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   **Sep. 2023 ~ present**
M.S. in Data Science (Advisor : Taesup Kim)   Seoul, South Korea

New York University   **Sep. 2021 ~ Dec. 2021**
Exchange Student, College of Arts and Science   New York, United States

Yonsei University   **Mar. 2018 ~ Feb. 2023**
B.A. in Political Science and International Studies   Seoul, South Korea

PUBLICATION (* : EQUAL CONTRIBUTION)

Conference

"Well, Keep Thinking": Enhancing LLM Reasoning with Adaptive Injection Decoding   ACL 2025
  Hyunbin Jin*, Je Won Yeom*, Seunghyun Bae*, Taesup Kim   Findings

ATAS: Any-to-Any Self-Distillation for Enhanced Open-Vocabulary Dense Prediction   ICCV 2025
  Juan Yeo*, Soonwoo Cha*, Jiwoo Song*, Hyunbin Jin, Taesup Kim

Preprint

From Threat to Tool: Leveraging Refusal-Aware Injection Attacks for Safety Alignment   arXiv 2025
  Kyubyung Chae*, Hyunbin Jin*, Taesup Kim

AWARDS AND HONORS

Support for Next-Generation Researchers Program   **Jul. 2024 ~ Jun. 2025**
Research Subsidy for Master's Students
Ministry of Education, South Korea

WORK EXPERIENCE

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 PROJECT

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.