

SIJIA LIU

+1 (412) 320-5164 | sijial@alumni.cmu.edu | [Website](#) | [Google Scholar](#) | [GitHub](#) | [LinkedIn](#)

EDUCATION

Carnegie Mellon University

M.S. in Statistics and Data Science

Pittsburgh, USA

August 2017 – May 2018

- Advisor: Howard J. Seltman, Joel Greenhouse
- GPA: 4.06/4.0

Peking University

B.Law. in Sociology, B.A. in Economics

Beijing, China

September 2013 – July 2017

- GPA: 3.84/4.0 (ranked 1/70)

RESEARCH INTEREST

- Building social/commonsense-aware and human-centered NLP systems.
- Aligning large language models with social values.

PUBLICATIONS

1. “Towards Credible Human Evaluation of Open-Domain Dialog Systems Using Interactive Setup”,
Sijia Liu, Patrick Lange, Behnam Hedayatnia, Alexandros Papangelis, Di Jin, Andrew Wirth, Yang Liu, Dilek Hakkani-Tür.
Proceedings of the AAAI Conference on Artificial Intelligence. **AAAI, 2023 (Oral and Poster)**; also presented orally at EMNLP GEM 2022 and NeurIPS HEGM 2022.
2. “Improving Bot Response Contradiction Detection via Utterance Rewriting”,
Di Jin, **Sijia Liu**, Yang Liu, and Dilek Hakkani-Tur.
The 23rd Annual Meeting of the Special Interest Group on Discourse and Dialogue. **SigDial, 2022**.
3. “MERCY: Multiple Response Ranking Concurrently in Realistic Open-Domain Conversational Systems”,
Sarik Ghazarian, Behnam Hedayatnia, Di Jin, **Sijia Liu**, Nanyun (Violet) Peng, Yang Liu, Dilek Hakkani-Tur.
The 24rd Annual Meeting of the Special Interest Group on Discourse and Dialogue. **SigDial, 2023**.
4. “DialGuide: Aligning Dialogue Model Behavior with Developer Guidelines”,
Prakhar Gupta, Yang Liu, Di Jin, Behnam Hedayatnia, Spandana Gella, **Sijia Liu**, Patrick Lange, Julia Hirschberg, Dilek Hakkani-Tur.
The 2023 Conference on Empirical Methods on Natural Language Processing. **EMNLP 2023**.

RESEARCH EXPERIENCE

Amazon Alexa AI

Research Scientist

Sunnyvale, USA

September 2021 – March 2023

- Supervisor: Yang Liu, Prof. Dilek Hakkani-Tur
- Research Topics:
 - Interactive Human Evaluation for Open-Domain Dialogue Systems.
 - Response Generation and Ranking for Conversational AI.
 - Contradiction Detection via Utterance Rewriting.
 - Situated Story Generation.

Peking University

Undergraduate Thesis

Beijing, China

August 2016 – May 2017

- Thesis Advisor: Prof. [Jianxin Li](#)

- Research Topic: Conducted multivariate logistic regression analysis to explore the impact of intergenerational relationship on fertility behavior in rural Chinese families through two major channels of time transfers and economic transfers with China Health and Retirement Longitudinal Study (CHARLS) 2011-2013 panel data.
- Thesis won **1st place in quantitative research** of the cohort and **2nd place over the past 5 years**.

Peking University

Undergraduate Researcher

Beijing, China

May 2016 – May 2017

- Advisor: Prof. [Lin Tao](#)
- Research Topic: Conducted field research and case studies on two well-known undergraduate startups to explore the impact of social recognition in entrepreneurial motivation dynamics from the perspective of startup lifecycle.
- Research was funded by **Peking University Principal Fellowship** (Top 5%).

Peking University

Co-Lead Undergraduate Researcher

Beijing, China

February 2016 – May 2016

- Advisor: Prof. [Yunfeng Lu](#)
- Research Topic: Conducted pluralistic discourse analysis on Chinese internet buzzwords to explore how pragmatics reflect speakers' social identity and contribute to the popularity of those internet buzzwords.
- Research won the **Outstanding Research Award** in the university-wide Multidisciplinary Challenge Cup Academic Competition – Special Contribution Track (Top 1%).

PROFESSIONAL EXPERIENCE

Amazon Alexa AI

Research Scientist @ AlexaLLM Team

Sunnyvale, USA

March 2023 – present

- Building ChatGPT-like Alexa experience with supervised fine-tuning, automatic and human evaluations.
- Improving API-aware model performance with multi-stage curriculum learning.

Amazon Web Services

Data Scientist @ EC2 Core Science Team

Seattle, USA

October 2018 – March 2021

- Built tree-based ensemble models to perform EC2 multi-class workload classification with high-dimensional data.
- Built causal inference methods to differentiate the cannibalization effects of new EC2 instance launch.
- Built time series models to predict EC2 growth with seasonality and moving holidays.

Carnegie Mellon University

Graduate Student Statistician

Pittsburgh, USA

August 2017 – May 2018

- Built hierarchical Zero-Inflated Negative Binomial regression models to predict baseball player injuries.
- Built a Shazam-like song identifier app to process and identify 200 songs' footprints using signal processing.
- Wrote a program to auto-complete a search entry given a prefix and 20k local records using prefix tree.

Tencent Research Institute

Data Scientist Intern

Beijing, China

September 2016 – February 2017

- Ran simulations to estimate digital lucky money spendings during Chinese spring festival.
- Analyzed WeChat user data to build user profiles and understand user behaviors.

Pfizer

Quantitative Consultant Intern

Beijing, China

March 2017 – June 2017

- Analyzed survey data and drug market scenarios under different projected quality evaluation policies.
- Analyzed big data applications for pharmaceutical companies, accelerated partnerships with 2 AI startups.

TEACHING EXPERIENCE

Department of Statistics, Carnegie Mellon University

Pittsburgh, PA

Graduate Teaching Assistant at Experimental Design

February 2017 – May 2018

- Held weekly programming labs and office hours to help 50 students understand the course materials better.
- Graded student homework and exams for 10 hours each week, advised over 20 students on their capstone projects.

SERVICE EXPERIENCE

Conference Reviewing Activities

- EMNLP GEM 2023

LEADERSHIP & COMMUNITY SERVICE

-
- Bay Area & Seattle Chinese Culture Salon, Volunteer
 - Seattle Theater Art Studio, Co-founder
 - China Social Enterprise and Investment Forum, Program manager.
 - Peking University Global Pioneer Spirit Walkathon, Co-organizer.
 - Peking University Student Association, Director of Department of Publicity.
 - Peking University Zhijiang Development Forum, Director of Department of Organization.
 - Peking University Volleyball Association, Leader of female volleyball team.

SELECTED AWARDS & SCHOLARSHIPS

-
- | | |
|--|-----------|
| • Statistical Practice Scholarship, Carnegie Mellon University. | 2018 |
| • Pacesetter to Merit Student, Peking University. | 2014-2017 |
| • Outstanding Undergraduate of the Class, Peking University. | 2017 |
| • Outstanding Undergraduate Thesis, Peking University. | 2017 |
| • Outstanding Research Award at Challenge Cup Academic Competition, Peking University. | 2016 |
| • Principal Fellowship for Undergraduate Research, Peking University. | 2016 |
| • Samsung Scholarship, Peking University. | 2016 |
| • National Scholarship, Peking University. | 2015 |
| • May 4 th Scholarship, Peking University. | 2014 |

ADDITIONAL INFORMATION

Technology and Language Skills:

- Programming Languages: Python, R, SQL, Matlab, Stata, SAS, SPSS
- Development Skills: Pytorch, Tensorflow, PySpark

Interests:

- Reading, Badminton, Volleyball