KIMHENG PENG

Phnom Penh, Cambodia

+855 77 457 727 | kimhengpengkh@gmail.com | GitHub: totifykh | https://www.linkedin.com/in/kimheng-peng/

A Software Engineer with a passion for learning and leveraging technology to drive innovation and efficiency. I have worked on projects that improve public services through modern software systems, and I am eager to continue applying my skills in software development, problem-solving, and system design.

EDUCATION

JANUARY 2019 – MAY 2021
COMPUTER SCIENCE AS. T, SKYLINE COLLEGE, CALIFORNIA | GPA: 3.93 | SUMMA CUM LAUDE

AUGUST 2021 – MAY 2023
COMPUTER SCIENCE BS, SAN FRANCISCO STATE UNIVERSITY, CALIFORNIA | GPA: 3.93 | SUMMA CUM LAUDE

RELEVANT COURSES

Data Structure, Software Engineering, Analysis of Algorithms, Operating System, Internet Application Design and Development, Unix Programming, Programming Methodology

EXPERIENCE

SOFTWARE ENGINEER, SEPTEMBER 2023 – PRESENT

GENERAL DEPARTMENT OF DIGITAL ECONOMY, MINISTRY OF FINANCE | PHNOMPENH, CAMBODIA

- Conduct in-depth research on global e-invoicing models and case studies, focusing on PEPPOL & UBL formats, and develop an XML invoice validator service to ensure compliance
- Tested and enhanced the efficiency of the e-invoicing system by increasing processing capacity by 150% while reducing resource utilization by over 60%, from 80% to 30%
- Contribute to the design of a scalable data lake architecture, integrating tools such as Trino, Apache Superset, Apache Airflow, and lead the backend development of the data sharing/data lake project.
- **Leveraged Knowledge:** JavaScript (NodeJS, Typescript), Python (Fast API), Microservices, Docker, PostgreSQL, Unix

PROJECTS

Portfolio Website: https://www.kimheng.dev

Scalable Data Lakehouse Ecosystem

- Designed and implemented a Data Lakehouse ecosystem using Trino, Apache Iceberg, MinIO, PostgreSQL, and Apache Hive Metastore, supporting scalable and efficient data storage.
- Integrated Apache Spark to build ETL pipelines, enabling the ingestion and transformation of tabular data formats (CSV, Parquet, ORC) into Iceberg tables, with asynchronous job queues for processing large files.
- Developed APIs and workflows in Apache Airflow to automate data ingestion and pipeline orchestration
- Connected Apache Superset to Trino for data visualization and OpenMetadata for metadata management and data governance across the Lakehouse ecosystem.
- *Technology*: Python, Docker, Trino, PostgreSQL, Apache Iceberg, Superset, Airflow, Spark

MERN Stack Bakery Web Application

- Developed a personal learning project with a focus on both frontend and backend development.
- Designed the backend using the MVC pattern to ensure a well-organized and scalable architecture.
- Implemented a fully functional shopping cart and checkout system, along with an admin dashboard for the website owner to add, update, or delete products
- Utilized AWS S3 for image storage, deployed the web application on AWS EC2, and utilized AWS Route 53 to create a hosted zone and assign a domain
- *Technology*: NodeJS, ExpressJS, MongoDB, ReactJS, Material UI, AWS

Raspberry Pi Autonomous Car

- Developed and optimized line sensor and obstacle detection functionalities, ensuring precise tracking and efficient navigation for robotic car systems.
- Implemented multi-threading for parallel processing, calibrated sensor data handling and obstacle avoidance routines using dual echo sensors.
- Diagnosed, tested, and resolved hardware issues, including wiring and power delivery challenges, to ensure reliable motor hat performance
- *Technology:* Programming C, RaspberryPi, PiGPIO, Threading, Unix

SKILLS

- **Programming**: Proficient in JavaScript, Python, Unix; familiar with C, C++, Java, SQL, HTML/CSS
- **Development**: Experienced in NodeJS, ExpressJS, Fast API, ReactJS, NextJS
- **Data Tools:** Familiar with Apache Superset, Trino, Iceberg, Airflow