박진영/JinYeong Bak/朴秦永

nosyu.kr jy.bak@skku.edu

Research Interests

My research interests lie at the intersection of machine learning and natural language processing, with a focus on understanding and modeling human and AI-driven communication. My work centers on developing methodologies to evaluate AI-generated content and simulate human behavior through computational models. I am also interested in uncovering latent patterns and extracting insights from large-scale datasets, including conversational corpora. [1]

Employment

Associate Professor — Sungkyunkwan University (SKKU)
Suwon, South Korea — Sep 2024 - Present

Assistant Professor — SKKU
Suwon, South Korea — Sep 2020 - Aug 2024
I am working as a professor in the College of Computing and Informatics at Sungkyunkwan University. I am the leader of the Human Language Intelligence Lab[2].

Junior Data Scientist — United Nations Global Pulse Lab
Jakarta, Indonesia — June 2016 - Aug 2016
I interned with Dr. Jonggun Lee in the United Nations Global Pulse Lab Jakarta. I developed a topic model-based methodology for expanding relevant keywords for specific subjects to identify related tweets on Twitter. This work is published at ICML 2017 workshop [50].

Research Intern — Microsoft Research Asia
Beijing, China — Sep 2013 - Feb 2014
I interned with Dr. Chin-Yew Lin in the Knowledge Mining group at Microsoft Research Asia. I developed a topic model to identify self-disclosure in Twitter conversations. This work is published at EMNLP 2014 as a full paper [52].

Education

KAIST, Ph.D., School of Computing, Mar 2013 - Aug 2020
- Dissertation: Speaker oriented Conversation Model and its Evaluation [43]
- Dissertation committee: Prof. Alice Oh, Dr. Chin-Yew Lin, Prof. Kee-Eung Kim, Prof. Meeyoung Cha, and Prof. Sung-Hyon Myaeng
- Outstanding PhD Thesis Award

KAIST, M.S., Computer Science, Feb 2011 - Feb 2013
- GPA: 3.98/4.30
- Thesis: Distributed Online Learning for Topic Models [53]
- Thesis committee: Prof. Alice Oh, Prof. Sung-Eui Yoon, and Prof. Sung-Hyon Myaeng

Sungkyunkwan University, B.S., Computer Engineering, Mar 2004 - Feb 2011
- GPA: 4.45/4.50
- Graduation project: On-Line LEcture Allocation[3], Worked with Byung Il Woo,
- Best Project Award
- SKKU Academic Excellence Scholarship (Fall 2004), Udeok Scholarship (Mar 2008 - Feb 2011)

Last updated: January, 2026

[1] Full research statement: https://nosyu.kr/assets/JinYeong_Bak_ResearchStatement.pdf

[2] https://hli.skku.edu

[3] Online system for lecture allocation and mileage management

Publications

Google Scholar: https://scholar.google.com/citations?user=oYK9Z_IAAAAJ
ORCiD: https://orcid.org/0000-0002-3212-5241
DBPia: https://www.dbpia.co.kr/author/authorDetail?ancId=5086204

[1] Taemin Yeom, Yonghyun Ryu, Yoonjung Choi, and **JinYeong Bak**. Tagged span annotation for detecting translation errors in reasoning LLMs. In *Proceedings of the Tenth Conference on Machine Translation*, pages 878–886, Suzhou, China, November 2025. Association for Computational Linguistics.

[2] Sooyung Choi, Jaehyeok Lee, Xiaoyuan Yi, Jing Yao, Xing Xie, and **JinYeong Bak**. Unintended harms of value-aligned LLMs: Psychological and empirical insights. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 31742–31768, Vienna, Austria, July 2025. Association for Computational Linguistics.

[3] JiWoo Kim, Minsuk Chang, and **JinYeong Bak**. Text overlap: An LLM with human-like conversational behaviors. In *Proceedings of the Third Workshop on Social Influence in Conversations (SICon 2025)*, pages 124–136, Vienna, Austria, July 2025. Association for Computational Linguistics.

[4] Jaehyeok Lee, Keisuke Sakaguchi, and **JinYeong Bak**. Self-training meets consistency: Improving LLMs' reasoning with consistency-driven rationale evaluation. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 10519–10539, Albuquerque, New Mexico, April 2025. Association for Computational Linguistics.

[5] YeongJun Hwang, Dongjun Kang, and **JinYeong Bak**. Dialogue response coherency evaluation with feature sensitive negative sample using multi list-wise ranking loss. In *Engineering Applications of Artificial Intelligence*, volume 150, page 110609, 2025.

[6] Xiaohong Yu, Jinyong Kim, Yoseop Ahn, Mose Gu, Jaehoon (Paul) Jeong, **JinYeong Bak**, and Jaemin Jo. An intelligent marketing platform with influencer classification in social networking services. In *Knowledge-Based Systems*, volume 310, page 112972, 2025.

[7] Dai Quoc Tran, Yuntae Jeon, Armstrong Aboah, **JinYeong Bak**, Minsoo Park, and Seunghee Park. Leveraging semisupervised learning for domain adaptation: Enhancing safety at construction sites through long-tailed object detection. In *Journal of Construction Engineering and Management*, volume 151, page 04024190, 2025.

[8] Heeyoung Lee, Hoyoon Byun, Changdae Oh, **JinYeong Bak**, and Kyungwoo Song. Perturb-and-compare approach for detecting out-of-distribution samples in constrained access environments. In *Proceedings of the 27th European Conference on Artificial Intelligence*, volume 392 of *Frontiers in Artificial Intelligence and Applications*. EurAI, October 2024.

[9] JiWoo Kim, Yunsu Kim, and **JinYeong Bak**. KpopMT: Translation dataset with terminology for kpop fandom. In *Proceedings of the The Seventh Workshop on Technologies for Machine Translation of Low-Resource Languages (LoResMT 2024)*, pages 37–43, Bangkok, Thailand, August 2024. Association for Computational Linguistics.

[10] Sangjun Park and **JinYeong Bak**. Memoria: Resolving fateful forgetting problem through human-inspired memory architecture. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 39587–39615. PMLR, 21–27 Jul 2024.

[11] Minsoo Park, Almo Senja Kulinan, Tran Quoc Dai, **JinYeong Bak**, and Seunghee Park. Preventing falls from floor openings using quadrilateral detection and construction worker pose-estimation. In *Automation in Construction*, volume 165, page 105536, 2024.

[12] Aron Berhanu Degefa, Geonyeol Jeon, Sooyung Choi, **JinYeong Bak**, Seunghee Park, Hyungchul Yoon, and Solmoi Park. Data-driven insights into controlling the reactivity of supplementary cementitious materials in hydrated cement. In *International Journal of Concrete Structures and Materials*, volume 18, page 39, Jun 2024.

[13] HyunJin Kim, Young Jin Kim, and **JinYeong Bak**. PEMA: An offsite-tunable plug-in external memory adaptation for language models. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 6045–6064, Mexico City, Mexico, June 2024. Association for Computational Linguistics.

[14] Youngjin Jo and **JinYeong Bak**. Emosum: Conversation summarization with emotional consistency. In *Proceedings of the 39th ACM/SIGAPP Symposium on Applied Computing*, SAC '24, page 723–730. Association for Computing Machinery, 2024.

[15] Aron Berhanu Degefa, Hokeun Yoon, Seunghee Park, Hyungchul Yoon, **JinYeong Bak**, and Solmoi Park. Machine learning applied to predicting phase assemblages of hardened cementitious systems. In *Ceramics International*, 2024.

[16] Sangjun Park and **JinYeong Bak**. Lengthy essay generation with summary-based memory system. In *Proceedings of the Korea Software Congress*, pages 1571–1573. The Korean Institute of Information Scientists and Engineers, December 2023.

[17] Dongjun Kang, Joonsuk Park, Yohan Jo, and **JinYeong Bak**. From values to opinions: Predicting human behaviors and stances using value-injected large language models. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 15539–15559, Singapore, December 2023. Association for Computational Linguistics.

[18] Hokeun Yoon and **JinYeong Bak**. Diversity enhanced narrative question generation for storybooks. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 465–482, Singapore, December 2023. Association for Computational Linguistics.

[19] Jiwoo Kim, Youngbin Kim, Ilwoong Baek, **JinYeong Bak**, and Jongwuk Lee. It ain't over: A multi-aspect diverse math word problem dataset. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 14984–15011, Singapore, December 2023. Association for Computational Linguistics.

[20] Van Vy, Yunwoo Lee, **JinYeong Bak**, Solmoi Park, Seunghee Park, and Hyungchul Yoon. Damage localization using acoustic emission sensors via convolutional neural network and continuous wavelet transform. In *Mechanical Systems and Signal Processing*, volume 204, page 110831, 2023.

[21] Minsoo Park, Dai Quoc Tran, **JinYeong Bak**, Almo Senja Kulinan, and Seunghee Park. Real-time monitoring unsafe behaviors of portable multi-position ladder worker using deep learning based on vision data. In *Journal of Safety Research*, 2023.

[22] Hyunjin Kim, **JinYeong Bak**, Kyunghyun Cho, and Hyungjoon Koo. A transformer-based function symbol name inference model from an assembly language for binary reversing. In *Proceedings of the 2023 ACM Asia Conference on Computer and Communications Security*, ASIA CCS '23, page 951–965, New York, NY, USA, 2023. Association for Computing Machinery.

[23] Dongjin Jeong and **JinYeong Bak**. Conversational emotion-cause pair extraction with guided mixture of experts. In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, pages 3280–3290, Dubrovnik, Croatia, May 2023. Association for Computational Linguistics.

[24] Minsoo Park, Dai Quoc Tran, **JinYeong Bak**, and Seunghee Park. Small and overlapping worker detection at construction sites. In *Automation in Construction*, volume 151, page 104856, 2023.

[25] Jihee Kim and **JinYeong Bak**. Increasing robustness of end-to-end speech translation using pitch and speed perturbation. In *Proceedings of the Korea Computer Congress*, pages 815–817. The Korean Institute of Information Scientists and Engineers, December 2022.

[26] Dongjun Kang and **JinYeong Bak**. Dialogue response evaluation model with conversational feature sensitive negative sampling. In *2023 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pages 183–186, 2023.

[27] Juhee Son, Jiho Jin, Haneul Yoo, **JinYeong Bak**, Kyunghyun Cho, and Alice Oh. Translating hanja historical documents to contemporary Korean and English. In *Findings of the Association for Computational Linguistics: EMNLP 2022*, pages 1260–1272, Abu Dhabi, United Arab Emirates, December 2022. Association for Computational Linguistics.

[28] Minsoo Park, Dai Quoc Tran, **JinYeong Bak**, and Seunghee Park. Advanced wildfire detection using generative adversarial network-based augmented datasets and weakly supervised object localization. In *International Journal of Applied Earth Observation and Geoinformation*, volume 114, page 103052, 2022.

[29] Haneul Yoo, Jiho Jin, Juhee Son, **JinYeong Bak**, Kyunghyun Cho, and Alice Oh. HUE: Pretrained model and dataset for understanding hanja documents of Ancient Korea. In *Findings of the Association for Computational Linguistics: NAACL 2022*, pages 1832–1844, Seattle, United States, July 2022. Association for Computational Linguistics.

[30] HyeJoon Jang and **JinYeong Bak**. Detoxifying toxic comments using text style transfer. In *Proceedings of the Korea Computer Congress*, pages 2081–2083. The Korean Institute of Information Scientists and Engineers, June 2022.

[31] ChaeYun Jang and **JinYeong Bak**. Multi-turn question generation using past and future information. In *Proceedings of the Korea Computer Congress*, pages 1976–1978. The Korean Institute of Information Scientists and Engineers, June 2022.

[32] Dai Quoc Tran, Minsoo Park, Yuntae Jeon, **JinYeong Bak**, and Seunghee Park. Forest-fire response system using deep-learning-based approaches with cctv images and weather data. In *IEEE Access*, volume 10, pages 66061–66071, 2022.

[33] Ida Ayu Putu Ari Crisdayanti, **JinYeong Bak**, YunSeok Choi, and Jee-Hyong Lee. Ia-bert: Context-aware sarcasm detection by incorporating incongruity attention layer for feature extraction. In *Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing*, SAC '22, page 1084–1091, New York, NY, USA, 2022. Association for Computing Machinery.

[34] JinUk Cho, MinSu Jeong, **JinYeong Bak**, and Yun-Gyung Cheong. Genre-controllable story generation via supervised contrastive learning. In *Proceedings of the ACM Web Conference 2022*, WWW '22, page 2839–2849, New York, NY, USA, 2022. Association for Computing Machinery.

[35] HyunJin Kim and **JinYeong Bak**. Function name prediction using binary code with transformer. In *Proceedings of the Korea Software Congress*, pages 449–451. The Korean Institute of Information Scientists and Engineers, December 2021.

[36] DongJin Jeong and **JinYeong Bak**. Extracting emotion-cause information from conversation data using the graph structure. In *Proceedings of the Korea Software Congress*, pages 515–517. The Korean Institute of Information Scientists and Engineers, December 2021.

[37] YeongJun Hwang and **JinYeong Bak**. User attribute inference using knowledge graph. In *Proceedings of the Korea Software Congress*, pages 551–553. The Korean Institute of Information Scientists and Engineers, December 2021.

[38] Ji Woo Kim and **JinYeong Bak**. Finding the hidden relation through text-mining the annals of the joseon dynasty. In *Proceedings of the Korea Software Congress*, pages 1473–1439. The Korean Institute of Information Scientists and Engineers, December 2021.

[39] Yohan Jo, Haneul Yoo, **JinYeong Bak**, Alice Oh, Chris Reed, and Eduard Hovy. Knowledge-enhanced evidence retrieval for counterargument generation. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 3074–3094, Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics.

[40] YunSeok Choi, **JinYeong Bak**, CheolWon Na, and Jee-Hyong Lee. Learning sequential and structural information for source code summarization. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, Online, August 2021. Association for Computational Linguistics.

[41] **JinYeong Bak** and Alice Oh. A Leader's Final Decision Classification Model Tested on Meeting Records with BERT. In *Journal of KIISE*, volume 48, 2021.

[42] Bonggun Shin, Sungsoo Park, **JinYeong Bak**, and Joyce C. Ho. Controlled molecule generator for optimizing multiple chemical properties. In *Proceedings of the Conference on Health, Inference, and Learning*, 2021.

[43] **JinYeong Bak**. Speaker oriented conversation model and its evaluation. PhD dissertation, KAIST, Daejeon, South Korea, 2020.

[44] **JinYeong Bak** and Alice Oh. Speaker sensitive response evaluation model. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020.

[45] **JinYeong Bak** and Alice Oh. Variational hierarchical user-based conversation model. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, 2019.

[46] **JinYeong Bak** and Alice Oh. Conversational decision-making model for predicting the king's decision in the annals of the Joseon dynasty. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2018.

[47] Gabriel Lima and **JinYeong Bak**. Speech emotion classification using raw audio input and transcriptions. In *Proceedings of the 2018 International Conference on Signal Processing and Machine Learning*, 2018.

[48] Sungjoon Park, **JinYeong Bak**, and Alice Oh. Rotated word vector representations and their interpretability. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 2017.

[49] Jongin Lee, Daeki Cho, Junhong Kim, Eunji Im, **JinYeong Bak**, Kwan Hong Lee, John Kim, and others. Itchtector: A wearable-based mobile system for managing itching conditions. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2017.

[50] **JinYeong Bak**, Imaduddin Amin, Jong Gun Lee, and Alice Oh. Keyword expansion for understanding crisis events in Indonesian tweets. In *ICML Workshop on Interactive Machine Learning and Semantic Information Retrieval*, 2017.

[51] **JinYeong Bak** and Alice Oh. Five centuries of monarchy in Korea: Mining the text of the annals of the Joseon dynasty. In *Proceedings of the 9th SIGHUM Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (LaTeCH)*, 2015.

[52] **JinYeong Bak**, Chin-Yew Lin, and Alice Oh. Self-disclosure topic model for classifying and analyzing Twitter conversations. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, 2014.

[53] **JinYeong Bak**. Distributed online learning for topic models. Master's thesis, KAIST, Daejeon, South Korea, 2012.

[54] **JinYeong Bak**, Dongwoo Kim, and Alice Oh. Distributed online learning for latent Dirichlet allocation. In *Proceedings of Workshop on Big Learning : Algorithms, Systems, and Tools at the Neural Information Processing Systems*, 2012.

[55] **JinYeong Bak**, Suin Kim, and Alice Oh. Self-disclosure and relationship strength in Twitter conversations. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*, 2012.

[56] Suin Kim, **JinYeong Bak**, and Alice Oh. Do you feel what I feel? social aspects of emotions in Twitter conversations. In *Proceedings of the AAAI International Conference on Weblogs and Social Media*, 2012.

[57] Suin Kim, **JinYeong Bak**, Yohan Jo, and Alice Oh. Do you feel what I feel? social aspects of emotions in twitter conversations. In *Proceedings of Workshop on Computational Social Science and the Wisdom of Crowds*, 2011.

[58] Rae Young Ko, Duk Sun Kim, **JinYeong Bak**, and Sang Gu Lee. Development of mobile sage-math and its use in linear algebra. In *J. Korea Soc. Math. Ed. Ser. E: Communications of Mathematical Education*, 2009.

[59] Duk Sun Kim, **JinYeong Bak**, and Sang Gu Lee. The educational models using enhanced mathematics ict in the Korean it environments. In *J. Korea Soc. Math. Ed. Ser. E: Communications of Mathematical Education*, 2008.

Manuscripts

Arxiv: https://arxiv.org/search/?query=JinYeong+Bak&searchtype=author

[1] Tarek Naous, Anagha Savit, Carlos Rafael Catalan, Geyang Guo, Jaehyeok Lee, Kyungdon Lee, Lheane Marie Dizon, Mengyu Ye, Neel Kothari, Sahajpreet Singh, Sarah Masud, Tanish Patwa, Trung Thanh Tran, Zohaib Khan, Alan Ritter, **JinYeong Bak**, Keisuke Sakaguchi, Tanmoy Chakraborty, Yuki Arase, and Wei Xu. Camellia: Benchmarking cultural biases in llms for asian languages, 2025. [link].

[2] HyunJin Kim, Xiaoyuan Yi, Jing Yao, Muhua Huang, **JinYeong Bak**, James Evans, and Xing Xie. Research on superalignment should advance now with parallel optimization of competence and conformity, 2025. [link].

[3] JiWoo Kim, Minsuk Chang, and **JinYeong Bak**. Beyond turn-taking: Introducing text-based overlap into human-llm interactions, 2025. [link].

[4] Jio Oh, Geon Heo, Seungjun Oh, Hyunjin Kim, **JinYeong Bak**, Jindong Wang, Xing Xie, and Steven Euijong Whang. Better think with tables: Tabular structures enhance llm comprehension for data-analytics requests, 2025. [link].

[5] HyunJin Kim, Xiaoyuan Yi, Jing Yao, Jianxun Lian, Muhua Huang, Shitong Duan, **JinYeong Bak**, and Xing Xie. The road to artificial superintelligence: A comprehensive survey of superalignment, 2024. [link].

[6] Yeonji Lee, Sangjun Park, Kyunghyun Cho, and **JinYeong Bak**. Mentalagora: A gateway to advanced personalized care in mental health through multi-agent debating and attribute control, 2024. [link].

[7] Kyumin Park, Myung Jae Baik, YeongJun Hwang, Yen Shin, HoJae Lee, Ruda Lee, Sang Min Lee, Je Young Hannah Sun, Ah Rah Lee, Si Yeun Yoon, Dong ho Lee, Jihyung Moon, **JinYeong Bak**, Kyunghyun Cho, Jong-Woo Paik, and Sungjoon Park. Harmful suicide content detection, 2024. [link].

Academic Services - Program Committee

AAAI	2020, 2021, 2022, 2023, 2024, 2025, 2026
AACL	2020, 2022, 2023, 2025
ACL	2015, 2016, 2017, 2018, 2019, 2020, 2021, 2023, 2024, 2025
ARR	2021 July, September, October, November 2022 January, July, October 2023 December 2024 February, April, June, August, October, December 2025 February, May[4], July, October
BigComp	2023, 2024, 2025, 2026
CHI	2025
CL	2024, 2025
COLING	2020, 2022, 2024, 2025
COLM	2024, 2025
EAAI	2024, 2025
EACL	2016, 2023, 2026

[4] Outstanding Area Chair Award

EMNLP	2015, 2016, 2019, 2020, 2021, 2022, 2023, 2024, 2025
HCLT	2023, 2024, 2025
ICLR	2021, 2022, 2023, 2024, 2025, 2026
ICML	2020, 2023, 2024
ICWSM	2015, 2016
IEIE	2024
JOK	2020, 2021, 2022, 2023, 2024, 2025
KCC	2025
KSC	2023, 2024
LaTeCH-CLfL	2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025
LREC	2022, 2024, 2026
NAACL	2019, 2021, 2024, 2025
NeurIPS	2021, 2022, 2023, 2024, 2025
NN	2024
RDGENAI	2025
SAC	2021, 2024, 2025, 2026
TASLP	2022, 2023
TheWebConf	2019, 2020, 2022
WiNLP	2024, 2025
W-NUT	2020, 2021, 2022, 2024, 2025

Academic Services - Others

Chair	IJCNLP-AACL 2023 Publicity and Social Media Chair ACM FAccT 2022 Virtual co-Chair
Session Chair	ACL 2025 KSC 2024 HCLT-KACL 2024 BigComp 2023 EMNLP 2022 TheWebConf 2022 POLTEXT 2019
Organizer	W-NUT 2024, 2025 IJCAI 2024 Workshop and Tutorial Coordinators

Volunteer	ACL 2020 Birds of a Feather Meetup ACL 2020 Group Mentor ACL 2020 Micro-blogging ICLR 2020 Student Volunteer EMNLP 2019 Student Volunteer EMNLP 2018 Student Volunteer ACL 2012 Student Volunteer

Projects

"AI Star Fellowship Support Program", IITP, Jul 2025 - Present

"AI Plus K-Construction Infrastructure Resilience (AI+KCIR) Global Research Center", NRF, Jun 2025 - Present

"Detection and Prediction of Emerging and Undiscovered Voice Phishing", IITP, Apr 2025 - Present

"Offsite AI Model Learning Algorithm for ensuring model and data confidentiality and performance", NRF, Mar 2025 - Present

"Enhancing Cultural Alignment in Large Language Models with Human Values and Safety", Microsoft Research Asia, Mar 2025 - Present

"AI Guardians: Development of Robust, Controllable, and Unbiased Trustworthy AI Technology", IITP, Aug 2024 - Present

"Demonstration and commercialization of a large-scale AI-based mental health service to eliminate blind spots in mental health", NIPA, Jul 2024 - Dec 2024

"Collaborative Research Projects with Microsoft Research", IITP, Apr 2024 - Present

"Enhancing Language Model Alignment with Human Values and Safety", Microsoft Research Asia, May 2024 - Present

"Research on the reliability and coherence of outcomes produced by Generative AI", IITP, Apr 2024 - Present

"Graduate Program of Next-Generation Computing for Sustainable Development", NRF, Mar 2024 - Present

"Development of an Artificial Intelligence Model for Integrated Depression Diagnosis Technology based on Depression Behavior Characteristics Dataset", NRF, Aug 2023 - Dec 2024

"A study on DetectGPT to detect plagiarism of codes and texts from large language models", Elice, June 2023 - June 2023

"Integrating Human Values into Language Models: Generating Human Value-Aligned Arguments", Microsoft Research Asia, June 2023 - Present

"Language Localization Neural Machine Translation Model for User Generated Text", NRF, June 2023 - May 2024

"A study on large AI de-biasing algorithms", KT, Dec 2022 - Sep 2023

"A study on explainable automatic evaluation model for generated natural language texts in terms of factual consistency", NRF, June 2022 - May 2023

"Abductive inference framework using omni-data for understanding complex causal relations", IITP, Apr 2022 - May 2025

"National Program for Excellence in SW", IITP, Apr 2021 - Present

"A study on automatic evaluation model for generated natural language texts in terms of topic consistency", NRF, June 2021 - May 2022

"Digital SOC D.N.A. based Health Care Laboratory for Carbon Neutralization", NRF, June 2021 - Feb 2024

"Development of Digital Therapeutics for Depression from COVID19", KEIT, June 2021 - Feb 2025

"Standard Development of Blockchain based Network Management Automation Technology", IITP, Sep 2020 - May 2021

"Artificial Intelligence Graduate School Program", IITP, Sep 2020 - Present

"Explainable Human-level Deep Machine Learning", IITP, Sep 2017 - Aug 2020

"Machine Learning Center", IITP, Mar 2013 - Mar 2017

"Distributed online topic modeling for big data analysis", Samsung Electronics, Dec 2012 - Nov 2013

Talks

"Reliability in Large Language Models: From Self-Training to Cultural Hallucination", Microsoft Research Asia, Beijing, China, 2025.09.16.

"Value Alignment and Superalignment in AI: Unpacking Safety Risks and Future Directions", New York University, New York, United States of America, 2025.08.25.

"Enhancing Memory in Neural Networks: Memory Retention, Efficient Adaptation, and Self-Training for Reasoning", Georgia Institute of Technology, Atlanta, United States of America, 2025.07.22.

"Enhancing Memory in Neural Networks: Memory Retention, Efficient Adaptation, and Self-Training for Reasoning", Tohoku University, Sendai, Japan, 2025.06.19.

"Value-Aligned Large Language Models, How & Unintended Harms", SKKU, Seoul, South Korea, 2025.05.30.

"PEMA & Memoria", Postech, Pohang, South Korea, 2025.04.04.

"PEMA & Memoria", DGIST, Daegu, South Korea, 2025.04.03.

"Memoria: Resolving Fateful Forgetting Problem through Human-Inspired Memory Architecture", Dongguk University, Seoul, South Korea, 2024.10.11.

"Memoria: Resolving Fateful Forgetting Problem through Human-Inspired Memory Architecture", RIKEN, Tokyo, Japan, 2024.08.19.

"Applying ML to Historical Corpora: Translating Classical Korean Texts to Contemporary Korean and Analyzing Monarchical Ruling Styles", Machine Learning for Ancient Languages Workshop, Bangkok, Thailand, 2024.08.15

"Life with human-like artificial intelligence", Korea National Diplomatic Academy, Seoul, South Korea, 2024.04.23

“At the King’s Command: A Historical Study of Extraordinary Natural Occurrences and Executive Orders”, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates, 2024.03.21

“From Human Values to Personal Opinions with LLM”, Adobe, San Jose, California, United States of America, 2023.08.07

“From Human Values to Personal Opinions”, Yonsei University, Seoul, South Korea, 2023.06.30

“From Human Values to Personal Opinions”, Microsoft Research Asia, Beijing, China, 2023.06.27

“Understanding ChatGPT”, Samsung Medical Center, Seoul, South Korea, 2023.06.09

“Understanding ChatGPT”, SKKU, Suwon, South Korea, 2023.04.26

“Transformer and Pretrained Language Model”, KICS, Seoul, South Korea, 2023.03.23

“Emotion-Cause Pair Extraction & Sociolect-to-Sociolect”, KRAFTON, Seoul, South Korea, 2023.02.16.

“Tips for graduates students and NLP Research Trends”, GIST, Online, 2023.01.13.

“NLP for Conversations”, Soongsil University, Seoul, South Korea, 2022.12.28.

“Natural Language Inference with Knowledge and Emotion-Cause Pair Extraction”, RIKEN, Tokyo, Japan, 2022.11.01.

“Mining the Text of the Annals of the Joseon Dynasty”, Seoul National University, Seoul, South Korea, 2022.10.28.

“NLP for Conversations”, Yonsei University, Seoul, South Korea, 2022.09.15.

“Introduction to NLP for Conversations”, University of Seoul, Seoul, South Korea, 2022.09.06.

“AI and Ethics”, The AI Korea 2022, Seoul, South Korea, 2022.08.17.

“Mining the Text of the Annals of the Joseon Dynasty”, KAIST, Daejeon, South Korea, 2021.12.14.

“Evaluation and Scalability of Conversation models”, Korea Institute of Science and Technology Information, Daejeon, South Korea, 2021.08.20.

“Automatic evaluation methods for conversation models”, SOCAR, Seoul, South Korea, 2021.07.28.

“Automatic evaluation methods for conversation models”, 2021 Summer Conference, Korean Artificial Intelligence Association, Online, 2021.07.09.

“2021 NLP Research Trends”, CJ Olive Networks, Seoul, South Korea, 2021.06.08.

“Automatic evaluation methods for conversation models”, AI Frontiers Summit 2021, The Korean Institute of Communications and Information Sciences, Seoul, South Korea, 2021.05.21.

“Scaling laws in Natural Language Processing”, 2021 Spring Colloquium, SKKU Department of Physics, Suwon, South Korea, 2021.03.03.

“Conversation Model and its Evaluation”, 2020 Fall Conference, Korean Artificial Intelligence Association, Online, 2020.11.20.

“Speaker-oriented Conversation Model and its Evaluation”, Online Seminar, Dongguk University, Seoul, South Korea, 2020.11.18.

"Speaker-oriented Conversation Model and its Evaluation", BK21 Artificial Intelligence Colloquium, Ajou University, Online, 2020.11.10.

"Speaker-oriented Conversation Model and its Evaluation", Fall 2020 CSE GSAI Seminar Series, Postech, Online, 2020.09.09.

"Speaker-oriented Conversation Model and its Evaluation", Machine Learning for Language, NYU, Online, 2020.06.16.

"At the King's command: rules and natural disasters in the Annals of Dynasty", 2nd Annual POLTEXT Conference 2019, Tokyo, Japan, 2019.09.14.

"Variational Hierarchical User-based Conversation Model", KAIST-NAVER Clova AI Workshop, KAIST, Daejeon, South Korea, 2019.06.19.

"Conversational decision-making model for predicting the kings decision in the annals of the Joseon dynasty", NAVER, Seongnam, South Korea, 2018.12.14.

"Conversational decision-making model for predicting the kings decision in the annals of the Joseon dynasty and Word vector Interpretability", Samsung Electronics, Suwon, South Korea, 2018.11.30.

"Studying Political Contention using Text as Data", GESIS, Cologne, Germany, 2015.12.02.

"Meaning Structures in Political Discourse: Measuring Institutional Dynamics of a Hybrid Democracy via the Topic Modeling from Contested Concepts in Newspaper Articles", GESIS, Cologne, Germany, 2014.12.01.

"Self-disclosure in Twitter conversations", Qatar Computing Research Institute, Doha, Qatar, 2014.10.23.

"Bayesian Nonparametric Topic Modeling", Korean machine learning summer school, Seoul, South Korea, 2013.08.22.

Teaching Experience

Lecturer, "Open Source Software Practice", SKKU, {Spring 2023, Fall 2023, Spring 2024, Fall 2024, Spring 2025, Fall 2025}

Lecturer, "Mathematics for Machine Learning", SKKU, {Fall 2022, Fall 2023, Fall 2024, Spring 2025}

Lecturer, "Machine Learning Algorithms and Applications", SKKU, {Spring 2022, Fall 2022, Spring 2023, Spring 2024}

Lecturer, "Smart Factory Application Programming", SKKU, {Fall 2021, Fall 2023}

Lecturer, "AI and Ethics", SKKU, {Fall 2021, Spring 2024, Fall 2025}

Lecturer, "Introduction to Artificial Intelligence", SKKU, {Fall 2021, Fall 2022, Fall 2023, Fall 2024, Fall 2025}

Lecturer, "Bayesian Learning", SKKU, {Spring 2021, Spring 2023}

Lecturer, "Natural Language Processing", SKKU, {Spring 2021, Fall 2021, Spring 2022, Spring 2023, Spring 2025}

Lecturer, "Natural Language Processing", DSME, {Oct 2020, June 2021, Dec 2021, Feb 2023}

Lecturer, “Probability and Random Process”, SKKU, {Fall 2020, Spring 2021, Spring 2022, Spring 2024, Spring 2025, Fall 2025}

Lecturer, “Artificial Intelligence for Data Scientists”, elice, {Apr 2017 - Aug 2017, Jun 2018 - Sep 2018}

Teaching Assistant, “Artificial Intelligence and Machine Learning”, KAIST, {Fall 2011, Spring 2013, Spring 2015, Fall 2018}

Teaching Assistant, “Introduction to Programming”, KAIST, {Spring/Fall 2014, Fall 2015, Spring/Fall 2016, Spring/Fall 2017}

Teaching Assistant, “IT Convergence User-centered Service Design”, KAIST - Microsoft Design Expo, Spring 2015

Teaching Assistant, “Operating Systems”, SKKU, Fall 2010

Teaching Assistant, “Discrete Mathematics”, SKKU, Spring 2010

Graduate Student Research Supervision

Academic Advisor

- Sunkyung Han, sunkyoung19@g.skku.edu, Master (Expected)
- HyunJin Hwang, nuyh2634@g.skku.edu, Master (Expected)
- San Kim, kimsan1120@g.skku.edu, Master (Expected)
- Taemin Yeom, taemin.yeom@g.skku.edu, Master (Expected)
- Jinsu Shin, jinsu0000@g.skku.edu, Master (Expected)
- Jaejun Shim, junshim@g.skku.edu, Master (Expected)
- Jaehyeok Lee, hjl8708@g.skku.edu, Ph.D (Expected)
- Gaeun Seo, gaeun0112@g.skku.edu, Master (Expected)
- Kyungdon Lee, leekd97@skku.edu, Master (Expected)
- Hyunbin Song, shbin05@g.skku.edu, Master (Expected)
- Suhyun Han, gkstngus01@g.skku.edu, Master (Expected)
- Nahyeon Park, nastela@g.skku.edu, Master (Expected)
- EunBeen Son, nabin111@g.skku.edu, Master, 2026
- YeongJun Hwang, hmtyj2@g.skku.edu, Ph.D (Expected)
- JiWoo Kim Belouadi, wldn9705@skku.edu, Master, 2026
- Yeonji Lee, yeonjilee@g.skku.edu, Master, 2025
- Jaehyeok Lee, hjl8708@g.skku.edu, Master, 2025
- HyunJin Kim, khyunjin1993@g.skku.edu, Ph.D (Expected)
- SooYung Choi, swimchoi@g.skku.edu, Master, 2025
- DongJun Kang, ehdwns2356@g.skku.edu, Master, 2024
- Taekhyun Kim, treecko.kth@g.skku.edu, Master, 2026
- HoKeun Yoon, hkyoon95@g.skku.edu, Master, 2024
- Youngjin Jo, yjspecial.jo@g.skku.edu, Master, 2023
- YeongJun Hwang, hmtyj2@g.skku.edu, Master, 2023
- DongJin Jeong, jdjin3000@g.skku.edu, Master, 2023
- HyunJin Kim, khyunjin1993@g.skku.edu, Master, 2023

Graduate Committee Member

- Jieun Woo, wjieun@g.skku.edu, Master, 2026
- Jiho Jang, zyoa@g.skku.edu, Master, 2026
- Dongjun Lim, djlim7@skku.edu, Master, 2026
- Hongjun Jeong, zun8861@g.skku.edu, Master, 2026
- Yoorhim Cho, yourmejo@skku.edu, Master, 2026
- Taehwan Kim, dmsdl5030@g.skku.edu, Master, 2026
- Minwoo Kang, skydnk4332@g.skku.edu, Master, 2026
- Dr. So-Eon Kim, sekim0211@khu.ac.kr, Ph.D., 2026
- Dr. WonJun Moon, wjun0830@gmail.com, Ph.D., 2026
- Dr. Eunseong Choi, eunseong@g.skku.edu, Ph.D., 2026
- Dr. Min Seok Choi, minseok.choi@kaist.ac.kr, Ph.D., 2026
- Dr. Bonggeun Choi, bonggeun.choi818@gmail.com, Ph.D., 2026
- Yumin Heo, ymheo1123@gmail.com, Master, 2026
- Sangwon Youn, mikeyoun2000@gmail.com, Master, 2026
- Minseok Kim, for8821@g.skku.edu, Master, 2026
- Jihyung Lee, jjklle@skku.edu, Master, 2025
- Jeong Woo Na, wjddn7946@g.skku.edu, Master, 2025
- Youngbin Kim, andyk3603@g.skku.edu, Master, 2025
- Yejin Do, dyj001213@gmail.com, Master, 2025
- Seokhyun Gong, kongsh95@skku.edu, Master, 2025
- Taewook Wi, dnlxodnr@skku.edu, Master, 2025
- Seongwan Park, waniboyy@gmail.com, Master, 2025
- Hyeonsu Cho, tacit3233@naver.com, Master, 2025
- Dr. Hoseung Kim, tree901024@skku.edu, Ph.D., 2025
- Dongjun Lim, flamecracker1220@gmail.com, Master, 2025
- Taewoo Yoo, woo990307@naver.com, Master, 2025
- Younjeong Lee, ioioiiopop@g.skku.edu, Master, 2025
- SungJoon Hwang, dbw02187@g.skku.edu, Master, 2025
- Seungmin Shin, seungminshin00@gmail.com, Master, 2025
- Suyoung Min, sujae9704@gmail.com, Master, 2025
- Dr. ChaeHun Park, ddehun@kaist.ac.kr, Ph.D., 2025
- Junehyung Kim, kalpa093@g.skku.edu, Master, 2025
- Jimin An, als398@skku.edu, Master, 2024
- Jiwon Jeong, jwjw9603@g.skku.edu, Master, 2024
- JunKoo Lee, dlwnsrn0727@g.skku.edu, Master, 2024
- HyunSung Kim, khs787621@gmail.com, Master, 2024
- Mingeun Kim, mgkim28@g.skku.edu, Master, 2024
- Junghun Kim, kjh9503@g.skku.edu, Master, 2024
- Hyo Jun Kim, rlagywns0213@skku.edu, Master, 2024
- Dong Yeop Han, hdy0159@naver.com, Master, 2024
- Yang Min Yeol, yangget@g.skku.edu, Master, 2024
- Kyuri Choi, gguriskku@gmail.com, Master, 2024
- Dr. YunSeok Choi, chldbstjr93@gmail.com, Ph.D., 2024

- Sookyung Kim, tnrud929@g.skku.edu, Master, 2023
- Jiwoo Kim, jindog1210@g.skku.edu, Master, 2023
- Olfa Jerbi, vania.848@g.skku.edu, Master, 2023
- Vania Miriam Ortiz Ramos, olfa19@g.skku.edu, Master, 2023
- Donghoon Jang, shings47@naver.com, Master, 2023
- Dr. Dai Quoc Tran, trandaign@gmail.com, Ph.D., 2023
- Dr. Byoungjoon Yu, mysinmu123@naver.com, Ph.D., 2023
- Dr. Minsoo Park, pmskku@naver.com, Ph.D., 2023
- Seoljun Go, gohseol@gmail.com, Master, 2023
- Haeun Yu, haeun.yu204@gmail.com, Master, 2023
- Sunkyung Lee, leesk1027@gmail.com, Master, 2023
- Eunchong Kim, prokkec@naver.com, Master, 2023
- HyunJu Kim, julia1028@skku.edu, Master, 2022
- Dr. Hwanhee Lee, wanted1007@snu.ac.kr, Ph.D., 2022
- Jong Hyeok Park, realkaya@g.skku.edu, Master, 2022
- Jiwung Hyun, kabbi159@gmail.com, Master, 2022
- Sanghee Park, parksangheeeee@naver.com, Master, 2022
- Hyeonjong Ha, wnajrqkq94@g.skku.edu, Master, 2022
- Dr. Byeongchang Kim, byeongchang.kim@gmail.com, Ph.D., 2022
- Dr. DongHyun Choi, cdh4696@gmail.com, Ph.D., 2022
- Donghyun Kim, donghyun.kim@g.skku.edu, Master, 2022
- Sangwoo Han, uoo723@g.skku.edu, Master, 2022
- Jeon Hyun-Kyu, hkjeon13@g.skku.edu, Master, 2022
- Min Su Jeong, omicro03@gmail.com, Master, 2022
- Hoseung Kim, tree901024@skku.edu, Master, 2022
- Samuel Kim, lemonl7@naver.com, Master, 2022
- Youngrok Song, id2thomas@gmail.com, Master, 2022
- Dr. Suin Kim, suin.kim@kaist.ac.kr, Ph.D., 2021
- Ida Ayu Putu Ari Crisdayanti, dayu.crish@gmail.com, Master, 2021
- Jaewoo Choi, yhjgoldhair@naver.com, Master, 2021
- Jinho Lee, most323@naver.com, Master, 2021
- Jung Hoon Lee, vhrehfdl@gmail.com, Master, 2021
- Won Kyu Lee, stbaker517@g.skku.edu, Master, 2021

Programming Skills

Programming languages that I've learned and used:
Bash, BASIC, Batch file, C, C++, C#, Go, Java, JavaScript, Maple, MATLAB, Perl, PHP, Prolog, Python, R, Rust, Scheme, SQL, TeX

GitHub Repositories: `https://github.com/NoSyu`

Extracurricular Activities

Advisory Committee Member, Suicide-Inducing Information Prevention Council, Ministry of Health and Welfare, 2025 - 2027

Panel, Inaugural Global AI Frontier Lab Workshop, Seoul, South Korea, 2025

Organizer, Korea & MSRA Societal AI Workshop, SKKU, 2025

Session Chair, 1st International NLP Workshop @ KAIST, KAIST, 2024

Organizer, Optimistic, Pessimistic and Realistic of Large Language Models, KOFST, 2023

Organizer, 2023 AI Doctoral Consortium, AIGS, 2023

Organizer, State, Limitations, and Future of Large Language Models, KOFST, 2022

Advisory Committee Member, Korea Tourism Organization, 2021

Students Representative of the Department of Computer Science in Student Council, KAIST, 2014

Social Service Personnel, Banyeo Library, Busan, South Korea, May 2005 - July 2007

Member and Representative, Linux & Open-source Learned Club (SKKULUG), SKKU, 2004 - 2010

References

Prof. Alice Oh, KAIST, alice.oh@kaist.edu

Mr. Byung Il Woo, SK C&C, outerlight@gmail.com

Dr. Chin-Yew Lin, Microsoft Research, cyl@microsoft.com

Prof. Eunseok Lee, SKKU, leees@skku.edu

Dr. Jonggun Lee, Forest Jalan, jonggunlee@gmail.com

Prof. Sang Gu Lee, SKKU, sglee@skku.edu