

# Zae Myung Kim

200 Union St SE, Minneapolis, MN 55455, USA

✉ zaemyung@gmail.com | 🏠 zaemyung.github.io | 📧 zaemyung | 🌐 zaemyung | 🎓 Google Scholar

## Education

### University of Minnesota Twin Cities

PH.D. IN COMPUTER SCIENCE

- Advisor: Dongyeop Kang

Minneapolis, MN

Sept. 2022 - Present

### KAIST (Korea Advanced Institute of Science and Technology)

M.S. IN COMPUTER SCIENCE

- Advisor: Ho-Jin Choi
- Thesis: "Classifying Travel-Related Intents in Textual Data Using Recurrent Neural Networks"

Daejeon, S.Korea

Sept. 2014 - Aug. 2016

### Imperial College London

B.ENG. IN COMPUTER SCIENCE

- Group Project: "Intention Recognition in Ambient Homes"
- Individual Project: "Computing Fictitious Plays in Graph-Based Games"

London, UK

Oct. 2008 - Aug. 2011

## Selected Publications

### CONFERENCE PAPERS

#### Threads of Subtlety: Detecting Machine-Generated Texts Through Discourse Motifs

ZAE MYUNG KIM, KWANG HEE LEE, PRESTON ZHU, VIPUL RAHEJA, DONGYEOP KANG

Feb. 2024

Under review

#### Under the Surface: Tracking the Artifactuality of LLM-Generated Data

DEBARATI DAS\*, KARIN DE LANGIS\*, ANNA MARTIN-BOYLE\*, JAEHYUNG KIM\*, MINHWA LEE\*, ZAE MYUNG KIM\*, SHIRLEY ANUGRAH HAYATI, RISAKO OWAN, BIN HU, RITIK SACHIN PARKAR, RYAN KOO, JONG INN PARK, AAHAN TYAGI, LIBBY FERLAND, SANJALI ROY, VINCENT LIU, DONGYEOP KANG

Jan. 2024

Under review

#### Improving Iterative Text Revision by Learning Where To Edit From Other Revision Tasks

ZAE MYUNG KIM, WANYU DU, VIPUL RAHEJA, DHRUV KUMAR, DONGYEOP KANG

Dec. 2022

EMNLP 2022

#### Read, Revise, Repeat: A System Demonstration for Human-in-the-loop Iterative Text Revision

WANYU DU\*, ZAE MYUNG KIM\*, VIPUL RAHEJA, DHRUV KUMAR, DONGYEOP KANG [BEST PAPER]

May 2022

In2Writing at ACL 2022

#### Understanding Iterative Revision from Human-Written Text

WANYU DU, VIPUL RAHEJA, DHRUV KUMAR, ZAE MYUNG KIM, MELISSA LOPEZ, DONGYEOP KANG

May 2022

ACL 2022

#### Visualizing Cross-Lingual Discourse Relations in Multilingual TED Corpora

ZAE MYUNG KIM, VASSILINA NIKOULINA, DONGYEOP KANG, DIDIER SCHWAB AND LAURENT BESACIER

Nov. 2021

CODI 2021 at EMNLP 2021

#### Do Multilingual Neural Machine Translation Models Contain Language Pair Specific Attention Heads?

ZAE MYUNG KIM, LAURENT BESACIER, VASSILINA NIKOULINA, DIDIER SCHWAB

Aug. 2021

Findings of ACL 2021

#### A Multilingual Neural Machine Translation Model for Biomedical Data

ALEXANDRE BÉRARD, ZAE MYUNG KIM, VASSILINA NIKOULINA, EUNJEONG PARK, MATTHIAS GALLÉ

Dec. 2020

NLP-COVID19 at EMNLP 2020

#### PATQUEST: Papago Translation Quality Estimation

YUJIN BAEK\*, ZAE MYUNG KIM\*, JIHYUNG MOON, HYUNJOONG KIM, EUNJEONG PARK

Nov. 2020

WMT 2020

#### An adaptive vocabulary learning application through modeling learner's linguistic proficiency and interests

ZAE MYUNG KIM, SUIN KIM, ALICE OH, HO-JIN CHOI

Feb. 2017

Int'l Conf. on Big Data and Smart Computing

#### TIMEX3 and Event Extraction Using Recurrent Neural Networks

ZAE MYUNG KIM, YOUNG-SEOB JEONG, AND HO-JIN CHOI

Jan. 2016

Int'l Conf. on Big Data and Smart Computing

#### An approach to spam comment detection through domain-independent features

JONG MYOUNG KIM, ZAE MYUNG KIM, AND KWANGJO KIM

Jan. 2016

Int'l Conf. on Big Data and Smart Computing

#### Temporal Information Extraction from Korean Texts

YOUNG-SEOB JEONG, ZAE MYUNG KIM, HYUN-WOO DO, CHAE-GYUN LIM, AND HO-JIN CHOI

Jul. 2015

CoNLL 2015

## JOURNAL PAPERS

### Modeling Long-Term Human Activeness Using Recurrent Neural Networks for Biometric Data

ZAE MYUNG KIM, HYUNGRAI OH, HAN-GYU KIM, CHAE-GYUN LIM, KYO-JOONG OH, AND HO-JIN CHOI

May 2017  
*BMC Med Inform Decis Mak [SCIE, IF:2.042]*

### Investigating the Impact of Possession-Way of a Smartphone on Action Recognition

ZAE MYUNG KIM\*, YOUNG-SEOB JEONG\*, KYO-JOONG OH, HYUNG-RAI OH, YOUSSEF IRAQI, AND HO-JIN CHOI

Jun. 2016  
*Sensors, 16(6) [SCIE, IF:2.437]*

## Awards & Accomplishments

2022	<b>Awarded</b> , Best Paper, In2Writing workshop at ACL 2022	Dublin, Ireland
2022	<b>Awarded</b> , 3M Science and Technology Fellowships	Minneapolis, MN
2018	<b>Runner-Up</b> , Elice Data Science Challenge (3rd place out of 300 participants)	Seongnam, S.Korea
2016	<b>Awarded</b> , Outstanding Teaching Assistant Award, 2016 Data Science Workshop & Challenge at Elice	Geumsan, S.Korea
2016	<b>Awarded</b> , Outstanding Graduate Teaching Assistant Award, School of Computing, KAIST	Daejeon, S.Korea
2012	<b>Runner-Up</b> , Republic of Korea Army Oratorical Contest, ROK Army	Incheon, S.Korea
2011	<b>Winning Team</b> , Robotics Localization Challenge, Imperial College London	London, UK
2009	<b>Winning Team</b> , The Deutsche Bank Prize for Topics in A.I. Group Project, Imperial College London	London, UK
2008	<b>Participation</b> , Northwestern European Regional Programming Contest, ACM-ICPC	Utrecht, NL
2008	<b>Bronze Medal</b> , Royal College of Science Union Science Challenge ( <i>a literary contest</i> )	London, UK

## Work Experience

### Salesforce AI Research

RESEARCH INTERN

Palo Alto, CA  
May - Sept. 2023

- Developed a large-scale realistic benchmarking system for scientific paper retrieval task.
- Developed a co-writing system with large language models in the scientific paper writing domain.

### Grammarly

PHD INTERN

Remote Work  
Apr. - Jul. 2022

- Developed human-in-the-loop iterative text revision system.
- Developed span-based text revision models with data augmentation.

### NAVER LABS Europe

RESEARCHER

Meylan, France  
Oct. 2020 - Feb. 2022

- Researched on discovering language-specific representation of multilingual NMT models.
- Researched cross-lingual discourse analysis and its application to document-level NMT.

### Papago, NAVER Corp.

RESEARCHER

Seongnam, S.Korea  
Mar. 2017 - Sept. 2020

- Developed machine translation (MT) service for Thai-Korean and German-Korean language pairs.
- Developed utility programs for MT, such as sentence segmenter, corpus preprocessor, corpus aligner/crawler, etc.
- Developed website translation service based on NMT models, utilizing encoder-decoder attentions and heuristics.
- Participated in Quality Estimation shared task at WMT20; came 4th for sentence-level, 1st for document-level tasks.

### Knowledge Engineering and Artificial Intelligence Lab, KAIST

RESEARCHER

Daejeon, S.Korea  
Dec. 2016 - Feb. 2017

- Developed an adaptive vocabulary learning application that incorporates a learner's linguistic proficiency and topical interests.
- Developed a health-care application that suggests *Pokémon Go* events when a user's activeness is predicted to be low.

### NAVER LABS Corp.

RESEARCH INTERN

Seongnam, S.Korea  
Sept. - Nov. 2016

- Researched on conducting extended named entity recognition on Korean texts that recognizes hierarchical categories of an entity.

### Samsung-KAIST Academic Initiatives

RESEARCH ASSISTANT

Daejeon, S.Korea  
Apr. 2014 - Jul. 2016

- [1st Yr.] Investigated intent classification in textual data in the context of providing well-suited services in travel and tourism domain.
- [2nd Yr.] Researched long-term human activeness modeling and prediction for wellness services using recurrent neural networks.

### Knowledge Engineering and Artificial Intelligence Lab, KAIST

RESEARCHER

Daejeon, S.Korea  
Dec. 2013 - Sept. 2014

- Surveyed on plan and intention recognition.
- Researched on temporal information extraction task on both English and Korean texts.

## International Peace Supporting Standby Force, Special Warfare Command, ROK Army

ARMY INTERPRETER & AVIATION IMAGE ANALYST

- Served as an army interpreter & translator during ROK-US combined training exercises.
- Acted as a 3D aviation image analyst and a squad leader.

Incheon, S.Korea

Jan. 2012 - Oct. 2013

## Department of Mechatronics, Chungnam National University

RESEARCH INTERN

- Designed and implemented a stair-walking robot using LEGO Mindstorms.
- Helped out with hosting Robotics Camp for primary school students.

Daejeon, S.Korea

Jul. - Aug. 2008

## Teaching Experience

### Natural Language Processing (CSCI 5541)

TEACHING ASSISTANT, GUEST LECTURER

- Grading assignments, mentoring group term projects, holding office hours.
- [Nov. 21, 2023; Apr. 2, 2024] Gave a lecture on instruction tuning and augmented language models.

Minneapolis, MN

Sep.-Dec. 2023, Feb.-Apr. 2024

### Elice DX (Digital Transformation) Training at Samsung Electronics

LECTURER

- Ran a data science lab session for corporate executives on predicting wafers with defects from tabular data.
- Taught an intensive introductory machine learning class to new employees (lectures and lab sessions).

Dongtan, S.Korea; Virtual

Sep. 10, 22 - 25, 2022

### Advanced Language Processing Winter School (ALPS) 2022

TEACHING ASSISTANT

- Participated in writing programming exercises on controllable NMT for Prof. Graham Neubig's lab. session.
- Helped out with the students' technical questions about the exercises.

Virtual

Jan. 2022

### Academic-Industrial Collaboration Project, Chungbuk National University

GROUP MENTOR

- Mentored three groups of undergraduate students on their group projects.
- The project topics included making use of public APIs, pre-trained language models, and training models for store sign classification.

Virtual

Oct. - Nov. 2021

### Advanced Language Processing Winter School (ALPS) 2021

TEACHING ASSISTANT

- Participated in writing programming exercises on multilingual NMT for Prof. Kyunghyun Cho's lab. session.
- Helped out with the students' technical questions about the exercises.

Virtual

Dec. 2020 - Jan. 2021

### Elice AI & Data Science Course, Daejeon Creative Economy Innovation Center

TEACHING ASSISTANT

- Wrote five programming exercises every week for the government-funded data science course.
- [Dec. 26-27, 2016] Participated as a T/A in the data science workshop & challenge—our team came second in the challenge.

Daejeon, S.Korea

Oct. 2016 - Apr. 2017

### Introduction to A.I.(CS470), School of Computing, KAIST

TEACHING ASSISTANT & GUEST SPEAKER

- Held office hours, designed and graded homeworks and programming assignments.
- [Sept. 2, 2015] As a guest speaker, presented my research work on extraction of travel-related intents from textual review data.

Daejeon, S.Korea

Sept. - Dec. 2015

### Machine Learning Summer Camp 2015 (Online), School of Computing, KAIST

TEACHING ASSISTANT

- Managed three discussion groups, helping out with the weekly tasks.

Daejeon, S.Korea

Aug. 3 - 28, 2015

### Programming Practice(CS109), School of Computing, KAIST

TEACHING ASSISTANT

- Helped out freshers with programming practical sessions taught in Scala language.

Daejeon, S.Korea

Mar. - Jun. 2015

## Talks

### Seminar talk at Dept. of Computer Engineering, Chungbuk National University

SPEAKER

- Talked about my recent work on iterative text revision, machine translation, and discourse relations.

Virtual

Jul. 28, 2022

### Colloquium talk at Graduate School of Software, KAIST

SPEAKER

- Talked about a research topic on "Cross-Lingual Discourse Analysis and Its Application".

Virtual

Dec. 2, 2021

## Invited lectures at a start-up company, Rye & Catchers

SPEAKER

- Gave a series of three seminars on recent trends on NLP and a few selected topics of the company's interests.
- Designed and provided three lab. sessions on practical NLP techniques.

Seoul, S.Korea

Aug. - Sept. 2021

## DEVVIEW 2018: "Papago Internals: Model Analysis and Development with Applied Technology"

SPEAKER

- Talked about developing a website translation service at Papago, highlighting research issues and technical difficulties.

Seoul, S.Korea

Oct. 12, 2018

## Invited talk at a start-up company, Rye & Catchers

GUEST SPEAKER

- Talked about how textual data can be processed and utilized by *Bibly*, the company's book recommendation service.

Seoul, S.Korea

Nov. 11, 2016

## Two invited lectures at the Dept. of Humanities, Chonbuk National University

GUEST LECTURER

- Talked about IT and its role in the performing arts in East Asia for BA students at the Department of Humanities.

Jeonju, S.Korea

May 27 and Jun. 3, 2016

## Patent

### Terminal, server and event suggesting methods thereof

HYUNGRAI OH, HO-JIN CHOI, ZAE MYUNG KIM, KYO-JOONG OH, CHAE-GYUN LIM

Jul. 2019

US Patent US10353975B2

## Academic Services

Jun. 2024 **Reviewer**, NAACL 2024

Mexico City, Mexico

Feb. 2024 **Program Committee**, The 7th International Workshop on Dialog Systems

Bangkok, Thailand

Feb. 2024 **Program Committee**, The 38th Annual AAAI Conference on Artificial Intelligence

Vancouver, Canada

Jan. 2024 **Reviewer**, Knowledge-Based Systems, Elsevier

Journal

Dec. 2023 **Reviewer**, EMNLP 2023

Singapore

Jan. 2023 **Program Committee**, The 6th International Workshop on Dialog Systems

Jeju, S.Korea

Jan. 2022 **Program Committee**, The 5th International Workshop on Dialog Systems

Virtual

Jan. 2021 **Program Committee**, The 4th International Workshop on Dialog Systems

Virtual

Dec. 2020 **Reviewer**, NLP-COVID19 at EMNLP 2020

Virtual

Mar. 2020 **Reviewer**, IEEE Access

Virtual

Feb. 2019 **Program Committee**, The 2nd International Workshop on Dialog Systems

Kyoto, Japan

## Skills

<b>Programming</b>	Python, Java, JavaScript, C++14
<b>ML</b>	TensorFlow, PyTorch, ONNX, Kubeflow
<b>Web</b>	Vue, React, HTML5, Flask, Streamlit, REST API
<b>DevOps</b>	SQLite, MongoDB, AWS, Docker, Kubernetes, Jenkins
<b>Languages</b>	Korean (Native), English (Proficient), French (Novice)