

Hyeon Hwang

AI Researcher

Email: hyeon-hwang@korea.ac.kr

Github: [/hwanghyeon1](https://github.com/hwanghyeon1)

Google Scholar: [\[Link\]](#)

Education

Korea University <i>Ph.D (MS Integrated) Student, Computer Science & Engineering</i> Advisor: Prof. Jaewoo Kang	Mar. 2023 – Present Seoul
Sogang University <i>B.S in Life Science & Artificial Intelligence (Double Major)</i>	Mar. 2016 – Aug. 2022 Seoul
Goyang Foreign Language High School <i>Major: English</i>	Mar. 2012 – Feb. 2015 Goyang

Experiences

Trillion Labs Research Intern <i>Focused on LLM mid-training to enhance biological capability [T2]</i>	Apr. 2026 – Jul. 2026
SK Telecom Research Intern <i>Focused on LLM post-training to enhance scientific reasoning [T1]</i>	Oct. 2025 – Dec. 2025
NVIDIA DLI University Ambassador <i>Certified Instructor: NVIDIA DLI 'Efficient LLM Customization'</i>	Apr. 2025 – Present

Awards & Honors

Curebench at NeurIPS 2025 <i>Our team achieved three Excellence Awards among the four eval categories</i>	Excellence 2025
ArchEHR-QA 2025 at ACL - BioNLP Workshop 2025 [W4] <i>A challenge on answering patient questions using clinical evidence from EHRs</i>	1st place 2025
MedHopQA (BioCreative IX Challenge) at IJCAI 2025 [W3] <i>A challenge on complex multi-hop medical question answering</i>	1st place 2025
MEDIQA-CORR at NAACL - Clinical NLP 2024 [W2] <i>A challenge on medical error correction in clinical notes</i>	4th place 2024
BioASQ 11B [W1] <i>A challenge on large scale biomedical semantic indexing and question answering</i>	2nd place 2023
SKT FlyAI Challenger <i>A hackathon focused on the development of AI services hosted by SKTelecom</i>	1st place 2022

Conference & Journal Publications

[C6] Minju Song, **Hyeon Hwang**, Junhyun Lee, Jaewoo Kang, "Enhancing Low-Resource Language Reasoning via High-Resource Language Feature Transfer", EMNLP 2026 Findings.

[C5] Yewon Cho, Hyeongsoon Hwang, Yein Park, **Hyeon Hwang**, Jaewoo Kang, "Adaptive Latent Block Insertion via Online Reward Optimization", EMNLP 2026 Main Conference.

[C4] Hyeon Hwang, Yewon Cho, Chanwoong Yoon, Yein Park, Minju Song, Kyungjae Lee, Gangwoo Kim, Jaewoo Kang, "Assessing LLM Reasoning Steps via Principal Knowledge Grounding", In: Proceedings of EMNLP 2025. Findings[[Paper](#)]

[J1] Hyunjae Kim, Hyeon Hwang, Jiwoo Lee, Sihyeon Park, Dain Kim, Taewhoo Lee, Chanwoong Yoon, Jiwoong Sohn, Donghee Choi, Jaewoo Kang, "Small language models learn enhanced reasoning skills from medical textbooks". npj Digit. Med. 8, 240 (2025) [IF: 15.1]. [[Huggingface](#)] [[Paper](#)]

[C3] Jungwoo Park, Taewhoo Lee, Chanwoong Yoon, Hyeon Hwang, Jaewoo Kang, "Outlier-Safe Pre-Training for Robust 4-Bit Quantization of Large Language Models", In: Proceedings of ACL 2025. [[Github](#)] [[Paper](#)]

[C2] Jiwoong Sohn, Yein Park, Chanwoong Yoon, Sihyeon Park, Hyeon Hwang, Mujeen Sung, Hyunjae Kim, Jaewoo Kang, "Rationale-Guided Retrieval Augmented Generation for Medical Question Answering", In: Proceedings of NAACL 2025. [[Github](#)] [[Paper](#)]

[C1] Chanwoong Yoon, Taewhoo Lee, Hyeon Hwang, Minbyul Jeong, Jaewoo Kang, "Compact: Compressing retrieved documents actively for question answering", In: Proceedings of EMNLP 2024 [[Github](#)] [[Paper](#)]

Workshop & Shared Task Papers

[W4] Hyeon Hwang et al., "DMIS Lab at ArchEHR-QA 2025: Evidence-Grounded Answer Generation for EHR-based QA via a Multi-Agent Framework", In: Proceedings of the 24th BioNLP Workshop at ACL 2025. [[Paper](#)]

[W3] Jongmyung Jung, Hyeongsoon Hwang, Yein Park, Minju Song, Jaehoon Yoon, Hyeon Hwang, Sanghoon Lee, Jiwoong Sohn and Jaewoo Kang, "DMIS Lab at MedHopQA-2025: Ensemble Multi-Retrieval Methodologies with Reasoning Language Model Decision.", In: Proceedings of the BioCreative IX Challenge (BC9) at IJCAI 2025. [[Paper](#)]

[W2] Hyeon Hwang, Taewhoo Lee, Hyunjae Kim, Jaewoo Kang, "Exploring the Reasoning Capabilities of Small Language Models in Medical Error Correction", In: Proceedings of The 6th Clinical NLP Workshop at NAACL 2024 [[Paper](#)]

[W1] Hyunjae Kim, Hyeon Hwang, Chaeun Lee, Minju Seo, Wonjin Yoon, Jaewoo Kang, "Exploring Approaches to Answer Biomedical Questions: From Pre-processing to GPT-4", In: Proceedings of CLEF 2023 [[Paper](#)]

Technical Reports

[T2] Hyunjin Seo*, Hyeon Hwang*, Gyubok Lee, Jay Shin, Jimin Park, Taesoo Kim, Sanghoon Lee, Hongjoon Ahn, Sungjun Han, Sangwon Jung, "TheBioCollection: Unified Pre-Training Scale LLM Corpus for Biology ". arXiv preprint. 2026. [[Paper](#)]

[T1] SK Telecom "A.X K1 Technical Report ". arXiv preprint. 2026.] [[Paper](#)]

Preprints

[P2] Hyeon Hwang, Jiwoo Lee, and Jaewoo Kang, "Teaching Language Models to Think in Code". arXiv preprint. 2026. (Under Review) [[Paper](#)]

[P1] Minbyul Jeong, Hyeon Hwang, Chanwoong Yoon, Taewhoo Lee, and Jaewoo Kang, "OLAPH: Improving Factuality in Biomedical Long-form Question Answering".

arXiv preprint. 2024. (Under Review)[[Github](#)] [[Paper](#)]

Academic Activities

Invited Talk, [Multiomics Reading Group](#) at [Mila](#) Aug. 19, 2026
"TheBioCollection: Unified Pre - Training Scale LLM Corpus for Biology" [[T2](#)]

Conference Reviewer 2026
ACL 2026, EMNLP 2026, NeurIPS 2026