

Hyunbin Jin

hyunbin.jin@snu.ac.kr | <https://hynbjn.github.io/>

RESEARCH INTERESTS

LLM Reasoning, LLM Safety, Computational Social Science

My research interests lie in LLM reasoning and safety, focusing on developing methods that enhance the capability and trustworthiness of LLMs for real-world deployment. I am also interested in leveraging LLMs to address interdisciplinary challenges at the intersection of AI and the social sciences.

EDUCATION

Seoul National University

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

Sep. 2023 ~ Feb. 2026

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 Diplomacy

Mar. 2018 ~ Feb. 2023

Seoul, South Korea

PUBLICATIONS (* : EQUAL CONTRIBUTION)

Conference

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

[Hyunbin Jin](#)*, Je Won Yeom*, Seunghyun Bae*, Taesup Kim

ACL 2025 Findings

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

Juan Yeo*, Soonwoo Cha*, Jiwoo Song*, [Hyunbin Jin](#), Taesup Kim

ICCV 2025

3. **Evaluating Structure-Aware Retrieval and Safety in Statute-Centric Legal QA**

Kyubyung Chae*, Je Won Yeom*, Jeongjae Park, Seunghyun Bae, Ijun Jang, [Hyunbin Jin](#), Jinkwan Jang, Taesup Kim

ACL 2026

Preprint

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

Kyubyung Chae*, [Hyunbin Jin](#)*, Taesup Kim

arXiv preprint

2. **What If TSF: A Benchmark for Reframing Forecasting as Scenario-Guided Multimodal Forecasting**

Jinkwan Jang*, [Hyunbin Jin](#)*, Hyungjin Park, Kyubyung Chae, Taesup Kim

arXiv preprint

Contribution

1. **HyperCLOVA X 8B Omni**

NAVER Cloud HyperCLOVA X Team - Contributor

Technical Report, 2026

WORK EXPERIENCE

Residency Program, NAVER Cloud Oct. 2025 ~ present

- Conducting research on LLM reasoning and safety as part of the Alignment Team for the Independent AI Foundation Model project, a national initiative led by government-affiliated agencies in South Korea.

NLP Research Engineer, Twigfarm Oct. 2022 ~ Feb. 2023

- Developed an image-guided machine translation (IMT) pipeline for Korean–Japanese webtoon translation and constructed a parallel dataset for video-guided machine translation (VMT).

Data Analyst Intern, UpennSolution Feb. 2021 ~ Jul. 2021

- Conducted quantitative analysis on real-world data to extract insights into social phenomena.

RESEARCH PROJECTS

Clinical Information Extraction Aug. 2025 ~ Feb. 2026

Research Collaboration with Seoul National University Bundang Hospital

- Constructed an LLM-based pipeline that extracts and structures key clinical information from epilepsy patient notes, ensuring robustness to heterogeneous physician documentation styles.

Legal RAG System Mar. 2025 ~ Aug. 2025

Research Collaboration with Naedam C&C

- Constructed a legal retrieval corpus and an evaluation QA dataset for Korean fire safety law, and developed a domain-specific RAG system validated on this benchmark.

Face Blur Estimation Jul. 2022 ~ Sep. 2022

Research Collaboration with Alchera

- Developed and evaluated a regression model estimating blur levels in facial images to filter motion-blurred inputs for 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

Fellowships/Scholarships

Support for Next-Generation Researchers Program – Ministry of Education, South Korea 2024 ~ 2025

Merit-based Scholarship – Seoul National University 2024

Jilli Scholarship (Merit-based Scholarship) – Yonsei University 2018, 2020

Awards

Honors (Top 10% in the College of Social Science) – Yonsei University 2018, 2020