

QINGCHENG ZENG

Ph.D. Candidate in Computational Linguistics · Northwestern University

2016 Sheridan Road, Evanston, IL · qcz@u.northwestern.edu · qcznlp.github.io · [Google Scholar](#)

SUMMARY

Ph.D. candidate working on large language models, with industry research experience at Snowflake, Tencent AI Lab, and Alibaba International. My research spans:

- **LLM Agents, Information Retrieval, and Reinforcement Learning** — search agents, process reward modeling, and retrieval robustness.
- **Trustworthy LLMs** — uncertainty quantification, calibration, and verbalized confidence across LLMs, reasoning models, vision–language models, and tool-use agents.
- **LLMs for Computational Social Science and Social Good** — including public-health and clinical NLP.
- **Understanding LLMs through Cognitive Science.**
- **Multilingual NLP and Machine Translation.**

INDUSTRY EXPERIENCE

Applied Scientist Intern, Alibaba International — Accio Agent Team June 2026 – Sep 2026
Working with [Kunyu Shi](#) and [Junbo Li](#) on long-horizon agentic reinforcement learning: trained agents with GRPO and PPO and helped deliver [Occamy-1.0](#).

Research Scientist Intern, Snowflake — Cortex Search AI Team, Remote March 2026 – June 2026
Working with [Puxuan Yu](#) and [Aman Mehta](#).
Research output: SQL schema retrieval at [arXiv 2026](#).

Research Scientist Intern, Snowflake — Cortex Search AI Team, Menlo Park, CA Sep 2025 – Dec 2025
Working with [Puxuan Yu](#) and [Aman Mehta](#).
Shipped instruction-following information retrieval for Cortex Search; research output accepted to [EMNLP 2026 \(Main\)](#).

Research Scientist Intern, Tencent AI Lab America, Bellevue, WA June 2025 – Aug 2025
Working with [Linfeng Song](#), [Xiaoyang Wang](#), [Hongming Zhang](#), and [Wenhao Yu](#) on process reward modeling for agents.
Research output: agent optimization at [Findings of ACL 2026](#).

EDUCATION

Ph.D. in Computational Linguistics 2022–2026 (expected)
Department of Linguistics, Northwestern University, Evanston, IL
Advisors: Professor [Rob Voigt](#) and Professor [Kaize Ding](#)

B.A. in Linguistics 2019–2021
University of Manchester, Manchester, UK

B.A. in Linguistics 2017–2021
Zhejiang University, Hangzhou, China

ACADEMIC RESEARCH EXPERIENCE

Visiting Graduate Student, WestlakeNLP, Westlake University Sep 2024 – Nov 2024
Research in multilingual calibration, supervised by Professor [Yue Zhang](#) and Dr. [Leyang Cui](#).
Research output: reasoning-model calibration at [EMNLP 2025](#); LLM evaluation at [NeurIPS 2025](#).

Visiting Graduate Student, Johns Hopkins University (CLSP)

Summer 2024

Research in AI-curated democratic discourse at the Center for Language and Speech Processing, supervised by Professor [Jason Eisner](#).

Research output: system demonstration at CSCW 2026 (Demo).

Research Assistant in Natural Language Processing, Zhejiang University, Hangzhou, China 2021–2022

Research in multilingual NLP and biomedical NLP, supervised by Professor [Jie Yang](#).

Research output: effective language model adaptation at IJCAI 2023.

PUBLICATIONS AND PREPRINTS

* denotes authors of equal contribution; # denotes corresponding author.

LLM Agents, Information Retrieval, and Reinforcement Learning

- [1] Zeqi Zhou*, **Qingcheng Zeng***, Rui Yang, Jiarui Liu, Heli Qi, Puxuan Yu, Weihao Xuan, Naoto Yokoya. [When Queries Crowd: Diagnosing and Improving Spread-Out Regularization for Text Embeddings](#). *EMNLP 2026 (Main)*.
- [2] Weihao Xuan*, **Qingcheng Zeng***, Heli Qi, Yunze Xiao, Junjue Wang, Naoto Yokoya. [The Confidence Dichotomy: Analyzing and Mitigating Miscalibration in Tool-Use Agents](#). *ACL 2026 (Oral)*.
- [3] **Qingcheng Zeng***, Yuheng Lu*, Zeqi Zhou, Heli Qi, Puxuan Yu, Fuheng Zhao, Hitomi Yanaka, Weihao Xuan, Naoto Yokoya. [Code-Switching Information Retrieval: Benchmarks, Analysis, and the Limits of Current Retrievers](#). *Findings of ACL 2026*.
- [4] Mukai Li, **Qingcheng Zeng**, Tianqing Fang, Zhenwen Liang, Linfeng Song, Qi Liu, Haitao Mi, Dong Yu. [Verified Critical Step Optimization for LLM Agents](#). *Findings of ACL 2026*.
- [5] Minghao Guo, **Qingcheng Zeng**, Xujiang Zhao, Yanchi Liu, Wenchao Yu, Mengnan Du, Haifeng Chen, Wei Cheng. [DeepSieve: Information Sieving via LLM-as-a-Knowledge-Router](#). *Findings of EACL 2026*.
- [6] Haoyu Wu*, **Qingcheng Zeng***, Kaize Ding. [Fact or Facsimile? Evaluating the Factual Robustness of Modern Retrievers](#). *CIKM 2025*.
- [7] Yuheng Lu*, **Qingcheng Zeng***, Heli Qi, Puxuan Yu, Fuheng Zhao, Rui Yang, Hitomi Yanaka, Naoto Yokoya, Weihao Xuan. [Beyond Monolingual Deep Research: Evaluating Agents and Retrievers with Cross-Lingual BrowseComp-Plus](#). *arXiv:2606.15345*, 2026.
- [8] **Qingcheng Zeng**, Puxuan Yu, Aman Mehta, Fuheng Zhao, Rajhans Samdani. [Dual-View Training for Instruction-Following Information Retrieval](#). *EMNLP 2026 (Main)*.
- [9] **Qingcheng Zeng**, Puxuan Yu, Aman Mehta, Fuheng Zhao, Rajhans Samdani. [Finding the Right Tables and Columns: A Benchmark and Corpus-Adaptive Embeddings for SQL Schema Retrieval](#). *arXiv:2607.13311*, 2026.
- [10] Fang Wu, Aaron Tu, Weihao Xuan, Heli Qi, Xu Huang, **Qingcheng Zeng**, ..., Yejin Choi. [Position: The Hidden Costs and Measurement Gaps of Reinforcement Learning with Verifiable Rewards](#). *arXiv:2509.21882*, 2025.
- [11] Shunyu Liu, Minghao Liu, Huichi Zhou, ..., **Qingcheng Zeng**, ..., Dacheng Tao. [VeriWeb: Verifiable Long-Chain Web Benchmark for Agentic Information-Seeking](#). *arXiv:2508.04026*, 2025.

Trustworthy LLMs and Uncertainty Quantification

- [12] **Qingcheng Zeng***, Weihao Xuan*, Leyang Cui, Rob Voigt. [Thinking Out Loud: Do Reasoning Models Know When They're Right?](#) *EMNLP 2025*.
- [13] Weihao Xuan*, **Qingcheng Zeng***, Heli Qi, Junjue Wang, Naoto Yokoya. [Seeing is Believing, but How Much? A Comprehensive Analysis of Verbalized Calibration in Vision-Language Models](#). *EMNLP 2025 (Oral)*.
- [14] **Qingcheng Zeng**, Mingyu Jin, Qinkai Yu, Zhenting Wang, Wenyue Hua, Guangyan Sun, Yanda Meng, Shiqing Ma, Qifan Wang, Felix Juefei-Xu, Fan Yang, Kaize Ding, Ruixiang Tang, Yongfeng Zhang. [Uncertainty Quantification for Multiple-Choice Questions is Just One-Token Deep](#). *CIKM 2025*.

- [15] Sijia Chen, Xiaomin Li, Mengxue Zhang, Eric Hanchen Jiang, **Qingcheng Zeng**, Chen-Hsiang Yu. [CARES: Comprehensive Evaluation of Safety and Adversarial Robustness in Medical LLMs](#). *NeurIPS 2025, Datasets & Benchmarks Track*.
- [16] Shulin Huang, Linyi Yang, Yan Song, Shuang Chen, Leyang Cui, Ziyu Wan, **Qingcheng Zeng**, Ying Wen, Kun Shao, Weinan Zhang, Jun Wang, Yue Zhang. [ThinkBench: Dynamic Out-of-Distribution Evaluation for Robust LLM Reasoning](#). *NeurIPS 2025, Datasets & Benchmarks Track*.
- [17] Jiayu Liu, Rui Wang, Qing Zong, Yumeng Wang, Cheng Qian, **Qingcheng Zeng**, Tianshi Zheng, Haochen Shi, Dadi Guo, Baixuan Xu, Chunyang Li, Yangqiu Song. [NOVA: Noise-Aware Verbal Confidence Calibration for Robust Large Language Models in RAG Systems](#). *Findings of EMNLP 2026*.

LLMs for Computational Social Science and Social Good

- [18] **Qingcheng Zeng***, Guanhong Liu*, Zhaoqian Xue, Diego Ford, Rob Voigt, Loni Hagen, Lingyao Li. [Sympathy over Polarization: A Computational Discourse Analysis of Social Media Posts about the July 2024 Trump Assassination Attempt](#). *Findings of AACL 2025*.
- [19] Grace LeFevre, **Qingcheng Zeng**, Adam Leif, Jason Jewell, Denis Peskoff, Rob Voigt. [Good Intentions Beyond ACL: Who Does NLP for Social Good, and Where?](#) *EMNLP 2025*.
- [20] **Qingcheng Zeng**, Mingyu Jin, Rob Voigt. [Adaptive Axes: A Pipeline for In-domain Social Stereotype Analysis](#). *EMNLP 2024*.
- [21] Zhaoqian Xue, Guanhong Liu, Chong Zhang, Kai Wei, **Qingcheng Zeng**, Songhua Hu, Wenyue Hua, Lizhou Fan, Yongfeng Zhang, Lingyao Li. [Toward Equitable Access: Leveraging Crowdsourced Reviews to Investigate Public Perceptions of Health Resource Accessibility](#). *International Conference on Social Computing (ICSC 2025)* (Oral).
- [22] Joseph C.Y. Lau, Emily Landau, **Qingcheng Zeng**, Ruichun Zhang, Stephanie Crawford, Rob Voigt, Molly Losh. [Pre-trained AI language model represents pragmatic language variability central to autism and genetically-related phenotypes](#). *Autism: International Journal of Research and Practice*, 2024.
- [23] Rui Yang*, **Qingcheng Zeng***, Keen You*, Yujie Qiao*, Chia-Chun Hsieh, Benjamin Rosand, Jeremy Goldwasser, Amisha Dave, Tiarnan Keenan, Yu He Ke, Chuan Hong, Nan Liu, Emily Chew, Dragomir Radev, Zhiyong Lu, Hua Xu, Qingyu Chen, Irene Li#. [Ascle—A Python Natural Language Processing Toolkit for Medical Text Generation: Development and Evaluation Study](#). *Journal of Medical Internet Research (JMIR)*, 26:e60601, 2024.
- [24] Jiageng Wu, Xiaocong Liu, Minghui Li, Wanxin Li, Zichang Su, Shixu Lin, Lucas Garay, Zhiyun Zhang, Yujie Zhang, **Qingcheng Zeng**, Jie Shen, Changzheng Yuan, Jie Yang#. [Clinical Text Datasets for Medical Artificial Intelligence and Large Language Models — A Systematic Review](#). *NEJM AI*, 1(6), 2024.
- [25] Mourad Heddaya, **Qingcheng Zeng**, Alexander Zentefis, Rob Voigt, Chenhao Tan. [Causal Micro-Narratives](#). *6th Workshop on Narrative Understanding (WNU)*, *EMNLP 2024*.
- [26] Rui Yang, Haoran Liu, Edison Marrese-Taylor, **Qingcheng Zeng**, Yu He Ke, Wanxin Li, Lechao Cheng, Qingyu Chen, James Caverlee, Yutaka Matsuo, Irene Li. [KG-Rank: Enhancing Large Language Models for Medical QA with Knowledge Graphs and Ranking Techniques](#). *BioNLP Workshop, ACL 2024*.
- [27] Rui Yang, Huitao Li, Weihao Xuan, ..., **Qingcheng Zeng**, ..., Nan Liu. [Toward Global Large Language Models in Medicine](#). *arXiv:2601.02186*, 2026.
- [28] Mourad Heddaya, Chenhao Tan, Rob Voigt, **Qingcheng Zeng**, Alexander Zentefis. [A Century of Inflation Narratives](#). *SSRN Working Paper 5259107*, 2025.
- [29] Philipp Schoenegger, Francesco Salvi, Jiacheng Liu, ..., **Qingcheng Zeng**, ..., Ezra Karger. [When Large Language Models are More Persuasive Than Incentivized Humans, and Why](#). *arXiv:2505.09662*, 2025.

Understanding LLMs: Cognitive Science and Interpretability

- [30] Suet-Ying Lam*, **Qingcheng Zeng***, Rob Voigt. [Encoder-only and decoder-only models exhibit distinct patterns of linguistic bias](#). *EMNLP 2026 (Main)*.
- [31] Ruoxi Ning, Yongpeng Zhu, **Qingcheng Zeng**, Tatsuki Kuribayashi, Freda Shi. [On the Effect of Hyperparameters in Language Modeling for Computational Linguistics](#). *ACL 2026*.
- [32] Kefan Yu*, **Qingcheng Zeng***#, Weihao Xuan, Wanxin Li, Jingyi Wu, Rob Voigt. [The Pragmatic Mind of Machines: Tracing the Emergence of Pragmatic Competence in Large Language Models](#). *EACL 2026 (Oral)*.
- [33] Suet-Ying Lam*, **Qingcheng Zeng***#, Jingyi Wu, Rob Voigt. [Leveraging Human Production-Interpretation Asymmetries to Test LLM Cognitive Plausibility](#). *ACL 2025 (Short Paper)*.
- [34] Mingyu Jin*, Qinkai Yu*, Jingyuan Huang*, **Qingcheng Zeng***, Zhenting Wang, Wenyue Hua, Haiyan Zhao, Kai Mei, Yanda Meng, Kaize Ding, Fan Yang, Mengnan Du, Yongfeng Zhang#. [Exploring Concept Depth: How Large Language Models Acquire Knowledge and Concept at Different Layers?](#) *COLING 2025*.
- [35] Suet-Ying Lam*, **Qingcheng Zeng***#, Kexun Zhang*, Chenyu You, Rob Voigt. [Large Language Models Are Partially Primed in Pronoun Interpretation](#). *Findings of ACL 2023*.
- [36] **Qingcheng Zeng**, An-Ran Li. [A Survey in Automatic Irony Processing: Linguistic, Cognitive, and Multi-X Perspectives](#). *COLING 2022*.
- [37] Mingyu Jin, Yutong Yin, Jingcheng Niu, **Qingcheng Zeng**, Wujiang Xu, Mengnan Du, Wei Cheng, Zhaoran Wang, Tianlong Chen, Dimitris N. Metaxas. [Farther the Shift, Sparser the Representation: Analyzing OOD Mechanisms in LLMs](#). *arXiv:2603.03415*, 2026.
- [38] Ziqi Gao, Yaotian Zhu, **Qingcheng Zeng**, Xu Zhao, Ziqing Wang, Feng Ruan, Kaize Ding. [RAPTOR: Ridge-Adaptive Logistic Probes](#). *arXiv:2602.00158*, 2026.

Multilingual NLP and Machine Translation

- [39] Xinhe Shi, **Qingcheng Zeng**, Weihao Xuan, Linchao Zhu. [How to Improve LLMs' Performance on Specific Languages: A Perspective on LLM-Derived Language Similarity](#). *ACL 2026 (Oral)*.
- [40] Weihao Xuan, Rui Yang, Heli Qi, **Qingcheng Zeng**, ..., Irene Li. [MMLU-ProX: A Multilingual Benchmark for Advanced Large Language Model Evaluation](#). *EMNLP 2025*.
- [41] Daoyang Li, Haiyan Zhao, **Qingcheng Zeng**, Mengnan Du. [Exploring Multilingual Probing in Large Language Models: A Cross-Language Analysis](#). *XLLM Workshop, ACL 2025*.
- [42] **Qingcheng Zeng***, Lucas Garay*, Peilin Zhou*, Dading Chong, Yining Hua, Jiageng Wu, Yikang Pan, Han Zhou, Rob Voigt, Jie Yang#. [GreenPLM: Cross-Lingual Transfer of Monolingual Pre-Trained Language Models at Almost No Cost](#). *IJCAI 2023*.

Other Publications

- [43] Fan Gao, Hang Jiang, Rui Yang, **Qingcheng Zeng**, Jinghui Lu, Moritz Blum, Dairui Liu, Tianwei She, Yuang Jiang, Irene Li. [Evaluating Large Language Models on Wikipedia-Style Survey Generation](#). *Findings of ACL 2024*.
- [44] Dading Chong, Helin Wang, Peilin Zhou, **Qingcheng Zeng**. [Masked Spectrogram Prediction for Self-Supervised Audio Pre-Training](#). *ICASSP 2023*.
- [45] **Qingcheng Zeng**, Dading Chong, Peilin Zhou, Jie Yang#. [Low-resource Accent Classification in Geographically-proximate Settings: A Forensic and Sociophonetics Perspective](#). *Interspeech 2022*.
- [46] Peilin Zhou, Dading Chong, Helin Wang, **Qingcheng Zeng**. [Calibrate and Refine! A Novel and Agile Framework for ASR Error Robust Intent Detection](#). *Interspeech 2022*.
- [47] Sheng Jin, Minghao Liu, Yunze Xiao, ..., **Qingcheng Zeng**, ..., Ge Zhang. [Knowledge Index of Noah's Ark](#). *arXiv:2606.05104*, 2026.

HONORS & AWARDS

- Gold Reviewer, ICML 2026
- Excellent Reviewer, KDD 2025 Applied Data Science Track
- TGS Conference Travel Grant (EMNLP 2024, \$800), Northwestern University
- Departmental Travel Grant (EMNLP 2024, \$400), Northwestern University
- TGS Conference Travel Grant (IJCAI 2023, \$850), Northwestern University
- Graduate Fellowship, Northwestern University
- Q-Step Center Best Dissertation Award (£100), University of Manchester
- The Dean Award for Achievement, University of Manchester

INVITED TALKS

On Verbalized Confidence and Factuality, LARK Lab, HKUST (Guangzhou)

Oct 2025

TEACHING EXPERIENCE

Teaching Assistant, LING 220: Language & Society, Northwestern University Spring 2025

Teaching Assistant, LING 331: Text Processing for Linguists, Northwestern University Winter 2025

Teaching Assistant, COGSCI 110: Introduction to Cognitive Science, Northwestern University Fall 2023

ACADEMIC SERVICE

Conference reviewer: ACL Rolling Review, NeurIPS, ICML, COLM, KDD, WWW, IJCAI, IC2S2, ICASSP

Journal reviewer: Artificial Intelligence; IEEE Journal of Selected Topics in Signal Processing; IEEE Transactions on Neural Networks and Learning Systems (TNNLS); Computer Speech and Language; Knowledge-Based Systems; Expert Systems With Applications; IEEE Access; Information Processing and Management; Neurocomputing; Neural Networks; Pattern Recognition; Applied Soft Computing; JMIR AI; JMIR Medical Informatics; International Journal of Medical Informatics; Telematics and Informatics Reports; Acta Psychologica

Departmental service: Colloquium Committee Member, Northwestern Linguistics 2023–2024

REFERENCES

Rob Voigt (Ph.D. Advisor)
Assistant Professor of Linguistics
University of California, Davis
robvoigt@ucdavis.edu

Lingyao Li (Ph.D. Committee)
Assistant Professor of Information
University of South Florida
lingyaol@usf.edu

Kaize Ding (Ph.D. Co-advisor)
Assistant Professor of Statistics and Data Science
Northwestern University
kaize.ding@northwestern.edu

Matt Goldrick (Ph.D. Committee)
Professor of Linguistics and Psychology (by courtesy)
Northwestern University
matt-goldrick@northwestern.edu