

Zae Myung Kim

CS PhD candidate at University of Minnesota Twin Cities

My research interests center on developing a “meta-scaffolding paradigm” that integrates discourse structures, dataset metadata, and metacognitive feedback into the training loop of large language models to stabilize learning and produce interpretable, coherent long-form text. I design disciplined training schemes that slash data and compute demands, curtail manual prompt engineering, and steer model reasoning toward human-like cognitive patterns.

I have collaborated with researchers at Amazon, Salesforce, Grammarly, and Naver Labs.

✉ kim01756@umn.edu

🏠 zaemyung.github.io

🐙 [GitHub](#)

🎓 [Google Scholar](#)

🌐 [LinkedIn](#)

🐦 [zaemyung](#)

Education

- Sept 2022 – Dec 2025 (expected) **University of Minnesota Twin Cities** Minneapolis, MN
PhD in Computer Science
Thesis: *Meta-Scaffolding for Large Language Models* (tentative)
Advisor: [Dongyeop Kang](#)
- Sept 2014 – Aug 2016 **Korea Advanced Institute of Science and Technology (KAIST)** Daejeon, Korea
MS in Computer Science
Thesis: *Classifying Travel-Related Intents in Textual Data Using Recurrent Neural Networks*
Advisor: [Ho-Jin Choi](#)
- Oct 2008 – Aug 2011 **Imperial College London** London, UK
BEng in Computer Science
Group Project: *Intention Recognition in Ambient Homes*
Individual Project: *Computing Fictitious Plays in Graph-Based Games*

Professional Employment

- Jun – Nov 2024 **Research Intern – Amazon AGI** Seattle, WA
Proposed and developed structural alignment of LLMs for long-form generation.
- May – Sept 2023 **Research Intern – Salesforce AI Research** Palo Alto, CA
Implemented a large-scale challenging benchmarking system for scientific-paper retrieval.
Built a co-writing system with large language models for scientific writing.
- Apr – Jul 2022 **Research Intern – Grammarly** Remote
Developed a human-in-the-loop iterative text-revision system.
Trained span-based text-revision models with data augmentation.
- Oct 2020 – Feb 2022 **ML Researcher – Naver Labs Europe** Meylan, France
Studied language-specific representations in multilingual NMT.
Conducted cross-lingual discourse analysis for document-level NMT.
- Mar 2017 – Sept 2020 **ML Researcher – Papago, Naver Corp.** Seongnam, Korea
Developed machine translation models for Thai–Korean and German–Korean pairs.
Implemented tools for NMT: [sentence-segmenter](#), corpus pre-processor, aligner, and crawler.
Created website-translation service using NMT + heuristic attention.
Ranked 1st (doc-level) and 4th (sent-level) at WMT20 quality estimation shared task.

Dec 2016 – Feb 2017	Invited Researcher – KEAI Lab, KAIST Built an adaptive vocabulary learning app using learner interest profiles. Developed a healthcare app that recommends Pokémon Go events by forecasting users’ physical activity patterns using recurrent neural network.	Daejeon, Korea
Sept – Nov 2016	Research Intern – Naver Labs Corp. Researched extending named entity recognition for Korean with hierarchical categories.	Seongnam, Korea
Apr 2014 – Jul 2016	Research Assistant – Samsung–KAIST Initiative [1 st Year] Intent classification for travel-service personalization. [2 nd Year] Modeled long-term activeness using RNNs for wellness prediction.	Daejeon, Korea
Dec 2013 – Sept 2014	Research Intern – KEAI Lab, KAIST Surveyed plan and intention recognition. Researched temporal information extraction for English and Korean.	Daejeon, Korea
Jan 2012 – Oct 2013	Military Service – Special Warfare Command, ROK Army Served as an army interpreter, aviation image analyst, and squad leader.	Incheon, Korea

Honors, Awards, and Accomplishments

- 2025 **Doctoral Dissertation Fellowship** (University of Minnesota)
- 2022 **Best Paper Award** (In2Writing workshop at ACL 2022)
- 2022 **3M Science and Technology Fellowship** (University of Minnesota)
- 2018 **Runner-Up**, Elice Data Science Challenge (3rd place out of 300 participants)
- 2016 **Outstanding Teaching Assistant Award** (2016 Elice Data Science Workshop & Challenge)
- 2016 **Outstanding Graduate Teaching Assistant Award** (KAIST)
- 2012 **Runner-Up**, Republic of Korea Army Oratorical Contest (ROK Army)
- 2011 **Winning Team**, Robotics Localization Challenge (Imperial College London)
- 2009 **Winning Team**, Deutsche Bank Prize for Topics in AI Group Project (Imperial College London)
- 2008 **Bronze Medal**, Royal College of Science Union Science Challenge (Imperial College London)

Publications

An asterisk (*) marks co-first authors.

- 2025 **Toward Evaluative Thinking: Meta Policy Optimization with Evolving Reward Models**
Zae Myung Kim, Chanwoo Park, Vipul Raheja, Dongyeop Kang.
- 2025 **Align to Structure: Aligning Large Language Models with Structural Information**
Zae Myung Kim, Anand Ramachandran, Farideh Tavazoei, Joo-Kyung Kim, Oleg Rokhlenko, Dongyeop Kang.
Under review.
- 2025 **anchors Aweigh! Sail for Optimal Unified Multi-Modal Representations**
Minoh Jeong*, Min Namgung*, Zae Myung Kim*, Dongyeop Kang, Yao-Yi Chiang, Alfred Hero.
Under review.
- 2024 **Threads of Subtlety: Detecting Machine-Generated Texts Through Discourse Motifs**
Zae Myung Kim, Kwang Hee Lee, Preston Zhu, Vipul Raheja, Dongyeop Kang.
ACL 2024.
- 2024 **Benchmarking Cognitive Biases in Large Language Models as Evaluators**
Ryan Koo, Minhwa Lee, Vipul Raheja, Jong Inn Park, Zae Myung Kim, Dongyeop Kang.
Findings of ACL 2024.

- 2024 **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.
Under review.
- 2022 **Improving Iterative Text Revision by Learning Where To Edit From Other Revision Tasks**
Zae Myung Kim, Wanyu Du, Vipul Raheja, Dhruv Kumar, Dongyeop Kang.
EMNLP 2022.
- 2022 **Read, Revise, Repeat: A System Demonstration for Human-in-the-loop Iterative Text Revision [Best Paper Awarded]**
Wanyu Du*, Zae Myung Kim*, Vipul Raheja, Dhruv Kumar, Dongyeop Kang.
In2Writing at ACL 2022.
- 2022 **Understanding Iterative Revision from Human-Written Text**
Wanyu Du, Vipul Raheja, Dhruv Kumar, Zae Myung Kim, Melissa Lopez, Dongyeop Kang.
ACL 2022.
- 2021 **Visualizing Cross-Lingual Discourse Relations in Multilingual TED Corpora**
Zae Myung Kim, Vassilina Nikoulina, Dongyeop Kang, Didier Schwab, Laurent Besacier.
CODI 2021 at EMNLP 2021.
- 2021 **Do Multilingual Neural Machine Translation Models Contain Language Pair Specific Attention Heads?**
Zae Myung Kim, Laurent Besacier, Vassilina Nikoulina, Didier Schwab.
Findings of ACL 2021.
- 2020 **A Multilingual Neural Machine Translation Model for Biomedical Data**
Alexandre Bérard, Zae Myung Kim, Vassilina Nikoulina, Eunjeong Park, Matthias Gallé.
NLP-COVID19 at EMNLP 2020.
- 2020 **PATQUEST: Papago Translation Quality Estimation**
Yujin Baek*, Zae Myung Kim*, Jihyung Moon, Hyunjoong Kim, Eunjeong Park.
WMT 2020.
- 2017 **An Adaptive Vocabulary Learning Application through Modeling Learner’s Linguistic Proficiency and Interests**
Zae Myung Kim, Suin Kim, Alice Oh, Ho-Jin Choi.
International Conference on Big Data and Smart Computing.
- 2016 **TIMEX3 and Event Extraction Using Recurrent Neural Networks**
Zae Myung Kim, Young-Seob Jeong, Ho-Jin Choi.
International Conference on Big Data and Smart Computing.
- 2016 **An Approach to Spam Comment Detection Through Domain-Independent Features**
Jong Myoung Kim, Zae Myung Kim, Kwangjo Kim.
International Conference on Big Data and Smart Computing.
- 2015 **Temporal Information Extraction from Korean Texts**
Young-Seob Jeong, Zae Myung Kim, Hyun-Woo Do, Chae-Gyun Lim, Ho-Jin Choi.
CoNLL 2015.
- 2017 **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, Ho-Jin Choi.
BMC Medical Informatics and Decision Making [SCIE, IF: 2.042].

- 2016 **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, Ho-Jin Choi.
Sensors 16(6) [SCIE, IF: 2.437].

Patent

- 2019 **Terminal, server and event suggesting methods thereof**
Hyungrai Oh, Ho-Jin Choi, Zae Myung Kim, Kyo-Joong Oh, Chae-Gyun Lim.
US Patent US10353975B2.

Teaching Experience

- Spring 2025 **Lecturer** – Natural Language Processing (CSCI 5541), University of Minnesota
[Apr 1, 2025] Lecture titled, *LLMs as Agents*.
- Fall 2024 **Head TA** – Intro. to Computer Programming in Java (CSCI 1103), University of Minnesota
- Fall 2024 **Lecturer** – Natural Language Processing (CSCI 5541), University of Minnesota
[Oct 31, 2024] Lecture titled, *LLMs as Agents*.
- Spring 2024 **TA, Lecturer** – Natural Language Processing (CSCI 5541), University of Minnesota
[Apr 2, 2024] Lecture titled, *Instructing and Augmenting LLMs*.
- Fall 2023 **TA, Lecturer** – Natural Language Processing (CSCI 5541), University of Minnesota
[Nov 21, 2023] Lecture titled, *Advanced Topics: Instructing and Augmenting LLMs*.
- Sept 2022 **Lecturer** – Elice Digital Transformation Training, Samsung Electronics
- Winter 2022 **TA** – Advanced Language Processing Winter School 2022
- Winter 2021 **TA** – Advanced Language Processing Winter School 2021
- Fall 2021 **Mentor** – Academic-Industrial Collaboration Project, Chungbuk National University
- Summer 2021 **Lecturer** – Three seminars on NLP trends and practical sessions, Rye & Catchers
- Fall 2016 **TA** – Elice AI & Data Science Course, Daejeon Creative Economy Innovation Center
- Fall 2015 **TA, Lecturer** – Introduction to AI (CS470), KAIST
[Sept 2, 2015] Lecture on extracting travel-related intents from textual review data.
- Summer 2015 **TA** – Machine Learning Summer Camp 2015 (Online), KAIST
- Spring 2015 **TA** – Programming Practice (CS109), KAIST

Mentorship

- Spring 2025 – Present **Yangfan He** (BS in Computer Science, University of Minnesota)
- Summer 2024 – **Bin Hu** (BS in Computer Science, University of Minnesota)
Spring 2025 Now: MS in Computer Science, University of Maryland.
- Spring 2024 – **Ryan Koo** (MS in Computer Science, University of Minnesota)
Spring 2025 Now: Software Engineer at Amazon
- Fall 2023 – Spring 2024 **Preston Zhu** (BS in Computer Science, University of Minnesota)
- Spring 2022 **Josh Spitzer-Resnick** (MS in Computer Science, University of Minnesota)
Now: Solventum Healthcare.

Invited Talks

- Jul 28, 2022 **Seminar Talk** – Dept. of Computer Engineering, Chungbuk Nat'l Univ.
Presented research on iterative text revision, machine translation, and discourse relations.

- Dec 2, 2021 **Colloquium Talk** – Graduate School of Software, KAIST
Talk titled, *Cross-Lingual Discourse Analysis and Its Application*.
- Oct 12, 2018 **Conference Talk** – DEVIEW 2018
Talk titled, *Papago Internals: Model Analysis and Development with Applied Technology*
- Nov 11, 2016 **Seminar Talk** – Rye & Catchers
Talked about how textual data can be processed and utilized by *Bibly*, the company's book recommendation service.

Academic Services

- 2025 **Reviewer** – ACL 2025
- 2024 **Reviewer** – NAACL 2024
- 2024 **Program Committee** – 7th International Workshop on Dialog Systems
- 2024 **Program Committee** – AAAI 2024
- 2024 **Reviewer** – Knowledge-Based Systems (Elsevier Journal)
- 2023 **Reviewer** – EMNLP 2023
- 2023 **Program Committee** – 6th International Workshop on Dialog Systems
- 2022 **Program Committee** – 5th International Workshop on Dialog Systems
- 2021 **Program Committee** – 4th International Workshop on Dialog Systems
- 2020 **Reviewer** – NLP-COVID19 Workshop at EMNLP 2020
- 2020 **Reviewer** – IEEE Access
- 2019 **Program Committee** – 2nd International Workshop on Dialog Systems