

# KANGHENG LIU

1111 Arlington Blvd, Arlington, VA 22209

☎ 202-725-1513

✉ [kl1099@georgetown.edu](mailto:kl1099@georgetown.edu)

🌐 [linkedin.com/in/kangheng-liu](https://www.linkedin.com/in/kangheng-liu)

🐙 [github.com/kanghengliu](https://github.com/kanghengliu)

## Education

### Georgetown University

*Master of Science in Data Science and Analytics*

Sep. 2023 – May. 2025

Washington D.C., USA

### Lanzhou University

*Bachelor of Science in Data Science*

Sep. 2019 – Jun. 2023

Lanzhou, China

## Experience

### Massive Data Institute, Georgetown University

*Research Fellow*

Jul. 2024 – Current

Washington D.C., USA

- Implemented R and Python data extraction, computation and tabulation modules for Secure Query System - a security-focused tax data handler in collaboration with IRS SOI department and multiple clients.

### School of Data Science and Analytics, Georgetown University

*Teaching Assistant*

Sept. 2023 – Jun. 2024

Washington D.C., USA

- Instructed and mentored 50 students in Neural Network and Deep Learning curriculum and Data Science BootCamp.

### School of Information Science and Engineering, Lanzhou University

*System Architect*

Jul. 2022 – May. 2023

Lanzhou, China

- Set up and optimized data communication by 10% with Kubernetes clustering containerized applications management system with Ceph distributed storage for university net-based services, hosted various Lanzhou University web portals. Significantly reduced cost (~20%) for hosting hardware and system maintenance while improving performance for scaling up applications and data trafficking.

### Huawei Inc. & Lanzhou University

*Teaching Assistant*

Jul. 2021 – Aug. 2021

Lanzhou, China

- Deployed extensive Linux based Data Mining development environment with OpenEuler<sup>®</sup> operating system.
- Tested OpenGauss<sup>®</sup> relational database system by Huawei Technologies Co., Ltd based on ARM CPU architecture.
- Adopted Mindspore<sup>®</sup> by Huawei Technologies Co., Ltd automatic differentiation technique, code transformation techniques for deep learning.

## Projects

### Code Analysis Agent with Graph Query Language | *Langchain, GraphQL, Generative AI*

Nov 2024

- Designed and developed a pipeline for repo code ingestion and analysis with Knowledge Graph construction with neo4j and Cypher graph query language.
- Architected micro/macro framework for repo code semantic and logical level deep analysis.

### Intelligent Campus Oriented Emotion Recognition & Depression Assistance System |

*Python, SQL, NLP*

Mar. 2022 – Nov. 2022

- Gathered data (information collection, text semantic analysis and classification), cleaned, and pipelined to application back-end database.
- Wrote handler on MySQL database with SQLAlchemy extension library.
- Performed semantic analysis with Natural Language Processing.

### Design and Implementation of Telemedicine System |

*Software Development Cycle Management, Data flow design, Database Architecture design*

Mar. 2021

- Evaluated the domestic and international developmental context as a precursor to the research background analysis.
- Designed the architectural framework for implementing a telemedicine system.

### Self-hosted homelab deployment |

*Linux, bash, Docker*

Nov 2023 – Present

- Converted old hardware to a fully functional homelab with multimedia, nginx, vpn, virtualization, jupyter, compute and storage services.
- Experimented with containerization, mocked AWS cloud sagemaker service architecture locally.
- Became proficient in Debian, Arch Linux distros, and neovim with an abundance of Linux CLI tools.

## Technical Skills

**Languages:** Python, Lua, R, bash/fish, HTML/CSS, JS, TS, SQL

**Developer Tools:** Neovim, VS Code, Huawei Cloud Platform, AWS, Azure, GCP bigquery, Xcode, nginx

**Technologies/Frameworks:** Linux, Tensorflow, Pytorch, Langchain, MCP, NextJS, ReactJS, GitHub, Gitee, Quarto,

L<sup>A</sup>T<sub>E</sub>X