

Hyunbin Jin

Seoul National University
hyunbin.jin@snu.ac.kr
<https://hynbjn.github.io/>

RESEARCH INTERESTS

LLM Reasoning, AI Safety, Computational Social Science

My research interest lies in investigating the underlying dynamics of LLMs as a foundation for advancing their reasoning capabilities and ensuring safer, more reliable behavior. I am also interested in applying LLMs to interdisciplinary challenges at the intersection of AI and the social sciences.

EDUCATION

Seoul National University

M.S. in Data Science (Advisor : Taesup Kim)

Sep. 2023 ~ present

Seoul, South Korea

New York University

Exchange Student, College of Arts and Science

Sep. 2021 ~ Dec. 2021

New York, United States

Yonsei University

B.A. in Political Science and International Studies

Mar. 2018 ~ Feb. 2023

Seoul, South Korea

PUBLICATIONS (* : EQUAL CONTRIBUTION)

"Well, Keep Thinking": Enhancing LLM Reasoning with Adaptive Injection Decoding

Hyunbin Jin*, Je Won Yeom*, Seunghyun Bae*, Taesup Kim

ACL 2025 Findings

ATAS: Any-to-Any Self-Distillation for Enhanced Open-Vocabulary Dense Prediction

Juan Yeo*, Soonwoo Cha*, Jiwoo Song*, Hyunbin Jin, Taesup Kim

ICCV 2025

From Threat to Tool: Leveraging Refusal-Aware Injection Attacks for Safety Alignment

Kyubyung Chae*, Hyunbin Jin*, Taesup Kim

Under Review

What If TSF: A Multimodal Benchmark for Conditional Time Series Forecasting with Plausible Scenarios

Jinkwan Jang*, Hyunbin Jin*, Hyungjin Park, Kyubyung Chae, Taesup Kim

Under Review

SearchFireSafety: A Retrieval-Augmented Legal QA Dataset for Fire Safety

Kyubyung Chae*, Je Won Yeom*, Jeongjae Park, Seunghyun Bae, Ijun Jang, Hyunbin Jin, Jinkwan Jang, Taesup Kim

Under Review

WORK EXPERIENCE

AI Research Engineer, Naver Cloud

Oct. 2025 ~ present

- Member of the Alignment Team for the Independent AI Foundation Model project, a national initiative led by government-affiliated agencies (NIPA, NIA, IITP) in South Korea, focusing on LLM alignment.

NLP Research Engineer, Twigfarm

Oct. 2022 ~ Feb. 2023

- Participated in an image-guided machine translation (IMT) project for Korean–Japanese webtoon translation, and constructed a Korean–English parallel dataset for video-guided machine translation (VMT).

Data Analyst Intern, UpennSolution

Feb. 2021 ~ Jul. 2021

- Participated in data analysis projects to extract insights on social phenomena using real-world data and statistical methods.

RESEARCH PROJECTS

Clinical Data Extraction

Aug. 2025 ~ present

Academic Research Collaboration with Seoul National University Bundang Hospital

- Developed LLM-based pipeline to automatically extract key clinical findings from unstructured clinical notes for epilepsy patients, optimized for physician-specific documentation patterns.

Legal RAG

Mar. 2025 ~ Aug. 2025

Academic Research Collaboration with Naedam C&C

- Constructed a QA dataset on Korean fire safety law, and developed a domain-specific chatbot using retrieval-augmented generation (RAG).

Blur Face Detection

Jul. 2022 ~ Sep. 2022

Academic Research Collaboration with Alchera

- Developed and evaluated a regression model to estimate blur degree in facial images, aiming to filter motion-blurred inputs for more reliable face recognition.

PATENTS

Method and Apparatus for Adaptive Injection Decoding to Prevent Incomplete Reasoning in Large Language Models

Patent No. 10-2025-0165083, South Korea (Patent Pending)

AWARDS AND HONORS

Support for Next-Generation Researchers Program

Jul. 2024 ~ Jun. 2025

Research Subsidy for Master's Students

Ministry of Education, South Korea