Uiji Hwang

Integrated M.S.–Ph.D. Student · Embodied AI / NLP Researcher
Email: willpower@hanyang.ac.kr · GitHub: github.com/hyuiji

ABOUT ME

Integrated M.S.–Ph.D. student at Hanyang University (HYU NLP Lab) focusing on Agent-based AI Systems, Embodied AI, and Reliable NLP. Experienced in machine unlearning, interpretability, and autonomous agent frameworks.

EDUCATION

Hanyang University — M.S.–Ph.D. in Computer Science (Feb 2025 – Feb 2027 expected)

- Advisor: Taeuk Kim (HYU NLP Lab), 4.5/4.5

Chung-Ang University — B.A. German Language & Literature / B.S. AI German Processing (Feb 2018 – Feb 2025)

Gisma University of Applied Sciences (Germany) — German Language School (Mar 2021 – Feb 2022)

Gimpo Foreign Language High School — English & Japanese (Mar 2015 – Feb 2018)

SELECTED PUBLICATIONS

"Query Lens: Interpreting Sparse Key-Value Features with Indirect Effects"

- Hwiyeong Lee, Ingyu Bang, **Uiji Hwang**, Hyelim Lim, Taeuk Kim. *ICML 2026 (Under Review)*.

"Development and Evaluation of a Dual-Expertise, Utterance-Level Framework for LLM-Based Science Classroom Discourse Analysis"

- Jin Eun Yoo, Nam-Hwa Kang, Suna Ryu, Jun-ki Lee, Youngsun Kwak, Taeuk Kim, Hyeong Gwan Kim, Youngwoo Shin, **Uiji Hwang**. *LAK 2026*.

"A Survey on Machine Unlearning for Large Language Models"

- **Uiji Hwang** , Taeuk Kim. *Communications of the KIISE, Vol. 43, No. 11.*

"Does Localization Inform Unlearning? A Rigorous Examination of Local Parameter Attribution for Knowledge Unlearning in Language Models"

- Hwiyeong Lee, **Uiji Hwang**, Hyelim Lim, Taeuk Kim. *EMNLP 2025*.

"Memorization or Reasoning? Exploring the Idiom Understanding of LLMs"

- Jisu Kim*, Youngwoo Shin*, **Uiji Hwang**, Jihun Choi, Richeng Xuan, Taeuk Kim. *EMNLP 2025*.

PROJECTS

Teaching Feedback Analytics — Developed a dual-expertise LLM framework for analyzing science classroom discourse. Collaboration with KNUE education researchers (Fall 2025)

TEACHING

Teaching Assistant, Natural Language Processing — Spring 2026

Teaching Assistant, Artificial Intelligence — Fall 2025

LANGUAGES

Korean — Native

English — Fluent

German — Fluent

Japanese — Basic

RESEARCH INTERESTS

• Embodied AI & Agent-based Systems

• Machine Unlearning & Interpretability

• Multilingual Safety