Na Min An

Ph.D. Student

+82-10-2673-6972 | namin0202@gmail.com | namin-an | namin-an-0202

Research Interests

My research aims to develop assistive and trustworthy AI and machine learning systems grounded in an understanding of human perception, cognition, and action. I am particularly interested in identifying and mitigating modality-specific and social biases in vision–language models, using a mechanistic interpretability framework to uncover the internal representations and decision-making processes that give rise to biased behaviors.

Education

Korea Advanced Institute of Science and Technology (KAIST) *Seoul, South Korea*

Doctor of Philosophy in Artificial Intelligence, GPA: 4.13/4.3 *March 2023 - PRESENT*

- Advisors: Prof. Hyunjung Shim and Dr. James Thorne

University of Seoul *Seoul, South Korea*

Bachelor of Science in Mathematics, GPA: 4.29/4.5 *Feb. 2022*

- Summa Cum Laude (Received academic excellence or Mayor's scholarship for every semester)

Skills

Language Python, MATLAB, R, C++, JavaScript

Certification Society of Actuaries (P, FM, and SRM)

Experience

KAIST *Seoul, South Korea*

Research Assistant *March 2023 - PRESENT*

- Nominated for Apple PhD Fellowhip program (3 students per institution)
- *Ministry of Science and ICT Project*: Contributed leadership to a national project aimed at developing socially unbiased generative models.
- *Basic Research Lab Project*: Co-led a national project on developing barrier-free AI systems in educational and mobility domains.
- Published six first-author and five coauthored works, with research centered on fairness and human-centered ML systems.

University of Copenhagen *Copenhagen, Denmark*

Visiting Researcher *March 2025 - May 2025*

- Received a Danish Data Science Academy grant to advance research on explainability and reliability in multimodal models to build fairness-aware ML systems.
- Developed a neuron-level bias mitigation framework for VLMs and LVLMs with Dr. Yoonna Jang, in collaboration with NVIDIA Taiwan team (Dr. Yusuke Hirota and Dr. Ryo Hachiuma).

Korea Institute of Science and Technology (KIST) *Seoul, South Korea*

Research Intern *Oct. 2022 - Feb. 2023*

- Analyzed human perceptual limits and ML model performance under severely low-resolution visual conditions.
- Developed a reinforcement learning–based framework that emulates the learning effects frequently occurring in human psychophysical tests.
- Built an automated imaging pipeline for detecting fluorescence-image edges and identifying hyperpolarization phenomena with high precision.

University of Seoul *Seoul, South Korea*

Research Intern *Sep. 2020 - Dec. 2020*

- Investigated district-level Covid-19 risk factors across Seoul and awarded 3rd place (out of 40+ teams) in a competitive big-data competition.
- Developed a hybrid CNN–RNN model to classify the progressive stages of pulmonary nodules from sequential CT scan slices.
- Analyzed YouTube channel performance to model subscriber trajectory and provided targeted strategies for subscriber acquisition.

Working Projects

Large Vision-Language Models for Blind and Low-Vision Individuals

- Evaluated BLV user preferences for LVLM-generated scene descriptions and built data-driven metrics to improve model outputs for better accessibility (*arXiv:2502.14883*)
- Designed MULTI-TAP, a robust multi-objective scoring framework leveraging LVLM hidden states and introduced the EYE4ALL benchmark to advance human-aligned and BLV-centered vision–language evaluation (*arXiv:2510.00766*).
- Planning to develop a lightweight, real-time LVLM evaluator that can be deployed in assistive applications for BLV users.

Multilingual Referring Image Segmentation (RIS)

- Developed a training-free framework leveraging untapped spatial knowledge in CLIP to enhance zero-shot RIS (*arXiv:2509.23098*).
- Extending the framework to a multilingual VLM encoder to enable zero-shot RIS across diverse languages and linguistic structures.

Patents

1. Im, M., Roh, H., and **An, N. M.**, "Artificial Vision Parameter Learning and Automating Method for Improving Visual Prosthetic Systems", *US Patent Application No. 18/075555*, 2022.
2. Im, M., Roh, H., and **An, N. M.**, "Artificial Vision Expression Parameter Automation Learning System Method for Improving Artificial Vision Device", *Korea Paper Granted No. 0172619*, 2021.

Publications

1. Waheed, S., **An, N. M.**, Milford, M., Ramchurn, S. D., and Ehsan, S., "VLM-Guided Visual Place Recognition for Planet-Scale Geo-Localization," *ACRA*, 2025.
2. Waheed, S.* and **An, N. M.***, "Image Embedding Sampling Method for Diverse Captioning," *EMNLP Main* (* denotes equal contribution), 2025.
3. **An, N. M.***, Kim, E.*, Thorne, J., and Shim, H., "I0T: Embedding Standardization Method Towards Zero Modality Gap," *ACL Outstanding* (Top 1%, 26 out of 3,000 accepted papers, * denotes equal contribution), 2025.
4. Bayramli, Z.*, Suleymanzade, A.*, **An, N. M.**, Ahmad, H., Kim, E., Kang, J., Thorne, J., and Oh, A., "Diffusion Models Through a Global Lens: Are They Culturally Inclusive?" *ACL Oral* (Top 8%, 243 out of 3,000 accepted papers, * denotes equal contribution), 2025.
5. Kang, W. J., Kim, E., **An, N. M.**, Kim, S., Choi, H., Kwak, K., and Thorne, J., "Sightation Counts: Leveraging Sighted User Feedback in Building a BLV-aligned Dataset of Diagram Descriptions," *ACL Main*, 2025.
6. Kim, J. S*, Yu, A.*, Ismayilzada J.*, Park, J., Kim, E., Ahmad, H., **An, N. M.**, Thorne, J., and Oh, A., "WHEN TOM EATS KIMCHI: Evaluating Cultural Awareness of Multimodal Large Language Models in Cultural Mixture Contexts," *NAACL C3NLP Workshop Outstanding* (* denotes equal contribution), 2025.
7. **An, N. M.**, Roh, H., Kim, S., Kim, J. H., and Im, M., "Machine Learning Techniques for Simulating Human Psychophysical Testing of Low-Resolution Phosphene Face Images in Artificial Vision," *Advanced Science* (Impact Factor: 14.3, JCR Ranking: 6.5%), 2025 - Featured in *Korean National News*.
8. Jung, W., Hong, J., **An, N. M.**, Thorne, J, and Yun, S., "Stable Language Model Pre-training by Reducing Embedding Variability," *EMNLP Main*, 2024.
9. **An, N. M.**, Weed, S., and Thorne, J., "Capturing the Relationship Between Sentence Triplets for LLM and Human-Generated Texts to Enhance Sentence Embeddings," *EACL Findings*, 2024.
10. Lee, N.*, **An, N. M.***, and Thorne, J. "Can Large Language Models Infer and Disagree Like Humans?" *EMNLP Main* (* denotes equal contribution), 2023.
11. **An, N. M.**, Roh, H., Kim, S., Kim, J. H., and Im, M. "Reinforcement Learning Framework to Simulate Short-Term Learning Effects of Human Psychophysical Experiments Assessing the Quality of Artificial Vision," *IJCNN Oral*, 2023.
12. **An, N. M.**, Roh. H., Jung, S., Kim, E. J., and Im, M. "Machine Learning Can Approaches as an Alternative to Human Psychophysical Tests of Prosthetic Vision," *EMBS* (1-page abstract), 2021.

Services

Talks

- AI Tutorial Series at KAIST AI (Winter 2025).

Volunteer

- Provided on-site and virtual support for EMNLP (2023), including registration, poster session management, and AV/technical support.
- Contributed to community support and educational outreach in Cambodia in Summer 2018.

Leader

- Lab representative student from March 2024 to Aug. 2025.
- Appeared in the KAIST AI promotional video (2024) and top-10 research achievements participant (2023).
- Lab research manager from Aug. 2023 to Dec. 2023.

Teaching Assistant

- Department Seminars & Colloquia, Fall 2023
- Computational Image Generation and Modification, Fall 2025

Mentor

- Led research mentorship for two Azerbaijan undergraduate students in South Korea from Sep. 2024 to Jan. 2025.
- Provided English and mathematics tutoring to Dongdaemun-gu (Seoul, South Korea) high school students in Fall 2019 and Winter 2020.

Reviewer

- IEEE Transactions on Audio, Speech and Language Processing 2025
- ACL ARR February, May, July, October 2025
- ACL Student Research Workshop 2025
- IJCNN 2024 and 2025

References

Hyunjung Shim — *March 2024 - PRESENT*

Associate Professor, KAIST AI — *Seoul, South Korea*

- Email: kateshim@kaist.ac.kr

James Thorne — *March 2023 - Aug. 2025*

Chief AI Scientist, Theia Insights (Past: Assistant Professor, KAIST AI) — *Cambridge, United Kingdom*

- Email: james@jamesthorne.com

Yoonna Jang — *March. 2025 - Dec. 2025*

Post-doc, University of Copenhagen — *Copenhagen, Denmark*

- Email: morelychee@gmail.com

Yusuke Hirota — *Aug. 2025 - Dec. 2025*

Research Scientist, NVIDIA Taiwan — *Taipei, Taiwan*

- Email: yusukeh@nvidia.com

Isabelle Augenstein — *March 2025 - Dec. 2025*

Professor, Department of Computer Science, University of Copenhagen — *Copenhagen, Denmark*

- Email: augenstein@di.ku.dk

Alice Oh — *Sep. 2024 - Dec. 2025*

Professor, KAIST School of Computing — *Daejeon, South Korea*

- Email: alice.oh@kaist.edu

Jeong-Rae Kim — *Sep. 2020 - Dec. 2020*

Professor, Department of Mathematics, University of Seoul — *Seoul, South Korea*

- Email: jrkim@uos.ac.kr

Maesoon Im — *Oct. 2021 - Feb. 2023*

Principal Research Scientist, Brain Science Institute, KIST — *Seoul, South Korea*

- Email: maesoon.im@kist.re.kr