

Email: dg.lee@postech.ac.kr Homepage: https://donggeon.github.io Google Scholar: /DongGeon Lee		
RESEARCH INTERESTS	Data-centric natural language processing (NLP). Building trustworthy and safe Large Language Models (LLMs) — safety & security oversight of language models, including safety evaluations, red teaming, guardrails.	
EDUCATION	M.S. student in Artificial Intelligence <i>Pohang University of Science and Technology (POSTECH)</i>	Feb 2024 - Present <i>Pohang, South Korea</i>
	B.S. in Information and Communication Engineering <i>Inha University</i>	Mar 2018 - Feb 2024 <i>Incheon, South Korea</i>
RESEARCH EXPERIENCES	Graduate Research Assistant <i>Data Intelligence Lab, POSTECH (Advisor: Prof. Hwanjo Yu)</i>	Feb 2024 - Present <i>Pohang, South Korea</i>
	<ul style="list-style-type: none">• Research on Vision-Language Model safety benchmarks and evaluation methodologies.• Research on data-centric approaches to improving the mathematical reasoning skills of low-resource LLMs.	
	Research Scientist (Freelance) <i>AIM Intelligence</i>	Jul 2025 - Present <i>Seoul, South Korea</i>
	<ul style="list-style-type: none">• Research on safety guardrails and robustness evaluations for multi-modal LLMs.	
	Research Intern <i>KT Corporation</i>	Jan 2025 - Feb 2025 <i>Seoul, South Korea</i>
	<ul style="list-style-type: none">• Research on mathematical data synthesis for pre-training Korea-centric LLM.	
	Undergraduate Research Assistant <i>Data Intelligence Lab, Inha University (Advisor: Prof. Wonik Choi)</i>	Nov 2022 - Nov 2023 <i>Incheon, South Korea</i>
	<ul style="list-style-type: none">• Research on post-training of Language Models (LMs) for domain adaptation.• Research on keyphrase extraction from aviation incident reports via fine-tuning LMs.	
	Undergraduate Research Assistant <i>Nursing Informatics Lab, Inha University (Advisor: Prof. Insook Cho)</i>	Jul 2021 - Jun 2023 <i>Incheon, South Korea</i>
PUBLICATIONS	International Publications	
	[14] Are Vision-Language Models Safe in the Wild? A Meme-Based Benchmark Study <u>DongGeon Lee*</u> , Joonwon Jang*, Jihae Jeong, Hwanjo Yu EMNLP'25 The 2025 Conference on Empirical Methods in Natural Language Processing	
	[13] Everyday Physics in Korean Contexts: A Culturally Grounded Physical Reasoning Benchmark Jihae Jeong*, DaeYeop Lee*, <u>DongGeon Lee</u> , Hwanjo Yu MRL @ EMNLP'25 5th Multilingual Representation Learning Workshop (Co-located with the 2025 Conference on Empirical Methods in Natural Language Processing)	
	[12] When Good Sounds Go Adversarial: Jailbreaking Audio-Language Models with Benign Inputs Bodam Kim*, Hiskias Dingeto*, Taeyoun Kwon*, Dasol Choi, <u>DongGeon Lee</u> , Haon Park, Jae-Hoon Lee, Jongho Shin arXiv Preprint , 2025.08	

- [11] Typed-RAG: Type-Aware Decomposition of Non-Factoid Questions for Retrieval-Augmented Generation
 DongGeon Lee*, Ahjeong Park*, Hyeri Lee, Hyeonseo Nam, Yunho Maeng
[XLLM @ ACL'25](#) | The First Workshop on Structure-aware Large Language Models (Co-located with the 63rd Annual Meeting of the Association for Computational Linguistics)
[NAACL'25 SRW](#) (Non-Archival) | 2025 Annual Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Student Research Workshop
- [10] REFIND at SemEval-2025 Task 3: Retrieval-Augmented Factuality Hallucination Detection in Large Language Models
 DongGeon Lee, Hwanjo Yu
[SemEval @ ACL'25](#) | The 19th International Workshop on Semantic Evaluation (Co-located with the 63rd Annual Meeting of the Association for Computational Linguistics)
- [9] Theme-Explanation Structure for Table Summarization using Large Language Models: A Case Study on Korean Tabular Data
 TaeYoon Kwack*, Jisoo Kim*, Ki Yong Jung, DongGeon Lee, Heesun Park
[TRL @ ACL'25](#) | The 4th Table Representation Learning Workshop (Co-located with the 63rd Annual Meeting of the Association for Computational Linguistics)
- [8] Enhancing Adverse Event Reporting With Clinical Language Models: Inpatient Falls
 Insook Cho, Hyunchul Park, Byeong Sun Park, DongGeon Lee
[Journal of Advanced Nursing](#) (SCIE; Q1), 2025.02
- [7] Effects of Language Differences on Inpatient Fall Detection Using Deep Learning
 Insook Cho, EunJu Lee, DongGeon Lee
[MedInfo'23](#) | The 19th World Congress on Medical and Health Informatics
- [6] Bridging the Reporting Gap of Inpatient Falls to Improve Safety Practices Using Deep-Learning-Based Language Models and Multisite Data
 DongGeon Lee, EunJu Lee, Insook Cho
[AMIA CIC'23](#) | AMIA 2023 Clinical Informatics Conference

Domestic Publications (written in *Korean*)

- [5] Synthesizing a Korean-centric Math Corpus to Enhance Math Problem-Solving Ability in Korean Large Language Models ([Excellent Paper Award](#))
 DongGeon Lee, Jisoo Baik, Wonjae Park, Hwanjo Yu
[HCLT'25](#) | The 37th Annual Conference on Human & Cognitive Language Technology
- [4] Designing Synthetic Data and Training Strategies for Multi-hop Retrieval-Augmented Generation
 Kyumin Lee, Minjin Jeon, Sanghwan Jang, DongGeon Lee, Hwanjo Yu
[KCC'25](#) | Korea Computer Congress 2025
- [3] Question Types Matter: An Analysis of Question-Answering Performance in Retrieval-Augmented Generation Across Diverse Question Types
 DongGeon Lee*, Ahjeong Park*, Hyeri Lee, Hyeonseo Nam, Yunho Maeng
[HCLT'24](#) | The 36th Annual Conference on Human & Cognitive Language Technology
- [2] Tabular-TX: Theme-Explanation Structure-based Table Summarization via In-Context Learning ([Excellent Paper Award](#))
 TaeYoon Kwack*, Jisoo Kim*, Ki Yong Jung, DongGeon Lee, Heesun Park
[HCLT'24](#) | The 36th Annual Conference on Human & Cognitive Language Technology

- [1] Through deep learning-based video processing, Design and implementation of Smart Port Parking Information System
 Changhun Koo*, Yoonjoo Jung*, DongGeon Lee*
[ACK'21](#) | Annual Conference of Korea Information Processing Society 2021

ACADEMIC SERVICES

Reviewer of AAAI'25 (The Association for the Advancement of Artificial Intelligence)	2025
Reviewer of International Journal of Nursing Studies	2025
Reviewer of MELT (Workshop on Multilingual and Equitable Language Technologies) at COLM'25	2025
Reviewer of SemEval (International Workshop on Semantic Evaluation) at ACL'25	2025
Student Volunteer of ACL'25 (Annual Meeting of the Association for Computational Linguistics)	2025

HONORS AND AWARDS

Excellent Paper Award <i>HCLT 2025 (The 37th Annual Conference on Human & Cognitive Language Technology)</i>	2025
NAACL 2025 Registration Grant <i>NAACL 2025 SRW (Student Research Workshop)</i>	2025
Gold Prize (Director's Award of the NIKL) <i>Korean AI Language Proficiency Challenge, NIKL (National Institute of Korean Language)</i>	2024
Excellent Paper Award <i>HCLT 2024 (The 36th Annual Conference on Human & Cognitive Language Technology)</i>	2024
Scholarship for Outstanding Graduate Students <i>POSTECH</i>	2024
Top Engineering Student Award <i>Inha University</i>	2024
Research Scholarship for Undergraduate Researchers <i>Inha University</i>	2023

TECHNICAL SKILLS

Professional working proficiency Python, PyTorch, transformers, vLLM, Git
Limited working proficiency Shell Script, Keras, \LaTeX
Elementary proficiency DeepSpeed, TensorFlow, C++, C, MySQL