

Joel Jang

joeljang.github.io ♦ joeljang@cs.washington.edu

EDUCATION

Ph.D. in Computer Science

University of Washington

Advisor: [Luke Zettlemoyer](#), [Dieter Fox](#)

Seattle, US

09/2023 -

M.S. in Artificial Intelligence

Korea Advanced Institute of Science and Technology (KAIST)

Advisor: [Minjoon Seo](#)

Seoul, Korea

03/2021 - 08/2023

B.S. in Computer Science and Engineering

Korea University

Seoul, Korea

03/2017 - 02/2021

PUBLICATIONS

Preprints

[P5] Latent Action Pretraining from Videos

Seonghyeon Ye*, **Joel Jang***, Byeongguk Jeon, Sejeun Joo, Jianwei Yang, Baolin Peng, Ajay Mandlekar, Reuben Tan, Yu-Wei Chao, Bill Yuchen Lin, Lars Liden, Kimin Lee, Jainfeng Gao, Luke Zettlemoyer, Dieter Fox, Minjoon Seo

Under Review at ICLR 2025,

Best Paper at CoRL 2024 LangRob Workshop (out of 75 accepted papers)

[P4] HAMSTER: Hierarchical Action Models for Open-World Robot Manipulation

Yi Li*, Yuquan Deng*, Jesse Zhang*, **Joel Jang**, Marius Memmel, Caelan Garrett, Fabio Ramos, Dieter Fox, Anqi Li, Abhishek Gupta, Ankit Goyal

Under Review at ICLR 2025

[P3] Knowledge Entropy Decay During Language Model Pretraining Hinders New Knowledge Acquisition

Jiyeon Kim, Hyunji Lee, Hyowon Cho, **Joel Jang**, Hyeonbin Hwang, Seungpil Won, Youbin Ahn, Dohaeng Lee, Minjoon Seo

Under Review at ICLR 2025

[P2] GRAPE: Generalizing Robot Policy via Preference Alignment

Zijian Zhang*, Kaiyuan Zheng*, Zhaorun Chen, **Joel Jang**, Yi Li, Chaoqi Wang, Minyu Ding, Dieter Fox, Huaxiu Yao

Under Review at CVPR 2025

[P1] Camels in a Changing Climate: Enhancing LM Adaptation with Tulu 2

Hamish Ivison, Yizhong Wang, Valentina Pyatkin, Nathan Lambert, Matthew Peters, Pradeep Dasigi, **Joel Jang**, David Wadden, Noah A. Smith, Iz Beltagy, Hannaneh Hajishirzi

Technical Report

Conference Papers

[C14] Semiparametric Token-Sequence Co-Supervision

Hyunji Lee*, Doyoung Kim*, Jihoon Jun, Sejeun Joo, **Joel Jang**, Kyoung-Woon Oh, Minjoon Seo
ACL 2024

[C13] LangBridge: Multilingual Reasoning Without Multilingual Supervision

Dongkeun Yoon, **Joel Jang**, Sungdong Kim, Seungone Kim, Sheikh Shafayat, Minjoon Seo
ACL 2024

[C12] Exploring the Practicality of Generative Retrieval on Dynamic Corpora

Chaeun Kim*, Soyoung Yoon*, Hyunji Lee, **Joel Jang**, Sohee Yang, Minjoon Seo
EMNLP 2024

[C11] How Well Do Large Language Models Truly Ground?

Hyunji Lee, SeJune Joo, Chaeun Kim, **Joel Jang**, Doyoung Kim, Kyoungwoon On, Minjoon Seo
NAACL 2024

[C10] Prometheus: Inducing Fine-grained Evaluation Capability in Language Models

Seungone Kim*, Jamin Shin*, Yejin Cho*, **Joel Jang**, Shayne Longpre, Hwaran Lee, Sangdoo Yun, Seongjin Shin, Sungdong Kim, James Thorne, Minjoon Seo
ICLR 2024

[C9] The CoT Collection: Improving Zero-shot and Few-shot Learning of Language Models via Chain-of-Thought Fine-Tuning

Seungone Kim*, Se June Joo*, Doyoung Kim, **Joel Jang**, Seonghyeon Ye, Jamin Shin, Minjoon Seo
EMNLP 2023

[C8] Retrieval of Soft Prompt Enhances Zero-shot Task Generalization

Seonghyeon Ye, **Joel Jang**, Doyoung Kim, Yongrae Jo, Minjoon Seo
EMNLP 2023 Findings [paper] [code]

[C7] Knowledge Unlearning for Mitigating Privacy Risks in Language Models

Joel Jang, Dongkeun Yoon, Sohee Yang, Sungmin Cha, Moontae Lee, Lajanugen Logeswaran, Minjoon Seo
ACL 2023

[C6] Gradient Ascent Post-training Enhances Language Model Generalization

Dongkeun Yoon*, **Joel Jang***, Sungdong Kim, Minjoon Seo
ACL 2023

[C5] Prompt Injection: Parameterization of Fixed Inputs

Eunbi Choi, Yongrae Jo, **Joel Jang**, Joonwon Jang, Minjoon Seo
ACL 2023 Findings

[C4] 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, Minjoon Seo
ICML 2023

[C3] Guess the Instruction! Making Language Models Stronger Zero-shot Learners

Seonghyeon Ye, Doyoung Kim, **Joel Jang**, Joongbo Shin, Minjoon Seo
ICLR 2023

[C2] TemporalWiki: A Lifelong Benchmark for Training and Evaluating Ever-Evolving Language Models

Joel Jang*, Seonghyeon Ye*, Changho Lee, Sohee Yang, Joongbo Shin, Janghoon Han, Gyeonghun Kim, Minjoon Seo
EMNLP 2022

[C1] Towards Continual Knowledge Learning of Language Models

Joel Jang, Seonghyeon Ye, Sohee Yang, Joongbo Shin, Janghoon Han, Gyeonghun Kim, Stanley Jungkyu Choi, Minjoon Seo
ICLR 2022

Workshop Papers

[W2] Personalized Soups: Personalized LARge Language Model Alignment via Post-hoc Parameter Merging

Joel Jang, Seungone Kim, Bill Yuchen Lin, Yizhong Wang, Jack Hessel, Luke Zettlemoyer, Hannaneh Hajishirzi, Yejin Choi, Prithviraj Ammanabrolu
NeurIPS 2024 Workshop on Adaptive Foundation Models (Oral)

[W1] Can Large Language Models Truly Follow Your Instructions? Case-study with Negated Prompts

Joel Jang*, Seonghyeon Ye*, Minjoon Seo
NeurIPS 2022 Workshop on Transfer Learning for NLP (TL4NLP)

Journal Papers

[J3] Improving Probability-based Prompt Selection Through Unified Evaluation and Analysis
Sohee Yang, Jonghyeon Kim, **Joel Jang**, Seonghyeon Ye, Hyunji Lee, Minjoon Seo
TACL 2024

[J2] Sequential Targeting: A Continual Learning Approach for Data Imbalance in Text Classification
Joel Jang, Yoonjeon Kim, Kyoungcho Choi, Sungho Suh
Expert Systems with Applications (2021)

[J1] Supervised Health Stage Prediction Using Convolution Neural Networks for Bearing Wear
Sungho Suh, **Joel Jang**, Seungjae Won, Mayank S. Jha, Yong Oh Lee
Sensors (2020)

EXPERIENCE

Nvidia Generalist Embodied Agent Research (GEAR) Lab San Jose, US
Research Intern (Mentors: : Scott Reed, Jim Fan, and Yuke Zhu) 10/2024- Present
Scaling Robot Foundation

Nvidia Robotics Lab Seattle, US
Research Intern (Mentors: : Dieter Fox and Ajay Mandlekar) 03/2024 - 09/2024
Scaling Robot Foundation

Allen Institute of AI (Ai2) Seattle, US
Research Intern (Mentors: : Prithviraj Ammanabrolu, Yuchen Lin, Yejin Choi) 06/2023 - 01/2024
(1) Personalized RLHF of LLMs and (2) Democratizing Robotic Foundational Models

LG AI Research Seoul, Korea
Research Intern (Mentors : Moontae Lee, Lajanugen Logeswaran) 07/2022 - 05/2023
(1) Knowledge Unlearning for LLMs and (2) Expert Language Models

Kakao Brain Seongnam, Korea
Research Intern (Mentor : Ildoo Kim) 12/2020 - 02/2021
Representation learning of vision-language models with weakly supervision

NAVER Seongnam, Korea
Software Engineer Intern 07/2020 - 09/2020
Worked on hate speech detection model, AI Clean Bot 2.0 (40+ million monthly users, >80% of Korean population)
Developed novel method of handling data imbalance using continual learning (*paper published under ESWA*)

KIST Europe Saarbrücken, Germany
Research Intern (Mentor : Sungho Suh) 08/2019 - 01/2020
Worked on anomaly detection & remaining useful life prediction of machinery (*paper published under Sensors*)
Gave an Oral Presentation at *PHM Korea 2020* (2020. 07. 23)

MENTORING

KAIST AI

Dongkeun Yoon (2022 - 2023), Konkuk University Undergrad. Published [C6, C7] Now Ph.D. at KAIST AI
Seungone Kim (2022 - 2023), Yonsei University Undergrad. Published [C4]. Now Ph.D. at CMU
Changho Lee (2021 - 2022), Korea University Undergrad. Published [C2]. Now Research Scientist at LG AI Research
Seonghyeon Ye (2021 - 2022), KAIST Undergrad. Published [C1,C2,W1]. Now Ph.D. at KAIST AI

TEACHING

(KAIST AI620) NLP Bias and Ethics <i>Teaching Assistant (TA) Instructor: James Throne</i>	<i>Spring 2023</i>
(KAIST AI599) AI for Law <i>Teaching Assistant (TA) Instructor: Minjoon Seo</i>	<i>Fall 2022</i>
(KAIST AI605) Deep Learning for NLP <i>Teaching Assistant (TA) Instructor: Minjoon Seo</i>	<i>Spring 2022</i>

INVITED TALKS

Latent Action Pretraining from Videos <i>OpenAI (Host: Adam Fishman)</i>	10/2024
Expert Language Models <i>UNC at Chapel Hill (Host: Colin Raffel)</i>	02/2023
Temporal Adaptation of Language Models <i>Korean AI Association Summer NLP Session (Host: Minjoon Seo)</i>	08/2022
Temporal Adaptation of Language Models <i>KAIST School of Computing (Host: Alice Oh)</i>	07/2022
Temporal Adaptation of Language Models <i>Hyperconnect (Host: Buru Chang)</i>	05/2022

HONORS AND AWARDS

Qualcomm Innovation Fellowship Korea (QIFK) 2022
Grand Prize in Undergraduate Capstone Competition (Best Paper Award), 2020 (*Advisor: [Jaewoo Kang](#)*)
Future Global Leader Scholarships, Korea University, 2019
Best Innovation Award, Intel AI Drone Hackathon, 2018

LANGUAGE PROFICIENCY

Bilingual in English (2004-2016 in US) and Korean (*native*)
GRE: 326 (Verbal, 157/170, 76th Percentile) | Quant, 169/170, 95th Percentile | Writing, 5.0/6.0, 92nd Percentile)
TOEFL: 119/120 (Reading, 30 | Listening, 30 | Speaking, 29 | Writing, 30)
SAT: 1530/1600 (Reading and Writing, 730 | Math, 800)
Conversational in Chinese