

# Yeongbin Seo

Seoul, South Korea

Email: suhcrates@yonsei.ac.kr , suhcrates111@gmail.com

Homepage: <https://ybseo-ac.github.io/>

## Research Interest : NLP, ML

My research is focused on building human-like self-evolving AI, and I believe this requires improving the architecture of LMs. For example, to accelerate RL in LMs, I believe the LMs need goal-orientation (bidirectional attention) and faster generation speed, which motivated my research on **diffusion LMs**. I also believe that, for RL of LMs, the token-level granularity of LMs should be compressed into concept-level units in order to avoid the long-horizon (i.e., sparse reward) problem, so I'm currently working on **hierarchical LMs** (i.e., **latent reasoning**).

Also, I conducted research on **continual knowledge learning**, which is essential in self-evolving AI. I also explored **efficient learning via data pruning**, which will be required for continual learning of a vast stream of text data. I am potentially open to research on **embodied agent**, to build the **emotion and personality of LM**.

## Education

### Yonsei University

Seoul, South Korea

*M.S. in Artificial Intelligence*

- **Advisor:** Prof. Jinyoung Yeo, Prof. Jaehyung Kim

### Yonsei University

Seoul, South Korea

*B.S. in Educational Sciences*

- **Coursework:** Educational Psychology, Neuro-cognitive Psychology

## Publications

\*= co-corresponding authors    †= main author

### Fast and Fluent Diffusion Language Models via Convolutional Decoding and Rejective Fine-tuning

**NeurIPS 2025**  
**Spotlight** (top 3.19%)

*Yeongbin Seo* †, Dongha Lee\*, Jaehyung Kim\*, Jinyoung Yeo\*

[Paper] [🔗](#)

### Train-Attention: Meta-Learning Where to Focus in Continual Knowledge Learning

**NeurIPS 2024**  
main

*Yeongbin Seo* †, Dongha Lee\*, Jinyoung Yeo\*

[Paper] [🔗](#)

### Prior-based Noisy Text Data Filtering: Fast and Strong Alternative For Perplexity

Under review.  
(4.75/6.0 at NeurIPS 2025)

*Yeongbin Seo* †, Gayoung Kim, Jaehyung Kim\*, Jinyoung Yeo\*

[Paper] [🔗](#)

### Quantifying Self-Awareness of Knowledge in Large Language Models

Under review.

*Yeongbin Seo* †, Dongha Lee\*, Jinyoung Yeo\*

[Paper] [🔗](#)

## Industrial Experience

### LLM-Powered Development-Assistant Tool with LG Electronics

2024 – 2025

- Developed a developer assistance tool commissioned by LG Electronics with the following features: (1) Automated assessing of SRS (Software Requirement Specification) documents (2) Conversion of SRS documents into Test Cases (TC) (3) Conversion of Test Cases (TC) into code scripts

- Tools Used: Python, Langchain, Pytorch

#### **RAG based chatbot with Donga-ILBO (One of the legacy media)**

2023

- Designed and developed “AskDonga”, a Retrieval-Augmented Generation (RAG)-based chatbot.
- Tools Used: Python, Langchain, Uvicorn, Faiss

#### **AI News Recommendation System with Donga-Ilbo**

2022

- Designed and developed an AI-based news recommendation system for the Dong-A Ilbo website.
- Tools Used: Python, SQL, redis, Nginx, AWS

### **Skills**

---

**Languages:** Python, Java, JavaScript, C/C++

**Deep Learning Frameworks and Tools:** PyTorch, Langchain, BitsAndBytes

### **Other Job Experience**

---

#### **Software Engineer**

*Seoul, South Korea*

*Donga-Ilbo*

*Jan 2023 – Aug 2024*

- Led the planning and development of AI services based on articles for Donga-Ilbo, South Korea’s top newspaper.

#### **News Journalist (Data Analysis Specialist)**

*Seoul, South Korea*

*News1, Yonhap News, Donga-Ilbo*

*Dec 2018 – Jan 2023*

- Worked as a journalist for major South Korean media outlets, News1, Yonhap News and Donga-Ilbo.
- Specialized in statistical analysis articles while covering the Ministry of Economy and Finance, Statistics Korea, and the financial investment industry.

#### **Rock Band Musician**

*Seoul, South Korea*

*Band Seoyeongbin*

*Jan 2015 – Dec 2017*

- Leader, guitarist, and composer of the underground rock band “Band Seoyeongbin”, a three-member group active in underground clubs of Hongdae.