

# Seongung Jo

## Personal Data

---

Phone	+82 10-4126-5266
Email	<a href="mailto:ung120266@gmail.com">ung120266@gmail.com</a> (or <a href="mailto:@kaist.ac.kr">@kaist.ac.kr</a> )
Research Interests	<ul style="list-style-type: none"><li>• Natural Language Processing (NLP)</li><li>• Human AI/Computer Interaction</li><li>• LLM operation system</li></ul>



## Education & Career

---

Bachelor's degree in computer science	Computer science dept., Korea University of Technology and Education (KOREATECH), Cheonan, Korea (Mar 2015 to Sept 2020, 7 semesters)
Master's degree in computer science	Computer science dept., KOREATECH, Cheonan, Korea (Sept 2020 to Dec 2022, 3 semesters)
Research Engineer	Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea (Feb 2024 to Sep 2024)
Research Internship	Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea (Sep 2024 to Jan 2025)

## Research Experience

---

Research Assistant (Jan 2019 to Dec 2022)	DICE-lab, KOREATECH, Cheonan, Korea <ul style="list-style-type: none"><li>- Mostly participating in the ML/AI projects during this period<ul style="list-style-type: none"><li>• Hierarchical Text Classification (HTC) model</li><li>• Medical LM (QA) model</li><li>• Document-level IE &amp; QA system</li><li>• Natural Question Generation (NQG) research</li></ul></li></ul>
Research Engineer (Feb 2024 to Sep 2024)	KAIST IT & Convergence Research Center (CRC), Daejeon, Korea <ul style="list-style-type: none"><li>- Mostly participating in the LLM application/Geospatial system projects during this period<ul style="list-style-type: none"><li>• Indoor Positioning System (IPS) relative tools and model</li></ul></li></ul>

---

	<ul style="list-style-type: none"> <li>Geospatial LLM research</li> </ul>
Research Internship (Sep 2024 to Now)	KAIST ICLab. (Prof. Uichin Lee), Daejeon, Korea - Mostly participating in the Human-Computer Interaction (HCI) / LLM agent projects during this period <ul style="list-style-type: none"> <li>LLM Ethics (privacy issues while using LLM) research</li> <li>LLM-based Data Agent system</li> </ul>

## Project Experience

---

Medical QA System (Mar 2019 to Feb 2022)	의료 인공지능을 위한 딥러닝 기반 의료 질의응답 기술 연구 (NLP, AI)
Document Intelligence (Mar 2020 to Sept 2020)	신기술 분야 기술적, 사회적 현안 추출 모델 연구(NLP, AI)
Natural Question Generation (NQG) (June 2022 to Dec 2022)	사전 학습 언어모델 기반 질의응답 데이터 셋 생성모델 개발 (NLP, AI)
Urban AI (Feb 2024 to Sep 2024)	대단위 입체공간 주소기반 실내 내비게이션 구현 (IoT, AI) 융-복합 데이터 활용 실감형 소방안전도시 구축(IoT, AI)
HCI/LLM Ethics (Sep 2024 to Dec 2024)	AI 위험 분야별 안정성 평가를 위한 데이터셋 구축(AI, HCI)
Others	(Detailed in the below docs) <ul style="list-style-type: none"> <li><a href="#">Research field</a> (Portfolio)</li> <li>Mainly participated in <b><u>NLP</u></b> (IE, Document Intelligence, QA system, LLM)</li> <li><b><u>AI Usage</u></b> (3D-Map-DT &amp; LLM application) and <b><u>Human-Computer Interaction</u></b> (HCI)</li> </ul>

## Publications

---

- Paper**

S.H. Im, G.B Kim, H.S. Oh, **S.U. Jo** and D.H. Kim, “Hierarchical Text Classification as Sub-Hierarchy Sequence Generation, AAI, now accepted, 2023 (Q1, SCI)

**S.U. Jo**, H.S. Oh, S.H. Im, G.B. Kim, and S.H. Kim, “Information Extraction Based on Multi-turn Question Answering for Analyzing Korean Research Trends”, Journal of Computers, Materials & Continua, no. 2. pp. 2967-2980, 2022 (Q2, SCI)

---

G.B. Kim, S.H. Im, **S.U. Jo**, D.H. Kim, and H.S. Oh, “Analysis of Knowledge Neurons in Text Classification using Pretrained Language Models”, 한국정보과학회, 학술발표 논문집 pp.729-731, 2022

**S.U. Jo**, H.S. Oh, S.H. Im, S.H. Kim, “CNN-based Novelty Detection with Effectively Incorporating Document-level Information”, KIPS Transactions on Computer and Communication Systems, vol. 9, no.10. pp. 231-238, 2020

K.M. Kim, D. H. Kim, **S.U. Jo**, H.S. Oh, M.H. Hwang, “TAGS: Text Augmentation with Generation and Selection”, KIPS Transactions on Software and Data Engineering, (accepted), 2023

**S.U. Jo**, H.S. Oh, “Multi-aspect Verifier (MaV): Enhancing Relation Extraction for Long Documents with Prompt Decomposition and Multi-task Learning”, JDCS Journal of Digital Contents Society, (accepted), 2023

- **Patent**

H.S. Oh, S.H. Im and **S.U Jo**, 계층적 다중 레이블 분류를 위한 데이터 분류 방법 및 장치 (DATA CLASSIFICATION METHOD AND APPARATUS FOR HIERARCHICAL MULTI-LABEL CLASSIFICATION)

## Challenges

---

- \* **Domestic challenge**

1. 인공지능 언어능력 평가, 국립국어원, 2021

- Won The 2<sup>nd</sup> Place.
- Developed language model & methodologies for Korean SuperGLUE
  - Investigating back translation (en to kr) based data augmentation methods.
  - Investigating novel dropout methods with embedding similarity for WiC.

2. AI Grand Challenge, 2019

- Track-1 17<sup>th</sup> Place.
- Developed language model & methodologies for Korean SuperGLUE
  - Investigating back translation (en to kr) based data augmentation methods.
  - Investigating novel dropout

3. 인공지능 문제해결 경진대회, 2020

- 민원 분류 모델(육군) 및 실시간 댓글 분류 모델(SBS) 개발

4. 한국어 문서 생성 및 추출 요약 경진대회, 2020

5. 관광데이터 AI 경진대회, 2022

- 7<sup>th</sup> Place.

**\* International challenge**

1. BioASQ-8b, PubMed, 2020

- Won 1<sup>st</sup> Place, 3<sup>rd</sup> Place on the sub-task (yes/no, factoid)
- Developed language model & methodologies for medical QA task
  - Investigating medical pre-trained language model based on ELECTRA.

2. Dialog System Technology Challenge (DSTC), AWS AI, 2020&2022

- DSTC9 Track-1, DSTC11 Track-2

2. Visual Question Answering Challenge, AWS, 2021