Joohyung **Yun**

PH.D. STUDENT IN COMPUTER SCIENCE

Data Systems Lab, POSTECH, Pohang, South Korea

jhyun@dblab.postech.ac.kr | joohyung00.github.io | joohyung00 | joohyung-yun-a1a881318

Summary

Hi! I'm Joohyung Yun, a 4th year integrated M.S./Ph.D. student (2nd year Ph.D.) in Computer Science at POSTECH, advised by Professor Wook-Shin Han in the Data Systems Lab. My research sits at the intersection of Natural Language Processing and Information Retrieval, with a focus on Retrieval-Augmented Generation (RAG) for complex, real-world documents.

I'm particularly interested in multimodal document RAG: collections that mix text, tables, and images. My recent projects explore multihop reasoning across these modalities and how to combine retrieval with LLM reasoning effectively. Currently, I'm extending this work toward Agentic RAG—building autonomous, data-centric systems that go beyond simple retrieval into complex reasoning with history-aware backtracking and adaptive traversal strategies.

Looking ahead, I aim to scale these ideas to Enterprise-level RAG: high-velocity, high-trust systems powered by Knowledge Graphs and Multi-Modal RAG that deliver verifiable, multi-modal insights. My long-term vision is to contribute to cost-efficient, data-centric platforms designed for AGI-level reasoning and agentic autonomy across enterprise ecosystems.

Education

Pohang University of Science and Technology (POSTECH) — *Pohang, South Korea*

INTEGRATED M.S./PH.D. IN COMPUTER SCIENCE — *Mar. 2023 - Present*

- Member of Data Systems Lab
- Advised by Professor Wook-Shin Han

Pohang University of Science and Technology (POSTECH) — *Pohang, South Korea*

B.S. IN COMPUTER SCIENCE — *Mar. 2019 - Feb. 2023*

- Summa Cum Laude, GPA 4.11 / 4.30

Hyundai Chungun High School — *Ulsan, South Korea*

HIGH SCHOOL DIPLOMA — *Mar. 2016 - Feb. 2019*

Publications

Failure is Feedback: History-Aware Backtracking for Agentic Traversal in Multimodal Graphs — *Seoul, South Korea*

ICML 2026 — *Jul. 2026*

- **Joohyung Yun**, Doyup Lee, and Wook-Shin Han
- In Proc., the 43rd International Conference on Machine Learning (ICML 2026)

LILaC: Late Interacting in Layered Component Graph for Open-domain Multimodal Multihop Retrieval — *Suzhou, China*

EMNLP MAIN 2025 — *Nov. 2025*

- **Joohyung Yun**, Doyup Lee, and Wook-Shin Han
- In Proc., the 2025 Conference on Empirical Methods in Natural Language Processing

HELIOS: Harmonizing Early Fusion, Late Fusion, and LLM Reasoning for Multi-Granular Table-Text Retrieval — *Vienna, Austria*

ACL MAIN 2025 — *Jul. 2025*

- Sungho Park, **Joohyung Yun**, Jongwuk Lee, and Wook-Shin Han
- In Proc., 63rd Annual Meeting of the Association for Computational Linguistics

ReCG: Bottom-Up JSON Schema Discovery Using a Repetitive Cluster-and-Generalize Framework — *Guangzhou, China*

VLDB 2024 — *Aug. 2024*

- **Joohyung Yun**, Byungchul Tak, and Wook-Shin Han
- In 50th Int'l Conf. on Very Large Data Bases (VLDB) / Proc. the VLDB Endowment (PVLDB)

KDD Cup Meta CRAG 2024 Technical Report: Three-step Question-Answering Framework — *Barcelona, Spain*

KDD Cup 2024 Workshop — *Aug. 2024*

- Sungho Park, Jeongeum Seok, Jooyoung Lee, **Joohyung Yun**, and Wonseok Lee
- 2024 KDD Cup Workshop for Retrieval Augmented Generation

Work Experience

New York University (NYU) — *New York, USA*

Visiting Researcher — *Apr. 2026 - Aug. 2026*

- Joint research collaboration
- Funded by Global AI Frontier Lab of IITP

KJMD 2025 (7th Joint Korea-Japan Workshop on Management of Data) — *Busan, South Korea*

Workshop Chair — *Oct. 2025*

- Managed the overall program
- Coordinated management and supervision of submissions
- `https://kjmd-po.github.io/kjmd2025/`

VLDB 2024: PhD Workshop — *Guangzhou, China*

Session Chair — *Aug. 2024*

- Moderated technical session: introduced speakers, managed timing, facilitated Q&A
- `https://vldb.org/2024/?program-schedule-phd-workshop`

SK Hynix Inc. — *Icheon, South Korea*

Full Stack Developer (Intern) — *Jun. 2021 - Jul. 2021*

- Web interface manipulation via Selenium, Python
- Server development using Springboot, Java
- Frontend development using NodeJS, JavaScript

Honors & Awards

2026	**2026 Overseas Dispatch Program for Graduate Students (Long-term)**, Funded for long-term international research visit program	*POSTECH*
2025	**2025 POSTECH Research Excellence Award**, Department of Computer Science and Engineering	*POSTECH*
2025	**2025 POSTECH CSE Outstanding Paper Award**, Department of Computer Science and Engineering	*POSTECH*
2025	**2025 BK POSTECHIAN Fellowship**, BK21 Fellowship for outstanding graduate students	*BK21*
2024	**1st Place (Comparison & Post-processing Tests)**, KDD Cup 2024: Meta CRAG Challenge (Team dRAGonRAnGers)	
2021	**National Science and Technology Scholarship**, Awarded to students with outstanding academic ability in science and engineering	*South Korea*
2021	**Global Leadership Program Scholarship**, Awarded to outstanding students majoring in computer science at POSTECH	*POSTECH*
2020	**Global Leadership Program Scholarship**, Awarded to outstanding students majoring in computer science at POSTECH	*POSTECH*

Patents

Subgraph Retrieval System Based on Layered Component Graph for Open-Domain Multimodal Document Retrieval — *Application No. 10-2025-0195617*

Korean Patent Application — *Dec. 2025*

- Wook-Shin Han and **Joohyung Yun**

Apparatus and Method for Table and Document Retrieval Based on Questions *Application No. PCT/KR2025/003037*

PCT International Application *Mar. 2025*

- Wook-Shin Han, Sungho Park, **Joohyung Yun**, and Jeonghoon Lee

AI Technology for Predicting On-site Safety Accidents in Hazardous Chemical Facilities *Application No. 10-2024-0167340*

Korean Patent Application *Nov. 2024*

- Wook-Shin Han, Jaehyun Ha, and **Joohyung Yun**

Teaching Experience

Design of Digital System (CSED273) *POSTECH*

Teaching Assistant *2024*

- Worked as an assistant for the Design of Digital System course

Advanced Machine Learning *Samsung Electronics Co.*

Teaching Assistant *2022*

- Worked as an assistant in Advanced Machine Learning class held under Data Systems Lab
- Guided Samsung Electronics employees to understand the implementation of machine learning concepts

Data Processing, Distributed Database & Deep Learning *Samsung Electronics Co.*

Teaching Assistant *2021*

- Worked as an assistant in Data Processing, Distributed Database & Deep Learning class held under Data Systems Lab
- Guided Samsung Electronics employees to understand the implementation of above concepts

Skills

Programming Python, C++, LaTeX

Languages Korean (native), English (fluent), Japanese (fluent)