

ZHENG LIU

🌐 <https://geog.umd.edu/gradprofile/liu/zheng> ✉ liuliu Zheng1208@gmail.com

EDUCATION

University of Maryland, College Park

Aug 2017 - May 2023

PhD in Geographical Information Science

Concentrations: statistical modeling; machine learning; spatial and temporal analysis

Peking University

Sep 2013 - July 2017

BSc in Geographical Information Science & Applied Math

SKILLS

Programming Languages: Python, SQL, Shell, JavaScript, R, C++, C#, Go, HTML, CSS, PHP

Frameworks & Tools: Git, Pandas, PostgreSQL, PyTorch, QGIS, ArcGIS, LaTeX

EXPERIENCES

EASIER Data Initiative

Aug 2022 - Present

Research Assistant and Developer

College Park, MD

- Investigated and developed data ETL pipeline for proof-of-concept in public spatial datasets access on decentralized cyberinfrastructure. See EASIER Data Initiative.
- Prototyped pipelines for remote sensing data acquisition, CAR file packing, and web3-enriched STAC API written in **Python** and **shell** script.

Global COVID-19 Trends and Impact Survey

May 2021 - June 2022

Data Engineer and Developer

College Park, MD

- **Led the data pipeline development** in the Global COVID-19 Trends and Impact Survey in partnership with Meta. Pivoted pipeline code development, coordinated with user demand and distributed tasks to the group. See project website.
- As a head developer, developed and parallelized data pipeline which is deployed on distributed cluster nodes with **Python, PostgreSQL and shell as back end, Python and Go as front end**. The daily pipeline processes more than two million survey responses every month. On-call data issues.
- Designed versatile tools and scripts to **monitor the data service**. Maintained internal and external documentation.

RaiseTech

Feb 2017 - April 2017

Software Engineer Intern

Beijing

- Built **Lua** script compiler monitor in C++ comparing difference between software versions
- Historical stock data cleaning using **Python** script

PROJECTS

Understanding and predicting bike-share trip demand

Sep 2019 - Present

PhD Dissertation research

- Built data pipeline from multi-source spatial datasets for **model training and visualization** as well as multiple **exploratory data analysis** based on **Python, SQLite**, with in-depth application of **GeoPandas, PySAL and PyTorch**.
- Proposed a novel methodology for modeling dynamic trip demands from a docked bike-sharing system with part of work published.

Time series classification of POI visits

Dec 2018

Graduate course project in Machine Learning

- Geo-tagged **microblogs collection with C#** and **big data management with MySQL**.
- Built a pipeline using Python to process places of interest (POI) data from microblogs and transform into time series data and topological images as model training input. Deployed **convolutional neural networks** in **TensorFlow** for time series classification.