

# Dong-Ho Lee

✉ dongho.lee@usc.edu 🏠 danny-lee.info 📱 danny911kr

## Education

### University of Southern California (USC)

Ph.D. in Computer Science

• Advisor: [Dr. Jay Pujara](#)

Los Angeles, United States

Aug. 2021 - Current

### University of Southern California (USC)

M.S. in Computer Science

• Advisor: [Dr. Xiang Ren](#)

Los Angeles, United States

Mar. 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

## Experience

### SoftlyAI

Founding Machine Learning Engineer

Seoul, Republic of Korea

Mar. 2022 - Current

### Snap Research

Research Intern

Los Angeles, United States

May. 2023 - Aug. 2023

### Microsoft Research

Research Intern

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

### Singapore University of Technology and Design (SUTD)

Visiting Researcher

Singapore

May. 2020 - Aug. 2021

## Publications

\* These authors contributed equally

### Preprints

#### [1] 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

[NLP+CSS@EMNLP 2022] EMNLP 2022 Workshop on Natural Language Processing and Computational Social Science.

#### [2] 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, Takashi Shibuya, Ryosuke Mitani, Toshiyuki Sekiya, Jay Pujara and Xiang Ren

#### [3] 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, Joongheon Kim

#### [4] Beyond the Imitation Game: Quantifying and extrapolating the capabilities of language models

442 authors including [Dong-Ho Lee](#)

### Research Papers in conference proceedings

#### [5] AutoTriggeER: 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)**

#### [6] 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.

#### [7] Good Examples Make A Faster Learner: Simple Demonstration-based Learning for Low-resource NER

[Dong-Ho Lee](#), Akshen Kadakia\*, Kangmin Tan\*, Mahak Agarwal, Xinyu Feng, Takashi Shibuya, Ryosuke Mitani, Toshiyuki Sekiya, Jay Pujara and Xiang Ren

[ACL 2022] Annual Meeting of the Association for Computational Linguistics (ACL), 2022.

#### [8] 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.

## [9] 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.

## [10] 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.

## [11] 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.

## [12] 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.

## [13] 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.

## [14] 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.

## [15] 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

### [16] 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.

### [17] 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.

### [18] AlpacaTag: An Active Learning-based Crowd Annotation Framework for Sequence Tagging

Bill Yuchen Lin\*, [Dong-Ho Lee\\*](#), Frank F. Xu, Ouyu Lan and Xiang Ren  
[ACL 2019] Annual Meeting of the Association for Computational Linguistics (ACL), 2019.

## Research papers in Journals

### [19] 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.

## Talks

---

### Explanation-based Learning

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

## Teaching Experience

---

2022 **Teaching Assistant**, CSCI 563: Building Knowledge Graphs (2022 Fall)

## Services

---

### Program Committee

International Conference on Machine Learning (**ICML**): 2022  
International Conference on Representation Learning (**ICLR**): 2022-2023  
Conference on Neural Information Processing Systems (**NeurIPS**): 2021-2022  
ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD**): 2022-2023  
Annual Meeting of the Association for Computational Linguistics (**ACL**): 2023  
Conference on Empirical Methods in Natural Language Processing (**EMNLP**): 2021-2022  
SIAM International Conference on Data Mining (**SDM**): 2022  
Conference of the European Chapter of the Association for Computational Linguistics (**EACL**): 2021, 2023  
Language Resources and Evaluation Conference (**LREC**): 2022

### External Reviewer

ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD**): 2020-2021  
International World Wide Web Conference (**WWW**): 2020-2021  
Annual Meeting of the Association for Computational Linguistics (**ACL**): 2020  
The SIGNLL Conference on Computational Natural Language Learning (**CONLL**): 2019

## 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

*Osan AB, Republic of Korea*

Sergeant

*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.

### Sangmoon High School

*Seoul, Republic of Korea*

Computer Science Teacher

*June. 2017 - Nov. 2017*

- Organized course curriculum and taught a computer science class for one semester.

### Sunrin Internet High school

*Seoul, Republic of Korea*

Teaching Intern

*May. 2017*

- Taught JAVA to students who showed promising talent in IT.
- Received K-12 teaching license.