

Qixiang Fang

PHD CANDIDATE IN METHODOLOGY & STATISTICS (SPECIALIZATION: NLP)

Padualaan 14, 3584CH Utrecht, the Netherlands

+31 684666253 | q.fang@uu.nl | github.com/fqixiang | [linkedin.com/in/fqixiang](https://www.linkedin.com/in/fqixiang) |  Qixiang Fang

Personal Profile

Pronunciation of my name: [ki-sjan fang].

Background in statistics, methodology, natural language processing (NLP), and social sciences. Hope to bring validity (e.g., measurement error), reliability (e.g., model robustness), and causal inference into NLP research. Looking for internships in NLP and computational social sciences.

Education

Department of Methods & Statistics, Utrecht University

Utrecht, The Netherlands

PhD Candidate in Statistics & NLP

2020 - Now

- Conducting research on topics related to validity and reliability in NLP and computational social sciences.
- Developing new software features for JASP (Open-source Statistical Software).
- Teaching methodology and text mining workshops to social science students.
- Providing statistical consultation to social science students and researchers.
- Assisting the The Interuniversity Graduate School of Psychometrics and Sociometrics (IOPS) with organisational issues.
- Supervisors: Prof. Dr. Daniel Oberski & Dr. Dong Nguyen.
- **Technical Skills:** Python, R, Git, QML.

Utrecht University

Utrecht, The Netherlands

MSc in Statistics & Methodology

2017 - 2019

- Cum laude and best thesis of the year.
- Education committee member & student representative.

Jacobs University

Bremen, Germany

Joint BA in Psychology & Social Sciences

2013 - 2016

- Cum Laude.

Work Experience

Department of Information and Computing Sciences, Utrecht University

Utrecht, The Netherlands

Research Assistant and Lecturer

2019 - 2020

- Analysed user process data in online educational tools to gain insights into students' learning behaviour.
- Published a paper in the European Conference on Technology Enhanced Learning.
- Taught three introductory statistics courses to bachelor computer science students (as the main lecturer).
- Supervised bachelor theses.
- **Technical Skills:** R, Python, SQL.

Department of Methodology & Data Collection, Statistics Netherlands

Heerlen, The Netherlands

Research Intern

2018 - 2019

- Used interpretable machine learning models to understanding and predict daily web survey response rates from contextual information like time, weather and Google Trends. This resulted in a journal publication.
- **Technical Skills:** R, LaTeX, GitHub.

Department of Methods & Statistics, Utrecht University

Utrecht, The Netherlands

Research Assistant

2018 - 2019

- Developed online statistical tutorials (e.g., *discrete-time survival analysis*, *multilevel GLM*).
- Collected datasets and preprocessed data for the *ASReview software team*.
- **Technical Skills:** R, Python, GitHub.

WINIT GmbH

Bremen, Germany

Data Analyst

2016 - 2017

- Automated the update and retrieval of warehouse data.
- Advised on warehouse storage and in-house logistics solutions based based on data-driven insights.
- **Technical Skills:** Python, JavaScript, web scraping.

Publications

- **Fang, Q.**, Nguyen, D. & Oberski, D.L. (2022) Evaluating the Construct Validity of Text Embeddings with Application to Survey Questions. *EPJ Data Science*, 11(39). [Link](#).

- **Fang, Q.**, Burger, J., Meijers, R., & van Berkel, K. (2021). The Role of Time, Weather and Google Trends in Understanding and Predicting Web Survey Response. *Survey Research Methods*, 15(1), 1-25. Link.
- Du, Y., **Fang, Q.**, & Nguyen, D. (2021). Assessing the Reliability of Word Embedding Gender Bias Measures. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Link.
- Arts, I., **Fang, Q.**, van de Schoot, R., & Meitinger, K. (2021). Approximate Measurement Invariance of Willingness to Sacrifice for the Environment Across 30 Countries: The Importance of Prior Distributions and Their Visualization. *Frontiers in Psychology*, 12. Link.
- Sosnovsky, S., **Fang, Q.**, de Vries, B., Luehof, S., & Wiegant, F. (2020). Towards Adaptive Social Comparison for Education. *Proceedings of the 15th European Conference on Technology Enhanced Learning*. Link.
- van de Schoot, R., de Bruin, J., Schram, R., Zahedi, P., de Boer, J., Weijdemans, F., Kramer, B., Huijts, M., Hoogerwerf, M., Ferdinands, G., Harkema, A., Willemsen, J., Ma, Y., **Fang, Q.**, Tummers, L., & Oberski, D. L. (2021). An open source machine learning framework for efficient and transparent systematic reviews. *Nature Machine Intelligence*, 3(2), 125-133. Link.

Manuscripts under Review

- **Fang, Q.**, Giachanou, A. & Bagheri, A. Modelling Stance Detection as Textual Entailment Recognition and Leveraging Measurement Knowledge from Survey Research.
- **Fang, Q.**, Giachanou, A. & Bagheri, A., Boeschoten, L., van Kesteren, E., Kamalabad, M., & Oberski, D.L. On Text-based Personality Computing: Challenges and Future Directions.

Skills

Programming	Python (pandas, PyTorch, scikit-learn, etc.), R (tidyverse, etc.), Git, SQL, JavaScript, QML, etc.
Statistics	Generalised linear models, latent variable models, multilevel models, survival analysis, experimental and survey design, etc.
Soft Skills	Clear communication, asking (relevant) questions, attention to details.
Languages	Fluent English, native Mandarin Chinese, basic German (B1) and Dutch (B1).

Other Achievements

2021	Winner Team (as Main Contributor) , The Complexaton Data Challenge at Utrecht University	<i>The Netherlands</i>
2020	Winner Team (as Main Contributor) , The Big Data Challenge at the BigSurv20 conference	<i>The Netherlands</i>

References available upon request.