Eunkyu (Eunice) Park

👩🏻‍💻 dbsltm.github.io 📧 eunkyu.park@vision.snu.ac.kr | eunkyup@andrew.cmu.edu
📍 Pittsburgh, PA

✨ RESEARCH STATEMENT

Working across NLP and HCI, my research explores human-centered approaches for AI models to learn Interactional Intelligence and help people make better sense of their social world. On the model side, I build benchmarks that evaluate social intelligence and methods that improve it, including structured reasoning traces and conversation policies for human–AI dialogue. On the human side, I run large-scale user studies to evaluate what users value in a model's reasoning and design interfaces to quantitatively measure perceived trust, reliance, and enjoyment.

EDUCATION

Seoul National University, *Seoul, South Korea* 2021–Present
Ph.D. Candidate, Interdisciplinary Program of Artificial Intelligence
Advisor: Gunhee Kim

Columbia University, *New York, NY* 2017–2020
B.S. Computer Science, Fu Foundation School of Engineering and Applied Sciences

Bard College, *Great Barrington, MA* 2014–2017
B.A. Mathematics

PUBLICATIONS

[1] Right Advice, Wrong Reasons: How Decision and Reasoning Alignment Shape Reliance on AI Advisors — Under Review
Eunkyu Park, Vasudha Varadarajan, Wesley Hanwen Deng, Anisha Reddy, Gunhee Kim, Motahhare Eslami, Jana Schaich Borg, Maarten Sap

[2] Who Does What in AI Auditing? Designing Human–AI Collaboration for Auditing Generative AI — Under Review
Eunkyu Park, Markelle Roesti, Wesley Hanwen Deng, Renata Barreto, Mohammad Tahaei, Ken Holstein, Jason I. Hong, Motahhare Eslami

[3] MM-SCALE: Evaluating Evidence-Grounded Moral Judgment in Vision-Language Models — NeurIPS 2026 Evaluations & Datasets
Eunkyu Park, Wesley Hanwen Deng, Cheyon Jin, Matheus Kunzler Maldaner, Jordan Wheeler, Jason I. Hong, Hong Shen, Adam Perer, Ken Holstein, Motahhare Eslami, Gunhee Kim

[4] Critical or Compliant? The Double-Edged Sword of Reasoning in Chain-of-Thought Explanations — EMNLP 2026 Main Conference
Eunkyu Park, Wesley Hanwen Deng, Vasudha Varadarajan, Mingxi Yan, Gunhee Kim, Maarten Sap*, Motahhare Eslami* (*Equal Advising)

[5] Jury Duty: Calibration and Orientation Failures in MLLM-as-a-Judge Under Cultural Ambiguity — EMNLP 2026 Main Conference
Daniel Lee, Harshit Sharma, **Eunkyu Park**, Pranav Narayanan Venkit, Jeonghwan Kim, Kah Mun Chia, Andreas Vlachos, Shafiq Joty

[6] Cognitive Chain-of-Thought (CoCoT): Structured Multimodal Reasoning about Social Situations — COLM 2026
Eunkyu Park, Wesley Hanwen Deng, Gunhee Kim, Motahhare Eslami, Maarten Sap

[7] HalLoc: Token-level Localization of Hallucinations for Vision Language Models — CVPR 2025
Eunkyu Park*, Minyeong Kim*, Gunhee Kim (*Co first-authors)

EXPERIENCE

Carnegie Mellon University, Pittsburgh PA — 2025.06–Present
Visiting Researcher at HCII
Hosted by: Professor Maarten Sap, Professor Motahhare Eslami

VISION & LEARNING LAB at Seoul National University — 2020.07–2021.03
Research Intern
Assisted research in developing models for Video Question and Answering benchmark that could expand to DramaQA, TV-QA domains, expanding to a multi-modal commonsense understanding framework.

PION CORPORATION, Seoul, South Korea — 2019.07–2019.09, 2020.04–2020.07
Machine Learning Engineer Intern
Developed a model extracting information such as an object's position, label, colors, coordinates from videos/images by masking objects with R-CNN and identifying them with YOLOv3 in a single frame

AWARDED GRANTS

Human-centered Approach for Measuring Societal Biases in Vision Language Models — 2025
Awarded ($55,000): SNU-CMU Joint Funding Proposal
Co-PI; Co-written with Wesley Hanwen Deng (CMU); Jason I. Hong, Adam Perer, Motahhare Eslami, Ken Holstein, Hong Shen (CMU team) and Gunhee Kim (SNU team)

PROFESSIONAL SERVICES

Reviewer for AAAI 2027, COLM 2026, TPAMI 2026, IJCV 2025

ACADEMIC SERVICES

Teaching Assistant at Seoul National University:
Computer Vision, Probabilistic Graphical Models, General Artificial Intelligence
Teaching Assistant for Corporate Training:
2025 LG Energy Deep Learning Course

TECHNICAL SKILLS

Experience	Post Training • Model Evaluation • NLP • Human-computer interaction • Crowdsourcing • User Studies • User Interfaces
Languages	English, Korean