

Wooseok Seo

✉ justin_seo@yonsei.ac.kr ⚡ Homepage ⚡ LinkedIn ⚡ Github

📍 Seoul, Korea

Research Interests

I am interested in pushing the boundaries of *foundational models* through better **evaluation frameworks** or effective **post-training**. I am currently excited about accurately evaluating & mitigating the hallucinations of LLMs, or improving models to generate faithful responses from expert-level knowledge sources (e.g., Search Agents, RAG, AI4Science, etc.) for practical use. I am also interested in defining, generating, and selecting *good* synthetic data for model generalization.

Education

Yonsei University Ph.D. in Artificial Intelligence Advisor: Prof. Youngjae Yu	Seoul, Korea Mar 2025 – Feb 2030 (Exp.)
Yonsei University B.S. in Nano Science and Engineering	Seoul, Korea Mar 2020 – Feb 2025

Publication

[C]: Conference, [P]: Preprint, * denotes equal contribution.

[P2] K-EXAONE Technical Report

LG AI Research

TLDR; We present K-EXAONE-236B-A23B, the best model in Korea. I contribute as a member of the post-training team, specifically working on synthetic data for reasoning.

[Technical Report](#), [Paper](#), [HuggingFace](#)

[C2] Verifying the Verifiers: Unveiling Pitfalls and Potentials in Fact Verifiers

Wooseok Seo*, Seungju Han*, Jaehun Jung, Benjamin Newman, Seungwon Lim, Seungbeen Lee, Ximing Lu, Yejin Choi, Youngjae Yu

TLDR; We systematically detect ambiguous & mislabeled examples in fact-verification benchmarks and introduce Clearfacts and Grayfacts, along with an SOTA 8B fact verifier and insights on building better fact verifiers.

[COLM 2025](#), [Paper](#), [Code](#)

[C1] V.I.P.: Iterative Online Preference Distillation for Efficient Video Diffusion Models

Jisoo Kim, Wooseok Seo, Junwan Kim, Seungho Park, Sooyeon Park, Youngjae Yu

TLDR; We integrate DPO and SFT loss for distillation to build an efficient video diffusion model, with an automatic pair curation pipeline and outperform the teacher only with the synthetic data generated from the teacher itself.

[ICCV 2025](#), [Paper](#)

[P1] Layout-and-Retouch: A Dual-stage Framework for Improving Diversity in Personalized Image Generation

Wooseok Seo*, Kangyeol Kim*, Sehyun Nam, Bodam Kim, Suhyeon Jeong, Wonwoo Cho, Jaegul Choo, Youngjae Yu

TLDR; We use a two-stage approach for personalized T2I generation, to first draw the context with step-blended denoising and enhance the context with multi-source attention swapping.

[Preprint](#), [Paper](#)

Experience

LG AI Research <i>Research Intern @ EXAONE Lab</i>	Seoul, Korea Sep 2025 - Jan 2026 (Exp.)
Seoul National University, PI LAB <i>Visiting Researcher</i>	Seoul, Korea Sep 2025 - Current
MAUM.AI <i>Research Scientist @ WoRV (World model for Robotics and Vehicle control) Team</i>	Seoul, Korea Jan 2024 - May 2024

Yonsei University, MIR Lab
Undergraduate Research Intern
Advisor: Prof. [Youngjae Yu](#)

Seoul, Korea
Sep 2023 - Feb 2025

BagelCode
Data Analyst Intern

Seoul, Korea
Apr 2022 - Oct 2022

Academic Services

Reviewer

ACL ARR
COLM

2025
2025

Awards and Honors

Thinking Machines Lab

\$5k Research Grant for research on synthetic data 2025