

SEOKHEE HONG

seokhee.hong@vision.snu.ac.kr hongcheki.github.io

RESEARCH INTEREST

My research interest is in the field of **Natural Language Processing (NLP)**, especially **Fairness, Safety, and Trustworthiness in NLP**. My ultimate research goal is to develop a fairer and safer NLP model.

EDUCATION

M.S. in Computer Science and Engineering, Seoul National University Mar. 2021 - Aug. 2023
GPA: 4.0/4.3; Advisor: Prof. Gunhee Kim

B.S. in Computer Science, Yonsei University Mar. 2014 - Feb. 2021
GPA: 4.13/4.3 (*Highest Honor*)

PUBLICATIONS

Who Wrote this Code? Watermarking for Code Generation

Taehyun Lee*, Seokhee Hong*, Jaewoo Ahn, Ilgee Hong, Hwaran Lee, Sangdoon Yun, Jamin Shin[†], Gunhee Kim[†]
arXiv preprint (under review)

SQuARe: A Large-Scale Dataset of Sensitive Questions and Acceptable Responses Created Through Human-Machine Collaboration

Hwaran Lee*, Seokhee Hong*, Joonsuk Park, Takyoung Kim, Meeyoung Cha, Yejin Choi, Byoung Pil Kim, Gunhee Kim, Eun-Ju Lee, Yong Lim, Alice Oh, Sangchul Park, Jung-Woo Ha
ACL 2023 (Oral)

KoSBI: A Dataset for Mitigating Social Bias Risks Towards Safer Large Model Application

Hwaran Lee*, Seokhee Hong*, Joonsuk Park, Takyoung Kim, Gunhee Kim, Jung-Woo Ha
ACL 2023 (Industry Track)

How Robust are Fact Checking Systems on Colloquial Claims?

Byeongchang Kim*, Hyunwoo Kim*, Seokhee Hong, Gunhee Kim
NAACL 2021

* : equal contribution / [†] : corresponding authors

RESEARCH AND WORK EXPERIENCES

EXAONE Lab, LG AI Research Seoul, Korea
• *AI Scientist* Nov. 2023 - Present

Vision&Learning Lab, Seoul National University Seoul, Korea
• *Research Intern* (Advisor: Gunhee Kim) Jun. 2020 - Feb. 2021

AWARDS AND HONORS

Graduated with Highest Honor, Yonsei University, Korea Feb 2021

National Science and Technology Scholarship (full tuition), Ministry of Science, Korea Fall 2019, Fall 2020

Academic Excellence Scholarship, 용인시인재육성재단, Korea Spring 2019, Spring 2020

TEACHING EXPERIENCE

Teaching Assistant of Samsung DS Reinforcement Learning
Teaching Assistant of Computer Vision (M1522.001000)
Teaching Assistant of Graduate Artificial Intelligence (M4190.569)

Spring 2022, Spring 2023
Fall 2021
Spring 2021