Chanjun Park

Technical Leader, Upstage
Email: bcj1210@naver.com
Home: https://parkchanjun.github.io/
LinkedIn: www.linkedin.com/in/bcj1210

LIFELONG RESEARCH OBJECTIVE

My lifelong research objective is to advance the development of machine translation systems that ensure consistent translation quality across diverse languages, thereby breaking down barriers to global communication. Through this research, I aim to contribute to creating an environment where linguistic diversity is honored, while enabling effective communication without cultural or linguistic impediments.

RESEARCH INTERESTS

My research philosophy is centered on service-driven research, aiming to bridge the gap between foundational theories in natural language processing (NLP) and their practical, real-world applications. My primary interests include the development of efficient, purpose-trained Large Language Models (LLMs), with a particular focus on their fundamental capabilities, rigorous evaluation, and addressing issues of recency and truthfulness. Additionally, I am deeply engaged in cross-lingual NLP and multidisciplinary research that integrates diverse fields to enhance NLP applications.

EDUCATION

- **Korea University** Sep. 2019 - Aug. 2023
  Ph.D. in Computer Science and Engineering (advisor: Heuiseok Lim)
  Dissertation: Data-Centric Neural Machine Translation - A Real-World Approaches
  GPA: 4.44/4.5
- **Busan University of Foreign Studies (BUFS)** Mar. 2015 - Feb. 2019
  B.S. in Natural Language Processing & Creative Convergence (advisor: Pum-mo Ryu)
  GPA: 4.12/4.5

WORK EXPERIENCES

- **Upstage** Mar. 2022 - Present
  Technical Leader (July. 2023 - Present)
  AI Research Engineer (July. 2022 - July. 2023)
  AI Research Engineer Intern (Mar. 2022 - July.2022)
- **Korea University** July. 2019 - July. 2022
  Graduate Research Assistant (advisor: Prof. Heuiseok Lim)
- **SYSTRAN** June. 2018 - July. 2019
  Research Engineer (Sep. 2018 - July. 2019)
  Research Engineer Intern (June. 2018 - Aug. 2018)

PUBLICATIONS

Equal contribution: *, *Corresponding author*: †

Google Scholar: https://scholar.google.com/citations?user=085jNAMAAAAJ&hl=en
Semantic Scholar: https://www.semanticscholar.org/author/Chanjun-Park/2115195904

Conferences

1. **Open Ko-LLM Leaderboard: Evaluating Large Language Models in Korean with Ko-H5 Benchmark**
   Chanjun Park, Hyeonwoo Kim, Dahyun Kim, SeongHwan Cho, Sanghoon Kim, Sukyung Lee, Yungi Kim, Hwalsuk Lee
   ACL 2024

2. **KoCommonGEN v2: A Benchmark for Navigating Korean Commonsense Reasoning Challenges in Large Language Models**
   Jaehyung Seo, Jaewook Lee, Chanjun Park, SeongTae Hong, Seungjun Lee, Heuiseok Lim
   ACL 2024-Findings

3. **Length-aware Byte Pair Encoding for Mitigating Over-segmentation in Korean Machine Translation**
   Jungseob Lee, Hyeonseok Moon, Seungjun Lee, Chanjun Park[†], Sugyeong Eo, Hyunwoong Ko, Jaehyung Seo, Seungyoon Lee, Heuiseok Lim[†]
   ACL 2024-Findings

4. **SOLAR 10.7B: Scaling Large Language Models with Simple yet Effective Depth Up-Scaling**
   Sanghoon Kim[*, †], Dahyun Kim[*], Chanjun Park[*, †], Wonsung Lee[*, †], Wonho Song[*], Yunsu Kim[*], Hyeonwoo Kim[*], Yungi Kim, Hyeonju Lee, Jihoo Kim, Changbae Ahn, Seonghoon Yang, Sukyung Lee, Hyunbyung Park, Gyoungjin Gim, Mikyoung Cha, Hwalsuk Lee[†], Sunghun Kim[†]
   NAACL 2024 - Industry Track

5. **Leveraging Pre-existing Resources for Data-Efficient Counter-Narrative Generation in Korean**
   Seungyoon Lee, Chanjun Park[†], DaHyun Jung, Hyeonseok Moon, Jaehyung Seo, Sugyeong Eo, Heuiseok Lim[†]
   LREC-COLING 2024, Oral

6. **Detecting Critical Errors Considering Cross-Cultural Factors in English-Korean Translation**
   Sugyeong Eo, Jungwoo Lim, Chanjun Park, Hyeonseok Moon, Jaehyung Seo, Heuiseok Lim
   LREC-COLING 2024, Oral

7. **Hyper-BTS Dataset: Scalability and Enhanced Analysis of Back TranScription (BTS) for ASR Post-Processing**
   Chanjun Park, Jaehyung Seo, Seolhwa Lee, Junyoung Son, Hyeonseok Moon, Sugyeong Eo, Chanhee Lee, Heuiseok Lim
   EACL 2024-Findings

8. **Generative Interpretation: Toward Human-Like Evaluation for Educational Question-Answer Pair Generation**
   Hyeonseok Moon, Jaewook Lee, Sugyeong Eo, Chanjun Park, Jaehyung Seo, Heuiseok Lim
   EACL 2024-Findings

9. **KEBAP: Korean Error Explainable Benchmark Dataset for ASR and Post-processing**
   Seonmin Koo[*], Chanjun Park[*], Jinsung Kim, Jaehyung Seo, Sugyeong Eo, Hyeonseok Moon, Heuiseok Lim
   EMNLP 2023

10. **CHEF in the Language Kitchen: A Generative Data Augmentation Leveraging Korean Morpheme Ingredients**
    Jaehyung Seo, Hyeonseok Moon, Jaewook Lee, Sugyeong Eo, Chanjun Park, Heuiseok Lim
    EMNLP 2023

11. **Informative Evidence-guided Prompt-based Fine-tuning for English-Korean Critical Error Detection**
    DaHyun Jung, Sugyeong Eo, Chanjun Park, Hyeonseok Moon, Jaehyung Seo, Heuiseok Lim
    IJCNLP-AACL 2023

12. **Improving Formality-Sensitive Machine Translation using Data-Centric Approaches and Prompt Engineering**
    Seugnjun Lee, Hyeonseok Moon, Chanjun Park, Heuiseok Lim
    IWSLT 2023

13. **PEEP-Talk: A Situational Dialogue-based Chatbot for English Education**
    Seugnjun Lee, Yoonna Jang, Chanjun Park, Jungseob Lee, Jaehyung Seo, Hyeonseok Moon, Sugyeong Eo, Seounghoon Lee, Bernardo Nugroho Yahya, Heuiseok Lim
    ACL 2023 - Demo Track

14. **PicTalky: Augmentative and Alternative Communication for Language Developmental Disabilities**
Chanjun Park*, Yoonna Jang*, Seolhwa Lee*, Jaehyung Seo*, Kisu Yang, Heuiseok Lim
AACL-IJCNLP 2022 - Demo Track

15. **KU X Upstage's submission for the WMT22 Quality Estimation: Critical Error Detection Shared Task**
Sugyeong Eo, Chanjun Park, Hyeonseok Moon, Jaehyung Seo, Heuiseok Lim
WMT 2022

16. **QUAK: A Synthetic Quality Estimation Dataset for Korean-English Neural Machine Translation**
Sugyeong Eo, Chanjun Park, Hyeonseok Moon, Jaehyung Seo, Gyeongmin Kim, Jungseob Lee, Heuiseok Lim
COLING 2022

17. **A Dog Is Passing Over The Jet? A Text-Generation Dataset for Korean Commonsense Reasoning and Evaluation**
Jaehyung Seo, Seounghoon Lee, Chanjun Park, Yoonna Jang, Hyeonseok Moon, Sugyeong Eo, Seonmin Koo, Heuiseok Lim
NAACL 2022-Findings

18. **Priming Ancient Korean Neural Machine Translation**
Chanjun Park*, Seolhwa Lee*, Hyeonseok Moon, Sugyeong Eo, Jaehyung Seo, Heuiseok Lim
LREC 2022, Oral

19. **FreeTalky: Don't Be Afraid! Conversations Made Easier by a Humanoid Robot using Persona-based Dialogue**
Chanjun Park*, Yoonna Jang*, Seolhwa Lee*, Sungjin Park*, Heuiseok Lim
LREC 2022

20. **Empirical Analysis of Synthetic Data Generation Using Noising Strategies for Automatic Post-editing**
Hyeonseok Moon, Chanjun Park, Seolhwa Lee, Jaehyung Seo, Jeongsub Lee, Sugyeong Eo, Heuiseok Lim
LREC 2022

21. **Should we find another model?: Improving Neural Machine Translation Performance with ONE-Piece Tokenization Method without Model Modification**
Chanjun Park*, Sugyeong Eo*, Hyeonseok Moon*, Heuiseok Lim
NAACL-HLT 2021 - Industry Track, Oral

Workshops

1. **Exploring Inherent Biases in LLMs within Korean Social Context: A Comparative Analysis of ChatGPT and GPT-4**
Seungyoon Lee, Dongjun Kim, Dahyun Jung, Chanjun Park†, Heuiseok Lim†
NAACL 2024 - Student Research Workshop

2. **Explainable CED: A Dataset for Explainable Critical Error Detection in Machine Translation**
Dahyun Jung, Sugyeong Eo, Chanjun Park†, Heuiseok Lim†
NAACL 2024 - Student Research Workshop

3. **Model-Based Data-Centric AI: Bridging the Divide Between Academic Ideals and Industrial Pragmatism**
Chanjun Park*, †, Minsoo Khang*, Dahyun Kim*
ICLR 2024 - Data-centric Machine Learning Research (DMLR) Workshop

4. **Proceedings of the Seventh Widening NLP Workshop (WiNLP 2023)**
Bonaventure F. P. Dossou, Isidora Tourni, Hatem Haddad, Shaily Bhatt, Fatemehsadat Mireshghallah, Sunipa Dev, Tanvi Anand, Weijia Xu, Atnafu Lambebo Tonja, Alfredo Gomez, Chanjun Park
EMNLP 2023-Widening NLP Workshop

5. **Alternative Speech: Complementary Method to Counter-Narrative for Better Discourse**
Seungyoon Lee*, DaHyun Jung*, Chanjun Park*, Seolhwa Lee, Heuiseok Lim
ICDM 2023 - The First Workshop on Data-Centric AI

6. **Synthetic Alone: Exploring the Dark Side of Synthetic Data for Grammatical Error Correction**
Chanjun Park*, Seonmin Koo*, Seolhwa Lee, Jaehyung Seo, Sugyeong Eo, Hyeonseok Moon, Heuiseok Lim
ICML 2023 - Data-centric Machine Learning Research (DMLR) Workshop

7. **DMOps: Data Management Operation and Recipes**
   Eujeong Choi, Chanjun Park[*, †]
   ICML 2023 - Data-centric Machine Learning Research (DMLR) Workshop

8. **Inter-Annotator Agreement in the Wild: Uncovering Its Emerging Roles and Considerations in Real-World Scenarios**
   NamHyeok Kim, Chanjun Park[*, †]
   ICML 2023 - Data-centric Machine Learning Research (DMLR) Workshop

9. **Transcending Traditional Boundaries: Leveraging Inter-Annotator Agreement (IAA) for Enhancing Data Management Operations**
   Damrin Kim, NamHyeok Kim, Chanjun Park[†], Harksoo Kim[†]
   ICML 2023 - Data-centric Machine Learning Research (DMLR) Workshop

10. **Data-Driven Approach for Formality-Sensitive Machine Translation: Language-Specific Handling and Synthetic Data Generation**
    Seugnjun Lee, Hyeonseok Moon, Chanjun Park, Heuiseok Lim
    ICML 2023 - Data-centric Machine Learning Research (DMLR) Workshop

11. **Toward Practical Automatic Speech Recognition and Post-Processing: a Call for Explainable Error Benchmark Guideline**
    Seonmin Koo[*], Chanjun Park[*], Jinsung Kim, Jaehyung Seo, Sugyeong Eo, Hyeonseok Moon, Heuiseok Lim
    ICML 2023 - Data-centric Machine Learning Research (DMLR) Workshop

12. **Knowledge Graph-Augmented Korean Generative Commonsense Reasoning**
    Dahyun Jung, Jaehyung Seo, Jaewook Lee, Chanjun Park, Heuiseok Lim
    ICML 2023 - Data-centric Machine Learning Research (DMLR) Workshop

13. **Focus on FoCus: Is FoCus focused on Context, Knowledge and Persona?**
    SeungYoon Lee, Jungseob Lee, Chanjun Park, Sugyeong Eo, Hyeonseok Moon, Jaehyung Seo, Jeongbae Park, Heuiseok Lim
    COLING 2022 - The 1st Workshop on Customized Chat Grounding Persona and Knowledge

14. **A Self-Supervised Automatic Post-Editing Data Generation Tool**
    Hyeonseok Moon, Chanjun Park, Sugyeong Eo, Jaehyung Seo, Seungjun Lee, Heuiseok Lim
    ICML 2022 - DataPerf Workshop

15. **FreeTalky: Don't Be Afraid! Conversations Made Easier by a Humanoid Robot using Persona-based Dialogue**
    Chanjun Park[*], Yoonna Jang[*], Seolhwa Lee[*], Sungjin Park[*], Heuiseok Lim
    AAAI 2022 -Artificial Intelligence for Education (AI4EDU)

16. **How should human translation coexist with NMT? Efficient tool for building high quality parallel corpus**
    Chanjun Park, Seolhwa Lee, Hyeonseok Moon, Sugyeong Eo, Jaehyung Seo, Heuiseok Lim
    NeurIPS 2021 - Data-centric AI (DCAI) workshop

17. **A New Tool for Efficiently Generating Quality Estimation Datasets**
    Sugyeong Eo, Chanjun Park, Jaehyung Seo, Hyeonseok Moon, Heuiseok Lim
    NeurIPS 2021 - Data-centric AI (DCAI) workshop

18. **Automatic Knowledge Augmentation for Generative Commonsense Reasoning**
    Jaehyung Seo, Chanjun Park, Sugyeong Eo, Hyeonseok Moon, Heuiseok Lim
    NeurIPS 2021 - Data-centric AI (DCAI) workshop

19. **Syntax-enhanced Dialogue Summarization using Syntax-aware information**
    Seolhwa Lee, Kisu Yang, Chanjun Park, João Sedoc, Heuiseok Lim
    NeurIPS 2021 - Women in Machine Learning workshop

20. **Towards Syntax-Aware Dialogue Summarization using Multi-task Learning**
    Seolhwa Lee, Kisu Yang, Chanjun Park, João Sedoc, Heuiseok Lim
    EMNLP 2021 -Widening NLP (WiNLP2021) workshop

21. **Two Heads are Better than One? Verification of Ensemble Effect in Neural Machine Translation**
    Chanjun Park, Sungjin Park, Seolhwa Lee, Taesun Whang, Heuiseok Lim
    EMNLP 2021 -The Second Workshop on Insights from Negative Results in NLP

22. **BTS: Back TranScription for Speech-to-Text Post-Processor using Text-to-Speech-to-Text**
Chanjun Park, Jaehyung Seo, Seolhwa Lee, Chanhee Lee, Hyeonseok Moon, Sugyeong Eo, Heuiseok Lim
ACL 2021 -Workshop on Asian Translation (WAT)

23. **Dealing with the Paradox of Quality Estimation**
Sugyeong Eo*, Chanjun Park*, Jaehyung Seo, Hyeonseok Moon, Heuiseok Lim
MT Summit 2021 - Workshop on Technologies for Machine Translation of Low-Resource Languages (LoResMT)

Selected Journals

1. **Uncovering the Risks and Drawbacks Associated with the Use of Synthetic Data for Grammatical Error Correction**
Seonmin Koo*, Chanjun Park*, Seolhwa Lee, Jaehyung Seo, Sugyeong Eo, Hyeonseok Moon, Heuiseok Lim
IEEE Access (SCIE), 2023

2. **Doubts on the Reliability of Parallel Corpus Filtering**
Hyeonseok Moon, Chanjun Park, Seonmin Koo, Jungseob Lee, Seungjun Lee, Jaehyung Seo, Sugyeong Eo, Yoonna Jang, Hyunjoong Kim, Hyoung-gyu Lee, Heuiseok Lim
Expert Systems With Applications (SCIE, Q1), 2023

3. **K-NCT: Korean Neural Grammatical Error Correction Gold-Standard Test Set Using Novel Error Type Classification Criteria**
Seonmin Koo*, Chanjun Park*, Jaehyung Seo, Seungjun Lee, Hyeonseok Moon, Jungseob Lee, Heuiseok Lim
IEEE Access (SCIE), 2022

4. **PU-GEN: Enhancing Generative Commonsense Reasoning for Language Models with Human-Centered Knowledge**
Jaehyung Seo, Dongsuk Oh, Sugyeong Eo, Chanjun Park, Kisu Yang, Hyeonseok Moon, Kinam Park, Heuiseok Lim
Knowledge-Based Systems (SCIE, Q1), 2022

5. **Word-level Quality Estimation for Korean-English Neural Machine Translation**
Sugyeong Eo*, Chanjun Park*, Hyeonseok Moon, Jaehyung Seo, Heuiseok Lim
IEEE Access (SCIE), 2022

6. **Mimicking Infants' Bilingual Language Acquisition for Domain Specialized Neural Machine Translation**
Chanjun Park*, Woo-Young Go*, Sugyeong Eo, Hyeonseok Moon, Seolhwa Lee, Heuiseok Lim
IEEE Access (SCIE), 2022

7. **An Automatic Post Editing with Efficient and Simple Data Generation Method**
Hyeonseok Moon*, Chanjun Park*, Jaehyung Seo, Sugyeong Eo, Heuiseok Lim
IEEE Access (SCIE), 2022

8. **An Empirical Study on Automatic Post Editing for Neural Machine Translation**
Hyeonseok Moon*, Chanjun Park*, Sugyeong Eo, Jaehyung Seo, Heuiseok Lim
IEEE Access (SCIE), 2021

9. **Ancient Korean Neural Machine Translation**
Chanjun Park, Chanhee Lee, YeongWookYang, Heuiseok Lim
IEEE Access (SCIE), 2020

10. **Comparison of the evaluation metrics for Neural Grammatical Error Correction with Overcorrection**
Chanjun Park, YeongWookYang, Chanhee Lee, Heuiseok Lim
IEEE Access (SCIE), 2020

11. **Neural Spelling Correction: Translating Incorrect sentences to Correct sentences for Multimedia**
Chanjun Park*, Kuekyeng Kim*, YeongWookYang, Minho Kang, Heuiseok Lim
Multimedia Tools and Applications (SCIE, Q1), 2020

In addition, I have published 31 papers in international journals (SCIE), 21 papers in domestic journals (KCI), 57 papers in domestic conferences, and 34 papers in other international conferences.

Research Projects

Upstage

- **LLM Projects**: Solar | sDPO | SAAS | Open Ko-LLM Leaderboard | Dataverse | Evalverse | Up 1 Trillion Token Club | MathGPT (with KT & Mathpresso)| E-Commerce LLM (with Connectwave) | News LLM (with BIG KINDS) | Law LLM (with Law&Company) | Medical LLM (with PHI Digital Healthcare) | Thai LLM (with KT & Jasmine Technology Solution) | AskUp Biz
- **Semantic Search Projects**: Semantic Search Pack (with LG U+ & Amorepacific respectively)
- **Education Projects**: Upstage AI Lab (with Fastcampus) | Boostcamp (with NAVER Connect Foundation) | Pretraining LLMs (with DeepLearning.AI)
- **National Projects**: Development of learning and utilization technology to reflect sustainability of generative language models and up-to-dateness over time (IITP) | Development of a Generative AI Agent-Based Legal Case Analysis Support Service (NIPA)

Korea University

- **National Projects**: A Neural-Symbolic Model for Knowledge Acquisition and Inference Techniques (IITP) | Research and Development of Human-inspired multiple intelligence (IITP) | ICT Creative Consilience Program (IITP) | Research on Korean/English Translation Technology in the Security Domain (National Security Research Institute (NRF)
- **Industry Projects**: Research on Utilizing Low to Medium Quality Data to Improve NMT Performance (Naver Papago) | Development of technology for extracting purpose sentences and key phrases from patent documents (LG innotek) | Research on Korean Prompt-Based Engineering for Few-Shot Learners (Hyundai Motors)

SYSTRAN

- **National Projects**: Core technology development of the real-time simultaneous speech translation based on knowledge enhancement (Ministry of Science and ICT)
- **Internal Projects**: Neural Machine Translation | Automatic Speech Recognition | Real-time Simultaneous Speech Translation | L Cloud (API Portal) | Grammar Error Correction (with Hello Talk) | AI Voice Assistant (with CJ Logistics)

BUFS

- **National Projects**: Development of multilingual word translation service based on Korean semantics analysis for multicultural family (National Research Foundation of Korea) | Exobrain (Ministry of Science and ICT, ETRI) | Research and Development of Korean Information Processing Technologies (Ministry of Science and ICT) | ICT VAEMS: Vertical Aquaculture Energy Management (Ministry of Trade, Industry and Energy)

Academic Services

Program Chair

- Widening NLP (WiNLP) Workshop (EMNLP 2023, EMNLP 2024)

Publication Chair

- ICLR 2024-Data-centric Machine Learning Research (DMLR) 2024

Virtual Social Chair

- COLING 2022

Program Committee

- EMNLP 2024-Industry Track, ACL 2024-ML4AL, NAACL 2024-NLP+CSS, NAACL 2024-Insights, NAACL 2024-BEA, ICML 2023-Data-centric Machine Learning Research (DMLR), ACL 2023-Industry Track, ACL 2023-BEA, ACL 2022-Insights, NAACL 2022-Industry Track

Conference Reviewer

- ACL, EMNLP, NAACL, EACL, COLING, ARR, NeurIPS, ICLR

Journal Reviewer

- ETRI Journal, IEEE Access, International Journal of Human-Computer Interaction, Cybernetics and Systems

External Activities

Committee Member

- AI Ethics and Trustworthiness Forum (2024.04-)
- AI Privacy Policy Council (2023.10-)

Advisor

- National Institute of Korean Language (2024.04-)
- Information Technology Committee of the Republic of Korea Army (2024.05-)
- National Information Society Agency (2024.01-)
- Ministry of Food and Drug Safety (2023.12)
- Yonsei University-Institute of Language and Information Studies (2023.12)
- Software Policy & Research Institutes (2023.08)

AWARDS/FELLOWSHIPS

Awards

- Forbes 30 Under 30 Korea (2024)
- Best Paper Award at Korea University (2023)
- Best Paper Award at Annual Conference on Human & Cognitive Language Technology (2019 (x1), 2020 (x1), 2021 (x3), 2022 (x1), 2023 (x2))
- 1st Place at WMT Quality Estimation Shared Task 2022 - Sentence-level Critical Error Detection (2022)
- 1st Place at Flitto Hackathon (2020)
- 1st Place at Microsoft AI Accessibility Hackathon in Korea (2019)
- Participation Award at Next Generation Information Processing NLP Competition (2018)
- Bit Computer Excellence Award (President Award) at Bit Computer (2017)

Fellowships

- Research Encouragement Scholarship by Korea University (2023)
- Naver Ph.D. Fellowship (2021)
- Scholarship for academic excellence by Sooyoungro Church (2016, 2017)
- Undergraduate Study Full Scholarship by Busan University of Foreign Studies (2015-2019)

PATENTS

International Patents

- **Method for generating training data and method for post-processing of speech recognition using the same**
  Apply for a patent (17/739,383)
- **Method of building training data of machine translation**
  Apply for a patent (PCT/KR2021/012195)

Domestic Patents

- **Device and method for generating of training data for quality estimation in machine translation**
  Granted Patent (10-2593447)
- **Apparatus for corpus processing, apparatus and method and machine translation**
  Granted Patent (10-2574167)
- **Device and method for generating training data for automatic post editing**
  Apply for a patent (10-2021-0118924)
- **Device and method for generating optimal translation subtitle using quality estimation**
  Granted Patent (10-2690953)
- **Improving speech recognition performance using TTS in domain-specific environment**
  Apply for a patent (10-2021-0028816)

- **Method for generating training data and method for post-processing of speech recognition using the same**
  Granted Patent (10-2557810)
- **Method of building training data of machine translation**
  Granted Patent (10-2409667)
- **Correction performance evaluation metrics of neural network machine translation and method of constructing the same**
  Granted Patent (10-2390154)
- **Apparatus and method for outputting image corresponding to language**
  Granted Patent (10-2476497)
- **Method of translating ancient Korean using machine translation**
  Granted Patent (10-2425922)
- **Device and method for correcting Korean spelling**
  Granted Patent (10-2430918)

TEACHING

- **Pretraining LLMs**, Content Planning and Production, DeepLearning.AI. (2024-)
- **Language Model to Large Language Model (LM to LLM)**, Instructor, Fast Campus. (2023-)
- **Natural Language Processing (NLP) Basic**, Instructor, Fast Campus. (2023-)
- **Learning ChatGPT Utilization and Service Construction with AskUP**, Instructor, Fast Campus. (2023-)
- **Finance Specialized Large Language Model for Everyone**, Instructor, Upstage Online Course. (2023)
- **Data-Centric NLP**, Instructor, BoostCamp - NAVER Connect Foundation. (2023-)
- **Introduction to Natural Language Processing in Big Data (BDC101)**, Teaching Assistant, Korea University (2021)
- **Introduction to Natural Language Processing in Big Data (BDC101)**, Head Teaching Assistant, Korea University (2020)
- **Natural Language Processing for Digital Finance Engineering (DFE610)**, Head Teaching Assistant, Korea University (2020)
- **Natural Language Processing (COSE461)**, Teaching Assistant, Korea University (2020)
- **Artificial Intelligence and Natural Language Processing (DFC615)**, Teaching Assistant, Korea University (2020)
- **Machine Translation for everyone**, Instructor, SK T Academy (2019)

TALKS

International Talks

- **Applied ML: LLMs and Knowledge Graphs- Tokyo**, The Ecosystem of LLMs from a Real-World Perspective, May 2024
- **The University of Tokyo - Center for Data-Driven Discovery**, SOLAR: The Next Frontier in Large Language Models by Upstage, January 2024
- **Korean-English Joint Seminar on Artificial Intelligence (AI) Safety and Reliability**, Upstage Vision for Ethical and Trustworthy Large Language Models, December 2023

Domestic Talks

- **Seoul National University**, Large Language Model in the Wild, September 2024
- **University of Ulsan**, Large Language Model in the Wild, September 2024
- **Embassy of the United States in Seoul**, Large Language Model in the Wild, September 2024
- **Sungkyunkwan University**, SOLAR: The Next Frontier in Large Language Models by Upstage and its Ecosystem, July 2024
- **TTA**, SOLAR: The Next Frontier in Large Language Models by Upstage and its Ecosystem, July 2024
- **KCC 2024**, SOLAR: The Next Frontier in Large Language Models by Upstage and its Ecosystem, June 2024
- **Kyungpook National University**, The Ecosystem of LLMs from a Real-World Perspective, June 2024

- **Konkuk University**, SOLAR: The Next Frontier in Large Language Models by Upstage and its Ecosystem, May 2024
- **Chungnam National University**, SOLAR: The Next Frontier in Large Language Models by Upstage and its Ecosystem, May 2024
- **AI Safety Compass Conference 2024**, Data and Evaluation Methods for Trustworthy AI, May 2024
- **Chonbuk National University**, SOLAR: The Next Frontier in Large Language Models by Upstage and its Ecosystem, May 2024
- **TTA**, LLM Evaluation and its Ecosystem, May 2024
- **ETRI**, SOLAR: The Next Frontier in Large Language Models by Upstage and its Ecosystem, May 2024
- **AI Tech 2024 - AI Frontier for AI Era**, SOLAR: The Next Frontier in Large Language Models by Upstage and its Ecosystem, May 2024
- **No code·Low code Hyper-automation Conference 2024**, The Ecosystem of LLMs from a Real-World Perspective, May 2024
- **Hallym University**, SOLAR: The Next Frontier in Large Language Models by Upstage and its Ecosystem, April 2024
- **Korean Society of Artificial Intelligence in Medicine**, Current state of Large-Scale Generative AI Development, April 2024
- **NIA**, Open Ko-LLM Leaderboard and its Ecosystem, April 2024
- **Korea University College of Medicine - Intelligent Medical Data Lab**, SOLAR: The Next Frontier in Large Language Models by Upstage, February 2024
- **Fast Campus**, Job Employment Special Lecture, February 2024
- **FuriosaAI**, SOLAR: The Next Frontier in Large Language Models by Upstage, February 2024
- **Defense Agency for Technology and Quality (DTaQ)**, Current State of Artificial Intelligence (AI) and Its Application Strategies in the Defense Sector, January 2024
- **National Library of Korea**, Upstage LLM One Pager, January 2024
- **Ministry of Science and ICT - Data Utilization Council**, Upstage LLM One Pager, December 2023
- **Seoul Metropolitan Office of Education-AI & Digital Education Conference**, Reset Moment by Large Language Model, December 2023
- **Digital Innovation Forum (Ministry of Culture, Sports and Tourism)**, Generative AI era: Copyright issues and countermeasures, November 2023
- **Kyungpook National University**, Upstage LLM One Pager, November 2023
- **Future Technology Exchange Conference in Southeast Region**, Upstage LLM One Pager, November 2023
- **Jeju National University - Guest lecture**, Upstage LLM One Pager, November 2023
- **Busan University of Foreign Studies**, From Language Model to Large Language Model, November 2023
- **KB Kookmin Bank**, Areas of Generative AI Application in Business, November 2023
- **Human-Inspired AI Research**, Upstage LLM One Pager, November 2023
- **NIA**, Upstage LLM and latest B2B LLM trends, October 2023
- **Woongjin Thinkbig**, Upstage LLM One Pager, September 2023
- **Kyobo Life Insurance**, Upstage and Private LLM, September 2023
- **HD Korea Shipbuilding & Offshore Engineering**, Upstage and Private LLM, September 2023
- **Korea University - Guest lecture**, From Language Model to Large Language Model, September 2023
- **HanYang University - Guest lecture**, From Language Model to Large Language Model, September 2023
- **Korea Institute of Patent Information**, Upstage LLM and latest B2B LLM trends, September 2023
- **AGI Town in Seoul**, The use case of solving customer problems using LLM, September 2023
- **Fast Campus**, Job Employment Special Lecture, September 2023
- **HD Korea Shipbuilding & Offshore Engineering**, Upstage LLM and latest B2B LLM trends, August 2023
- **Upstage Webinar**, The Future of Finance/Insurance Transformed by Generative AI, August 2023
- **Google I/O Extended 2023 Incheon**, From Language Model to Large Language Model, August 2023
- **POSCO RIST**, Data-Centric AI in Real-World, July 2023

- **Google I/O Extended 2023 Seoul**, From Language Model to Large Language Model, July 2023
- **TensorFlow Korea LLM Day**, Language Model to Large Language Model, July 2023
- **Jeju National University**, Real-World Artificial Intelligence and Large Language Model for Everyone, July 2023
- **AI EDucation Alliance Policy lab (AIEDAP)**, Deep Learning Understanding and Practice, June 2023
- **KIDA (Korea Institute for Defense Analysis)**, Data-Centric AI in the Large Language Model Era, May 2023
- **Upstage**, NLP based Large Language Model for All, April 2023
- **AI·DATA SUMMIT 2023**, Real-World Centric AI, February 2023
- **Sunmoon University**, Real-World Centric AI, December 2022
- **Kyungsung University**, Language and Information Studies and the Future of Artificial Intelligence, August 2022
- **Hankuk University of Foreign Studies**, Basic practice of natural language processing for everyone, July 2022
- **Dongguk University**, Artificial intelligence and Machine Translation, January 2022
- **Busan Social Welfare Development Group**, Attending advisory meetings and Focus Group Interview, July 2021
- **LLsoLLu**, Latest natural language processing Research, March 2020
- **NC SOFT**, Technology Transfer Seminar, February 2020
- **Dongguk University**, A.I - NLP - MT for Liberal Arts, January 2020
- **SKC**, Text Preprocessing, Machine Translation, Language Embedding, October-November 2019
- **NAVER**, Machine Translation for everyone, August 2019