

# Yohan Lee

AI Research Engineer (NLP Specialist)

☎ (+82) 10-6323-6913 | ✉ yhlee.nlp@gmail.com | 📱 I-yohai | 🌐 I-yohai

## Professional Experience

### WRTN Technologies / AI Research Engineer (NLP Specialist)

Seoul, S.Korea

Research on LLM agent evaluation

Mar, 2024 - Present

- Design and implement a benchmark system for in-the-wild human-assistant dialogues
- Develop comprehensive evaluation framework for assessing the performance of human-assistant interactions in real-world scenarios

### Riiid / Research Scientist (NLP)

Seoul, S.Korea

Research on Large Language Models for Education

Jul, 2023 - Feb, 2024

- Competed on the Huggingface Open LLM Leaderboard, achieving 1st place on Oct, 2023
- Explored the effects of instruction tuning from data (quantity, quality, diversity) and model (scale, efficiency, objective) perspectives
- Implemented diverse optimization techniques (ZeRO, FSDP, and FlashAttention) for training/inference with single 8xA100 machine

Automated Essay Scoring

- Achieved state-of-the-art on public essay scoring benchmarks which scored 0.53 to 0.75 on QWK
- Conducted “Bar exam” scoring which performs better than GPT-4

### TUNiB / NLP Engineer

Seoul, S.Korea

Korean Open-domain Chatbot Service

Dec, 2021 - Feb, 2023

- Directed dialogue data collection and quality filtering using advanced LLMs (HyperCLOVA and ChatGPT)
- Developed an in-house 1.3B Korean LM for multi-persona chatbot with self-collected datasets (20B/4B tokens for pre-training and fine-tuning)
- Operated a Kakaotalk-based chatbot service with DAU ~ 2,000
- Utilized Triton Inference Server and AWS ECS for efficient model serving
- Enhanced model generation and API performance, increasing RPS from 4 to 13 and from 24 to 110, respectively

AI Grand Challenge: Policy Support AI

- Awarded Ministry of Science and ICT Minister's Award
- Orchestrated TableQA data collection with policy domain experts
- Developed continual learning framework with OCR-based parsing and additional table data
- Developed an integrated QA system for processing texts, tables, and charts (450M LM fine-tuning)

## Research Projects

### Story Completion Competition (2nd Place)

[Link]

- Instruction fine-tuned with diverse data augmentation, achieving 2nd place on public/private leaderboard in 3 days

Oct, 2023

### Achieved SOTA in KLUE Benchmark

[Github]

- Implemented R-BERT and Retrospective Reader models, and enhanced their structures and learning methods

Dec, 2021

### Open Domain Question Answering Competition (Gold Medal)

[Github] [Blog]

- Improved top-1 accuracy of document retrieval from 32% to 78% by hybrid retrieval techniques
- Improved EM score of question answering from 62.7 to 79.9 by effective methods

Nov, 2021

### Relation Extraction Competition (Silver Medal)

[Github] [Blog]

- Trained custom embeddings, layers, and loss function with diverse augmented dataset

Oct, 2021

### Image Classification Competition (Silver Medal)

[Github] [Blog]

- Utilized multi-task learning and test time augmentation techniques

Sep, 2021

## Development Projects

### Korean College Scholastic Ability Test

[Github]

- Compared the performance of LLMs in solving the 2024 CSAT

Nov, 2023

### Daily Papers

[Github]

- Developed a tool for auto-translating and summarizing Huggingface's daily papers into Korean using ChatGPT

Nov, 2023

### PFP Story Generation

[Link]

- Completed story for 5,000 pfp characters using GPT-3

Sep, 2022

## “Look, Attend, and Generate Poem”

- Developed a web service for generating poetry from user-uploaded photos

[\[Github\]](#)

**Dec, 2021**

## Movie Review Rating Service

- Developed a web service for auto-rating and archiving key movie reviews

[\[Github\]](#) [\[Report\]](#)

**Sep, 2020**

## Education

---

### Yonsei University

Bachelor of Arts

- German Language and Literature
- Cognitive Science

*Seoul, S.Korea*

*Mar. 2015 - Feb. 2022*

## Honors & Awards

---

Sep, 2023 **2nd place, CJ Logistics CEO's Award**, The 3rd Future Tech Challenge

*CJ Logistics*

Jan, 2023 **3rd place, Ministry of Science and ICT Minister's Award**, AI Grandchallenge: Policy Support AI

*IITP*

Nov, 2021 **1st place, grand prize**, The 3rd Annual University Student AI x Bookathon

*SKKU, Seoul*

Nov, 2021 **3rd place / 19 teams**, Open Domain Question Answering Competition

*Naver Connect*

Oct, 2021 **5th place / 19 teams**, Relation Extraction Competition

*Naver Connect*

Sep, 2021 **13th place / 38 teams**, Mask Image Classification Competition

*Naver Connect*

Aug, 2019 **Scholarship**, German Language and Literature Department

*Yonsei University*