

---

## SUMMARY

Strong background in natural language processing (NLP), language analysis, information retrieval (IR), information extraction (IE), and machine learning (ML). Skilled in data pre-processing and managing large datasets. Eager to contribute to innovative research and development by applying NLP and ML expertise to develop scalable, responsible AI systems that improve real-world user experiences.

---

## PROFESSIONAL & RESEARCH EXPERIENCE

### NLP Research Lab, NHN diquest, Inc.

Senior Researcher, Foundational Technologies Team

Associate Researcher, Foundational Technologies Team

Seoul, South Korea

Mar. 2022 - Jun. 2023

Feb. 2019 - Feb. 2022

#### ◦ LANGUAGE ANALYSIS

- Worked on high-throughput part-of-speech (POS) tagger, named entity recognition (NER) model, and speech act (SA) and sentiment classifier optimized for search engines, large-scale document processing, and chatbots.
- Built a customizable and robust hybrid NER model (pattern matching + BERT). *1st Place, U+ BMT NER*
- Worked on spelling and spacing correction models using Bi-LSTM-CRF with statistical post-processing.

#### ◦ INFORMATION RETRIEVAL & EXTRACTION

- (Project Manager, PM) Developed a system for suggesting trending and related content based on user behavior.
- (PM) Led conversation summarization in the legal domain using BART and reinforcement learning (RL).
- Built a recommendation system for course and job opportunity matching for students.
- Worked on machine reading comprehension (MRC) for Korean question answering (QA) using ELECTRA.

#### ◦ OTHERS

- Worked on an MLOps platform for training, deploying, and managing machine learning (ML) models.
- Developed tools for personal data extraction/masking and automating sentence completion to support criminal investigations.

### University of Washington / Dong-A University

#### ◦ LANGUAGE ANALYSIS

- Co-developed DL models for POS tagging, NER, dependency parsing (DP), and semantic role labeling (SRL) using Bi-LSTM-CRF. *Bronze Prize, Korean Information Processing System Competition*
- Built a SA classifier combining hand-crafted rules and a CNN-based classification model. *Proceedings of the 30th Annual Conference on Human and Cognitive Language Technology*
- (PM) Led creation of a NER dataset for extracting time entities from text. *TimeWise NER*

#### ◦ INFORMATION RETRIEVAL & EXTRACTION

- Built a lay summarization system for medical texts by fine-tuning Llama 3. *1st Place, BioNLP @ ACL 2025*
- Created a tool to generate emergency alerts in 15 languages by fine-tuning and constraining Llama 3.
- Developed a Chrome extension to improve web accessibility for people with visual impairments using Llama 4.
- Worked on MRC for English QA using language analysis results. *Master's Thesis*
- Contributed to the "WiseWordNet" project by improving entity linking performance using Wikipedia data.

---

## EDUCATION

### Master of Science in Computational Linguistics (NLP) (4.0/4.0)

University of Washington

Seattle, USA

Sep. 2024 - Present

### Master of Science in Computer Engineering (3.91/4.0)

Dong-A University

Busan, South Korea

Sep. 2017 - Feb. 2019

### Bachelor of Engineering in Computer Engineering (3.97/4.0)

Dong-A University

Busan, South Korea

Mar. 2014 - Aug. 2017

---

## SKILLS

Python • Java • C/C++ • Bash • TensorFlow • PyTorch • scikit-learn • NumPy • NLTK • spaCy • Regular Expressions • Git/GitFlow • Docker • SQL • Databases (MariaDB, MySQL, Oracle DB)