxena, Zheng Zhao, Yifu Qiu, Mirella Lapata, and Vera Demberg. “What Is That Talk About? A Video-to-Text Summarization Dataset for Scientific Presentations”. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistic: ACL 2025*.
7. Zhangdie Yuan, **Chenxi Whitehouse**, Eric Chamoun, Rami Aly, and Andreas Vlachos. “PRobELM: Plausibility Ranking Evaluation for Language Models”. In *First Conference on Language Modeling: CoLM 2024*.
8. David Romero, Chenyang Lyu, Haryo Akbarianto Wibowo, ..., **Chenxi Whitehouse**, ..., Thamar Solorio, and Alham Fikri Aji. “CVQA: Culturally-diverse Multilingual Visual Question Answering Benchmark”. In *Advances in Neural Information Processing Systems: NeurIPS 2024*.
9. **Chenxi Whitehouse**, Fantine Huot, Jasmijn Bastings, Mostafa Dehghani, Chu-Cheng Lin, and Mirella Lapata. “Low-Rank Adaptation for Multilingual Summarisation: An Empirical Study”. In *Findings of the Association for Computational Linguistics: NAACL 2024*. (*Work conducted as research intern at Google DeepMind*)
10. Yuxia Wang, Jonibek Mansurov, Petar Ivanov, Jinyan Su, Artem Shelmanov, Akim Tsvigun, **Chenxi Whitehouse**, Osama Mohammed Afzal, Tarek Mahmoud, Alham Fikri Aji, and Preslav Nakov. “M4: Multi-generator, Multi-domain, and Multi-lingual Black-Box Machine-Generated Text Detection”. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: EACL 2024*. (*Best Resource Paper Award*)
11. Chu-Cheng Lin, Xinyi Wang, Jonathan H. Clark, Yun Zhu, Han Lu, **Chenxi Whitehouse**, and Hongkun Yu. “Inducing Generalization across Languages and Tasks using Featureized Low-Rank Mixtures”. (2024, *Work conducted as research intern at Google DeepMind*)
12. **Chenxi Whitehouse**, Monojit Choudhury, and Alham Fikri Aji. “LLM-powered Data Augmentation for Enhanced Cross-lingual Performance”. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: EMNLP 2023*. (*Work conducted as visiting researcher at MBZUAI*)
13. **Chenxi Whitehouse**, Clara Vania, Alham Fikri Aji, Christos Christodoulopoulos, and Andrea Pierleoni. “WebIE: Faithful and Robust Information Extraction on the Web”. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistic: ACL 2023*. (*Work conducted as research intern at Amazon*)
14. **Chenxi Whitehouse**, Tillman Weyde, and Pranava Madhyastha. “Towards a Unified Model for Generating Answers and Explanations in Visual Question Answering”. In: *Findings of the Association for Computational Linguistics: EACL 2023*.
15. **Chenxi Whitehouse**, Fenia Christopoulou, and Ignacio Iacobacci. “EntityCS: Improving Zero-Shot Cross-lingual Transfer with Entity-Centric Code Switching”. In: *Findings of the Association for Computational Linguistics: EMNLP 2022*. (*Work conducted as research intern at Huawei*)
16. **Chenxi Whitehouse**, Tillman Weyde, Pranava Madhyastha, and Nikos Komninos. “Evaluation of Fake News Detection with Knowledge-Enhanced Language Models”. In: *Proceedings of the Sixteenth International AAAI Conference on Web and Social Media: AAAI-ICWSM 2022*.

Awards & Honors

- 2024 **Best Resource Paper Award**: EACL 2024
- 2020-2023 **PhD Studentship**: City, University of London
- 2016-2017 **Erasmus Scholarship**: University College London
- 2012-2015 **Outstanding Student Scholarship**: Xi'an Jiaotong University

Professional Service

- 2025 **Diversity and Inclusion Chair**: EMNLP 2025
- 2024 **Area Chair**: COLING 2025
- 2024-2025 **Workshop Organisation**: FEVER @ EACL 2026, FEVER @ ACL 2025, FEVER @ EMNLP 2024
- 2022-2025 **Reviewer**: ACL, EMNLP, NeurIPS, ICLR, CoLM