

Hwanhee Lee (이환희)

ASSISTANT PROFESSOR @ CHUNG-ANG UNIVERSITY

+82-10-2621-9532 | hwanheelee@cau.ac.kr | hwanheelee1993.github.io | hwanheelee1993

Employment

Chung-Ang University

ASSISTANT PROFESSOR

- Department of Artificial Intelligence/Graduate School of Artificial Intelligence (Language Intelligence Lab.)

Seoul, Korea

Mar. 2023 - Current

Seoul National University

SENIOR RESEARCHER

- Postdoctoral Researcher, Automation and Systems Research Institute

Seoul, Korea

Sep. 2022 - Feb. 2023

NAVER CLOVA AI Research

RESEARCH INTERN

- NAVER AI LAB Language Research Team
- Working with Dr. Hwaran Lee, Dr. Kang Min Yoo, and Prof. Joonsuk Park
- Theme: Factual Consistency Evaluation and Optimization for Abstractive Summarization
- Published a paper at NAACL 2022

Seongnam, Korea

Aug. 2021 - Jan. 2022

Research Interests Natural Language Processing - Truthfulness and Safety of Language Models

Current Focus Safety and Policy Alignment in Language Agents, Context Engineering and Agentic Memory, Multi-Agent Systems

Education

Seoul National University

PH.D OF ELECTRICAL AND COMPUTER ENGINEERING

Seoul, Korea

Aug. 2022

- Machine Intelligence Laboratory (Advisor: Prof. Kyomin Jung)
- Thesis: Factual Consistency Evaluation for Conditional Text Generation Systems
(Committee: Kyuseok Shim, Kyomin Jung, Taesup Moon, Seung-won Hwang (SNU), and JinYeong Bak (SKKU))
- Distinguished Ph.D. Dissertation Award

Seoul National University

BACHELOR OF ELECTRICAL AND COMPUTER ENGINEERING

Seoul, Korea

Feb. 2017

Scholarships & Awards

Outstanding Area Chair

AWARDED BY EMNLP 2024

Nov. 2024

Adobe Research Gift

GRANTED BY ADOBE RESEARCH (\$28,000)

May. 2024, June. 2023

Distinguished Ph.D. Dissertation Award

AWARDED BY DEPT. OF ELECTRICAL AND COMPUTER ENGINEERING, SEOUL NATIONAL UNIVERSITY

Aug. 2022

Youlchon AI STAR Scholarship

GRANTED BY NONGSIM YOULCHON FOUNDATION

Sep. 2021

Scholarship of Electronics, Electrical and Computer Engineering

GRANTED BY FOUNDATION FOR ELECTRONICS, ELECTRICAL AND COMPUTER ENGINEERING AT SEOUL NATIONAL UNIVERSITY

Mar. 2018

National Science & Technology Scholarship (Full-tuition)

GRANTED BY KOREA STUDENT AID FOUNDATION

Mar. 2013 - Feb. 2017

Publications

International Conferences

Keep Security! Benchmarking Security Policy Preservation in Large Language Model Contexts Against Indirect Attacks in Question Answering

HWAN CHANG*, YUMIN KIM*, YONGHYEON JUN, HWANHEE LEE†

- The 2025 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2025**)

ToDi: Token-wise Distillation via Fine-Grained Divergence Control

SEONGRYONG JUNG, SUWAN YOON, DONGGEON KIM, HWANHEE LEE†

- The 2025 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2025**)

How Do Large Vision-Language Models See Text in Image? Unveiling the Distinctive Role of OCR Heads

INGEOL BAEK, HWAN CHANG, SUNGHYUN RYU, HWANHEE LEE†

- The 2025 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2025**)

SAFE-SQL: Self-Augmented In-Context Learning with Fine-grained Example Selection for Text-to-SQL

JIMIN LEE, INGEOL BAEK, BYEONGJEONG KIM, HYUNKYUNG BAE, HWANHEE LEE†

- The 2025 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2025**)

Event-Driven Storytelling with Multiple Lifelike Humans in a 3D scene

DONGGEUN LIM, JINSEOK BAE, INWOOW Hwang, SEUNGMIN LEE, HWANHEE LEE, YOUNG MIN KIM

- The 2025 International Conference on Computer Vision (**ICCV 2025**)

Exploring Persona Sentiment Sensitivity in Personalized Dialogue Generation

YONGHYUN JUN, HWANHEE LEE†

- The 63rd Annual Meeting of the Association for Computational Linguistics (**ACL 2025**)

Dynamic Label Name Refinement for Few-Shot Dialogue Intent Classification

GYUTAE PARK, INGEOL BAEK, BYEONGJEONG KIM, JOONGBOK SHIN, HWANHEE LEE†

- The 63rd Annual Meeting of the Association for Computational Linguistics (**ACL 2025**)

Dynamic Order Template Prediction for Generative Aspect-Based Sentiment Analysis

YONGHYUN JUN, HWANHEE LEE†

- The 63rd Annual Meeting of the Association for Computational Linguistics (**ACL 2025**)

Which Retain Set Matters for LLM Unlearning? A Case Study on Entity Unlearning

HWAN CHANG, HWANHEE LEE†

- Findings of the Association for Computational Linguistics: ACL 2025 (**Findings of ACL 2025**)

Investigating Language Preference of Multilingual RAG systems

JEONGHYUN PARK, HWANHEE LEE†

- Findings of the Association for Computational Linguistics: ACL 2025 (**Findings of ACL 2025**)

AdvisorQA: Towards Helpful and Harmless Advice-seeking Question Answering with Collective Intelligence

MINBEOM KIM, HWANHEE LEE, JOONSUK PARK, HWARAN LEE, KYOMIN JUNG

- 2025 Annual Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL 2025**)

Probing-RAG: Self-Probing to Guide Language Models in Selective Document Retrieval

INGEOL BAEK, HWAN CHANG, BYEONGJEONG KIM, JIMIN LEE, HWANHEE LEE†

- Findings of the Association for Computational Linguistics: NAACL 2025 (**Findings of NAACL 2025**)

FIZZ: Factual Inconsistency Detection by Zoom-in Summary and Zoom-out Document

JOONHO YANG, SEUNGHYUN YOON, BYEONGJEONG KIM, HWANHEE LEE†

- Conference on Empirical Methods in Natural Language Processing (**EMNLP 2024**)

IterCQR: Iterative Conversational Query Reformulation with Retrieval Guidance

YUNAH JANG, KANG-IL LEE, HYUNKYUNG BAE, HWANHEE LEE†, KYOMIN JUNG†

- 2024 Annual Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL 2024**)

LifeTox: Unveiling Implicit Toxicity in Life Advice

MINBEOM KIM, JAHYUN KOO, **HWANHEE LEE**, JOONSUK PARK†, HWARAN LEE, KYOMIN JUNG†

- 2024 Annual Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL 2024**)

KoCoSa: Korean Context-aware Sarcasm Detection Dataset

YUMIN KIM*, HEEJAE SUH*, MINGI KIM*, DONGYEON WON*, **HWANHEE LEE**†

- The 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (**COLING 2024**)

Kosmic: Korean Text Similarity Metric Reflecting Honorific Distinctions

YERIN HWANG, YONGIL KIM, HYUNKYUNG BAE, JEESOO BANG, **HWANHEE LEE**†, KYOMIN JUNG†

- The 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (**COLING 2024**)

Asking Clarification Questions to Handle Ambiguity in Open-Domain QA

DONGRYEOL LEE, SEGWANG KIM, MINWOO LEE, **HWANHEE LEE**, JOONSUK PARK, SANG-WOO LEE, KYOMIN JUNG

- Findings of the Association for Computational Linguistics: EMNLP 2023 (**Findings of EMNLP 2023**)

Dialogizer: Context-aware Conversational-QA Dataset Generation from Textual Sources

YERIN HWANG*, YONGIL KIM*, HYUNKYUNG BAE, **HWANHEE LEE**†, JEESOO BANG, KYOMIN JUNG†

- The 2023 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2023**)

Critic-Guided Decoding for Controlled Text Generation

MINBEOM KIM, **HWANHEE LEE**, KANG MIN YOO, JOONSUK PARK, HWARAN LEE, KYOMIN JUNG

- Findings of the Association for Computational Linguistics: ACL 2023 (**Findings of ACL 2023**)

Task-specific Compression for Multi-task Language Models using Attribution-based Pruning

NAKYEONG YANG, YUNAH JANG, **HWANHEE LEE**, SEOHYEONG JUNG, AND KYOMIN JUNG

- Findings of the Association for Computational Linguistics: EACL 2023 (**Findings of EACL 2023**)

Masked Summarization to Generate Factually Inconsistent Summaries for Improved Factual Consistency Checking

HWANHEE LEE, KANG MIN YOO, JOONSUK PARK, HWARAN LEE, AND KYOMIN JUNG

- Findings of the Association for Computational Linguistics: NAACL 2022 (**Findings of NAACL 2022**)

QACE: Asking Questions to Evaluate an Image Caption

HWANHEE LEE, THOMAS SCIALOM, SEUNGHYUN YOON, FRANK DERNONCOURT, AND KYOMIN JUNG

- Findings of the Association for Computational Linguistics: EMNLP 2021 (**Findings of EMNLP 2021**)

CrossAug: A Contrastive Data Augmentation Method for Debiasing Fact Verification Models

MINWOO LEE, SEUNGIL WON, JUAE KIM, **HWANHEE LEE**, CHEONUM PARK, AND KYOMIN JUNG

- 30th ACM International Conference on Information and Knowledge Management (**CIKM 2021**)

UMIC: An Unreferenced Metric for Image Captioning via Contrastive Learning

HWANHEE LEE, SEUNGHYUN YOON, FRANK DERNONCOURT, TRUNG BUI, AND KYOMIN JUNG

- The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (**ACL 2021**)

Entangled Bidirectional Encoder to Autoregressive Decoder for Sequential Recommendation

TAEGWAN KANG, **HWANHEE LEE**, BYEONGJIN CHOE, AND KYOMIN JUNG

- The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (**SIGIR 2021**)

KPQA: A Metric for Generative Question Answering Using Keyphrase Weights

HWANHEE LEE, SEUNGHYUN YOON, FRANK DERNONCOURT, DOO SOON KIM, TRUNG BUI, JOONGBO SHIN, AND KYOMIN JUNG

- 2021 Annual Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL 2021**)

Attentive Modality Hopping Mechanism for Speech Emotion Recognition

SEUNGHYUN YOON, SUBHADEEP DEY, **HWANHEE LEE**, AND KYOMIN JUNG

- IEEE International Conference on Acoustics, Speech and Signal Processing (**ICASSP 2020**)

Improving Neural Question Generation Using Answer Separation

YANGHOON KIM, **HWANHEE LEE**, JOONGBO SHIN, AND KYOMIN JUNG

- Thirty-Third AAAI Conference on Artificial Intelligence (**AAAI 2019**)

International Journal

Crafting the Path: Robust Query Rewriting for Information Retrieval

INGEOL BAEK, JIMIN LEE, JOONHO YANG, HWANHEE LEE†

- IEEE Access, 2025

Cross-Lingual Summarization for Low-Resource Languages using Multilingual Retrieval-Based In-Context Learning

GYETA PARK, JEONGHYUN PARK, HWANHEE LEE†

- Applied Sciences, 2025

REMDoc: Reference-free Evaluation for Medical Document Summaries via Contrastive Learning

JIMIN LEE, INGEOL BAEK, HWANHEE LEE†

- IEEE Access, 2024

Improving Abstractive Dialogue Summarization Using Keyword Extraction

CHONGJAE YOO AND HWANHEE LEE†

- Applied Sciences, 2023

Gated Relational Target-aware Encoder and Local Context-aware Decoder for Target-oriented Opinion Words Extraction

TAEGWAN KANG, SEGWANG KIM, HYEONGU YUN, HWANHEE LEE, AND KYOMIN JUNG

- IEEE Access, 2022

Preprint/Under Review

ChatInject: Abusing Chat Templates for Prompt Injection in LLM Agents

HWAN CHANG*, YONGHYUN JUN*, HWANHEE LEE†

- arXiv (2509.22830)

MIRAGE: Multi-hop Reasoning with Ambiguity Evaluation for Illusory Questions

JEONGHYUN PARK*, INGEOL BAEK*, SEUNGHYUN YOON, HAEUN JANG, APARNA GARIMELLA, AKRITI JAIN, NEDIM LIPKA, HWANHEE LEE†

- arXiv (2509.22750)

Chronological Passage Assembling in RAG framework for Narrative Text

BYEONGJEONG KIM, JEONGHYUN PARK, JOONHO YANG, HWANHEE LEE†

- arXiv (2508.18748)

Hallucinate at the Last in Long Response Generation: A Case Study on Long Document Summarization

JOONHO YANG, SEUNGHYUN YOON, HWAN CHANG, BYEONGJEONG KIM, HWANHEE LEE†

- arXiv (2505.15291)

Personality Editing for Language Models through Adjusting Self-Referential Queries

SEOJIN HWANG, YUMIN KIM, BYEONGJEONG KIM, HWANHEE LEE†

- arXiv (2502.11789)

MultiLexNorm++: Lexical Normalization Beyond Indo-European Languages

WEERAYUT BUAPHET, THANH-NHI NGUYEN, RISA KONDO, TOMOYUKI KAJIWARA, YUMIN KIM, JIMIN LEE, HWANHEE LEE, HOLY LOVENIA, PEERAT LIMKONCHOTIWAT, SARANA NUTANONG, ROB VAN DER GOOT

- Under-review

Selective Demonstration Retrieval for Improved Implicit Hate Speech Detection

YUMIN KIM, HWANHEE LEE†

- arXiv (2504.12082)

Leveraging Fine-grained Error Correction in Korean Speech Recognition for Consultation Services

YONGHYUN JUN, JIMIN LEE, HWAN CHANG, DONGHO SHIN, SEOLAH KIM, HWANHEE LEE†

- Under Review

Conversational Query Reformulation with the Guidance of Retrieved Documents

JEONGHYUN PARK, HWANHEE LEE†

- Under Review - arXiv (2407.12363)

Workshop Papers & Other International Conferences

Logistic Regression-Based Example Selection for Enhanced Few-Shot Learning in Intent Classification

GYUTAE PARK, HWANHEE LEE†

- IEEE International Conference on Consumer Electronics (IEEE ICCE 2025)

Dynamically Managing a Prompt Pool via Self-Enhancement in Continual Learning

HAYUN LEE, KISEONG HONG, HWANHEE LEE, SUNGHO SUH, EUNWOO KIM

- NeurIPS 2024 Workshop on Adaptive Foundation Models (NeurIPS-W 2024)

Low-Resource Cross-Lingual Summarization through Few-Shot Learning with Large Language Models

GYUTAE PARK, SEOJIN HWANG, HWANHEE LEE†

- ACL Workshop on Technologies for Machine Translation of Low-Resource Languages (ACL-W 2024)

Factual Error Correction for Abstractive Summaries Using Entity Retrieval

HWANHEE LEE, CHEONEUM PARK, SEUNGHYUN YOON, TRUNG BUI, FRANCK DERNONCOURT, JUAE KIM AND KYOMIN JUNG

- EMNLP Workshop on Generation, Evaluation & Metrics (EMNLP-W 2022)

Improving Multiple Documents Grounded Goal-Oriented Dialog Systems via Diverse Knowledge Enhanced Pre-trained Language Model

YUNAH JANG, DONGRYEOL LEE, HYUNG JOO PARK, TAEGWAN KANG, HWANHEE LEE, HYUNKYUNG BAE, AND KYOMIN JUNG

- ACL Workshop on Document-grounded Dialogue and Conversational Question Answering (ACL-W 2022)

Learning to Select Question-Relevant Relations for Visual Question Answering

JAEWOONG LEE, HEEJOON LEE, HWANHEE LEE, AND KYOMIN JUNG

- NAACL Workshop on Multimodal Artificial Intelligence Workshop (NAACL-W 2021)

Relevance Similarity Scorer and Entity Guided Reranking for Knowledge Grounded Dialog System

HYUNKYUNG BAE, MINWOO LEE, AHHYEON KIM, AND HWANHEE LEE, CHEONGJAE LEE, CHEONUM PARK, DONGHYEON KIM, AND KYOMIN JUNG

- AAAI Workshop on Dialog System Technology Challenge (AAAI-W 2021)

ViLBERTScore: Evaluating Image Caption Using Vision-and-Language BERT

HWANHEE LEE, SEUNGHYUN YOON, FRANK DERNONCOURT, DOO SOON KIM, TRUNG BUI, AND KYOMIN JUNG

- EMNLP Workshop on Evaluation and Comparison of NLP Systems (EMNLP-W 2020)

Multimodal Semantic Transformer Network with Retrieval Style Word Generator

HWANHEE LEE, SEUNGHYUN YOON, FRANK DERNONCOURT, DOO SOON KIM, TRUNG BUI, AND KYOMIN JUNG

- AAAI Workshop on Dialog System Technology Challenge (AAAI-W 2020)

Attention-based Convolutional Neural Networks for Multi-label Emotion Classification

YANGHOON KIM, HWANHEE LEE, AND KYOMIN JUNG

- NAACL Workshop on Semantic Evaluation 2018 (NAACL-W 2018)

Research Projects

Plug-in External-Memory Models for Augmenting Knowledge of Large Language Models

FUNDED BY THE NATIONAL RESEARCH FOUNDATION OF KOREA, PI

Sep. 2025 - Aug. 2026

- Research on Next-generation RAG via Plug-in External Memory

Next-Generation O2O Fashion Retail System: Enhancing Sales Through Integration of Personal Shoppers and Immersive Virtual Fitting

FUNDED BY THE MINISTRY OF TRADE, INDUSTRY AND ENERGY, CO-PI

July. 2025 - Dec. 2028

- Developing Conversational Fashion AI Agent

Development of an LLM System for Financial Transactions

FUNDED BY SEOUL BUSINESS AGENCY, Co-PI

Sep. 2025 - Aug. 2026

- Developing Transaction Classification System with RAG

LLM based Claim Report Generation

FUNDED BY DOOSAN ENERBILITY, PI

Jun. 2025 - May. 2026

- Claim Report Generation with Agentic RAG System

Development of a LLM System for Financial Transaction Categorization to Enhance Risk and Customer Management in Financial Institutions

FUNDED BY THE MINISTRY OF SMEs AND STARTUPS, Co-PI

May. 2025 - Dec. 2025

- Developing Transaction Classification System with RAG

Risk Assessment of Construction Site based on LLMs

FUNDED BY DOOSAN ENERBILITY, PI

Jun. 2024 - May. 2025

- Risk Assessment and Prediction with Large Language Models

Improvement of Speech Recognition Rate Using LLMs

FUNDED BY SEOUL 120 DASAN CALL FOUNDATION, PI

April. 2024 - Dec. 2024

- Detecting Errors of ASR System for Dialogue Data with Large Language Models

RAG based Conversational Question Answering System

FUNDED BY URP, Co-PI

Mar. 2024 - Dec. 2024

- Developing ConvQA System for Real Estate Domain

Anomaly Detection in Fuel Cell

FUNDED BY DOOSAN FUEL CELL, PI

June. 2023 - May. 2024

- Developing Anomaly Detection System for Full Cell

Factual Error Correction on Grounded Text Generation Systems

FUNDED BY ADOBE RESEARCH (RESEARCH GIFTS, 2023-2024)

May. 2023 - Current

- Developed Factual Error Detection and Correction System for Abstractive Summarization Systems
- Published a Paper at EMNLP 2024

Academic Services

Organizing Committee

IJCNLP-AACL 2023 (DIVERSITY & INCLUSION CHAIR)

Area Chair

ACL 2024-2025, EMNLP 2024-2025, NAACL 2024-2025, COLING 2025

Reviewer

ACL 2021-2023, EMNLP 2020-2023, COLING 2022-2024, AAAI 2021-2025, NEURIPS 2023-2025

Advisor

SOFTWARE POLICY & RESEARCH INSTITUTE

Aug. 2022 - Mar. 2023

- Technical Advising on Hyperscale Language Models

Selected Invited Talks

Korea Computer Congress 2025

Jeju, Korea

FAITHFULNESS IN LANGUAGE MODELS

July. 2025

Korea Advanced Institute of Science & Technology

Daejeon, Korea

COGNITIVE AND REASONING CAPABILITIES OF HYPERSCALE AI (NEXTGENAI-X WORKSHOP)

Mar. 2025

Adobe Research

San Jose, CA, USA

TOWARD TRUTHFUL AND SAFE LANGUAGE MODELS

Jan. 2025

The 36th Annual Conference on Human & Cognitive Language Technology	<i>Seongnam, Korea</i>
TOWARD TRUTHFUL AND SAFE LANGUAGE MODELS	<i>Oct. 2024</i>
Hanyang University	<i>Seoul, Korea</i>
HALLUCINATION OF LARGE LANGUAGE MODELS	<i>March. 2024</i>
Korea Institute of Science and Technology Information	<i>Daejeon, Korea</i>
RECENT RESEARCH ON FACTUAL CONSISTENCY AND APPLICATIONS TO THE SCIENTIFIC DOMAIN	<i>January. 2024</i>
Korea Advanced Institute of Science & Technology	<i>Daejeon, Korea</i>
RESEARCH ON TEXT GENERATION SYSTEMS	<i>Oct. 2023</i>
Korea Electronics Technology Institute	<i>Seongnam, Korea</i>
LARGE LANGUAGE MODELS CHATGPT: HOW IT WORKS AND HOW TO USE	<i>July. 2023</i>
Eötvös Loránd University	<i>Budapest, Hungary</i>
RESEARCH ON TEXT GENERATION SYSTEMS	<i>July. 2023</i>
Chung-Ang University	<i>Seoul, Korea</i>
LARGE LANGUAGE MODELS CHATGPT: HOW IT WORKS AND HOW TO USE	<i>April. 2023</i>
LG AI Research	<i>Seoul, Korea</i>
FACTUAL CONSISTENCY OF TEXT GENERATION SYSTEMS	<i>Jan. 2023</i>
BrainLink 2022	<i>Incheon, Korea</i>
FACTUAL CONSISTENCY OF TEXT GENERATION SYSTEMS	<i>Oct. 2022</i>
Korea Electronics Technology Institute	<i>Seongnam, Korea</i>
FACTUAL CONSISTENCY OF TEXT GENERATION SYSTEMS	<i>Aug. 2022</i>

Teaching Experiences

Advanced Natural Language Processing: Large Language Models	<i>Chung-Ang University</i>
LECTURER	<i>Spring 2025, Spring 2024, Spring 2023</i>
Objected Oriented Programming	<i>Chung-Ang University</i>
LECTURER	<i>Spring 2025, Spring 2024</i>
Embedded Artificial Intelligence	<i>Chung-Ang University</i>
LECTURER	<i>Fall 2025, Fall 2024, Fall 2023</i>
Artificial Intelligence Design	<i>Chung-Ang University</i>
LECTURER	<i>Fall 2025, Fall 2024, Fall 2023</i>
Probability and Random Process	<i>Chung-Ang University</i>
LECTURER	<i>Spring 2023</i>
Machine Learning	<i>Seoul National University</i>
TEACHING ASSISTANT	<i>Spring 2019, Spring 2018</i>
Programming Methodology	<i>Seoul National University</i>
TEACHING ASSISTANT	<i>Spring 2022, Spring 2020, Spring 2017</i>
Seminar on Deep Learning	<i>Seoul National University</i>
TEACHING ASSISTANT	<i>Fall 2019</i>
Introduction to IoT, Artificial Intelligence, Big Data	<i>Seoul National University</i>
TEACHING ASSISTANT	<i>Fall 2019, Spring 2019, Fall 2018</i>
Electrical Engineering Research Project	<i>Seoul National University</i>
TEACHING ASSISTANT	<i>Fall 2020</i>