

Sejoon Oh, Ph.D.

Senior AI Research Scientist @ Netflix | EB-1B Green Card Holder
Santa Clara, CA | 404-889-1929 | sejun6431@gmail.com
sejoonoh.github.io | linkedin.com/in/sejoon-oh

RESEARCH INTERESTS

Primary Focus: Agents, Post Training, Personalization, Multimodality, Safety and Interpretability.

Applied Domains: Semantic ID, Ads & Commerce, Search & Retrieval, Physical AI, World Models.

RESEARCH SUMMARY

AI Researcher specializing in Agents, Personalization, Multimodaloty, and Safety, Robustness, and Interpretability. Proven track record of bridging the gap between research and production at Netflix, developing foundation models that drove **\$10M+ in Y-Y revenue lift** and optimized inference costs by **30%**. Expert in personalizing LLMs via Post-training methods (SFT, DPO, GRPO, etc.), designing universal safety guardrails for Multimodal LLMs, and integrating agents and semantic IDs into LLMs for hyper-personalized content discovery.

EXPERIENCE

Netflix

Senior AI Foundation Research Scientist

Los Gatos, CA

June 2024 – Present

LLM Post Training: Main “chef” of Netflix LLM post-training using multi-task Supervised Fine-Tuning (SFT) and Reinforcement Learning (DPO, GRPO, etc.) to enhance personalization and content understanding.

Semantic ID Integration: Lead researcher of semantic ID integration into LLMs, enabling the unified recommendation of diverse content types (videos, games, live events, podcasts).

User Intent Modeling (FM-Intent): Invented a hierarchical multi-task learning FM to predict user session intent. Successfully productized the model, achieving an impressive revenue lift across millions of users.

Efficient Inference (FM-light): Developed a lightweight foundation model framework. Achieved **30% faster runtime** and **50% memory savings** while increasing recommendation freshness by 50% through dynamic response to users.

Discovery Optimization: Created a specialized fine-tuning mechanism for the Netflix foundation model focused on content discovery, resulting in a significant revenue lift in global A/B testing across millions of users.

Georgia Institute of Technology

Graduate Research Assistant

Atlanta, GA

Aug 2019 – May 2024

Multimodal LLMs and their Safety: Developed “UniGuard,” a universal safety guardrail to defend Multimodal LLMs against jailbreak attacks (AAAI 2025). Demonstrated vulnerabilities in text-aware recommender systems via adversarial text rewriting (CIKM 2024).

Recommender Systems: Authored multiple papers on stabilizing recommendations (TKDD 2024) and rank list sensitivity (CIKM 2022). Investigated cross-modal projection in Multimodal LLMs (ACL 2024).

Netflix

Machine Learning Research Intern

Los Gatos, CA

May 2023 – Aug 2023

Project IntentRec: Developed “IntentRec” to predict user session intent using deep learning.

The Home Depot

Data Science Research Intern

Atlanta, GA

May 2021 – Aug 2021

Real-time Recommendation: Researched and implemented "M2TRec," a metadata-aware multi-task transformer for session-based recommendations. Solved cold-start problems and improved real-time personalization accuracy.

Adobe Research

Data Science Research Intern

San Jose, CA

May 2020 – Aug 2020

Neural Tensor Completion: Developed influence-guided data augmentation techniques for neural tensor completion. Work resulted in a publication at CIKM 2021.

EDUCATION

Georgia Institute of Technology

Ph.D. in Computer Science (Advisor: Prof. Srijan Kumar)

Atlanta, GA

Aug 2019 – May 2024

Seoul National University

B.S. in Computer Science and Engineering (Advisor: Prof. U Kang)

Seoul, South Korea

Mar 2012 – Aug 2018

RECENT PUBLICATIONS

Conferences

A. Nam, **S. Oh**, E. Kong, Y. Feng, M. Bhattacharya, "Modeling diverse preferences in movie artwork personalization with large language models." *Pluralistic Alignment Workshop at ICML*, 2026.

S. Oh, Y. Jin, M. Sharma, et al. "UniGuard: Towards Universal Safety Guardrails for Jailbreak Attacks on Multimodal Large Language Models." *Deployable AI Workshop at AAAI*, 2025.

S. Oh, G. Verma, S. Kumar. "Adversarial Text Rewriting for Text-aware Recommender Systems." *ACM CIKM*, 2024.

G. Verma... **S. Oh**, et al. "Cross-Modal Projection in Multimodal LLMs Doesn't Really Project Visual Attributes to Textual Space." *ACL*, 2024.

S. Oh, J. McAuley, B. Ustun, S. Kumar. "Rank List Sensitivity of Recommender Systems to Interaction Perturbations." *ACM CIKM*, 2022.

Technical Blogs

S. Oh, M. Bhattacharya, Y. Feng, S. Lamkhede. "FM-Intent: Predicting User Session Intent with Hierarchical Multi-Task Learning." *Netflix Tech Blog*, 2025.

TECHNICAL SKILLS

Core AI: Agents, Multimodal LLMs, Post-training, Foundation Models, Adversarial ML.

Model Architecture: LLMs, VLMs, Semantic IDs, Multi-task Learning, Model Distillation, High-Performance Computing, Tensor Analysis.

Languages & Tools: Python, PyTorch, TensorFlow, Hugging Face, C, Scala, TorchTune, Qwen, Llama.

PROFESSIONAL SERVICES

Conference Reviewer: AAAI 2024, KDD 2023, CIKM 2021-2022.

Journal Reviewer: ACM Transactions on Knowledge Discovery from Data (TKDD), IEEE TPDS.

HONORS & AWARDS

EB-1B Green Card Holder | Outstanding Professors and Researchers Classification Jan 2025

Kwanjeong Ph.D. Fellowship | \$150K total funding, awarded to top Korean students 2019 – 2024

Best Thesis Award | Seoul National University, 1st among all CSE undergrads 2018

Samsung Humantech Paper Award | Gold Prize (1st in Computer Science) 2018