

Qingcheng Zeng

Ph.D. Student at Northwestern
University

2016 Sheridan Road, Evanston, IL

✉ qcz@u.northwestern.edu

🌐 [qcznlp.github.io](https://github.com/qcznlp)

Education

- 2022–2027 **Ph.D. in Computational Linguistics**
(Expected) Department of Linguistics, Northwestern University, Evanston, IL
- Advisor: Professor [Rob Voigt](#) and Professor [Kaize Ding](#)
- 2019–2021 **B.A. in Linguistics**
University of Manchester, Manchester, UK
- B.A. Dissertation: "*Automatic Sociophonetics: Investigating Language Variation in General Northern English with Machine Learning Based Classification Methods*," supervised by [Dr. Patrycja Strycharczuk](#).
- 2017–2021 **B.A. in Linguistics**
Zhejiang University, Hangzhou, China

Research Interests

- **NLP for Social Applications**
I am curious about the potential of large language models (LLMs) to enhance our comprehension of political, economic, and social trends, promote more meaningful discussions, and how social-oriented benchmarks can assist in evaluating these models.
- **Trustworthy LLMs**
My goal is to evaluate and improve the calibration of LLMs.
- **Multilingual NLP**
I am interested in evaluating multilingual LLMs, creating multilingual benchmarks, and understanding the inner architecture of these models to help build a fairer language technology world.

Research Appointment

- Sep 2024 - **Visiting Graduate Student**
Nov 2024 WestlakeNLP, Westlake University
- Conducted research in multilingual calibration, supervised by Professor [Yue Zhang](#) and Dr [Leyang Cui](#).
- 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*.

Publications

Referred Conference/Journal Articles

* denotes authors of equal contribution; # denotes corresponding author

- EMNLP 2024 **Qingcheng Zeng**, Mingyu Jin, Rob Voigt, 2024, Adaptive Axes: A Pipeline for In-domain Social Stereotype Analysis, Accepted in *EMNLP 2024 Main Conference*
- 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**, Accepted in *Journal of Medical Internet Research*
- 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, Xiacong 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)
- 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.
- 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.
- 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.

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.

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.

Articles in Peer-reviewed Workshop Proceedings

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.

W-NUT 2020 **Qingcheng Zeng**, Xiaoyang Fang, Zhixin Liang, Haoding Meng, 2020, **Fancy Man Launches Zippo at WNUT 2020 Shared Task-1: A Bert Case Model for Wet Lab Entity Extraction**, In *Proceedings of the Sixth Workshop on Noisy User-generated Text (W-NUT 2020)*, pages 299-304, Online. Association for Computational Linguistics.

Abstracts and Conference Posters

LabPhon 19 Seung-Eun Kim, **Qingcheng Zeng**, Bronya R. Chernyak, Joseph Keshet, Matthew Goldrick, Ann R. Bradlow, 2024, Quantifying perceptual similarity of connected speech, In *the 19th Conference on Laboratory Phonology (LabPhon 19)*, Hanyang University, June 2024

NWAV 49 Fenqi Wang, Delin Deng, **Qingcheng Zeng**, Ratrete Wayland, Kevin Tang, 2021, A corpus study of the production of syllable-initial /l/ and /n/ in Gan-accented Mandarin: Random forest classification, In *New Ways of Analyzing Variation*, University of Texas at Austin, October 2021

UKLVC 13 **Qingcheng Zeng**, 2021, Machine learning in sociophonetics: Comparison of different approaches to accent classification and feature extraction, In *UK Language Variation and Change 13*, University of Glasgow, September 2021

Preprints

* denotes authors of equal contribution; # denotes corresponding author

arXiv 2024 **Qingcheng Zeng***, Mingyu Jin*, Qinkai Yu*, Zhenting Wang, Wenyue Hua, Zihao Zhou, Guangyan Sun, Yanda Meng, Shiqing Ma, Qifan Wang, Felix Juefei-Xu, Kaize Ding, Fan Yang, Ruixiang Tang, Yongfeng Zhang, 2024, **Uncertainty is Fragile: Manipulating Uncertainty in Large Language Models**, In arXiv.

- arXiv 2024 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, 2024, **Exploring Concept Depth: How Large Language Models Acquire Knowledge at Different Layers?**, In arXiv.
- arXiv 2022 Zezhong Jin, Dading Chong, Xiao Song, Zhaoyi Liu, Naipeng Ye, **Qingcheng Zeng**, 2022, **Filter and evolve: progressive pseudo label refining for semi-supervised automatic speech recognition**, In arXiv.

Honors & Awards

- Q-Step Center Best Dissertation Award (£100) University of Manchester
- The Dean Award for Achievement University of Manchester
- Graduate Fellowship Northwestern University

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 **ICTAI 2022, ACL 2022, ICASSP 2024/2025, ACL Rolling Review 2024**
Reviewer
- Journal **Computer Speech and Language, Knowledge-Based System**
Reviewer
- Departmental **Colloquium Committee Member 2023-2024**
Service

Skills

- **NLP and LLMs**
Huggingface Transformers, Deepspeed
- **Machine Learning and Deep Learning**
PyTorch, TensorFlow, Scikit-learn
- **Statistics**
Python, R

References

- **Rob Voigt**
Department of Linguistics
Assistant Professor of Linguistics and Computer Science (by courtesy)
Northwestern University
Evanston, IL
robvoigt@northwestern.edu

- **Kaize Ding**
Department of Statistics and Data Science
Assistant Professor of Statistics and Data Science
Northwestern University
Evanston, IL
kaize.ding@northwestern.edu

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