

Tak-Sung Heo

CONTACT INFORMATION 546-16 Mok-dong, Yangcheon-gu, Seoul **MILITARY SERVICE** 2014.07.15. ~ 2016.04.14. Army Sergeant, Honorable Discharge

DATE OF BIRTH 1994.02.21. **E-MAIL** gjxkrtjd221@gmail.com

PHONE (+82)-10-3997-4664 **GITHUB** <https://github.com/HeoTaksung>

COMPANY/ POSITION NHN Diquest (2021.02.01 ~), NLP AI Researcher [Part Leader]



INTERESTS Interested in Natural Language Processing, Biomedical Natural Language Processing, and Deep Learning, especially about Natural Language Understanding.

EDUCATION **Hallym University**, Chuncheon-si, Gangwon-do, Republic of Korea **Mar. 2019 – Feb. 2021**
Department of Convergence Software
Master of Science (CGPA: 4.25/4.5)
Advisor: Professor Yu-Seop Kim
Research Area: Natural Language Processing, Biomedical Natural Language Processing and Deep Learning

Hallym University, Chuncheon-si, Gangwon-do, Republic of Korea **Mar. 2013 – Feb. 2019**
Major Department of Life Science and Double Major Department of Convergence Software
Bachelor of Science (CGPA: 3.53/4.5)

EXPERIENCED TASK

- Document/Sentence Classification
- Dialogue State Tracking
- Sentence Similarity Measuring
- Sentiment Analysis
- Data Augmentation

CAREER **NHN DIQUEST** **Feb.2021 ~**
• Performance of research papers related to assignments **Feb.2021 ~ Dec.2021**
Aim: Achievement of research performance target

- Paper performance of “Development of data augmentation technology by using heterogeneous information and data fusion”
- Paper performance of “Development of deep tagging and 2D virtual try on for fashion online channels to provide mixed reality visualized service based on fashion attributes”
- Paper performance of “Development of Self-Detection Technology for Online Grooming in Social Networks”
- Paper performance of “Development of causal reasoning and expression technology to enhance the proof of digital evidence”

• Development of Self-Detection Technology for Online Grooming in Social Networks **Apr.2022 ~[Dec.2024]**
Aim: Development of online grooming detection intelligence to understand the semantics of media services of SNS and chatting apps

- Establishment and distribution of guidelines for data building related to online grooming
- Training for cloud worker to build online grooming data and Conducting of data quality inspection
- Building of Pre-trained Language Model specialized for online grooming (KConvo-RoBERTa)
- Development of deep learning model for online grooming semantic analysis
- “Development of Self-Detection Technology for Online Grooming in Social Networks” Patent

● **Korean SNS multi-turn conversation data**

May.2023 ~ Dec.2023

Aim: Building Korean SNS multi-turn conversation data

- Definition of multi-turn conversation data labeling
- Dialogue State Tracking modeling for multi-turn conversation data validation

PROJECTS

● **Building language resources and developing deep learning/natural language processing for automatic language disorder diagnosis** Apr.2019 ~ Feb.2021

Aim: Automating language analysis using conversation data from early childhood through high school.

Role: Used LDA and Sent2Vec to automate topic change detection, which is one of the methods of measuring the development of conversational ability.

● **Development of an automatic prognosis prediction system for cerebral infarction through natural language processing based on deep learning** Jan.2019 ~ Feb.2021

Aim: Development of a deep learning algorithm that can predict important clinical outcomes using text from electronic medical records.

Role: Applied a deep learning algorithm (CNN, LSTM, BERT) to the text of electronic medical record.

● **Reliability and validity verification of automatic evaluation of machine translation and application to the evaluation of human translation** Jan.2019 ~ Jun.2019

Aim: Establishing a system to verify the reliability and validity of automatic translation evaluation.

Role: Direct implementation of BLEU and METEOR, the metrics for evaluating the quality of machine translation.

AWARDS

● **Excellence Award at KSC 2019 Undergraduate/Junior Paper Contest (2019)**

Korean Institute of Information Scientists and Engineers (KIISE)

● **2018 SW WEEK Contest Code Ground Gold Award (2018) – 1st place**

Hallym University, Chuncheon-si, Gangwon-do, Republic of Korea

SKILLS

- **Language:** Python, Java
- **Frameworks:** TensorFlow, Keras, Pytorch
- **Others:** Pycharm, Jupyter notebook

CONFERENCES

[HCLT '23] [Keyword Based Conversation Generation using Large Language Model](#)

Juhwan Lee, [Tak-Sung Heo](#), Jisu Kim, Minsu Jeong, Kyounguk Lee, Kyoungsun Kim

The 35th Annual Conference on Human & Cognitive Language Technology (HCLT), 2023 [KIISE]

[HCLT '23] [Methodology for Overcoming the Problem of Position Embedding Length Limitation in Pre-training Models](#)

Minsu Jeong, [Tak-Sung Heo](#), Juhwan Lee, Jisu Kim, Kyounguk Lee, Kyoungsun Kim

The 35th Annual Conference on Human & Cognitive Language Technology (HCLT), 2023 [KIISE]

[ICMLA '21] [Medical Code Prediction from Discharge Summary: Document to Sequence BERT using Sequence Attention](#)

[Tak-Sung Heo](#)*, Yongmin Yoo*, Yeongjoon Park*, Byeong-Cheol Jo*, Kyoungsun Kim

The 20th IEEE International Conference on Machine Learning and Applications (ICMLA), 2021 [IEEE]

*These authors contributed equally

[ClinicalNLP '20] [Various Levels of Representation for Predicting Stroke Prognosis using Text Records of Magnetic Resonance Imaging](#)

[Tak-Sung Heo](#), Chulho Kim, Jeong-Myeong Choi, Yeong-Seok Jeong, Yu-Seop Kim

The 3rd Clinical Natural Language Processing Workshop (ClinicalNLP), 2020 [EMNLP]

[HCLT '20] Korean sentence spacing correction model using syllable and morpheme information

Jeong-Myeong Choi, Byoung-Doo Oh, Tak-Sung Heo, Yeong-Seok Jeong, Yu-Seop Kim

The 32nd Annual Conference on Human & Cognitive Language Technology (HCLT), 2020 [KIISE]

[HCLT '20] Attention based multimodal model for Korean speech recognition post-editing

Yeong-Seok Jeong, Byoung-Doo Oh, Tak-Sung Heo, Jeong-Myeong Choi, Yu-Seop Kim

The 32nd Annual Conference on Human & Cognitive Language Technology (HCLT), 2020 [KIISE]

[KSC '19] Depression Judgment System based on Deep Neural Network

Seok-Ju Park, Byoung-Doo Oh, Tak-Sung Heo, Yu-Seop Kim

Proceedings of Korea Software Congress (KSC), 2019 [KIISE]

[KSC '19] The performance comparison of Korean text tokenizing and defining stopwords for sentiment analysis

Yeong-Seok Jeong, Tak-Sung Heo, Yu-Seop Kim

Proceedings of Korea Software Congress (KSC), 2019 [KIISE]

[KSC '19] Measurement of the number of topics in children's speech using LDA and Affinity propagation algorithm

Se-Eun Oh, Tak-Sung Heo, Yoonkyoung Lee, Yu-Seop Kim

Proceedings of Korea Software Congress (KSC), 2019 [KIISE]

[HCLT '19] Detection of Topic Changes in Child Speech Using Sent2Vec

Tak-Sung Heo, Yoonkyoung Lee, Yu-Seop Kim

The 31st Annual Conference on Human & Cognitive Language Technology (HCLT), 2019 [KIISE]

[HCLT '18] Prediction of the age of speakers based on Convolutional Neural Networks and polarization model

Tak-Sung Heo, Ji-Soo Kim, Byoung-Doo Oh, Yu-Seop Kim

The 30th Annual Conference on Human & Cognitive Language Technology (HCLT), 2018 [KIISE]

[HCLT '18] Automatic Analysis Service for Korean Speaking by Age

Ji-Eun Choi, Byoung-Doo Oh, Tak-Sung Heo, Yu-Seop Kim

The 30th Annual Conference on Human & Cognitive Language Technology (HCLT), 2018 [KIISE]

JOURNALS

[(KCI), Journal of Translation Studies '22] Significance of Recall in Automatic Metrics for HT Evaluation

Hyeyeon Chung, Jisoo Choi, Tak-Sung Heo, Soo-Young Seo

Journal of Translation Studies, Mar.2022 [KATS]

[(KCI), zfdsl '21] Die Applikabilität der automatischen Evaluation von Humanübersetzungen (English: The applicability of the automatic evaluation of human translations)

Hyeyeon Chung, Hye-jeong Myeong*, Hye-Rim Choi*, Tak-Sung Heo*

*These authors contributed equally

Zfdsl, Aug.2021 [KDSDL]

[(SCIE), Symmetry '21] A Novel Hybrid Methodology of Measuring Sentence Similarity

Yongmin Yoo*, Tak-sung Heo*, Yeongjoon Park*, Kyoungsun Kim

*These authors contributed equally

Symmetry, Aug.2021 [MDPI]

[(SCIE), Appl. Sci. '21] Global and Local Information Adjustment for Semantic Similarity Evaluation

Tak-Sung Heo, Jong-Dae Kim, Chan-Young Park, Yu-Seop Kim

Applied Sciences, Mar.2021 [MDPI]

[(SCIE), Sens. Mater. '21] Prediction of Atrial Fibrillation Cases: Convolutional Neural Networks using the Output Texts of Electrocardiography

Tak-Sung Heo, Chulho Kim, Jong-Dae Kim, Chan-Young Park, Yu-Seop Kim

Sensors and Materials, Jan.2021 [MYU]

[(SCIE), *Sci. Rep.* '21] [Deep learning based prediction of prognosis in nonmetastatic clear cell renal cell carcinoma](#)

Seok-Soo Byun, Tak-Sung Heo, Jeong-Myeong Choi, Yeong-Seok Jeong, Yu-Seop Kim, Won-Ki Lee, Chulho Kim

Scientific Reports, Jan.2021 [Nature]

[(SCIE), *J. Intell. Fuzzy Syst.* '21] [Sentence Similarity Evaluation using Sent2Vec and Siamese Neural Network with Parallel Structure](#)

Tak-Sung Heo, Jong-Dae Kim, Chan-Young Park, Yu-Seop Kim

Journal of Intelligent and Fuzzy Systems, Jan.2021 [IOS Press]

[(SCIE), *J. Pers. Med.* '20] [Prediction of Stroke Outcome Using Natural Language Processing-Based Machine Learning of Radiology Report of Brain MRI](#)

Tak-Sung Heo, Yu-Seop Kim, Jeong-Myeong Choi, Yeong-Seok Jeong, Soo-Young Seo, Jun-Ho Lee, Jin-Pyeong Jeon, Chulho Kim

Journal of Personalized Medicine, Dec.2020 [MDPI]

[(KCI), *Journal of Translation Studies* '20] [Application of Automatic Evaluation to Human Translation](#)

Bo-Young Kim, Yeon-Joo Kim, Seung-Hee Seo, Shin-Ae Song, Jin-Hyun Lee, Kyoung-Ah Jeon, Ji-Soo Choi, Seung-Bin Hong, Hye-yeon Chung, Tak-Sung Heo

Journal of Translation Studies, Mar.2020 [KATS]

PENDING PAPER

[Arxiv'23] [Multi label classification of Artificial Intelligence related patents using Modified D2SBERT and Sentence Attention mechanism](#)

Yongmin Yoo, Tak-Sung Heo*, Dongjin Lim, Deaho Seo

[Arxiv'22] [DAGAM: Data Augmentation with Generation And Modification](#)

Byeong-Cheol Jo*, Tak-Sung Heo*, Yeongjoon Park, Yongmin Yoo, Won Ik Cho, Kyungsun Kim

*These authors contributed equally

SUMMARY

Tak-Sung Heo is a research and development engineer specializing in natural language processing. He has a particular interest in automating various problems through natural language understanding and wishes to develop and enhance natural language processing models' performance. He has conducted research in various fields, including text classification, dialogue state tracking, disease prediction, estimation of conversational ability development, sentence similarity evaluation, and data augmentation, using text data. Through his extensive research experience, he believes that he can identify crucial points when solving new problems and has the ability to quickly explore and decide on methods to address them, thus possessing numerous advantages.