

# Junmo Kang

Ph.D. Student at Georgia Tech

✉ junmo.kang@gatech.edu | 🌐 jm-kang.github.io

## RESEARCH INTERESTS

---

### NATURAL LANGUAGE PROCESSING (NLP)

LARGE LANGUAGE MODELS, RETRIEVAL-BASED NLP, EFFICIENCY, ROBUSTNESS

The long-term goal of my research is to enhance the *practicality* of NLP systems (e.g., Large Language Models) so that they can be widely used *in real-world scenarios*. In particular, my research focuses on *i) how to make NLP models cheaper (i.e., efficient) in terms of data, parameters, or compute*, and *ii) how to build NLP models that are robust to unseen cases in the wild*.

## EDUCATION

---

### GEORGIA INSTITUTE OF TECHNOLOGY

PH.D. IN COMPUTER SCIENCE

Atlanta, GA, USA

Aug. 2022 - Present

- Research Assistant at NLP Lab (Advisors: Dr. Alan Ritter, Dr. Wei Xu)
- GPA: 4.0 / 4.0

### KOREA ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY (KAIST)

M.S. IN COMPUTER SCIENCE

Daejeon, Republic of Korea

Feb. 2019 - Feb. 2021

- Research Assistant at IR&NLP Lab (Advisor: Dr. Sung-Hyon Myaeng)
- Thesis committee: Dr. Sung-Hyong Myaeng, Dr. Hojin Choi, Dr. Alice Oh
- GPA: 4.03 / 4.30

### CHUNGNAM NATIONAL UNIVERSITY

B.E. IN COMPUTER SCIENCE & ENGINEERING

Daejeon, Republic of Korea

Mar. 2012 - Feb. 2019

- Summa Cum Laude
- GPA: 4.30 / 4.50 (Rank: 1/125 in CSE, Major GPA: 4.41)

## PUBLICATIONS

---

- [1] Discern and Answer: Mitigating the Impact of Misinformation in Retrieval-Augmented Models with Discriminators  
Giwon Hong\*, Jeonghwan Kim\*, Junmo Kang\*, Sung-Hyon Myaeng, Joyce Jiyoung Whang [pdf] Preprint
- [2] Distill or Annotate? Cost-Efficient Fine-Tuning of Compact Models  
Junmo Kang, Wei Xu, Alan Ritter [pdf] ACL 2023
- [3] Graph-Induced Transformers for Efficient Multi-Hop Question Answering  
Giwon Hong, Jeonghwan Kim, Junmo Kang, Sung-Hyon Myaeng [pdf] EMNLP 2022
- [4] Exploiting Numerical-Contextual Knowledge to Improve Numerical Reasoning in Question Answering  
Jeonghwan Kim, Junmo Kang, Giwon Hong, Kyung-min Kim, Sung-Hyon Myaeng [pdf] Findings of NAACL 2022
- [5] Ultra-High Dimensional Sparse Representations with Binarization for Efficient Text Retrieval  
Kyoung-Rok Jang, Junmo Kang, Giwon Hong, Sung-Hyon Myaeng, Joohee Park, Taewon Yoon, Heecheol Seo [pdf] EMNLP 2021

- [6] **Leveraging Order-Free Tag Relations for Context-Aware Recommendation** EMNLP 2021  
Junmo Kang, Jeonghwan Kim, Suwon Shin, Sung-Hyon Myaeng [pdf]
- [7] **Have You Seen That Number? Investigating Extrapolation in Question Answering Models** EMNLP 2021  
Jeonghwan Kim, Giwon Hong, Kyung-min Kim, Junmo Kang, Sung-Hyon Myaeng [pdf]
- [8] **Can You Distinguish Truthful from Fake Reviews? User Analysis and Assistance Tool for Fake Review Detection** HCI+NLP@EACL 2021  
Jeonghwan Kim\*, Junmo Kang\*, Suwon Shin\*, Sung-Hyon Myaeng [pdf]
- [9] **Regularization of Distinct Strategies for Unsupervised Question Generation** Findings of EMNLP 2020  
Junmo Kang\*, Giwon Hong\*, Haritz Puerto San Roman\*, Sung-Hyon Myaeng [pdf]
- [10] **Handling Anomalies of Synthetic Questions in Unsupervised Question Answering** COLING 2020  
Giwon Hong\*, Junmo Kang\*, Doyeon Lim\*, Sung-Hyon Myaeng [pdf]
- [11] **Let Me Know What to Ask: Interrogative-Word-Aware Question Generation** MRQA@EMNLP 2019  
Junmo Kang\*, Haritz Puerto San Roman\*, Sung-Hyon Myaeng [pdf]

\* indicates equal contribution.

## EXPERIENCES

---

### MIT-IBM WATSON AI LAB

RESEARCH INTERN

*Cambridge, MA, USA*

May. 2023 - Aug. 2023 (Expected)

- Working on large language models.

### GEORGIA TECH NLP LAB

GRADUATE RESEARCH ASSISTANT

*Atlanta, GA, USA*

Aug. 2022 - Present

- Worked on cost-efficiency analysis of small and large language models [2].

### KAIST IR&NLP LAB

RESEARCH ASSOCIATE

*Daejeon, Republic of Korea*

Mar. 2021 - Jul. 2022

- Worked on robust retrieval-augmented LMs [1].
- Worked on efficient methods for multi-hop QA [3].

### KAIST IR&NLP LAB

GRADUATE RESEARCH ASSISTANT

*Daejeon, Republic of Korea*

Feb. 2019 - Feb. 2021

- Worked on question generation and unsupervised question answering for data-efficiency [9,10,11].
- Presented sample-efficient and robust number representations for question answering [7,4].
- Proposed a novel generation model that takes into account the inter-dependency of tags while alleviating the order sensitivity [6].
- Proposed a novel sparse representation model for passage retrieval that can take advantage of an efficient inverted index and symbolic IR techniques [5].

### REPUBLIC OF KOREA ARMY

HONORABLY DISCHARGED AS SERGEANT

*Republic of Korea*

Apr. 2013 - Jan. 2015

- Compulsory military service.

## HONORS & AWARDS

---

Graduated with Highest Honor in CSE, Chungnam National University

2019

Grand Prize, Business ICT Competition	2018
Excellence Award, Startup Competition	2018
Finalist, NAVER AI Hackathon	2018
Grand Prize, Daejeon Startup School	2017
Best Excellence Award, Startup Picnic	2016
Finalist, Microsoft Imagine Cup Korea	2016