

SIJIA LIU

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

RESEARCH INTERESTS

Social Commonsense Reasoning, Human-Centered NLP, Large Language Models, AI Alignment, AI for Social Good

EDUCATION

CARNEGIE MELLON UNIVERSITY

August 2017 – May 2018

M.S., Statistics and Data Science, GPA: 4.06/4.0 (Top 10%)

PEKING UNIVERSITY

Sept 2013 – July 2017

B.Law., Sociology, B.A., Economics, GPA: 3.84/4.0 (Top 1%)

SELECTED PUBLICATIONS

- **Sijia Liu***, Patrick Lange, Behnam Hedayatnia, Alexandros Papangelis, Di Jin, Andrew Wirth, Yang Liu, Dilek Hakkani-Tür, “[Towards Credible Human Evaluation of Open-Domain Dialog Systems Using Interactive Setup](#)”. AACL, 2023 (Oral and poster presentations; also presented in EMNLP GEM and NeurIPS HEGM workshops).
- Di Jin, **Sijia Liu***, Yang Liu, and Dilek Hakkani-Tur. 2022. “[Improving Bot Response Contradiction Detection via Utterance Rewriting](#)”. **SigDial, 2022**.
- Sarik Ghazarian, Behnam Hedayatnia, Di Jin, **Sijia Liu***, Nanyun Peng, Yang Liu and Dilek Hakkani-Tur. “[MERCY: Multiple Response Ranking Concurrently in Realistic Open-Domain Conversational Systems](#)”. **SigDial, 2023**.
- Prakhara Gupta, Yang Liu, Di Jin, Behnam Hedayatnia, Spandana Gella, **Sijia Liu***, Patrick Lange, Julia Hirschberg, Dilek Hakkani-Tur. “[DialGuide: Aligning Dialogue Model Behavior with Developer Guidelines](#)”. In arXiv preprint: 2212.10557.

AWARDS AND RECOGNITIONS

- Statistical Practice Scholarship 2018 (Top 10%), Carnegie Mellon University
- National Scholarship 2015, 2016 (Top 1%), Peking University
- Samsung Scholarship 2014, 2017 (Top 5%), Peking University
- Pacesetter to Merit Student 2014, 2015, 2016, 2017 (Top 1%), Peking University
- First Prize in Challenge Cup Academic Competition 2016 (Top 1%), Peking University

PEER REVIEW SERVICE

- EMNLP GEM 2022, reviewer
- EMNLP GEM 2023, program committee

PROFESSIONAL EXPERIENCES

AMAZON ALEXA AI

Sunnyvale, CA

Research Scientist at Alexa AI, with Dr. [Yang Liu](#) and Dr. [Dilek Hakkani-Tur](#)

September 2021 – present

- Building ChatGPT-like Alexa LLMs with supervised fine-tuning and human evaluation.
- Led an interactive human evaluation project for open-domain dialog systems with sample size estimation framework.
- Built GPT-2 based response generators for Alexa Prize SocialBot with unsupervised topic segmentation and filtering.
- Led multiple data collections (e.g., contradiction detection, utterance rewriting, situated storytelling).

AMAZON WEB SERVICES

Seattle, WA

Data Scientist, EC2 Core Product Management

October 2018 – March 2021

- Built tree-based ensemble models to perform EC2 multi-class workload classification.

- 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.

CMU DATA SCIENCE CAPSTONE PROJECT

Pittsburgh, PA

Computational Statistician Team co-Lead, with Prof. [Howard J. Seltman](#)

August 2017 – May 2018

- Built hierarchical Zero-Inflated Negative Binomial regression models to predict baseball player injuries.
- Developed 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

Beijing, China

Data Scientist Intern

September 2016 – February 2017

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

PFIZER

Beijing, China

Quantitative Strategy Consultant

March 2017 – June 2017

- Analyze applications of big data for pharmaceutical companies, accelerated partnerships with 2 AI-based startups.
- Analyzed survey data and drug market scenarios under different projected Quality Consistency Evaluation policies.

TEACHING EXPERIENCE

Experimental Design

Pittsburgh, PA

Graduate Teaching Assistant

February 2017 – May 2018

- Held weekly programming labs and office hours to help a class of 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.

LEADERSHIP & COMMUNITY SERVICE

- 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
- China Social Enterprise and Investment Forum, program manager.

COURSES

Undergraduate Level Courses

- Fundamentals of Computer
- Applied Mathematics
- Probability and Statistics
- Linear Algebra
- Calculus
- Social Statistics
- Population Statistics
- Socio-Economic Survey Data Analysis
- Methods of Social Survey and Social Research
- Econometrics
- Game Theory
- Social Psychology
- Thinking in Programming (Python)
- German (A2)

Graduate Level Courses

- Statistical Computing (Python)
- Statistical Learning
- Data Mining
- Data Management & Distributed Environments
- Applied Linear Models (Regression)
- Time Series
- Experimental Design

