

# Joel Jang

Last updated on Oct 05, 2022

[joeljang.github.io](https://joeljang.github.io) ♦ [joeljang@kaist.ac.kr](mailto:joeljang@kaist.ac.kr)

## RESEARCH INTERESTS

---

Temporal/Domain Adaptation of Large Language Models, Continual Learning, Knowledge-Intensive Tasks

## EDUCATION

---

### M.S. & Ph.D. (Integrated) in Artificial Intelligence

KAIST, Graduate School of AI | [Language & Knowledge Lab](#)

Advisor: [Minjoon Seo](#)

Seoul, Korea

March 2021 – Present

### B.S. in Computer Science and Engineering

Korea University

Advisors: [Jaewoo Kang](#), [Heuiseok Lim](#)

Seoul, Korea

March 2017 – February 2021

## PUBLICATIONS

---

### Preprints

[P6] Guess the Instruction! Making Language Models Stronger Zero-shot Learners

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

Submitted ICLR 2023

[P5] Knowledge Unlearning for Mitigating Privacy Risks in Language Models

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

Submitted to ICLR 2023

[P4] Prompt Injection: Parameterization of Fixed Inputs

Eunbi Choi, Yongrae Jo, **Joel Jang**, Minjoon Seo

Submitted to ICLR 2023, [\[paper\]](#)[\[code\]](#)

[P3] Can Language Models Truly Follow Your Instructions? Case-study with Negated Prompts

**Joel Jang\***, Seonghyeon Ye\*, Minjoon Seo

Submitted to NeurIPS 2022 Workshop

[P2] Retrieval of Prompt Tuning Enhances Zero-shot Task Generalization

Seonghyeon Ye, **Joel Jang**, Doyoung Kim, Yongrae Jo, Minjoon Seo

Submitted to EMNLP 2022

[P1] TemporalWiki: A Lifelong Benchmark for Training and Evaluation Ever-Evolving Language Models

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

Submitted to EMNLP 2022, [Spa-NLP @ ACL 2022 \(poster\)](#) [\[paper\]](#)[\[code\]](#)

### Peer-Reviewed Conference Papers

[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 (*poster*), [Spa-NLP @ ACL 2022 \(oral\)](#) [\[paper\]](#) [\[code\]](#)

### Peer-Reviewed Journal Papers

[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) [[paper](#)] [[code](#)]

[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) [[paper](#)] [[code](#)]

(\* denotes equal contribution)

## WORK IN PROGRESS

---

[WIP4] Generating the Rationale through Intermediate Instructions Amplifies the Emergent Reasoning Capabilities of Large Language Models  
Seungone Kim, Hyungjoo Chae, SeJune Joo, Doyoung Kim, **Joel Jang**, Yongho Song, Jinyoung Yeo  
*Targeting ACL 2023*

[WIP3] Do you remember me? Conversation Injection for Continually Learning Chat-agent  
Eunbi Choi, Joonwon Jang, **Joel Jang**, Minjoon Seo  
*Targeting ACL 2023*

[WIP2] Retrieval of Experts for Zero-shot Task Generalization  
**Joel Jang**, Seungone Kim, Seonghyeon Ye, Kyungjae Lee, Moontae Lee, Minjoon Seo  
*Targeting ACL 2023*

[WIP1] Why doesn't your prompt work?  
Sohee Yang, Jonghyeon Kim, **Joel Jang**, Seonghyeon Ye, Hyunji Lee, Sangwoo Lee, Minjoon Seo  
*Targeting ICLR 2023*

(\* denotes equal contribution)

## EXPERIENCES

---

### LG AI Research

Seoul, Korea

Research Intern (Mentor : [Moontae Lee](#), [Lajanugen Logeswaran](#))

July 2022 – Present

Working on developing (1) unlearning for LMs and (2) instruction following LMs that can continually learn new tasks.

### Kakao Brain

Seongnam, Korea

Research Intern (Mentor : [Ildoo Kim](#))

December 2020 – February 2021

Worked on large-scaled representation learning with weakly supervision of images and caption data using TPUs.

### NAVER Corp. | Media Tech Group

Seongnam, Korea

Software Engineer Intern

July 2020 – September 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*)

### Korea Institute of Science and Technology European Research Centre

Saarbrücken, Germany

Research Intern (Mentor : [Yong Oh Lee](#))

August 2019 – January 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)

## HONORS AND AWARDS

---

Grand Prize in Graduation Capstone Competition (Best Paper Award), 2020 (*Advisor: [Jaewoo Kang](#)*)  
4<sup>th</sup> place, AI NLP Challenge Enliple Cup, 2020  
3<sup>rd</sup> place, HAAFOR Challenge 2019  
Future Global Leader Scholarships, Korea University, 2019  
Best Innovation Award, Intel AI Drone Hackathon, 2018

## SERVICES

---

### Conference Reviewer

*COLING 2022, EMNLP 2022, AKBC 2022*

### Journal Reviewer

*Journal of Artificial Intelligence Research (JAIR)*

## TEACHING

---

(KAIST AI599) AI for Law  
*Teaching Assistant (TA)*

*Fall 2022*

(KAIST AI605) Deep Learning for NLP  
*Teaching Assistant (TA)*

*Spring 2022*

## INVITED TALKS

---

Temporal Adaptation of Language Models  
*Korean AI Association Summer NLP Session (Host: Minjoon Seo)*

*August 2022*

Temporal Adaptation of Language Models  
*KAIST School of Computing (Host: Alice Oh)*

*July 2022*

Temporal Adaptation of Language Models  
*Hyperconnect (Host: Buru Chang)*

*May 2022*

## TECHNICAL STRENGTHS

---

Coding	Tensorflow, Pytorch, Huggingface, Pytorch-Lightning, Deepspeed, Wandb
Others	Large-scale models, Multi-node parallel training, Spot VM instances, Amazon Mechanical Turk

## LANGUAGE PROFICIENCY

---

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