

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	<div><div>M.S. student in Artificial Intelligence <i>Pohang University of Science and Technology (POSTECH)</i></div><div>Feb 2024 - Present <i>Pohang, South Korea</i></div></div> <div><div>B.S. in Information and Communication Engineering <i>Inha University</i></div><div>Mar 2018 - Feb 2024 <i>Incheon, South Korea</i></div></div>
RESEARCH EXPERIENCES	<div><div>Graduate Research Assistant <i>Data Intelligence Lab, POSTECH (Advisor: Prof. Hwanjo Yu)</i></div><div>Feb 2024 - Present <i>Pohang, South Korea</i></div><div><ul style="list-style-type: none">Research on Vision-Language Model safety benchmarks and evaluation methodologies.Research on knowledge conflicts of LLMs between external and internal knowledge.</div></div> <div><div>Research Intern <i>AIM Intelligence</i></div><div>Jul 2025 - Present <i>Seoul, South Korea</i></div><div><ul style="list-style-type: none">Research on safety of Vision-Language-Action Models.</div></div> <div><div>Research Intern <i>KT Corporation</i></div><div>Jan 2025 - Feb 2025 <i>Seoul, South Korea</i></div><div><ul style="list-style-type: none">Research on mathematical data synthesis for pre-training Korea-centric LLM, Mi:dm 2.0.</div></div> <div><div>Undergraduate Research Assistant <i>Data Intelligence Lab, Inha University (Advisor: Prof. Wonik Choi)</i></div><div>Nov 2022 - Nov 2023 <i>Incheon, South Korea</i></div><div><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.</div></div> <div><div>Undergraduate Research Assistant <i>Nursing Informatics Lab, Inha University (Advisor: Prof. Insook Cho)</i></div><div>Jul 2021 - Jun 2023 <i>Incheon, South Korea</i></div><div><ul style="list-style-type: none">Research on detecting fall events in clinical notes via fine-tuning LMs.</div></div>
PUBLICATIONS	<div>International Publications</div> <div><div>[11] Are Vision-Language Models Safe in the Wild? A Meme-Based Benchmark Study DongGeon Lee*, Joonwon Jang*, Jihae Jeong, Hwanjo Yu arXiv Preprint, 2025.05</div><div>[10] 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) Annual Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Student Research Workshop</div><div>[9] 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)</div></div>

- [8] 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)
- [7] 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
- [6] Effects of Language Differences on Inpatient Fall Detection Using Deep Learning
Insook Cho, EunJu Lee, DongGeon Lee
[MedInfo'23](#) | World Congress on Medical and Health Informatics
- [5] 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 Clinical Informatics Conference

Domestic Publications (written in Korean)

- [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
- [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](#) | 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](#) | 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

HONORS AND AWARDS

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

ACADEMIC SERVICES	Reviewer of MELT (Workshop on Multilingual and Equitable Language Technologies) at COLM'25	2025
	Student Volunteer of ACL'25 (Annual Meeting of the Association for Computational Linguistics)	2025
	Secondary Reviewer of ACL ARR (ACL Rolling Review) February	2025
	Reviewer of SemEval (International Workshop on Semantic Evaluation) at ACL'25	2025
TECHNICAL SKILLS	Professional working proficiency	
	Python, PyTorch, transformers, vLLM, Git	
	Limited working proficiency	
	Shell Script, Keras, \LaTeX	
	Elementary proficiency	
	DeepSpeed, TensorFlow, C++, C, MySQL	