

# MOONTAE LEE

[mu:nte li:]

University Hall #2425  
601 South Morgan St.  
Chicago, IL 60607

+1 312.543.6290  
moontae@uic.edu  
<https://moontae.github.io>

## RESEARCH INTEREST

Large Language Models

- Data Generation/Evaluation
- Reasoning: Retrieval, Search, Thinking, Verification
- Preference/Value Alignment, Moral Reasoning

Foundation Models

- Multimodal Reasoning, Time-series Forecasting
- Computational Creativity, Algorithmic Awareness
- Machine Unlearning, Geometric Deep Learning

World Models

- Ranking Monopoly, Voting Fairness
- Counterfactual Evaluation, Bayesian/Causal Inference
- Economy/Market Uncertainty, Multi-agent Simulation

## EDUCATION

- 08/2012 – 05/2018 PhD in *Computer Science*  
Cornell University, Ithaca, NY  
Advisors: David Mimno & David Bindel
- 09/2009 – 03/2012 MS in *Computer Science*  
Stanford University, Stanford, CA  
Advisors: Scott Klemmer & Daphne Koller
- 03/2001 – 02/2006 BS in *Computer Science*, BS in *Mathematics*, BA in *Psychology*  
Sogang University, Seoul  
Advisors: Sungyong Park & Jungyun Seo

## EMPLOYMENT

- 01/2023 – PRESENT LG AI Research  
*Senior Research Fellow*, Head of Superintelligence Lab  
Supervisor: Kyunghoon Bae
- 08/2018 – PRESENT University of Illinois Chicago  
*Assistant Professor*, Department of Information and Decision Sciences  
Supervisor: Ali Tafti
- 12/2019 – 06/2021 Microsoft Research Redmond  
*Research Scientist*, Deep Learning Group  
Supervisor: Jianfeng Gao
- 01/2019 – 08/2019 Microsoft Research Redmond  
*Visiting Scholar*, Deep Learning Group  
Supervisor: Susan Dumais
- 08/2012 – 05/2018 Cornell University

*Research Assistant, Department of Computer Science*  
Advisors: David Mimno | David Bindel

06/2016 – 08/2016 Microsoft Research Redmond  
*Research Intern, Deep Learning Group*  
Advisor: Paul Smolensky | Supervisor: Li Deng

06/2015 – 08/2015 Microsoft Research Redmond  
*Research Intern, Deep Learning Group*  
Advisor: Paul Smolensky | Supervisor: Li Deng

04/2012 – 07/2012 Neustar Research  
*Research Intern, IP Geolocation Team*  
Supervisor: Armand Frieditis

04/2011 – 03/2012 Stanford University  
*Research Assistant, Human-Computer Interaction Group*  
Advisor: Scott Klemmer

03/2006 – 03/2008 Korean Army  
*Military Service*

## CONCURRENT POSITIONS

05/2021 – PRESENT KAIST  
*Board Committee, Kim Jaechul Graduate School of Artificial Intelligence*

08/2020 – PRESENT Yukyoung Scholarship  
*Independent Director, Dongwon Corporation*

## REFEREED PUBLICATIONS

- [C53] *Shifting from Ranking to Set Selection for Retrieval Augmented Generation*  
Dahyun Lee, Yongrae Jo, Haeju Park, and **Moontae Lee**  
Association of Computational Linguistics, [ACL 2025](#)
- [C52] *One Missing Piece for Open-Source Reasoning Models: A Dataset to Mitigate Cold-Starting Short CoT LLMs in RL*  
Hyunjoo Chae, Dongjin Kang, Jihyuk Kim, Beong-woo Kwak, Sunghyun Park, Haeju Park, Jinyoung Yeo, **Moontae Lee**, and Kyungjae Lee  
Association of Computational Linguistics, [ACL 2025](#)
- [C51] *Overlapping Context with Variable-Length Stride Increases Diversity when Training Large Language Model for Code*  
Geonmo Gu, Jaeho Kwak, Haksoo Moon, Hyun Seung Shim, Yu Jin Kim, Byoungjip Kim, **Moontae Lee**, and Hyejeong Jeon  
Association of Computational Linguistics, [ACL 2025](#)
- [C50] *Finding Information Quality: Counterfactual Voting Adjustment for Quality Assessment and Fairer Voting in Online Platforms with Helpfulness Evaluation*  
Chang Liu, Yixin Wang, **Moontae Lee**  
International Conference on Machine Learning, [ICML 2025](#)
- [C49] *From Popularity to Meritocracy: Evolution of Excellence and Information Monopoly in Online Communities*  
**Moontae Lee**, Chang Liu, and Ali Tafti  
International Conference for Computational Social Science, [IC2S2 2025](#)
- [C48] *Towards Robust and Parameter-Efficient Knowledge Unlearning for Large Language Models*  
Sungmin Cha, Sungjun Cho, Dasol Hwang, **Moontae Lee**

- [C47] *Learning to Explore and Select for Coverage-Conditioned Retrieval-Augmented Generation*  
Takyoung Kim, Kyungjae Lee, Young Rok Jang, Ji Yong Cho, Gangwoo Kim, Minseok Cho, and **Moontae Lee**  
North American Association of Computational Linguistics, [NAACL 2025](#), [Findings](#)
- [C46] *The BiGGen Bench: A Principled Benchmark for Fine-grained Evaluation of Language Models with Language Models* **BEST PAPER AWARD**  
Seungone Kim, Juyoung Suk, Ji Yong Cho, Shayne Longpre, Chaeun Kim, Dongkeun Yoon, Guijin Son, Yejin Cho, Sheikh Shafayat, Jinheon Baek, Sue Hyun Park, Hyeonbin Hwang, Jinkyung Jo, Hyowon Cho, Haebin Shin, Seongyun Lee, Hanseok Oh, Noah Lee, Namgyu Ho, Se June Joo, Miyoung Ko, Yoonjoo Lee, Hyungjoo Chae, Jamin Shin, Joel Jang, Seonghyeon Ye, Bill Yuchen Lin, Sean Welleck, Graham Neubig, **Moontae Lee**, Kyungjae Lee, Minjoon Seo  
North American Association of Computational Linguistics, [NAACL 2025](#)
- [C45] *Hierarchical Decomposition Framework for Steiner Tree Packing Problem* **BEST STUDENT PAPER AWARD**  
Hanbum Ko, Minu Kim, Han-Seul Jeong, Sunghoon Hong, Deunsol Yoon, Youngjoon Park, Woohyung Lim, Honglak Lee, **Moontae Lee**, Kanghoon Lee, Sungbin Lim, Sungryull Sohn  
International Conference on Operations Research and Enterprise Systems, [ICORES 2025](#)
- [C44] *3D Denoisers are Good 2D Teachers: Molecular Pretraining via Denoising and Cross-Modal Distillation*  
Sungjun Cho, Dae-Woong Jeong, Sung Moon Ko, Jinwoo Kim, Sehui Han, Seunghoon Hong, Honglak Lee, and **Moontae Lee**  
Association for the Advancement of Artificial Intelligence, [AAAI 2025](#)
- [W15] *Towards Robust and Cost-Efficient Knowledge Unlearning for Large Language Models*  
Sungmin Cha, Sungjun Cho, Dasol Hwang, **Moontae Lee**  
Neural Information Processing Systems, [NeurIPS Workshop 2024](#)
- [C43] *Prometheus 2: An Open Source Language Model Specialized in Evaluating Other Language Models*  
Seungone Kim, Juyoung Suk, Shayne Longpre, Bill Yuchen Lin, Jamin Shin, Sean Welleck, Graham Neubig, **Moontae Lee**, Kyungjae Lee, Minjoon Seo  
Empirical Methods in Natural Language Processing, [EMNLP 2024](#)
- [C42] *When “A Helpful Assistant” Is Not Really Helpful: Personas in System Prompts Do Not Improve Performances of Large Language Models*  
Mingqian Zheng, Jiaxin Pei, Lajanugen Logeswaran, **Moontae Lee**, David Jurgens  
Empirical Methods in Natural Language Processing, [EMNLP Findings 2024](#)
- [C41] *Prospector: Improving LLM Agents with Self-Asking and Trajectory Ranking*  
Byoungjip Kim, Youngsoo Jang, Lajanugen Logeswaran, Geon-Hyeong Kim, Yu Jin Kim, Honglak Lee, and **Moontae Lee**  
Empirical Methods in Natural Language Processing, [EMNLP Findings 2024](#)
- [C40] *Degeneration-free Policy Optimization: RL Fine-Tuning for Language Models without Degeneration*  
Youngsoo Jang, Geon-Hyeong Kim, Byeongjip Kim, Honglak Lee, and **Moontae Lee**  
International Conference on Machine Learning, [ICML 2024](#)
- [C39] *Semantic Skill Grounding for Embodied Instruction-Following in Cross-Domain Environments*  
Sangwoo Shin, SeungHyun Kim, Youngsoo Jang, **Moontae Lee**, and Honguk Woo  
Association of Computational Linguistics, [ACL 2024](#)
- [C38] *Self-Improving Small Language Models for Self-Correction in Reasoning*  
Yunxiang Zhang, Muhammad Khalifa, Lajanugen Logeswaran, Jaekyeom Kim, **Moontae Lee**, Honglak Lee, and Lu Wang  
Association of Computational Linguistics, [ACL 2024](#)
- [C37] *Co-Creating Question-and-Answer Style Articles with Large Language Models for Research Promotion*  
Hyunseung Lim, Ji Yong Cho, Taewan Kim, Jeongeon Park, Hyungyu Shin, Seulgi Choi, Sunghyun Park, Kyungjae Lee, Juho Kim, **Moontae Lee**, and Hwajung Hong  
ACM Conference on Designing Interactive Systems, [DIS 2024](#)

- [W14] *LG AI Research & KAIST at EHRSQ 2024: Self-Training Large Language Models with Pseudo-Labeled Unanswerable Questions for a Reliable Text-to-SQL System on EHRs*  
Yongrae Jo, Seongyun Lee, Minju Seo, Sung Ju Hwang, and **Moontae Lee**  
North American Association of Computational Linguistics, [NAACL 2024](#)
- [C36] *Towards More Diverse Evaluation of Class Incremental Learning: Representation Learning Perspective*  
Sungmin Cha, Jihwan Kwak, Dongsub Shim, Hyunwoo Kim, **Moontae Lee**, Honglak Lee, and Taesup Moon  
Conference on Lifelong Learning Agents, [CoLLA 2024](#)
- [W13] *Towards an Evaluation of LLM-Generated Inspiration by Developing and Validating Inspiration Scale*  
Hyungyu Shin, Seulgi Choi, Ji Yong Cho, Sahar Admoni, Hyunseung Lim, Taewan Kim, Hwajung Hong, **Moontae Lee**, Juho Kim  
Conference on Human Factors in Computing Systems, [CHI Workshop 2024](#)
- [C35] *You don't need a personality test to know these models are unreliable: Assessing the Reliability of Large Language Models on Psychometric Instruments*  
Bangzhao Shu, Lechen Zhang, Minje Choi, Lavinia Dunagan, Lajanugen Logeswaran, **Moontae Lee**, Dallas Card, David Jurgens  
North American Association of Computational Linguistics, [NAACL 2024](#)
- [C34] *Understanding the Capabilities and Limitations of Large Language Models for Cultural Commonsense*  
**SOCIAL IMPACT AWARD**  
Siqi Shen, Lajanugen Logeswaran, **Moontae Lee**, Honglak Lee, Soujanya Poria, Rada Mihalcea  
North American Association of Computational Linguistics, [NAACL 2024](#)
- [C33] *Code Models are Zero-shot Precondition Reasoners*  
Lajanugen Logeswaran, Sungryull Sohn, Yiwei Lyu, Anthony Zhe Liu, Dong-Ki Kim, Dongsub Shim, **Moontae Lee**, Honglak Lee  
North American Association of Computational Linguistics, [NAACL 2024](#)
- [W12] *YTCommentQA: Video Question Answerability in Instructional Video*  
Saelyne Yang, Sunghyun Park, Yunseok Jang, and **Moontae Lee**  
Computer Vision and Pattern Recognition Conference, [CVPR Workshop 2024](#)
- [W11] *Degeneration-free Policy Optimization: RL Fine-Tuning for Language Models without Degeneration*  
Youngsoo Jang, Geon-Hyeong Kim, Byoungjip Kim, Yu Jin Kim, Honglak Lee, and **Moontae Lee**  
International Conference on Learning Representations, [ICLR Workshop 2024](#)
- [C32] *Learning to Unlearn: Instance-wise Unlearning for Pre-trained Classifiers*  
Sungmin Cha, Sungjun Cho, Dasol Hwang, Honglak Lee, Taesup Moon, and **Moontae Lee**  
Association for the Advancement of Artificial Intelligence, [AAAI 2024](#)
- [C31] *YTCommentQA: Video Question Answerability in Instructional Video*  
Saelyne Yang, Sunghyun Park, Yunseok Jang, and **Moontae Lee**  
Association for the Advancement of Artificial Intelligence, [AAAI 2024](#)
- [C30] *SafeDICE: Offline Safe Imitation Learning with Non-Preferred Demonstrations*  
Youngsoo Jang, Geon-Hyeong Kim, Jongmin Lee, Sungryull Sohn, Byeongjip Kim, Honglak Lee, and **Moontae Lee**  
Neural Information Processing Systems, [NeurIPS 2023](#)
- [C29] *Projection Regret: Reducing Background Bias for Novelty Detection via Diffusion Models*  
Sungik Choi, Hankook Lee, Honglak Lee, and **Moontae Lee**  
Neural Information Processing Systems, [NeurIPS 2023](#)
- [W10] *Prospector: Improving LLM Agents with Self-Asking and Trajectory Ranking*  
Byeongjip Kim, Youngsoo Jang, Lajanugen Logeswaran, Geon-Hyeong Kim, Yu Jin Kim, Honglak Lee, and **Moontae Lee**  
Neural Information Processing Systems, [NeurIPS Workshop 2023](#)
- [W9] *Reasoning about Action Preconditions with Programs*

- Lajanugen Logeswaran, Sungryull Sohn, Yiwei Lyu, Anthony Zhe Liu, Dong-Ki Kim, Dongsub Shim, **Moontae Lee**, and Honglak Lee  
Neural Information Processing Systems, [NeurIPS Workshop 2023](#)
- [C28] *From Heuristic to Analytic: Cognitively Motivated Strategies for Coherent Physical Commonsense Reasoning*  
Zheyuan Zhang, Shane Storks, Fengyuan Hu, Sungryull Sohn, **Moontae Lee**, Honglak Lee, and Joyce Chai  
Empirical Methods in Natural Language Processing, [EMNLP 2023](#)
- [C27] *Merging Generated and Retrieved Knowledge for Open-Domain QA*  
Yunxiang Zhang, Muhammad Khalifa, Lajanugen Logeswaran, **Moontae Lee**, Honglak Lee, and Lu Wang  
Empirical Methods in Natural Language Processing, [EMNLP 2023](#)
- [C26] *GRACE: Discriminator-Guided Chain-of-Thought Reasoning*  
Muhammad Khalifa, Lajanugen Logeswaran, **Moontae Lee**, Honglak Lee, and Lu Wang  
Empirical Methods in Natural Language Processing, [EMNLP Findings 2023](#)
- [C25] *On Sample-Efficient Code Generation*  
Hojae Han, Yu Jin Kim, Byeongjip Kim, Youngwon Lee, Kyungmin Lee, Kyungjae Lee, **Moontae Lee**, Kyunghoon Bae, and Seung-won Hwang  
Empirical Methods in Natural Language Processing, [EMNLP Industry Track 2023](#)
- [C24] *When to Read Documents or QA History: On Unified and Selective Open-domain QA*  
Kyungjae Lee, Sang-eun Han, Seung-won Hwang, and **Moontae Lee**  
Association for Computational Linguistics, [ACL Findings 2023](#)
- [C23] *Few-shot Reranking for Multi-hop QA via Language Model Prompting*  
Muhammad Khalifa, Lajanugen Logeswaran, **Moontae Lee**, Honglak Lee, and Lu Wang  
Association for Computational Linguistics, [ACL 2023](#)
- [C22] *Unsupervised Task Graph Generation from Instructional Video Transcripts*  
Lajanugen Logeswaran, Sungryull Sohn, Yunseok Jang, **Moontae Lee**, and Honglak Lee  
Association for Computational Linguistics, [ACL Findings 2023](#)
- [C21] *Knowledge Unlearning for Mitigating Privacy Risks in Language Models*  
Joel Jang, Dongkeun Yoon, Sohee Yang, Sungmin Cha, **Moontae Lee**, Lajanugen Logeswaran, and Minjoon Seo  
Association for Computational Linguistics, [ACL 2023](#)
- [W8] *Mixed-Curvature Transformers for Graph Representation Learning*  
Sungjun Cho, Seunghyuk Cho, Sungwoo Park, Hankook Lee, Honglak Lee, and **Moontae Lee**  
International Conference on Machine Learning, [ICML Workshop 2023](#)
- [C20] *QASA: Advanced Question Answering on Scientific Articles*  
Yoonjoo Lee, Kyungjae Lee, Sunghyun Park, Dasol Hwang, Jaehyeon Kim, Hong-in Lee, and **Moontae Lee**  
International Conference on Machine Learning, [ICML 2023](#)
- [C19] *Neural Stochastic Differential Games for Time-series Analysis*  
Sungwoo Park, Byoung Woo Park, **Moontae Lee**, and Changhee Lee  
International Conference on Machine Learning, [ICML 2023](#)
- [C18] *Exploring the Benefits of Training Expert Language Models over Instruction Tuning*  
Joel Jang, Seungone Kim, Seonghyeon Ye, Doyoung Kim, Lajanugen Logeswaran, **Moontae Lee**, Kyungjae Lee, and Minjoon Seo  
International Conference on Machine Learning, [ICML 2023](#)
- [W7] *Exploring Demonstration Ensembling for In-context Learning*  
Muhammad Khalifa, Lajanugen Logeswaran, **Moontae Lee**, Honglak Lee, and Lu Wang  
International Conference on Learning Representations, [ICLR Workshop 2023](#)

- [C17] *Rebalancing Batch Normalization for Exemplar-based Class-Incremental Learning*  
Sungmin Cha, Sungjun Cho, Dasol Hwang, Sunwon Hong, **Moontae Lee**, and Taesup Moon  
Computer Vision and Pattern Recognition Conference, [CVPR 2023](#)
- [C16] *Grouping-matrix based Graph Pooling with Adaptive Number of Clusters*  
Sung Moon Ko, Sungjun Cho, Dae-Woong Jeong, Sehui Han, **Moontae Lee**, and Honglak Lee  
Association for the Advancement of Artificial Intelligence, [AAAI 2023](#)
- [C15] *Pure Transformers are Powerful Graph Learners*  
Jinwoo Kim, Tien Dat Nguyen, Seonwoo Min, Sungjun Cho, **Moontae Lee**, Honglak Lee, and Seunghoon Hong  
Neural Information Processing Systems, [NeurIPS 2022](#)
- [C14] *Transformers meet Stochastic Block Models: Attention with Data-Adaptive Sparsity and Cost*  
Sungjun Cho, Seonwoo Min, Jinwoo Kim, **Moontae Lee**, Honglak Lee, and Seunghoon Hong  
Neural Information Processing Systems, [NeurIPS 2022](#)
- [C13] *Transferring Pre-trained Multimodal Representations with Cross-modal Similarity Matching*  
Byoungjip Kim, Sungik Choi, Dasol Hwang, **Moontae Lee**, and Honglak Lee  
Neural Information Processing Systems, [NeurIPS 2022](#)
- [C12] *CEDe: A collection of expert-curated datasets with atom-level entity annotations for Optical Chemical Structure Recognition*  
Rodrigo Hormazabal, Changyoung Park, Soonyoung Lee, Sehui Han, Yeonsik Jo, Jaewan Lee, Ahra Jo, Seung Hwan Kim, Jaegul Choo, **Moontae Lee**, and Honglak Lee  
Neural Information Processing Systems, [NeurIPS Dataset & Benchmark 2022](#)
- [W6] *ReSPack: A Large-Scale Rectilinear Steiner Tree Packing Data Generator and Benchmark*  
Kanghoon Lee, Youngjoon Park, Han-Seul Jeong, Sunghoon Hong, Deunsol Yoon, Sungryull Sohn, Minu Kim, Hanbum Ko, **Moontae Lee**, Honglak Lee, Kyunghoon Kim, Euihyuk Kim, Seonggeon Cho, Jaesang Min, and Woohyung Lim  
Neural Information Processing Systems, [NeurIPS Workshop 2022](#)
- [C11] *Symbolic Music Loop Generation with Neural Discrete Representations*  
Sangjun Han, Hyeongrae Ihm, **Moontae Lee**, and Woohyung Lim  
International Society for Music Information Retrieval, [ISMIR 2022](#)
- [C10] *Path-aware and Structure-preserving Generation of Synthetically Accessible Molecules*  
Juhwan Noh, Dae-Woong Jeong, Kiyoun Kim, Sehui Han, **Moontae Lee**, Honglak Lee, and Yousung Jung  
International Conference on Machine Learning, [ICML 2022](#)
- [C9] *Few-shot Subgoal Planning with Language Models*  
Lajanugen Logeswaran, Yao Fu, **Moontae Lee**, and Honglak Lee  
North American Association of Computational Linguistics, [NAACL 2022](#)
- [W5] *Few-shot Subgoal Planning with Language Models*  
Lajanugen Logeswaran, Yao Fu, **Moontae Lee**, and Honglak Lee  
Association of Computational Linguistics, [ACL Workshop 2022](#)
- [J2] *Social Media Users' Perceptions of a Wearable Mixed Reality Headset During the COVID-19 Pandemic: Aspect-Based Sentiment Analysis*  
Heejin Jeong, Allison Bayro, Sai Patipati Umesh, Kaushal Mamgain, and **Moontae Lee**  
Journal of Medical Internet Research, [JMIR 2022](#)
- [W4] *Task-Balanced Batch Normalization for Exemplar-based Class-Incremental Learning*  
Sungmin Cha, Sunwon Hong, **Moontae Lee**, and Taesup Moon  
Computer Vision and Pattern Recognition Conference, [CVPR Workshop 2023](#)
- [C8] *On-the-Fly Rectification for Robust Large-Vocabulary Topic Inference*  
**Moontae Lee**, Sungjun Cho, Kun Dong, David Mimno, and David Bindel



- International Conference on Machine Learning, [ICML 2021](#)
- [C7] *Prior-aware Composition Inference for Spectral Topic Models*  
**Moontae Lee**, David Bindel, and David Mimno  
 Artificial Intelligence and Statistics, [AISTATS 2020](#)
- [C6] *Practical Correlated Topic Modeling and Analysis via Rectified Anchor Word Algorithm*  
**Moontae Lee**, Sungjun Cho, David Bindel, and David Mimno  
 Empirical Methods in Natural Language Processing, [EMNLP 2019](#)
- [W3] *The Chinese Voting Process: The Evolution of Online Helpfulness Evaluations*  
**Moontae Lee**, Seok Hyun Jin, and David Mimno  
 Institute for Operations Research and the Management Sciences, [INFORMS Workshop on DMDA 2017](#)
- [W2] *From Correlation to Hierarchy: Practical Spectral Topic Modeling* BEST STUDENT PAPER AWARD  
**Moontae Lee**, David Bindel, and David Mimno  
 Institute for Operations Research and the Management Sciences, [INFORMS Workshop on DMDA 2017](#)
- [C5] *Beyond Exchangeability: The Chinese Voting Process*  
**Moontae Lee**, Seok Hyun Jin, and David Mimno  
 Neural Information Processing Systems, [NeurIPS 2016](#)
- [C4] *Reasoning in Vector Space: An Exploratory Study of Question Answering*  
**Moontae Lee**, Xiaodong He, Wen-tau Yih, Jianfeng Gao, Li Deng, and Paul Smolensky  
 International Conference on Learning Representations, [ICLR 2016](#)
- [C3] *Robust Spectral Inference for Joint Stochastic Matrix Factorization*  
**Moontae Lee**, David Bindel, and David Mimno  
 Neural Information Processing Systems, [NeurIPS 2015](#)
- [C2] *Low-dimensional Embeddings for Interpretable Anchor-based Topic Inference*  
**Moontae Lee** and David Mimno  
 Empirical Methods in Natural Language Processing, [EMNLP 2014](#)
- [J1] *TaskGenies: Providing Action Plans Helps People Complete Tasks*  
 Nicolas Kokkalis, Thomas Kohn, Johannes Huebner, **Moontae Lee**, Florian Schulze, and Scott Klemmer  
 Transactions on Computer-Human Interaction, [TOCHI 2013](#)
- [C1] *When Classification Becomes a Problem: Using Branch-and-Bound to Improve Classification Efficiency*  
 Armand Frieditis and **Moontae Lee**  
 International Conference on Machine Learning and Data Mining, [MLDM 2013](#)
- [W1] *Automatically Providing Action Plans Helps People Complete Tasks*  
 Nicolas Kokkalis, Johannes Huebner, Steven Diamond, Dominic Becker, Michael Chang, **Moontae Lee**, Florian Schulze, Thomas Koehn, and Scott R Klemmer  
 Association for the Advancement of Artificial Intelligence, [AAAI Workshop 2012](#)

## PREPRINTS

- PREPRINTED *Curve Your Attention: Mixed-Curvature Transformers for Graph Representation Learning*  
 Sungjun Cho, Seunghyuk Cho, Sungwoo Park, Hankook Lee, Honglak Lee, and **Moontae Lee**
- PREPRINTED *Practical Reconstruction of Structural Embeddings for Symbolic Music Generation with Large Language Models*  
 Seungyeon Rhyu, Sungjun Cho, Kichang Yang, Jaehyeon Kim, and **Moontae Lee**
- PREPRINTED *Multimodal Subtask Graph Generation from Instructional Videos*  
 Yunseok Jang, Sungryull Sohn, Lajanugen Logeswaran, Tiange Luo, **Moontae Lee**, and Honglak Lee
- PREPRINTED *Basic Reasoning with Tensor Product Representations*  
 Paul Smolensky, **Moontae Lee**, Xiaodong He, Wen-tau Yih, Jianfeng Gao, and Li Deng

## WORK IN PROGRESS (JOURNAL)

- DRAFTED *From Popularity to Meritocracy: Evolution of Excellence and Information Monopoly in Online Communities*  
**Moontae Lee**, Chang Liu, and Ali Tafti
- DRAFTED *Finding Information Quality: Counterfactual Voting Adjustment for Quality Assessment and Voting Fairness in Online Platforms with Helpfulness Evaluation*  
Chang Liu, Yixin Wang, and **Moontae Lee**
- DRAFTED *Robust and Scalable Spectral Inference for Bayesian Mixed-Membership Models in Latent Topic Discovery*  
**Moontae Lee**, Sungjun Cho, David Bindel, and David Mimno
- DRAFTED *Towards Interpretable Spectral Inference: Rectified Anchor-Word Algorithms for Correlated, Large-Vocabulary, and Hierarchical Topic Modeling*  
**Moontae Lee**, Kun Dong, David Bindel, and David Mimno
- DRAFTED *Reinforcement Learning from Reflective Feedback: Aligning & Improving LLMs via Fine-Grained Self-Reflection*  
Kyungjae Lee, Dasol Hwang, Sunghyun Park, Youngsoo Jang, Kyunghoon Bae, and **Moontae Lee**
- DRAFTED *Supervised Quantile-Conditioned Policy Learning for Offline Reinforcement Learning*  
Geon-Hyeong Kim, Youngsoo Jang, Byeongjip Kim, Honglak Lee, and **Moontae Lee**  
*(About Minsinformation and Explanation)*  
Elina H. Hwang, **Moontae Lee**, Stephanie Lee  
*(About Organizations and Evaluations)*  
**Moontae Lee**

## WORK IN PROGRESS (CONFERENCE)

- SUBMITTED *(About AI-generated Image Detection without Training or Pretraining Data)*  
Sungik Choi, Hankok Lee, Jaehoon Lee, SeungHyun Kim, Stanley Jungkyu Choi, and **Moontae Lee**
- SUBMITTED *(About the Probing Strategies of Large Language Models)*  
Siqi Shen, Mehar Singh, Lajanugen Logeswaran, **Moontae Lee**, Honglak Lee, and Rada Mihalcea
- SUBMITTED *(About Paper Reviewing with Large Language Models)*  
Hyungyu Shin, Jingyu Tang, Yoonjoo Lee, Nayoung Kim, Hyunseung Lim, Ji Yong Cho, Hwajung Hong, **Moontae Lee**, and Juho Kim
- SUBMITTED *(About the System Prompt Optimization for Large Language Models)*  
Lechen Zhang, Tolga Ergen, Lajanugen Logeswaran, **Moontae Lee**, and David Jurgens
- SUBMITTED *(About Implicit Policy Regularization for Preference Optimization)*  
Youngsoo Jang, Yu Jin Kim, Geon-Hyeong Lee, Honglak Lee, and **Moontae Lee**
- SUBMITTED *(About Better Multivariate Time-Series Forecasting with Transformer)*  
Jaehoon Lee, Hankook Lee, Sungik Choi, Sungjiun Cho, and **Moontae Lee**
- SUBMITTED *(About a New Time-Series Foundation Model with Mean-Field Stochastic Differential Equations)*  
Sungwoo Park, Jaehoon Lee, Honglak Lee, and **Moontae Lee**
- SUBMITTED *(About Calibration via Langevin Diffusion)*  
Sungwoo Park, SeungHyun Kim, and **Moontae Lee**
- SUBMITTED *(About Machine Learning Benchmark)*  
Yunxiang Zhang, Muhammad Khalifa, Shitanshu Bhushan, Grant D Murphy, Lajanugen Logeswaran, Jaekyeom Kim, **Moontae Lee**, Honglak Lee, Lu Wang
- SUBMITTED *(About Simultaneously Aligning Safety and Usefulness for Large Language Models)*



Geon-Hyeong Kim, Youngsoo Jang, Yu Jin Kim, Byoungjip Kim, Honglak Lee, Kyunghoon Bae, **Moontae Lee**

*Leveraging Large Language Models to Create Question-and-Answer Style Promotional Content for Research Papers*

Ji Yong Cho, Hyunseung Lim, Haeju Park, Taewan Kim, Sunghyun Park, Yongrae Jo, Hyungyu Shin, Seulgi choi, Kyungjae Lee, Juho Kim, **Moontae Lee**, and Hwajung Hong

*LLoRA: Locally Low-rank Random Matrix Adaptation for Diffusion and Large Language Models*

Sungjun Cho, Sunghyun Park, Hyunsoo Lee, Hwigeon Oh, Inhyuk Cho, and **Moontae Lee**

*On-the-fly Adaptation of Instruction-tuned Language Models through Parameter Merging*

Yu Jin Kim, Joel Jang, Kyungjae Lee, Minjoon Seo, Lajanugen Logeswaran, and **Moontae Lee**

*Trinity of Large Language Models: Generation, Evaluation, and Verification for Knowledge Mastery and Reasoning Accuracy*

**Moontae Lee**

*Language of Evaluations, and the Evaluations of Language: Consistency is the Hallmark of Expertise*

Ji Yong Cho, Bumsoo Kang, and **Moontae Lee**

*Geometry of Words: Meaning Continuum and Computational Creativity*

**Moontae Lee**, Sungjun Cho, and Jianfeng Gao

## TEACHING EXPERIENCE

UIC IDS594: Machine Learning Applications with Python

Main Instructor

IDS576: Deep Learning and Modern Applications

Main Instructor

IDS575: Machine Learning and Statistical Methods

Main Instructor

IDS566: Advanced Text Analytics

Main Instructor

CORNELL INFO5306: Crowdsourcing and Human Computation

Teaching Assistant (Supervisor: Haym Hirsh, haym.hirsh@cornell.edu)

INFO3300: Data-driven Web Applications

Teaching Assistant (Supervisor: David Mimno (mimno@cornell.edu))

CS4780: Machine Learning

Teaching Assistant (Supervisor: Thorsten Joachims, tj@cs.cornell.edu)

CS4740: Natural Language Processing

Teaching Assistant (Supervisor: Claire Cardie, cardie@cs.cornell.edu)

STANFORD CS261: Optimization, Algorithmic Paradigms

Teaching Assistant (Supervisor: Serge Plotkin, plotkin@cs.stanford.edu)

CS161: Design and Analysis of Algorithms

Teaching Assistant (Supervisor: Tim Roughgarden, tim@cs.stanford.edu)

## HONORS AND AWARDS

*Best Paper Award*

Received the best paper award among 1400 submissions at NAACL 2025

*Best Student Paper Award*

Received the best student paper award at ICORES 2025

*Social Impact Award*

Received a special award at NAACL 2024

*Top 33% High-scoring Reviewers of the Year*

Certificate received at ICML 2020

*Top 50% High-scoring Reviewers of the Year*

Received the registration support at NeurIPS 2019

*Best Student Paper Award*

Received the best student paper award at DMDA, INFORMS 2017

*Graduate Research Symposium Award*

Selected as one of the two student speakers at Amazon Graduate Research Symposium 2017

*Travel Award*

Received the travel award at NeurIPS 2016

*Research Award*

Won the research travel grant competition for graduate students, Cornell University 2016

*Travel and Registration Awards*

Received both travel and registration awards, ICLR 2016

*Outstanding TA Award*

Selected as an outstanding teaching assistant for CS4740, Cornell University 2012

*Summa Cum Laude*

Second best in the class of 2006, Sogang University 2006

*Best Undergraduate Thesis Award*

Third best in the thesis competition, Sogang University 2006

*Bronze Medal in Korean Mathematical Competition*

Received from the final nationwide competition, KMC 2000

## INVITED TALKS AND MEDIA APPEARANCES

- 07/2025 *Learning to Unlearn: Robust and Efficient Machine Unlearning for Large Foundation Models*  
Invited talk in CAIDA/TrustML Research in University of British Columbia (Vancouver, Canada)
- 06/2025 *Counterfactual Voting Adjustment for Quality Assessment at Information Monopoly*  
Invited keynote in Data Intelligence Workshop (Jeju, South Korea)
- 05/2025 *From Text to Reality: Frontiers of Generative AI and Financial Application*  
Invited lecture at National Tax Service (Seoul, South Korea)
- 05/2025 *From Text to Reality: Frontiers of Generative AI*  
Invited keynote in the 220th Morning Lecture Series in Korean Nuclear Society (Seoul, South Korea)
- 04/2025 *Finding Information Quality: Counterfactual Voting Adjustment for Quality Assessment and Voting Fairness in Online Platforms with Helpfulness Evaluation*  
Invited talk in School of Management at KAIST (Seoul, South Korea)
- 04/2025 *AI in Action: Optimize your AI Infrastructure*  
Panel discussion in Google Next 2025 (Las Vegas, NV)
- 03/2025 *Computing Infrastructure and AI Models*  
Panel discussion in AI Global Conference by the Ministry of Science and ICT (Seoul, South Korea)
- 03/2025 *Atlas of Intelligence: How Generative AI Reshapes Knowledge, Creativity, and Society*

- Invited Seminar at Cornell Tech (New York, NY)
- 03/2025 *Atlas of Intelligence: How Generative AI Reshapes Knowledge and Creativity*  
Invited Talk at Stony Brook University (Stony Brook, NY)
- 03/2025 *The Art of AI: Past Present, Present Past, and Future*  
Invited Talk at Guggenheim (Manhattan, NY)
- 12/2024 *Leading Talent for the AI Era*  
Panel discussion in Future Society Education Conference at Chosun Ilbo (Seoul, South Korea)
- 11/2024 *Learning to Unlearn: Robust and Efficient Machine Unlearning for Large Foundation Models*  
Invited talk in AI Fairness Research Center at KAIST (Daejeon, South Korea)
- 11/2024 *Learning to Unlearn: Robust and Efficient Machine Unlearning for Large Foundation Models*  
Invited talk in Computer Science Department at Sungkyunkwan University (Suwon, South Korea)
- 11/2024 *Frontiers of Generative AI: From Text to Reality*  
1st Future Economic Forum at Yonhapnews (Seoul, South Korea)
- 10/2024 *From Text to Reality: Frontiers of Generative AI*  
Keynote talk and panel discussion in Gwangju AI Conference, AICON 2024 (Seoul, South Korea)
- 10/2024 *Generative AI: Power, Perils, and Possibilities*  
Keynote talk in Future of Data & AI Research Summit 2024 at Foundry/IDG (Seoul, South Korea)
- 09/2024 *AI Special Session: From Text to Reality*  
Keynote talk in WebOS Conference 2024 (Seoul, South Korea)
- 09/2024 *From Text to Reality: Frontiers of Large Language Models and Generative AI*  
Keynote talk in Data & Artificial Intelligence Symposium (DAISY) 2024 Workshop (Jacksonville, FL)
- 08/2024 *Probing and Proving the Complexities of LLMs: From Personal Inconsistencies to Cultural Biases*  
Invited talk in ACL 2024 C3NLP Workshop (Bangkok, Thailand)
- 06/2024 *The Art of Generative AI and Downstream Impacts*  
Invited talk in Business School at Sogang University (Seoul, South Korea)
- 06/2024 *AI Impact and New Revolution*  
Panel discussion in Munhwa Industry Forum 2024 (Seoul, South Korea)
- 05/2024 *AI Orientation, Evolution towards Super Intelligent Agent*  
Panel discussion in MIT Tech Review 2024 (Seoul, South Korea)
- 05/2024 *Technology, Business, and Ethics in Generative AI*  
Invited talk in Frontiers of AI: Business and Society 2024 (Chicago, IL)
- 05/2024 *Dilemma of Breakthrough, AI Safety and Ethics*  
Invited talk and panel discussion in Asian Leadership Conference 2024 (Seoul, South Korea)
- 04/2024 *The Art of Generative AI: Technology, Ethics, and Morality*  
Invited talk in Law School at the University of Illinois Chicago (Chicago, IL)
- 04/2024 *Recent Advances in Large Language Models and Generative AI*  
Invited talk in Kellogg Business School at Northwestern University (Evanston, IL)
- 04/2024 *Recent Advances in Large Language Models and Generative AI*  
Invited talk in Discovery Partners Institute (Chicago, IL)
- 02/2024 *The Art of Generative AI: Recent Advances in Large Language Models*  
Invited talk in Math, Stat, & Comp Sci at the University of Illinois Chicago (Chicago, IL)
- 02/2024 *The Art of Generative AI*  
Invited talk in Chancellor's Spark Talk at the University of Illinois (Chicago, IL)
- 01/2024 *The Art of Generative AI*

- Invited talk in Information and Decision Sciences in UIC Business School (Chicago, IL)
- 12/2023 *Disruption of Generative AI*  
Invited talk in Dongwon Corporation (Seoul, South Korea)
- 12/2023 *Large Language Models for Artificial Expert Intelligence*  
Invited talk in Expo Talk Panel at NeurIPS 2023 (New Orleans, LA)
- 11/2023 *The Art of Generative AI*  
Invited talk in Young Engineers Honor Society (Seoul, South Korea)
- 10/2023 *The Art of Generative AI: Text, Code, and Beyond*  
Invited talk in Graduate School of Data Science at Seoul National University (Seoul, South Korea)
- 10/2023 *AI Revolution*  
Invited Speaker in 30th Conference at Korea Veterans Hospital (Seoul, South Korea)
- 10/2023 *Welcome to the Wonderland of Generative AI*  
Invited Speaker in Gwacheon National Science Museum (Gwacheon, South Korea)
- 09/2023 *Large Language Models, Paradigm Shift by Generative AI*  
Keynote talk/Panelist in AI Summit Seoul 2023 (Seoul, South Korea)
- 08/2023 *Generative AI: Current Trends and Future Directions*  
Invited Speaker in LG Open Forum (Seoul, South Korea)
- 08/2023 *Hidden Heroes*  
Guest Speaker in JTBC Educational TV Show (Seoul, South Korea)
- 07/2023 *Welcome to the Wonderland of Generative AI*  
Invited talk in Summer 2023 at Korea Foundation for Advanced Studies (Pyungchang, South Korea)
- 06/2023 *Large-scale Generative AI and Our Future*  
Guest Speaker in 11th Anniversary Forum of Maeil Business (Seoul, South Korea)
- 05/2023 *Welcome to the Wonderland of AI*  
Invited talk at Korea University (Seoul, South Korea)
- 04/2023 *Unlocking Your Creativity with Generative AI*  
Invited talk/Panelist in Digital Marketing Summit 2023 (Seoul, South Korea)
- 03/2023 *The Era of AI: How AI Will Change Our Lives*  
Guest Speaker in MBC Economy Concert Radio/TV Show (Seoul, South Korea)
- 03/2023 *Better than ChatGPT: Providing Factual Answers with References*  
Article at Donga Business Review (DBR) (Seoul, South Korea)
- 12/2022 *Beyond the Limits: Be careful but not fearful*  
Invited talk in the 5th Global Ethics Forum at Sogang University (Seoul, South Korea)
- 09/2022 *Large Models Are Coming: Deep Dive into the New AI Paradigm*  
Invited talk in Changbal Conference 2022 (Seattle, WA)
- 08/2022 *Co-evolutionary AI Business Ecosystem*  
Invited talk in LG Academy (Seoul, South Korea)
- 06/2022 *Beyond the Limits: Be careful but not fearful*  
Invited talk/Panelist in Fairness, Accountability, and Transparency FAccT 2022 (Seoul, South Korea)
- 04/2022 *Large-scale AI*  
Guest Speaker in 3Pro TV Show (Seoul, South Korea)
- 05/2022 *Training Secret of AI Models*  
Guest Speaker in 3Pro TV Show (Seoul, South Korea)
- 04/2022 *Can You Measure Quality of Information?*

- Invited talk in Graduate School of Data Science at Seoul National University (Seoul, South Korea)
- 12/2021 *Spectral Inference and Beyond*  
Invited talk in Graduate School of AI at KAIST (Seoul, South Korea)
- 11/2021 *Large-scale AI: Present, Visions, and Future*  
Invited talk/Panelist in AI Summit Seoul 2021 (Seoul, South Korea)
- 11/2021 *The Future of AI Healthcare by Materials Informatics*  
Invited talk in MBC Global AI Summit 2021 (Seoul, South Korea)
- 09/2021 *Co-evolutionary AI Business Ecosystem*  
Invited talk in LG's CEO Workshop 2021 (Seoul, South Korea)
- 08/2021 *Co-evolutionary AI Business Ecosystem*  
Invited talk in LG Biz Informal 2021 (Seoul, South Korea)
- 06/2021 *Veils of Distortion: A Bird's Eye View of my Research*  
Invited talk in LG AI Research (Seoul, South Korea)
- 06/2021 *Artificial Intelligence and our Future*  
Invited talk in Business School at Sogang University (Seoul, South Korea)
- 05/2021 *Representation Learning and Beyond*  
Invited talk in Computer Science Department at Sogang University (Seoul, South Korea)
- 01/2021 *Life with Artificial Intelligence*  
Invited talk at Dongwon Corporation (Seoul, South Korea)
- 07/2020 *Artificial Intelligence: Past, Present, and Future*  
Invited talk in the National Assembly of the Republic of Korea (Seoul, South Korea)
- 07/2020 *AI Startup Hackathon*  
Guest Speaker in Jayang Life Academy (Seoul, South Korea)
- 07/2020 *Well-Tempered Artificial Intelligence in the Information Age*  
Invited talk in Jayang Life Academy (Seoul, South Korea)
- 11/2019 *Modeling the World × Modeling the Bias: The Chinese Voting Process and Beyond*  
Invited talk in Rennes School of Business (Rennes, France)
- 05/2019 *Modeling the World × Modeling the Bias: The Chinese Voting Process and Beyond*  
Invited talk in Korean Advanced Institute of Science and Technology (Daejeon, South Korea)
- 05/2018 *Toward Interpretable Machine Learning*  
Invited talk in Microsoft Research (Redmond, WA)
- 02/2018 *Toward Transdisciplinary Machine Learning – Scalable Text Mining and Social Influence Modeling*  
Invited talk at the University of Illinois Chicago (Chicago, IL)
- 09/2017 *Spectral Topic Modeling and the Chinese Voting Process*  
Invited talk at Chungnam National University (Daejeon, South Korea)
- 01/2017 *The Chinese Voting Process: The Evolution of Helpfulness in Product Reviews and Question Answers*  
Invited Best Paper Talk in Amazon Research Symposium (Seattle, WA)
- 04/2016 *Robust Spectral Inference for Joint Stochastic Matrix Factorization*  
Invited Talk at seminar series in Scientific Computing and Numerics (Ihtaca, NY)
- 11/2015 *Robust Spectral Inference for Joint Stochastic Matrix Factorization and Topic Modeling*  
Invited Talk in Microsoft Research (Redmond, WA)

## PROFESSIONAL SERVICE

Neural Information Processing Systems (NeurIPS 2015-2024)  
International Conference on Machine Learning (ICML 2020-2024)  
Association for Computational Linguistics (ACL 2016-2025)  
North American Chapter of the Association for Computational Linguistics (NAACL 2015-2025)  
Empirical Methods in Natural Language Processing (EMNLP 2014-2025)  
International Conference on Learning Representation (ICLR 2020-2024)  
Conference on Large Language models (COLM 2024-2025)  
Journal of Machine Learning Research (JMLR 2015-2019)  
European Association for Computational Linguistics (EACL 2016)  
Transactions of the Association for Computational Linguistics (TACL 2015)  
Association for the Advancement of Artificial Intelligence (AAAI 2020-2022)  
Artificial Intelligence and Statistics (AISTATS 2020-2021)  
Management Information Systems (MIS Quarterly 2018-2020)  
Management of Science (MS 2018-2019)  
Transaction on Social Computing (TSC 2019)

## PATENTS

- 11/2024 *Method and System for Providing Specialized Document Sharing Platform*  
US Patent 0148506A1, CN Patent 119234229, KR Patent 016289
- 01/2023 *Auto-encoding Device for Synthesizable Molecule Generation Model Taking into Consideration Molecular Structure Conditions and Molecule Generation Method Using Same*  
US Patent 0077867A1, EP Patent 04506950, KR Patent 016186
- 11/2022 *Learning Processing Device and Learning Processing Method for Pooling Hierarchically Structured Graph Data on Basis of Grouping Matrix, and Method of for Training Artificial Intelligence Model*  
US Patent 0077621A1, CN Patent 119234229, KR Patent 016289
- 03/2016 *Computational-model Operation Using Multiple Subject Representations*  
US Patent 10592519B2, US Patent 0286494A1

## REFERENCES

David Mimno (mimno@cornell.edu)  
Professor | Information Science | Cornell University

David Bindel (bindel@cs.cornell.edu)  
Professor | Computer Science | Cornell University

Peter Frazier (pf98@cornell.edu)  
Professor | Operations Research and Information Engineering | Cornell University

Jianfeng Gao (jfgao@microsoft.com)  
Distinguished Scientist and Vice President | Microsoft Research

Sungyong Park (parksy@sogang.ac.kr)  
Professor | Computer Science | Sogang University

Jungyun Seo (seojoy@sogang.ac.kr)  
Professor | Computer Science | Sogang University

Kyunghoon Bae (email me for reference)  
Minister of Science and ICT | Government of the Republic of Korea