

Dong-Ho Lee

✉ dongho.lee@usc.edu

🏠 dongholee.com

📱 danny911kr

Employment History

Microsoft AI

Member of Technical Staff

- Copilot's Personality & Personalization

Redmond, United States
(Incoming) Aug. 2025 -

Softly AI

Founding Engineer

- Built AI agent platform for dermatology/plastic surgery clinics in Korea, automating reservations and FAQs.
- Led AI engineering team, focusing on agent framework development, evaluation workflows, and prompt debugging.

Seoul, Republic of Korea
Mar. 2022 - Apr. 2025

Google Deepmind & Google Play

Research Intern

- Mentors: Adam Kraft, Long Jin, Xinyang Yi / Collaborators: Nikhil Mehta, Taibai Xu, Lichan Hong, Ed H. Chi
- Developed a training-free recommendation system using Gemini, outperforming Google's TIGER baseline by over 20%.

Mountain View, United States
May. 2024 - Aug. 2024

Snap Inc.

Research Intern

- Mentors: Francesco Barbieri / Collaborators: Adyasha Maharana, Yuwei Fang
- Created and evaluated LoCoMo and REALTALK benchmarks for LLM memory, now widely used in LLM personalization research.

Santa Monica, United States
May. 2023 - Aug. 2023

Microsoft Research

Research Intern

- Mentors: Sujay Kumar Jauhar / Collaborators: Mohit Sewak, Ryen W. White
- Developed prompt-based data augmentation for various tasks, improving robustness in low-resource settings.

Redmond, United States
May. 2022 - Aug. 2022

Upstage AI

AI Research Engineer

Seoul, Republic of Korea
Aug. 2021 - Mar. 2022

Information Sciences Institute, USC

Research Programmer I

Los Angeles, United States
Sep. 2020 - Aug. 2021

Education

University of Southern California (USC)

Ph.D. Candidate in Computer Science

- Dissertation: Improving Language Models through Context
- Advisor: Jay Pujara
- Committee Members: Xiang Ren, Robin Jia, Fred Morstatter, Meisam Razaviyayn

Los Angeles, United States
Aug. 2021 - Aug. 2025

University of Southern California (USC)

M.S. in Computer Science

- Advisor: Xiang Ren

Los Angeles, United States
Aug. 2018 - May. 2020

Sungkyunkwan University (SKKU)

B.S. in Computer Education

Seoul, Republic of Korea
Mar. 2012 - Feb. 2018

Sheffield University

Exchange program in Computer Science

Sheffield, United Kingdom
Sep. 2016 - Feb. 2017

Services

Area Chair

ACL Rolling Review (ARR): Oct 2023 (EACL 2023), Dec 2023 (NAACL 2024), Feb 2024 (ACL 2024), Apr 2024, June 2024 (EMNLP 2024),
Aug 2024, Feb 2025 (ACL 2025), May 2025 (EMNLP 2025)

Program Committee

International Conference on Machine Learning (ICML): 22'
International Conference on Representation Learning (ICLR): 22', 23', 24', 25'
Conference on Neural Information Processing Systems (NeurIPS): 21', 22', 23', 24'
ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD): 22', 23', 24'
The Web Conference (WWW): 25'
Annual Meeting of the Association for Computational Linguistics (ACL): 23'
Conference on Empirical Methods in Natural Language Processing (EMNLP): 21', 22', 23'

Conference on Language Modeling (**COLM**): 24'
SIAM International Conference on Data Mining (**SDM**): 22'
Conference of the European Chapter of the Association for Computational Linguistics (**EACL**): 21', 23'
Language Resources and Evaluation Conference (**LREC**): 22'

External Reviewer

ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD**): 20', 21'
International World Wide Web Conference (**WWW**): 20', 21'
Annual Meeting of the Association for Computational Linguistics (**ACL**): 20'
The SIGNLL Conference on Computational Natural Language Learning (**CONLL**): 19'

Publications

* These authors contributed equally

Preprints

- [1] **STAR: A Simple Training-free Approach for Recommendations using Large Language Models**
Dong-Ho Lee*, Adam Kraft*, Long Jin, Nikhil Mehta, Taibai Xu, Lichan Hong, Ed H. Chi and Xinyang Yi ([w/ Google Deepmind](#))
- [2] **REALTALK: A 21-Day Real-World Dataset for Long-Term Conversation**
Dong-Ho Lee*, Adyasha Maharana*, Jay Pujara, Xiang Ren, Francesco Barbieri ([w/ Snap Inc.](#))
- [3] **What is a Good Question? Utility Estimation with LLM-based Simulations**
Dong-Ho Lee*, Hyundong J. Cho*, Jonathan May, Jay Pujara
- [4] **Harmful Suicide Content Detection**
16 authors including Dong-Ho Lee and Kyunghyun Cho

Research Papers in Conference Proceedings

- [5] **PLUG: Leveraging Pivot Language in Cross-Lingual Instruction Tuning**
Zhihan Zhang, Dong-Ho Lee, Yuwei Fang, Wenhao Yu, Mengzhao Jia, Meng Jiang and Francesco Barbieri ([w/ Snap Inc.](#))
[ACL 2024] Annual Meeting of the Association for Computational Linguistics (ACL), 2024.
- [6] **Evaluating Very Long-Term Conversational Memory of LLM Agents**
Adyasha Maharana, Dong-Ho Lee, Sergey Tulyakov, Mohit Bansal, Francesco Barbieri and Yuwei Fang ([w/ Snap Inc.](#))
[ACL 2024] Annual Meeting of the Association for Computational Linguistics (ACL), 2024.
- [7] **Making Large Language Models Better Data Creators**
Dong-Ho Lee, Jay Pujara, Mohit Sewak, Ryen White and Sujay Kumar Jauhar ([w/ Microsoft Research](#))
[EMNLP 2023] Conference on Empirical Methods in Natural Language Processing (EMNLP), 2023.
- [8] **Temporal Knowledge Graph Forecasting Using In-Context Learning**
Dong-Ho Lee*, Kian Ahrabian*, Woojeong Jin, Fred Morstatter and Jay Pujara
[EMNLP 2023] Conference on Empirical Methods in Natural Language Processing (EMNLP), 2023.
- [9] **Analyzing Norm Violations in Real-Time Live-Streaming Chat**
Jihyung Moon*, Dong-Ho Lee*, Hyundong J. Cho, Woojeong Jin, Chan Young Park, Minwoo Kim, Jay Pujara and Sungjoon Park
[EMNLP 2023] Conference on Empirical Methods in Natural Language Processing (EMNLP), 2023.
- [10] **AutoTriggER: Named Entity Recognition with Auxiliary Trigger Extraction**
Dong-Ho Lee*, Ravi K. Selvam*, Sheikh M. Sarwar, Bill Y. Lin, Mahak Agarwal, Fred Morstatter, Jay Pujara, Elizabeth Boschee, James Allan and Xiang Ren
[EACL 2023] Conference of the European Chapter of the Association for Computational Linguistics (EACL), 2023.
[WeaSul@ICLR 2021] ICLR 2021 Workshop on Weakly Supervised Learning.
[TrustNLP@NAACL 2021] NAACL 2021 Workshop on Trustworthy Natural Language Processing. (**Best Paper Award**)
- [11] **Reflect, Not Reflex: Inference-Based Common Ground Improves Dialogue Response Quality**
Pei Zhou, Hyundong J. Cho, Pegah Jandaghi, Dong-Ho Lee, Bill Yuchen Lin, Jay Pujara and Xiang Ren
[EMNLP 2022] Conference on Empirical Methods in Natural Language Processing (EMNLP), 2022.
- [12] **Good Examples Make A Faster Learner: Simple Demonstration-based Learning for Low-resource NER**
Dong-Ho Lee, Akshen Kadakia*, Kangmin Tan*, M. Agarwal, X. Feng, T. Shibuya, R. Mitani, T. Sekiya, Jay Pujara and Xiang Ren ([w/ Sony AI](#))
[ACL 2022] Annual Meeting of the Association for Computational Linguistics (ACL), 2022.
- [13] **Leveraging Visual Knowledge in Language Tasks: An Empirical Study on Intermediate Pre-training for Cross-Modal Knowledge Transfer**
Woojeong Jin*, Dong-Ho Lee*, Chenguang Zhu, Jay Pujara and Xiang Ren
[ACL 2022] Annual Meeting of the Association for Computational Linguistics (ACL), 2022.
- [14] **Assessing Scientific Research Papers with Knowledge Graphs**
Kexuan Sun, Zhiqiang Qiu, Abel Salinas, Yuzhong Huang, Dong-Ho Lee, Daniel Benjamin, Fred Morstatter, Xiang Ren, Kristina Lerman and Jay Pujara
[SIGIR 2022] ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), 2022.
- [15] **Perhaps PTLMs should go to School – A Task to Assess Open Book and Closed Book QA**
Manuel Ciosici, Joe Cecil, Alex Hedges, Dong-Ho Lee, Marjorie Freedman and Ralph Weischedel
[EMNLP 2021] Conference on Empirical Methods in Natural Language Processing (EMNLP), 2021.
- [16] **Improving Text Auto-Completion with Next Phrase Prediction**
Dong-Ho Lee, Zhiqiang Hu and Roy Ka-Wei Lee
[EMNLP 2021 (Findings)] Conference on Empirical Methods in Natural Language Processing (EMNLP), Findings, 2021.
- [17] **ForecastQA: A Question Answering Challenge for Event Forecasting**
Woojeong Jin, Suji Kim, Rahul Khanna, Dong-Ho Lee, Fred Morstatter, Aram Galstyan and Xiang Ren
[ACL 2021] Annual Meeting of the Association for Computational Linguistics (ACL), 2021.

- [18] **RiddleSense: Answering Riddle Questions as Commonsense Reasoning**
Bill Yuchen Lin, Ziyi Wu, Yichi Yang, **Dong-Ho Lee** and Xiang Ren
[ACL 2021 (Findings)] Annual Meeting of the Association for Computational Linguistics (ACL), Findings, 2021.
- [19] **Pre-training Text-to-Text Transformers for Concept-centric Common Sense**
Wangchunshu Zhou*, **Dong-Ho Lee***, Ravi Kiran Selvam, Seyeon Lee, Bill Yuchen Lin and Xiang Ren
[ICLR 2021] International Conference on Representation Learning (ICLR), 2021.
[SSL@NIPS 2020] NeurIPS 2020 Workshop on Self-Supervised Learning: Theory and Practice.
- [20] **TriggerNER: Label-Efficient Learning for Named Entity Recognition via Annotating Entity Triggers**
Bill Yuchen Lin*, **Dong-Ho Lee***, Ming Shen, Ryan Moreno, Xiao Huang, Prashant Shiralkar and Xiang Ren
[ACL 2020] Annual Meeting of the Association for Computational Linguistics (ACL), 2020.

System Demonstration Papers in Conference Proceedings

- [21] **XMD*: An End-to-End Framework for Interactive Explanation-Based Debugging of NLP Models**
Dong-Ho Lee*, Akshen Kadakia*, Brihi Joshi, Aaron Chan, Ziyi Liu, T. Shibuya, R. Mitani, T. Sekiya, Jay Pujara and Xiang Ren ([w/ Sony AI](#))
[ACL 2023] Annual Meeting of the Association for Computational Linguistics (ACL), 2023.
- [22] **Machine-Assisted Script Curation**
Manuel R. Ciosici, Joseph Cummings, Mitchell DeHaven, Alex Hedges, Yash Kankanampati, **Dong-Ho Lee**, Ralph Weischedel and Marjorie Freedman
[NAACL-HLT 2021] The Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT), 2021.
- [23] **LEAN-LIFE: A Label-Efficient Annotation Framework Towards Learning from Annotator Explanation**
Dong-Ho Lee*, Rahul Khanna*, Bill Yuchen Lin, Jamin Chen, Seyeon Lee, Qinyuan Ye, Elizabeth Boschee, Leonardo Neves and Xiang Ren
[ACL 2020] Annual Meeting of the Association for Computational Linguistics (ACL), 2020.
- [24] **AlpacaTag: An Active Learning-based Crowd Annotation Framework for Sequence Tagging**
Bill Yuchen Lin*, **Dong-Ho Lee***, Frank F. Xu, Ouyi Lan and Xiang Ren
[ACL 2019] Annual Meeting of the Association for Computational Linguistics (ACL), 2019.

Research papers in Journals

- [25] **Beyond the Imitation Game: Quantifying and extrapolating the capabilities of language models**
442 authors including **Dong-Ho Lee**
[TMLR] Transactions on Machine Learning Research, 2023. (**Outstanding Paper Finalist**)
- [26] **Fake news detection using deep learning**
Dong-Ho Lee, Yu-Ri Kim, Hyeong-Jun Kim, Seung-Myun Park, Yu-Jun Yang
Journal of Information Processing Systems, vol.15, no.5, pp.1119-1130, 2019.

Research papers in Other Fields

- [27] **Integrating Pre-Trained Language Model with Physical Layer Communications**
Ju-Hyung Lee, **Dong-Ho Lee**, Joohan Lee, Jay Pujara
[IEEE TWC] IEEE Transactions on Wireless Communications, 2024.
- [28] **Seq2Seq-SC: End-to-End Semantic Communication Systems with Pre-trained Language Model**
Ju-Hyung Lee*, **Dong-Ho Lee***, Eunsoo Sheen, Thomas Choi, Jay Pujara
[IEEE Asilomar 2023] Asilomar Conference on Signals, Systems, and Computers (ACSSC), 2023.

Research papers in Workshops

- [29] **Improving Covert Toxicity Detection by Retrieving and Generating References**
Dong-Ho Lee, Hyundong J. Cho, Woojeong Jin, Jihyung Moon, Sungjoon Park, Paul Rottger, Jay Pujara and Roy Ka-Wei Lee
[WOAH@NAACL2024] The 8th Workshop on Online Abuse and Harms (WOAH), 2024.

Patents

- [1] **A Recommendation System using a Language Model Neural Network**
Dong-Ho Lee, Adam Kraft, Long Jin, Nikhil Mehta, Taibai Xu, Lichan Hong, Ed H. Chi and Xinyang Yi
Patent Application (Submitted). October 2024. Assigned to Google LLC.

Talks

Improving Language Models through Context

Google Deepmind Tech Talk, Feb 2025
Microsoft Research Seminar, Feb 2025, May 2025
Sierra AI Tech Talk, Feb 2025

Explanation-based Learning

Microsoft Research AI Power Lunch, Jul 2022
POSTECH AI Seminar, Aug 2022

Mentoring

- 2023 **Kiran Narahari**, MS student, USC (Now @ USC ISI)
- 2022 **Jae Young Kim**, MS student, USC (Now @ Walmart)
- 2021 **Kangmin Tan**, Undergraduate, USC (Now @ Citadel)
- 2021 **Ravi Kiran Selvam**, MS student, USC (Now @ Amazon)

Teaching Experience

- 2022 **Teaching Assistant**, CSCI 563: Building Knowledge Graphs (2022 Fall)
- 2025 **Teaching Assistant**, CSCI 572: Information Retrieval and Web Search Engines (2025 Spring)

Awards

- 2021 **Best Paper Award**, NAACL 2021 TrustNLP Workshop
- 2018 **Best paper award**, KIPS 2018 Spring Conference
- 2016 **Academic Scholarship**, SKKU (Sungkyunkwan University)

Extracurricular

Republic of Korea Air Force

Sergeant

Osan AB, Republic of Korea

Aug. 2013 - Aug. 2015

- Led a team at the Master Control and Reporting Center (MCRC) at Osan Air Base
- Team managed radars and consoles that monitored possible activities from North Korean aircraft and ballistic missiles.