
SUMMARY

Strong background in natural language processing (NLP), information retrieval (IR), information extraction (IE), and machine learning (ML). Focused on practical solutions that enhance productivity, accessibility, and real-world user experience.

PROFESSIONAL & RESEARCH EXPERIENCE

University of Washington

Seattle, USA

*Graduate Research Assistant, eScience Institute**Dec. 2025 - Present*

- Accelerating scientific discovery by building sustainable, high-quality open-source software and platforms for open science.

*Graduate Student**Sep. 2024 - Present*

- Built a **lay summarization** system for medical texts by fine-tuning **Llama 3 8B & 70B LLMs**. **1st Place, BioNLP @ ACL 2025**
- Created a tool to generate **emergency alerts** in 15 languages by fine-tuning and constraining **Llama 3 1B**.
- Developed a Chrome extension to improve **web accessibility** for people with visual impairments using **Llama 4 Maverick**.
- Currently collaborating with Microsoft researchers on medical open response **question answering**.

NLP Research Lab, NHN Diquet | #1 search engine and chatbot provider in Korea

Seoul, South Korea

*Senior Researcher, Foundational Technologies Team**Mar. 2022 - Jun. 2023*

- (Project Manager) Developed a content **recommendation** system based on user behavior, **serving 5M users**.
- (Project Manager) Led conversation **summarization** for the legal domain using deep learning (DL) and reinforcement learning (RL), a system **adopted by the Korea Legal Aid Corporation** to assist legal consultations.
- Developed a self-service **MLOps** platform for managing diverse ML models, abstracting complex infrastructure to enable non-expert teams to train and deploy solutions on site.
- Developed data security and NLP tools for **personal data extraction and masking**, and sentence completion, **used by the Korean National Police Agency** to ensure compliance and accelerate criminal investigations.
- Worked on **machine reading comprehension** for **question answering (QA)**, **used by millions of users**.

*Associate Researcher, Foundational Technologies Team**Feb. 2019 - Feb. 2022*

- Developed and maintained high-throughput **language analyzers** comprising part-of-speech (POS) taggers, named entity recognition (NER) model, and speech act (SA) and sentiment classifiers **used by all core company products**, including a search engine that processes **millions of queries per day** and high-traffic chatbots.
- Built a **recommendation** system for course and job opportunity matching **used by 35,000 university students**.
- Built a customizable, **hybrid NER** model using deep learning and pattern matching. **1st Place, U+ BMT NER**
- Improved **autocorrect** models using statistical post-processing, **reducing errors by 47%**.

Intelligent Systems Lab, Dong-A University

Busan, South Korea

*Researcher**Dec. 2015 - Feb. 2019*

- Conducted research on **machine reading comprehension** for English **QA** using language analysis results. **Master's Thesis**
- Built deep learning-based **language analyzers** for POS tagging, NER, dependency parsing, and semantic role labeling, creating a comprehensive and streamlined NLP pipeline. **Bronze Prize, Korean Information Processing System Competition**
- Presented research on **speech act classifier** fusing hand-crafted rules with deep learning. **Poster @ HCLT 2018**
- (Project Lead) Led creation of an **NER dataset** for extracting time entities from text. **TimeWise NER**

EDUCATION

Master of Science in Computational Linguistics (AI / NLP) (3.97/4.0)

Seattle, USA

University of Washington

Sep. 2024 - Present

Master of Science in Computer Engineering (3.91/4.0)

Busan, South Korea

Dong-A University

Sep. 2017 - Feb. 2019

Bachelor of Engineering in Computer Engineering (3.97/4.0)

Busan, South Korea

Dong-A University

Mar. 2014 - Aug. 2017

SKILLS

NLP • Information Retrieval • Information Extraction • Text Mining • Machine Learning • Deep Learning • Parameter Efficient Fine-tuning (PEFT) • Supervised Fine-tuning (SFT) • Python • Java • Go • C • C++ • Bash • TensorFlow • PyTorch • MLX • scikit-learn • Regular Expressions • Git/Gitflow • CI/CD • Docker • SQL • Databases (MariaDB, MySQL, Oracle DB)