

VINH DUY LE

650.556.3917 | Leduyvinh91@gmail.com | Santa Clara, CA | [linkedin.com/in/vinhduyle](https://www.linkedin.com/in/vinhduyle) | github.com/VinhDuyLe

EXPERIENCE

Software Developer, Frugal Innovation Hub, Santa Clara

June 2024-present

- Designed and implemented a **Spring Boot** search service, leveraging **Hibernate Search** to integrate **Hibernate ORM** with **Elasticsearch**, achieving robust data persistence, full-text search capabilities, and query latencies under 50ms.
- Developed a scalable data ingestion pipeline for a books catalog using **Java**, leveraging **Kafka Producers** to stream data to logs topics, Consumers to process and batch-index messages into **Elasticsearch**, and enhancing data integrity through verification, filtering, and indexing accuracy.

Software Engineer Intern, OpenKnect Inc.

March 2023 – Sep 2023

- Leveraged **Resilience4j** with **Spring Boot** in the Media Setup Service to implement a circuit breaker for monitoring remote calls, along with retries and rate limiters, preventing cascading failures and boosting system **availability to 99%**.
- Developed and executed 100+ unit and integration tests across five microservices using **Java**, **JUnit**, **Mockito**, achieving 90% test coverage and ensuring system reliability.

Software Engineer, Khanh Loi JSC, Vietnam

Jan 2019 – May 2022

Khanh Loi JSC is the start-up company in Viet Nam that manages hotels, restaurants, and tourism subsidiaries

- Designed and developed 20+ **RESTful APIs** using **Spring Boot**, enhancing error management with custom exception handling, strengthening authentication with **JWT** and **Spring Security**, and providing comprehensive documentation with **Swagger** for hotel management, booking services, and discovery platforms.
- Implemented a shared database repository library across microservices using **Spring Data JPA**, **PostgreSQL**, and **Cassandra** to standardize data access with reusable operations, enforce strict data encapsulation, and ensure API-driven access to maintain modularity and privacy.
- Implemented **Apache Kafka** as an intermediary communication layer between booking services and discovery platforms, utilizing **Kafka topics** for asynchronous messaging, **Protobuf** for efficient data serialization, and **Spring Boot Kafka** integration, enabling effective decoupling between the systems and reducing data synchronization latency by 90%.

PROJECTS

Researchify, TreeHacks 2024 Hackathon at Stanford University (github.com/VinhDuyLe/researchify)

- Deployed Together AI's text summarization model on **AWS Elastic Beanstalk**, using Flask to provide APIs for summarization and media services, with **EC2 load balancing and auto-scaling**, efficiently handling 500 requests/second with sub-300ms latency.
- Built an automated pipeline for arXiv articles by integrating **Flask** and Together AI for text summarization, **Google Cloud TTS** for audio generation, and **AWS S3** for storage, leveraging **Celery** and **Redis** for asynchronous tasks, reducing audio generation time from 12 seconds to 5 seconds.

EDUCATION

Master of Science, Computer Science at Santa Clara University

August 2022-Dec 2024

- GPA: 3.8

SKILLS

Java, Python, C++, JavaScript, HTML/CSS, RESTful Services, Spring Boot, React, MySQL, PostgreSQL, Cassandra, Kafka, AWS, GCP, Git, Linux, Logstash, Docker, Kubernetes