

Jiwoong (Andrew) Sohn

PROFILE

PhD Candidate at ETH Zurich working on large language models and agentic systems for knowledge-intensive reasoning. Work spans retrieval-augmented generation, post-training, process reward modeling, and test-time strategies.

EDUCATION

ETH Zurich Doctor of Science (PhD) in Artificial Intelligence	Switzerland 2025 – Present
• Advisors: Michael Moor (Medical AI Lab), Torsten Hoeffler (Scalable Parallel Computing Laboratory)	
Korea University Master of Science (MSc) in Computer Science	South Korea 2025
• GPA: 4.44/4.5 (Merit-based Scholarship, Early Completion)	
• Integrated Bachelor's and Master's Program	
Korea University Bachelor of Science (BSc), Bachelor of Engineering (BEng)	South Korea 2023
• GPA: 4.0/4.5	

RESEARCH EXPERIENCE

ETH Zurich Medical AI Lab ETH AI Center  Doctoral Researcher, Medical AI Lab (Advisors: Michael Moor , Torsten Hoeffler) Associated Researcher, ETH AI Center	Switzerland May 2025 – Present
Yale University Gerstein Lab  Research Assistant (Advisor: Mark Gerstein , Mentor: Xiangru Tang)	United States Aug 2024 – Apr 2025
Yale University Qingyu Chen's Lab  Research Assistant (Advisor: Qingyu Chen , Mentor: Hyunjae Kim)	United States Dec 2024 – Apr 2025
Korea University Data Mining and Information Systems Lab  Research Assistant (Advisor: Jaewoo Kang)	South Korea Mar 2023 – Apr 2025

SELECTED PUBLICATIONS

[†] indicates equal contribution; [‡] indicates co-corresponding authors.

[9] Process Reward Agents for Steering Knowledge-Intensive Reasoning

Jiwoong Sohn[†], Tomasz Sternal[†], Kenneth Styppa[†], Torsten Hoeffler[‡], Michael Moor[‡]
ICML 2026, Main Track

[8] Med-PRM: Medical Reasoning Models with Stepwise, Guideline-verified Process Rewards

Jaehoon Yun[†], **Jiwoong Sohn**[†], Jungwoo Park[†], Hyunjae Kim, Xiangru Tang, Daniel Shao, Yonghoe Koo, Minhyeok Ko, Qingyu Chen, Mark Gerstein, Michael Moor[‡], Jaewoo Kang[‡]
EMNLP 2025, Main Track (Oral Presentation)

[7] Rationale-Guided Retrieval Augmented Generation for Medical Question Answering

Jiwoong Sohn, Yein Park, Chanwoong Yoon, Sihyeon Park, Hyeon Hwang, Mujeen Sung, Hyunjae Kim[‡], Jaewoo Kang[‡]
NAACL 2025, Main Track (Oral Presentation)

[6] MedAgentsBench: Benchmarking Thinking Models and Agent Frameworks for Complex Medical Reasoning

Xiangru Tang[†], Daniel Shao, **Jiwoong Sohn**[†], Jiapeng Chen, Jiayi Zhang, Jinyu Xiang, Fang Wu, Yilun Zhao, Chenglin Wu, Wenqi Shi, Arman Cohan, Mark Gerstein
Patterns

[5] Improving Medical Reasoning through Retrieval and Self-Reflection with Retrieval-Augmented Large Language Models

Minbyul Jeong, **Jiwoong Sohn**, Mujeen Sung[‡], Jaewoo Kang[‡]
ISMB 2024, Bioinformatics 2024 (Oral Presentation)

[4] Rethinking Retrieval-Augmented Generation for Medicine: A Large-Scale, Systematic Expert Evaluation and Practical Insights

Hyunjae Kim, **Jiwoong Sohn**, Aidan Gilson, Nicholas Cochran-Caggiano, Serina Applebaum, Heeju Jin, Seihee Park, Yujin Park, Jiyeong Park, Seoyoung Choi, Brittany Alexandra Herrera Contreras, Thomas Huang, Jaehoon Yun, Ethan F. Wei, Roy Jiang, Leah Colucci, Eric Lai, Amisha Dave, Tuo Guo, Maxwell B. Singer, Yonghoe Koo, Ron A. Adelman, James Zou, Andrew Taylor, Arman Cohan, Hua Xu, Qingyu Chen
arXiv:2511.06738 (2025)

[3] Small Language Models Learn Enhanced Reasoning Skills from Medical Textbooks

Hyunjae Kim, Hyeon Hwang, Jiwoo Lee, Sihyeon Park, Dain Kim, Taewhoo Lee, Chanwoong Yoon, **Jiwoong Sohn**, Jungwoo Park, Olga Reykhart, Thomas Fetherston, Donghee Choi, Soo Heon Kwak, Qingyu Chen, Jaewoo Kang
NPJ Digital Medicine

[2] RadAgent: A tool-using AI agent for stepwise interpretation of chest computed tomography

Mélanie Roschewitz, Kenneth Styppa, Yitian Tao, **Jiwoong Sohn**, Jean-Benoit Delbrouck, Benjamin Gundersen, Nicolas Deperrois, Christian Bluethgen, Julia Vogt, Bjoern Menze, Farhad Nooralahzadeh, Michael Krauthammer, Michael Moor
arXiv:2604.15231 (2026)

[1] Med-CoReasoner: Reducing Language Disparities in Medical Reasoning via Language-Informed Co-Reasoning

Fan Gao, Sherry T. Tong, **Jiwoong Sohn**, Jiahao Huang, Junfeng Jiang, Ding Xia, Piyalitt Ittichaiwong, Kanyakorn Veerakanjana, Hyunjae Kim, Qingyu Chen, Edison Marrese Taylor, Kazuma Kobayashi, Akiko Aizawa, Irene Li
ACL 2026, Main Track

INVITED TALKS

6th SFCNS Congress 2025 | Swiss Federation of Clinical Neuro-Societies

Oct. 2025

Invited Talk

Krauthammer Lab | University of Zurich

Jul. 2025

Invited Talk

HONORS AND AWARDS

— *Competition Awards*

[IJCAI 2025 BioCreative IX] Track 1: MedHopQA

Aug 2025

First Prize

[ACL 2024 BioNLP] Shared Task on Clinical Text Generation

May 2024

Top-performing Teams

[LG Electronics] iPBL Competition Second Prize (USD 5,000)	Feb 2024
[AMIA 2023 Biocreative VIII Track 3] Genetic Phenotype Extraction and Normalization from Dysmorphology Physical Examination Entries Third Prize	Sep 2023
SDGs Innovation Arduino-based Technology Competition Second Prize	Dec 2020
— Scholarships	
Integrated Bachelor's and Master's Scholarship (50% Tuition)	Fall 2024
Integrated Bachelor's and Master's Scholarship (50% Tuition)	Spring 2024
Integrated Bachelor's and Master's Scholarship (50% Tuition)	Fall 2023
SW Talent Training Scholarship	Spring 2024
Teaching Assistant Scholarship (ICT Talent Training Program)	Fall 2023
Departmental Special Scholarship	Spring 2023
Departmental Special Scholarship	Fall 2022
Departmental Special Scholarship	Spring 2022
Departmental Special Scholarship	Fall 2021
— Academic Distinctions	
[ICML 2026] Gold Reviewer	May 2026
Early Completion in Master's Program (3 terms)	Sep 2023 – Feb 2025
Semester High Honors	Fall 2020
Semester High Honors	Fall 2019

TEACHING EXPERIENCE

Multimodal Medical AI (262-0201-00L) <i>Teaching Assistant</i>	Fall 2025
SW Programming Basics (GECT002) <i>Teaching Assistant</i>	Spring 2024
Machine Learning (COSE362) <i>Teaching Assistant</i>	Fall 2023
Advanced Data Science for Finance (DFE620) <i>Teaching Assistant</i>	Fall 2023
Python Programming for Everybody (COSE156) <i>Teaching Assistant</i>	Fall 2020
Private Academies and Private Tutoring <i>Math, Physics, Chemistry</i>	2016 - 2023

TECHNICAL SKILLS

Programming Languages: Python, C, MATLAB
Machine Learning: Deep learning with PyTorch and the Hugging Face ecosystem (Transformers, Datasets, Accelerate, PEFT, TRL); distributed and parameter-efficient training (DeepSpeed, LoRA/QLoRA)
Large Language Models: Post-training and instruction tuning; process reward modeling; test-time scaling methods including beam search and rejection sampling; agentic systems and tool use
Retrieval-Augmented Generation: Dense, sparse, and hybrid retrieval (FAISS, Milvus, BM25); cross-encoder reranking
Model Serving & Inference: vLLM, FastAPI, quantization, tensor parallelism, stepwise batching
Research Infrastructure: Linux, Git, Docker, SLURM, AWS, GCP, Weights & Biases
Data Analysis: NumPy, Pandas, scikit-learn, Matplotlib, Seaborn

LEADERSHIP AND EXTRACURRICULAR ACTIVITIES

AIVLE School AI Developer Track KT	Jul 2022 – Jan 2023
Buddy Assistants for Exchange Students Korea University <i>Group Leader</i>	Aug 2022 – Dec 2023
Korea University Rowing Team Korea University	Jul 2020 – Jan 2021
Disease Study Society Korea University <i>Treasurer, Academic Member</i>	Mar 2019 – Dec 2019
Compulsory Military Service Republic of Korea Air Force <i>Staff Sergeant (honorable discharge)</i>	Apr 2017 – Apr 2019

ACADEMIC SERVICE

- Conference Reviewer**
- NeurIPS (2026, Reviewer & Emergency Reviewer)
 - ICML (2026, Gold Reviewer)
 - AAAI (2027)
 - ACL (2025, SRW)
 - EMNLP (2024)
- Journal Reviewer**
- PLOS Digital Health (Q1, IF 7.7)
 - PLOS ONE (IF 2.6)

LANGUAGES

- Korean** (Native)
- English** (Native-level)
- TOEFL: 115/120 (MyBest®)
 - TOEIC: 985/990
 - TOEIC Speaking: 200/200
- German** (Beginner)