

Qingcheng Zeng

*Ph.D. Student at Northwestern
University*

2016 Sheridan Road, Evanston, IL

✉ qcz@u.northwestern.edu

🌐 qcznlp.github.io

Google Scholar

Education

- 2022–2026 **Ph.D. in Computational Linguistics**
(Expected) Department of Linguistics, Northwestern University, Evanston, IL
○ Advisors: Professor [Rob Voigt](#) and Professor [Kaize Ding](#)
- 2019–2021 **B.A. in Linguistics**
University of Manchester, Manchester, UK
- 2017–2021 **B.A. in Linguistics**
Zhejiang University, Hangzhou, China

Research Interests

- LLMs for Computational Social Science and Social Good
- Information Retrieval and Search Agent
- Understanding LLMs through Cognitive Science
- Trustworthy LLMs, Uncertainty Quantification, Verbalized Confidence
- Multilingual NLP, Machine Translation

Research Appointment

- Sep 2025 - **Research Scientist Intern**
Dec 2025 Snowflake, Menlo Park, CA
○ Working with [Puxuan Yu](#) and [Aman Mehta](#) in Cortex Search AI team.
- June 2025 - **Research Scientist Intern**
Aug 2025 Tencent AI Lab America, Bellevue, WA
○ Working with [Linfeng Song](#), [Xiaoyang Wang](#), [Hongming Zhang](#), and [Wenhao Yu](#), focusing on process reward modeling for agents.
- Sep 2024 - **Visiting Graduate Student**
Nov 2024 WestlakeNLP, Westlake University
○ Conducted research in multilingual calibration, supervised by Professor [Yue Zhang](#) and Dr [Leyang Cui](#).
○ Research Output: reasoning model calibration research at [EMNLP 2025](#), LLM evaluation at [NeurIPS 2025](#).
- 2024 **Visiting Graduate Student**
Summer Center for Language and Speech Processing, Johns Hopkins University

- Conducted research in AI-curated democratic discourse, supervised by Professor **Jason Eisner**.
- 2021–2022 **Research Assistant in Natural Language Processing**
Zhejiang University, Hangzhou, China
- Conducted research in multilingual NLP and biomedical NLP, supervised by Professor **Jie Yang**.
- Research Output: effective language model adaptation at *IJCAI 2023*.

Main Publications

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

LLMs for Computational Social Science and Social Good

- ICSC 2025 Zhaoqian Xue, Guanhong Liu, Chong Zhang, Kai Wei, **Qingcheng Zeng**,
Oral Songhua Hu, Wenyue Hua, Lizhou Fan, Yongfeng Zhang, Lingyao Li, 2025, *Toward Equitable Access: Leveraging Crowdsourced Reviews to Investigate Public Perceptions of Health Resource Accessibility*, Accepted at International Conference on Social Computing 2025 (*Oral*)
- Findings of ***Qingcheng Zeng**, *Guanhong Liu, Zhaoqian Xue, Diego Ford, Rob Voigt,
AAACL 2025 Loni Hagen, Lingyao Li, 2025, *Sympathy over Polarization: A Computational Discourse Analysis of Social Media Posts about the July 2024 Trump Assassination Attempt*, In *Proceedings of the 14th International Joint Conference on Natural Language Processing and the 4th Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics*, pages 56–68, Mumbai, India. The Asian Federation of Natural Language Processing and The Association for Computational Linguistics.
- EMNLP Grace LeFevre, **Qingcheng Zeng**, Adam Leif, Jason Jewell, Denis Peskoff, Rob
2025 Voigt, 2025, *Good Intentions Beyond ACL: Who Does NLP for Social Good, and Where?*, In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 5138–5150, Suzhou, China. Association for Computational Linguistics.
- Autism Joseph C.Y. Lau, Emily Landau, **Qingcheng Zeng**, Ruichun Zhang, Stephanie
Crawford, Rob Voigt, Molly Losh, 2024, *Pre-trained AI language model represents pragmatic language variability central to autism and genetically-related phenotypes*, In *Autism: International Journal of Research and Practice*
- EMNLP **Qingcheng Zeng**, Mingyu Jin, Rob Voigt, 2024, *Adaptive Axes: A Pipeline for
2024 In-domain Social Stereotype Analysis*, In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 15576–15593, Miami, Florida, USA. Association for Computational Linguistics.
- JMIR 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#, 2024, *Ascle—A Python Natural Language Processing Toolkit for Medical Text Generation: Development and Evaluation Study*, In *Journal of Medical Internet Research* 2024;26:e60601

- WNU Mourad Heddaya, **Qingcheng Zeng**, Alexander Zentefis, Rob Voigt, Chenhao Tan, 2024, **Causal Micro-Narratives**, In *Proceedings of The 6th Workshop on Narrative Understanding*, pages 67–84, Miami, Florida, USA. Association for Computational Linguistics.
- Trustworthy LLMs and Uncertainty Quantification
- EMNLP 2025 **Qingcheng Zeng***, Weihao Xuan*, Leyang Cui, Rob Voigt, 2025, **Thinking Out Loud: Do Reasoning Models Know When They’re Right?**, In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 1394–1407, Suzhou, China. Association for Computational Linguistics.
- EMNLP 2025 Oral Weihao Xuan*, **Qingcheng Zeng***, Heli Qi, Junjue Wang, Naoto Yokoya, 2025, **Seeing is Believing, but How Much? A Comprehensive Analysis of Verbalized Calibration in Vision-Language Models**, In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 1408–1450, Suzhou, China. Association for Computational Linguistics.
- CIKM 2025 **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, 2025, **Uncertainty Quantification for Multiple-Choice Questions is Just One-Token Deep**, In *Proceedings of the 34th ACM International Conference on Information and Knowledge Management (CIKM ’25)*. Association for Computing Machinery, New York, NY, USA, 5474–5478.
- Information Retrieval and Search Agent
- Findings of EACL 2026 Minghao Guo, **Qingcheng Zeng**, Xujiang Zhao, Yanchi Liu, Wenchao Yu, Mengnan Du, Haifeng Chen, Wei Cheng, 2026, **DeepSieve: Information Sieving via LLM-as-a-Knowledge-Router**, Accepted at Findings of EACL 2026
- CIKM 2025 Haoyu Wu*, **Qingcheng Zeng***, Kaize Ding, 2025, **Fact or Facsimile? Evaluating the Factual Robustness of Modern Retrievers**, In *Proceedings of the 34th ACM International Conference on Information and Knowledge Management (CIKM ’25)*. Association for Computing Machinery, New York, NY, USA, 5351–5355.
- Understanding LLMs through Cognitive Science
- EACL 2026 Kefan Yu*, **Qingcheng Zeng*#**, Weihao Xuan, Wanxin Li, Jingyi Wu, Rob Voigt, 2026, **The Pragmatic Mind of Machines: Tracing the Emergence of Pragmatic Competence in Large Language Models**, Accepted at EACL 2026
- ACL 2025 Suet-Ying Lam*, **Qingcheng Zeng*#**, Jingyi Wu, Rob Voigt, 2025, **Leveraging Human Production-Interpretation Asymmetries to Test LLM Cognitive Plausibility**, In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 158–171, Vienna, Austria. Association for Computational Linguistics.

- COLING 2025 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#, 2025, **Exploring Concept Depth: How Large Language Models Acquire Knowledge and Concept at Different Layers?**, In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 558–573, Abu Dhabi, UAE. Association for Computational Linguistics.
- Findings of ACL 2023 Suet-Ying Lam*, **Qingcheng Zeng*#**, Kexun Zhang*, Chenyu You, Rob Voigt, 2023, **Large Language Models Are Partially Primed in Pronoun Interpretation**, In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 9493–9506, Toronto, Canada. Association for Computational Linguistics.
- COLING 2022 **Qingcheng Zeng** and An-Ran Li, 2022, **A Survey in Automatic Irony Processing: Linguistic, Cognitive, and Multi-X Perspectives**, In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 824–836, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.
- Multilingual NLP and Machine Translation**
- EMNLP 2025 Weihao Xuan, Rui Yang, Heli Qi, **Qingcheng Zeng**, ..., Irene Li, 2025, **MMLU-ProX: A Multilingual Benchmark for Advanced Large Language Model Evaluation**, In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 1513–1532, Suzhou, China. Association for Computational Linguistics.
- IJCAI 2023 **Qingcheng Zeng***, Lucas Garay*, Peilin Zhou*, Dading Chong, Yining Hua, Jiageng Wu, Yikang Pan, Han Zhou, Rob Voigt, Jie Yang#, 2023, **GreenPLM: Cross-Lingual Transfer of Monolingual Pre-Trained Language Models at Almost No Cost**, In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, pages 6290–6298, Macau, China. International Joint Conferences on Artificial Intelligence.
- XLLM Daoyang Li, Haiyan Zhao, **Qingcheng Zeng**, Mengnan Du, 2025, **Exploring Multilingual Probing in Large Language Models: A Cross-Language Analysis**, In *Proceedings of the 1st Joint Workshop on Large Language Models and Structure Modeling (XLLM 2025)*, pages 61–70, Vienna, Austria. Association for Computational Linguistics.
- Other Publications**
- NeurIPS 2025 Sijia Chen, Xiaomin Li, Mengxue Zhang, Eric Hanchen Jiang, **Qingcheng Zeng**, Chen-Hsiang Yu, 2025, **CARES: Comprehensive Evaluation of Safety and Adversarial Robustness in Medical LLMs**, Accepted at NeurIPS 2025 D&B Track
- NeurIPS 2025 Shulin Huang, Linyi Yang, Yan Song, Shuang Chen, Leyang Cui, Ziyu Wan, **Qingcheng Zeng**, Ying Wen, Kun Shao, Weinan Zhang, Jun Wang, Yue Zhang, 2025, **ThinkBench: Dynamic Out-of-Distribution Evaluation for Robust LLM Reasoning**, Accepted at NeurIPS 2025 D&B Track

- BioNLP Rui Yang, Haoran Liu, Edison Marrese-Taylor, **Qingcheng Zeng**, Yu He Ke, Wanxin Li, Lechao Cheng, Qingyu Chen, James Caverlee, Yutaka Matsuo, Irene Li, 2024, **KG-Rank: Enhancing Large Language Models for Medical QA with Knowledge Graphs and Ranking Techniques**, In *Proceedings of the 23rd Workshop on Biomedical Natural Language Processing*, pages 155–166, Bangkok, Thailand. Association for Computational Linguistics.
- Findings of ACL 2024 Fan Gao, Hang Jiang, Rui Yang, **Qingcheng Zeng**, Jinghui Lu, Moritz Blum, Dairui Liu, Tianwei She, Yuang Jiang, Irene Li, 2024, **Evaluating Large Language Models on Wikipedia-Style Survey Generation**, In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 5405–5418, Bangkok, Thailand. Association for Computational Linguistics.
- NEJM AI 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#, 2024, **Clinical Text Datasets for Medical Artificial Intelligence and Large Language Models — A Systematic Review**, In *New England Journal of Medicine AI*, 1(6)
- ICASSP 2023 Dading Chong, Helin Wang, Peilin Zhou, **Qingcheng Zeng**, 2023, **Masked Spectrogram Prediction for Self-Supervised Audio Pre-Training**, In *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1-5, Rhodes Island, Greece. IEEE.
- IS 2022 **Qingcheng Zeng**, Dading Chong, Peilin Zhou, Jie Yang#, 2022, **Low-resource Accent Classification in Geographically-proximate Settings: A Forensic and Sociophonetics Perspective**, In *Proceedings of Interspeech 2022*, pages 5308-5312, Incheon, Korea. International Speech Communication Association.
- IS 2022 Peilin Zhou, Dading Chong, Helin Wang, **Qingcheng Zeng**, 2022, **Calibrate and Refine! A Novel and Agile Framework for ASR Error Robust Intent Detection**, In *Proceedings of Interspeech 2022*, pages 1096-1100, Incheon, Korea. International Speech Communication Association.

Honors & Awards

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

Invited Talks

- Oct 17, 2025 **On Verbalized Confidence and Factuality**, **LARK Lab**, HKUST (GZ)

Teaching Experiences

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

2025 Winter **Teaching Assistant**, *LING331*, Text Processing for Linguists, Northwestern University

2025 Spring **Teaching Assistant**, *LING220*, Language & Society, Northwestern University

Service

Conferences Reviewer **ICTAI 2022, ACL 2022, ICASSP 2024/2025, ACL Rolling Review, IC2S2 2024, WWW 2025/2026, KDD 2025/2026, IJCAI 2025, NeurIPS 2025**

Journal Reviewer **Computer Speech and Language, Knowledge-Based Systems, Expert Systems With Applications, IEEE Access, International Journal of Medical Informatics, Information Processing and Management, Neurocomputing, Neural Networks, Telematics and Informatics Reports, Pattern Recognition, Acta Psychologica**

Departmental Service **Colloquium Committee Member 2023-2024**

References

- **Rob Voigt** (PhD Advisor)
Department of Linguistics
Assistant Professor of Linguistics
UC Davis
Davis, CA
robvoigt@ucdavis.edu
- **Kaize Ding** (PhD Co-advisor)
Department of Statistics and Data Science
Assistant Professor of Statistics and Data Science
Northwestern University
Evanston, IL
kaize.ding@northwestern.edu
- **Lingyao Li** (PhD Committee)
School of Information
Assistant Professor of Information
University of South Florida
Tampa, FL
lingyaol@usf.edu

- **Matt Goldrick** (PhD Committee)
Department of Linguistics
Professor of Linguistics and Psychology (by courtesy)
Northwestern University
Evanston, IL
matt-goldrick@northwestern.edu