

Dongsuk Jang

✉ jamesjang26@snu.ac.kr 📞 +82-10-5105-3184 🔗 [LinkedIn](#) 📄 [Google Scholar](#) 🌐 [Webpage](#)

EDUCATION

Seoul National University, Seoul, Korea | [MedInfoLab](#) Mar 2023 – Feb 2026
M.S. in Interdisciplinary Program in Bioengineering (Advisor: [Jinwook Choi](#))

Yale University, New Haven, CT, USA | [YaleNLP](#) Oct 2024 – Sep 2025
Visiting Student, Department of Computer Science (Advisor: [Arman Cohan](#))

Sungkyunkwan University, Suwon, Korea Mar 2015 – Feb 2023
B.S. in Integrative Biotechnology (Micro Degree in Core-Bio A-School Track)

PUBLICATIONS

- [1] **Dongsuk Jang**, Kyle Tegtmeier, Muhammad Qureshi, Anurag Gupta, Ziyao Shangguan, Sophie Chheang, Arman Cohan. Retrieval-Augmented Structured Reporting for Radiology Workflow Integration. SIIM-CAIMI 2025.
- [2] **Dongsuk Jang**, Ziyao Shangguan, Kyle Tegtmeier, Anurag Gupta, Jan Czereminski, Sophie Chheang, Arman Cohan. MedTutor: A Retrieval-Augmented LLM System for Case-Based Medical Education. EMNLP 2025 System Demo. [\[PDF\]](#)
- [3] **Dongsuk Jang**, Alan Li, Arman Cohan. Multi-Perspective Integration via Mixture-of-Agents for Enhanced Healthcare QA Summarization. CL4Health Workshop @ NAACL 2025. [\[PDF\]](#)
- [4] Hyeryun Park, Jeongwon Kwak, **Dongsuk Jang**, Sumin Park, Jinwook Choi. LoraMap: Harnessing the Power of LoRA Connections. arXiv preprint 2024. [\[PDF\]](#)
- [5] **Dongsuk Jang**[†], Hyeryun Park[†], Jiye Son, Hyeonuk Hwang, Su-jin Kim, Jinwook Choi. Automated Information Extraction from Thyroid Operation Narrative: A Comparative Study of GPT-4 and Fine-Tuned KoELECTRA. AMIA Informatics Summit 2024. [\[PDF\]](#)

PROJECTS

Evidence-Aware Reranker for Reasoning Intensive Medical Tasks Apr 2026 – Present
Designed and developed a specialized, evidence-aware generative reranker to meet the critical needs of medical and scientific RAG systems. The model aims to improve the relevance and accuracy of retrieved documents, significantly enhancing the overall performance and reliability of RAG pipelines.

Clinical NL2SQL Agent Evaluation for Large-Scale Hospital Databases Apr 2026 – Present
Analyzing the feasibility of NL2SQL agents for retrieving clinically relevant information from hospital CDW and EMR systems, with planned expert evaluation in real-world Oracle database environments containing hundreds to thousands of tables.

Multi-Agent System for Automated Clinical Data Preprocessing Mar 2026 – Present
Developing a multi-agent system that automates preprocessing from raw clinical data, including imputation, and evaluating how closely the resulting datasets support downstream ML tasks using an expert-built benchmark on MIMIC-IV, with planned validation on internal Chung-Ang University Hospital data.

VLN-Hybrid: Advancing Visual Imagination for VLN Agents with RAG System [\[Link\]](#) Jan 2026 – Feb 2026
Proposed VLN-Hybrid to mitigate visual hallucination errors in imagination-based agents by integrating Retrieval-Augmented Generation. Engineered a Dual-Key RAG system to resolve feature mismatch between CLIP retrieval and ViT navigation encoders, and implemented an adaptive soft fusion strategy. Achieved a +0.75% SPL improvement on the R2R benchmark.

Table-to-Text Generation for Clinical Decision Support [\[Link\]](#) Dec 2023 – Jul 2024
Led a full-cycle project to mitigate clinician cognitive load by generating natural language summaries from complex prescription tables. Directed the creation of a ground-truth dataset using Seoul National University Hospital data and systematically evaluated traditional Seq2Seq architectures against modern LLMs fine-tuned with PEFT.

Automatic Generation of Thyroid Operation Records on Web [\[Video\]](#) Mar 2023 – Feb 2024
Led the model development lifecycle to automate the creation of structured thyroid operation records from voice-recorded surgical notes. Trained and rigorously evaluated a fine-tuned KoELECTRA model to ensure high-accuracy clinical information extraction.

Evaluating Transformers for Clinical Natural Language Inference [\[GitHub\]](#) Dec 2022 – Feb 2023
Conducted a comparative performance analysis of various models, including m-LSTM and multiple BERT-based architectures, on the MedNLI task with the MIMIC-III dataset. Identified and implemented optimal parameter settings to achieve the best performance.

EXPERIENCE

Student Intern, Samsung Advanced Institute for Health Sciences and Technology	Mar 2021 – Dec 2022
Acquired foundational skills in Python and Natural Language Processing, applying them to conduct preliminary research on analyzing biomedical text data under expert mentorship.	
Military Service, Korean Augmentation to the U.S. Army (KATUSA)	Jan 2018 – Sep 2019
Managed human resources and administrative operations for a multinational military unit, developing strong organizational and cross-cultural communication skills.	

HONORS AND AWARDS

BK21 Research Scholarship, Seoul National University	2025
A competitive, government-funded scholarship for outstanding graduate students.	
2nd Place in PerAnsSumm Shared Task at CL4Health @ NAACL [Link]	2025
Biomedical Global Talent Nurturing Program [Link]	2024
Awarded a \$30,000 grant by the Korea Health Industry Development Institute for visiting research opportunity.	
2nd Place, Seocho YES 5th Startup Pitching Competition [Link]	2024
Selected for the competitive startup incubation program and secured grant funding for project development.	
Selected Team, 6th SNU Haedong Junior Startup Program	2024
Next Generation BioHealth Leader Award, Sungkyunkwan University	2023
Best Research Intern Award, Samsung Advanced Institute for Health Sciences and Technology	2021
Commandant's List, KATUSA Training Academy	2018

PROFESSIONAL SERVICES

Reviewer for the Oct ARR 2025	2025
Lead Graduate Student, MEDINFO Lab, Seoul National University	Sep 2025 – Feb 2026
Local organizer for NENLP 2025 [Link]	2025
Reviewer for the AMIA 2025 Annual Symposium, Informatics Summit	2024

SKILLS SUMMARY

Languages

Python, SQL, R, Matlab

NLP Libraries

PyTorch, Transformers, vLLM, Accelerate, bitsandbytes, LangChain

Data Science & Visualization

Scikit-learn, Pandas, NumPy, SciPy, Matplotlib, Seaborn, OpenCV

Developer Tools & Platforms

Git, Docker, SLURM, Linux (Ubuntu), WandB, TensorBoard, JupyterLab

REFERENCES

Arman Cohan

Assistant Professor, Department of Computer Science, Yale University
arman.cohan@yale.edu

Jinwook Choi

Professor, Interdisciplinary Program in Bioengineering, Seoul National University
jinchoi@snu.ac.kr