Gao Zhiyuan

+82 1072198956
alapha23@gmail.com

Work Experiences

Seoul National University   City Energy Lab
Research Intern   *Sep. 2023 - Now.*
LLM-based QA system for Urban Planning with RAG
- Apply Retrieval Augmented Generation by enable index searching / generation for domain-specific knowledge
- Designed a chatbot with NextJS, Typescript, Prisma, FAISS and OpenAI API

Trilogy
Software Engineer   *Aug. 2022 - Aug. 2023*
1. AWS Cost Optimization project
- Lowered AWS Service Costsfrom 25 million USD annually down to 5 million USD in a team of 3.
- Designed non-intrusive, sustainable plans to guarantee resource sharing, user resource awareness, automated anomaly detection and permission rules.

2. Social Platform for Education
A social platform that enables 4 12 grade students to memorize curricular knowledge through making posts with the help of AI and memorization techniques.
The tech stack includes Next.js, Prisma/MySQL, GPT-3, Stable-Diffusion, Docker and AWS Devops.

3. Data Platform based on Azure AD
Construct a graph-based data platform and use it to calculate company employee interaction intensity. Extract data from MS 365 through Azure Graph API, engineer the data into graph relationships and maintain the in the graph database.

4. Automate Company ERP operations on NetSuite
In charge of project ideation, development and delivery and was solely responsible for communication with the clients. Tech stack involves Selenium, Chrome Plugin, Python and Javascript

MISE
Devops Engineer   *Jun. 2022 - Sep. 2022*
- Solve system performance bottleneck by analysing AWS Lambda, API Gateway and MySQL RDS
- Apply Terraform IaaS for fast deployment
- Enabled development environment vs. production environment split to stablise our services
- Lead the team to learn to apply Github Best practices (PRs, code reviews)
- Implement the development/staging/release processes of our product

MetaMUI-SovereignWallet
Python SDK Developer   *Dec. 2021 - Mar. 2022*
Migrate Rust SDK to Python SDK for MetaMUI blockchain built with Substrate framework

Alibaba Cloud
Intern for Summer of Code   *Jun. 2021 - Sep. 2021*

Nodejs Dev/QA intern for Serverless Devs. Developed unit tests, integration tests, and APIs and Git workflow automation with Typescript, Jest and Github workflow

Chinese Academy of Science   PLCT Lab
QEMU Researcher   *May. 2020 - Nov. 2020*
One patch to QEMU upstream to emulate Nuclei RISC-V SoCs with customised interrupt controllers and registers,
Delivered an oral presentation at CRVA 2020

Google   Summer of Code
Java Developer   *Jun. 2019 - Sep. 2019*
Engineered Apache Nemo to process single-stage batch data with AWS Lambda Functions

Seoul National University   Software Platform Lab
Research Intern   *Mar. 2019 - Nov. 2019*
Enable distributed dataflow system to benefit from serverless computing using AWS Lambda Functions, Java and Apache Nemo
Two published domestic papers, and one pending research for international conferences

SUSE   Beijing, China
QA Intern   *Oct. 2018 - Jan. 2019*
Part of the Dev&QA team working on openSUSE and SUSE Linux Enterprise, with git, openQA and perl

Research Project

Accomodate Input Spikes with AWS Lambda Functions
A hyrbid system of Dockers and AWS Lambda Functions, where stable workloads are running on docker and input rate spikes are offloaded to AWS Lambda Functions,
Implemented job scheduling/balancing algorithm to coordinate Lambda Functions' life cycle, data relay, and system-level dynamic scaling

Education

National Cheng Kung University   Taiwan
B.S., Political Science   *Sep. 2015 - May. 2017*
Seoul National University   Korea
B.S., Computer Science and Engineering   *Sep. 2019 - Aug. 2024 (Expected)*

Skills: Languages: Python, Java, Typescript, C/C++
     Backend Development: NextJS, Typescript, Prisma, MySQL
     Devops: AWS, Redis, K8S, Docker, Azure AD Graph APIs, CI/CD
     Others: Serverless development, QEMU, React, Selenium, GPT, Stable Diffusion

Languages: Chinese, English (TOEFL 100), Japanese (JLPT N1) and Korean