Intern Application

Jirui Qi

betswish.github.io | j.qi@rug.nl

Current Employment

PhD Candidate, University of Groningen, Netherlands   **04/2023 – 03/2027**
- Research area: alignment in generative LM, reinforcement learning, reasoning, multilinguality, RAG, explainability
- Supervisors: Prof. Arianna Bisazza and Prof. Raquel Fernández
- Funding: NWA-ORC project LESSEN (grant nr. NWA.1389.20.183)

Incoming Academic Guest, ETH Zürich, Switzerland   **03/2026 – 06/2026**
- Research area: reinforcement learning, multilingual LMs
- Advisors: Prof. Ryan Cotterell and Prof. Mrinmaya Sachan

Honors and Grants

- **2025 State of AI Report**: Work on multilingual reasoning featured among the year's key research contributions
- **2023 EMNLP Outstanding Paper Award**: Top 1.5% of 2,011 accepted papers
- **2023 GenBench-23 Workshop**: Best Data Award

Publications

[Google Scholar] (Citations:143) (*=Equal Contribution)

[0] Mitigating Language Mismatch Issue of Thinking Traces in Large Reasoning Models with GRPO
Shan Chen*, **Jirui Qi*,** Zidi Xiong, Raquel Fernández, Arianna Bisazza, Danielle S. Bitterman
On-going project

[1] Optimizing Language Models for Crosslingual Knowledge Consistency via Reinforcement Learning
Tianyu Liu*, **Jirui Qi*,** Raquel Fernández, Mrinmaya Sachan, Arianna Bisazza, Ryan Cotterell
Paper submitted to ICLR in September 2025, under review

[2] When Models Reason in Your Language: Controlling Thinking Language Comes at the Cost of Accuracy (Link)
Jirui Qi*, Shan Chen*, Zidi Xiong, Raquel Fernández, Danielle S. Bitterman, Arianna Bisazza
*EMNLP 2025 Findings   /   **Featured in State of AI Report 2025 (Research section highlight)***

[3] On the consistency of multilingual context utilization in retrieval-augmented generation (Link)
Jirui Qi, Raquel Fernández, Arianna Bisazza
MRL 2025 Workshop

[4] Likelihood as a Performance Gauge for Retrieval-Augmented Generation (Link)
Tianyu Liu*, **Jirui Qi*,** Paul He, Arianna Bisazza, Mrinmaya Sachan, Ryan Cotterell
NAACL 2025 Main

[5] Model Internals-based Answer Attribution for Trustworthy Retrieval-Augmented Generation (Link)
Jirui Qi*, Gabriele Sarti*, Raquel Fernández, Arianna Bisazza
EMNLP 2024 Main

[6] SIFo Benchmark: Investigating Sequential Instruction Following Ability of Large Language Models (Link)
Xinyi Chen, Baohao Liao, **Jirui Qi**, Panagiotis Eustratiadis, Christof Monz, Arianna Bisazza, Maarten de Rijke
Findings of EMNLP 2024

[7] Democratizing Advanced Attribution Analyses of Generative Language Models with Inseq Toolkit (Link)
Gabriele Sarti, Nils Feldhus, **Jirui Qi**, Malvina Nissim, Arianna Bisazza
The 2nd World Conference on eXplainable Artificial Intelligence, 2024

[8] Cross-Lingual Consistency of Factual Knowledge in Multilingual Language Models (Link)
Jirui Qi, Raquel Fernández, Arianna Bisazza
EMNLP 2023 Outstanding Award (Top 1.5% of 2,011 accepted papers)

[9] Multi-Mask Label Mapping for Prompt-Based Learning (Link)
Jirui Qi, Richong Zhang, Jaein Kim, Junfan Chen, Wenyi Qin, Yongyi Mao
AAAI-23

[10] Parameter-free Automatically Prompting: Latent Pseudo Label Mapping Model for Prompt-based Learning (Link)
Jirui Qi, Richong Zhang, Junfan Chen, Jaein Kim, Yongyi Mao
Findings of EMNLP 2022

[11] Cross Domain Few-Shot Learning via Meta Adversarial Training (Link)
Jirui Qi, Richong Zhang, Chune Li, Yongyi Mao
arXiv Preprint, 2022

Industrial Experience

Research Intern   **03/2025 – 03/2026**
KPN Research
Amsterdam, North Holland, Netherlands

Teaching Experience

LIX030B05: Introduction to Neural Network (Bachelor Course, Fall 2025)   Teaching assistant
Instructor: Prof. Arianna Bisazza, University of Groningen

LIX006M05: Advanced Topics in Natural Language Processing (Master Course, Spring 2025)   Teaching assistant
Instructor: Prof. Arianna Bisazza, University of Groningen

LIX030B05: Introduction to Neural Network (Bachelor Course, Fall 2024)   Teaching assistant
Instructor: Prof. Arianna Bisazza, University of Groningen

LIX001M05: Natural Language Processing (Master Course, Spring 2024)   Teaching assistant
Instructor: Prof. Arianna Bisazza, University of Groningen

Reviewing Experience

ICLR 2025   Reviewer
ACL ARR - 2025 May   Reviewer
ACL ARR - 2025 February   Reviewer
ACL ARR - 2024 October   Reviewer
ACL ARR - 2024 August   **Great Reviewer**
GenBench - 2024 Workshop   Reviewer

Past Education

Master: Beihang University   09/2020 - 01/2023
Bachelor: Beijing Jiaotong University   09/2016 - 07/2020
Senior High: The High School Affiliated to Renmin University of China (RDFZ)   09/2013 - 07/2016

Skill & Behavior

Computer Skills: Microsoft Azure, TRL, Python, PyTorch, Linux, C++
Personal Behavior: Self-motivated, efficiency under pressure, curiosity, fast learner, effective time management