Seokwon Song

Seoul, Korea
seokwon.song@vision.snu.ac.kr
(+82) 010-8078-1178
https://seokwon99.github.io/

About Me

I am a Ph.D. student in the Interdisciplinary Program in Artificial Intelligence at Seoul National University, advised by Prof. Gunhee Kim in the Vision and Learning Lab. I am also currently a Research Intern at LG AI Research.

My research goal is to build multimodal intelligence systems that search, read, and reason over real-world knowledge in collaboration with humans. I focus on how such systems acquire information from web sources and visually rich documents, and how multimodal large language models (MLLMs) ground their answers in that evidence. My recent research interests include:

- **Multimodal Retrieval**: searching web and document sources for evidence that covers the diverse perspectives of open-ended queries
- **Visual Document Understanding & RAG**: reading visually rich documents and grounding MLLMs in multimodal evidence for long-form, cited generation
- **Document Agent** (ongoing): generalist agents that plan and orchestrate multiple tools over documents

Education

Seoul National University — Seoul, Korea
Combined M.S. / Ph.D. of Interdisciplinary Program in Artificial Intelligence — *Mar 2024 – Present*

Korea University — Seoul, Korea
B.S. in Computer Science and Engineering — *Mar 2018 – Feb 2024*

Research Experiences

Language Lab @ LG AI Research — Feb *2026 – Present*
Research Intern

Vision & Learning Lab @ Seoul National University — Mar *2024 – Present*
Ph.D. Student (Sep 2025 – Present), M.S. Student (Mar 2024 – Sep 2025) (Advisor: Gunhee Kim)

Language Lab @ LG AI Research — Jan *2023 – Sep 2023*
Software Engineer Intern

Machine Learning & Vision Lab @ Korea University — *May 2022 – Jan 2023*
Undergraduate Research Intern (Advisor: Hyunwoo J. Kim)

Publications (*: equal contribution)

Multi3IR: A Benchmark for Multi-perspective Multi-domain Multi-modal Information Retrieval — 2026 EMNLP
Seokwon Song, *Sohyeon Kim, Gunhee Kim* — *Main (Acceptance = 2,719/17,669 ~ 15.4%)*

MAVIS: A Benchmark for Multimodal Source Attribution in Long-form Visual Question Answering — 2026 AAAI
Seokwon Song, *Minsu Park, Gunhee Kim* — *Poster (Acceptance = 4,167/23,680 ~ 17.6%)*

Is a Peeled Apple Still Red? Evaluating LLMs' Ability for Conceptual Combination with Property Type *2025 NAACL*
Seokwon Song*, *Taehyun Lee**, Jaewoo Ahn, Jae Hyuk Sung, Gunhee Kim* *Oral (245/3099 ~ 7.9%)*

Read-only Prompt Optimization for Vision-Language Few-shot Learning *2023 ICCV*
*Dongjun Lee**, ***Seokwon Song****, Jehee Suh, Junmyeong Choi, Sanghyeok Lee, Hyunwoo J. Kim* *Poster (2,160/8,260 ~ 26.2%)*

Preprints & Work in Progress

Every Tool Has Its Blind Spots: Building a Generalist Document Agent through Multi-Tool Orchestration In preparation
Seokwon Song*, Min Hyung Park, Young Rok Jang, Gyeonghun Kim, Gunhee Kim, Jinsik Lee*

Academic Services

Conference Reviewer

- *ACL Rolling Review (2026.08)*
- *NeurIPS (2026)*
- *CVPR (2026)*
- *AAAI (2026)*

Workshop Reviewer

- *ACL-SRW (2025)*

Teaching Assistant

- *Seoul National University - Probabilistic Graphical Models (2024)*

Technical Strengths

Coding Python, C++

Others