Seungbeen Lee

seungblee@yonsei.ac.kr
seunbite.github.io
MS Student @Yonsei University

Education

2023-2028[*]	**Integrated Course**, Artificial Intelligence, Yonsei University
2018-2023	**B.A**, Psychology and Economics, Yonsei University

Research Experience

2023-	Graduate Student, MIR Lab, Yonsei University, (Advisor: Youngjae Yu)

Publications

G Google Scholar

† → Equal contribution

Conference Proceedings

C1. Chung, J., Lim, S., Jeon, J., **Seungbeen Lee** & Yu, Y. *Can visual language models resolve textual ambiguity with visual cues? Let visual puns tell you!* in (2024). arXiv: 2410.01023 [cs.CV]. https://arxiv.org/abs/2410.01023.

C2. Lee, S., Kim, S., Kim, M., Kang, D., Yang, D., Kim, H., Kang, M., Jung, D., Kim, M. H., **Seungbeen Lee**, Chung, K.-M., Yu, Y., Lee, D. & Yeo, J. *Cactus: Towards Psychological Counseling Conversations using Cognitive Behavioral Theory* in (2024). arXiv: 2407.03103 [cs.CL]. https://arxiv.org/abs/2407.03103.

C3. **Lee, Seungbeen**, Lim, S., Han, S., Oh, G., Chae, H., Chung, J., Kim, M., Kwak, B.-w., Lee, Y., Lee, D., *et al. Do llms have distinct and consistent personality? trait: Personality testset designed for llms with psychometrics* in (2024). https://arxiv.org/abs/2406.14703.

Working papers

W1. Lim, S., **Seungbeen Lee**, Min, D. & Yu, Y. *Persona Dynamics: Unveiling the Impact of Persona Traits on Agents in Text-Based Games* 2025.

W2. Seo, W., Han, S., Jung, J., Newman, B., Lim, S., **Seungbeen Lee**, Lu, X., Choi, Y. & Yu, Y. *Verifying the Verifiers: Unveiling Pitfalls and Potentials in Fact Verifiers* 2025.

W3. **Seungbeen Lee**, Jeong, J., Kim, D., Son, Y. & Yu, Y. *Connecting the Dots from Data: LLM-driven Tree-search Career Cartographies of Career Pathways as Your AI Career Explorer* 2025.

W4. **Seungbeen Lee**, Jeong, J., Kim, D., Son, Y. & Yu, Y. *Wait... Was That a Sign? Reading Minds Through Actions: Observable Theory of Mind with Nonverbal Cues* 2025.

---

[*]Expected.

W5. Yousefpour, A., Kim, T., Kwon, R. S., **Seungbeen Lee**, Jeung, W., Han, S., Wan, A., Ngan, H., Yu, Y. & Choi, J. *Representation Bending for Large Language Model Safety* 2025.

Tools & Software

Software skills

Python

Language

English (Proficient), Spanish (Intermediate)

Presentations

Posters

P1. **Lee, Seungbeen**, Lim, S., Han, S., Oh, G., Chae, H., Chung, J., Kim, M., Kwak, B.-w., Lee, Y., Lee, D., Yeo, J. & Yu, Y. *Do LLMs Have Distinct and Consistent Personality? TRAIT: Personality Testset designed for LLMs with Psychometrics* 2025 NAACL (Albuquerque, New Mexico, USA). Apr. 2025.

P2. Han, S., **Seungbeen Lee**, Kim, J., Hessel, J., Jiang, L., Chung, J., Son, Y., Choi, Y. & Yu, Y. *Reading Books is Great, But Not if You Are Driving! Visually Grounded Reasoning about Defeasible Commonsense Norms* 2023 EMNLP NILLI Workshop (Singapore). Dec. 2023.

Selected Media Coverage

2025   ScienceNews, Are AI chatbot 'personalities' in the eye of the beholder?

Teaching

Assistant Instructor

2025   MULTIMODAL DEEP LEARNING (AAI5010) (Prof. Youngjae Yu)
2024   INTRODUCTION TO AI RESEARCH (CCO2105) (Prof. Youngwoon Lee)

Other Experience

2020   HR Internship., Human Metrics, Seoul (Human Intelligence Evaluation)

Last updated: April 2, 2025