

# 허탁성



주소 서울특별시 양천구 목동 546-16 병역사항 2014.07.15. ~ 2016.04.14.  
병장 만기 제대  
생년월일 1994.02.21. 이메일 gjxkrtjd221@gmail.com  
휴대전화 (+82)-10-3997-4664 GITHUB <https://github.com/HeoTaksung>

현재 회사/직책 NHN Diquest (2021.02.01 ~ ), NLP AI Researcher [Part Leader]

관심 분야 Natural Language Processing, Biomedical Natural Language Processing, and Deep Learning, especially about Natural Language Understanding.

학력 사항 한림대학교 2019.03. – 2021.02.

융합소프트웨어학과

공학 석사 (평균 학점: 4.25/4.5)

지도 교수: 김 유 섭

연구 분야: Natural Language Processing, Biomedical Natural Language Processing and Deep Learning

한림대학교 2013.03. – 2019.02.

주전공: 생명과학과, 복수전공: 융합소프트웨어학과

이학 학사 (평균 학점: 3.53/4.5)

연구 경험

- 문서/문장 분류
- 대화 상태 추적
- 문장 유사도 측정
- 감성 분석
- 데이터 증강

경력 사항

NHN DIQUEST 2021.02.01. ~

• 과제 관련 연구 논문 실적

2021.02. – 2023.12.

목표: 과제 연구 실적 목표량 달성

- 이종 정보 활용 및 데이터 융합을 통한 데이터 증식 논문 실적
- 패션속성기반 혼합현실 시각화 서비스 제공을 위한 패션 온라인채널용 딥태깅 기술 및 2D기반 가상피팅 기술 개발 논문 실적
- 소셜 네트워크에서 온라인 그루밍을 위한 자가 탐지 기술 개발 논문 실적
- 디지털 증거의 증명력 제고를 위한 인과관계 추론 및 표현 기술 개발 논문 실적

• 소셜 네트워크에서 온라인 그루밍을 위한 자가 탐지 기술 개발

2022.04. – [2024.12.]

목표: SNS 및 채팅 앱의 미디어 서비스의 의미를 이해하기 위한 온라인 그루밍 탐지 기술 개발

- 온라인그루밍 관련 데이터 구축에 관한 가이드라인 구축 및 배포
- 온라인그루밍 데이터 구축을 위한 클라우드 워커 교육 실시 및 데이터 검수
- 온라인그루밍에 특화된 Pretrained Language Model 구축 (KConvo-RoBERTa)
- 온라인그루밍 의미 분석을 위한 딥러닝 모델 개발
- “소셜 네트워크에서 온라인 그루밍을 위한 자가 탐지 기술 개발” 특허 출원

- 한국어 SNS 멀티턴 대화 데이터

**2023.05. – 2023.12.**

**목표:** 한국어 SNS 멀티턴 대화 데이터 구축

- 멀티턴 대화 데이터 라벨링 정의
- 멀티턴 대화 데이터 유효성 검증을 위한 Dialogue State Tracking 모델링

## 주요 연구 과제 및 프로젝트

- 자동 언어 장애 진단을 위한 언어 자원 구축 및 딥 러닝/자연어처리 개발

**2019.04. – 2021.02.**

**목표:** 유아기부터 고등학교까지의 대화 데이터를 사용하여 언어 분석을 자동화

**역할:** 각 연령에 따른 발화 데이터를 사용하여 대화 능력 발달을 측정하는 방법 중 하나인 화제 변경 탐지 자동화를 위한 딥 러닝 모델 개발

- 딥 러닝 기반 자연어처리를 통한 뇌경색 자동 예후 예측 시스템 개발

**2019.01. – 2021.02.**

**목표:** 전자의무기록의 텍스트를 이용하여 중요한 임상 결과를 예측할 수 있는 딥 러닝 알고리즘 개발

**역할:** MRI 텍스트 기록지, 심전도 텍스트 기록지를 사용하여 뇌경색 예후 및 심방세동 발생 가능성 예측을 위한 딥러닝 모델 개발

- 기계번역 자동평가의 신뢰도, 타당도 검증 및 인간번역 평가에의 적용

**2019.01. – 2019.06.**

**목표:** 자동번역평가의 신뢰성과 유효성 검증 시스템 구축

**역할:** 기계 번역 품질을 평가하기 위한 평가 지표인 BLEU 및 METEOR를 직접 구현

## 수상 내역

● 2019 KSC2019 학부생/주니어논문경진대회 우수상(2019)

한국정보과학회

● 2018 SW WEEK 경진 대회 Code Ground (2018) – 1st place

강원도 춘천시 한림대학교

## 보유 기술

- **Language:** Python, Java
- **Frameworks:** TensorFlow, Keras, Pytorch
- **Others:** Visual Studio Code, Jupyter Notebook

## 학회 논문

## [HCLT '23] Large Language Model을 활용한 키워드 기반 대화 생성

이주환, 허탁성, 김지수, 정민수, 이경욱, 김경선

제 35회 한글 및 한국어 정보처리 학술대회 (HCLT), 2023 [한국정보과학회]

## [HCLT '23] 사전 학습 모델의 위치 임베딩 길이 제한 문제를 극복하기 위한 방법론

정민수, 허택성, 이주환, 김지수, 이경욱, 김경선

제 35회 한글 및 한국어 정보처리 학술대회 (HCLT), 2023 [한국정보과학회]

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

허탁성\*, 유용민\*, 박영준\*, 조병철\*, 김경선

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

\*공동저자

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

허탁성, 김철호, 최정명, 정영석, 김유섭

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

**[HCLT '20] 음절과 형태소 정보를 이용한 한국어 문장 띄어쓰기 교정 모델**

최정명, 오병두, 허탁성, 정영석, 김유섭

제 32회 한글 및 한국어 정보처리 학술대회 (HCLT), 2020 [한국정보과학회]

**[HCLT '20] 한국어 음성인식 후처리를 위한 주의 집중 기반의 멀티모달 모델**

정영석, 오병두, 허탁성, 최정명, 김유섭

제 32회 한글 및 한국어 정보처리 학술대회 (HCLT), 2020 [한국정보과학회]

**[KSC '19] 심층 신경망 기반의 우울증상 판단 시스템**

박석주, 오병두, 허탁성, 김유섭

한국소프트웨어종합학술대회 (KSC), 2019 [한국정보과학회]

**[KSC '19] 한국어 감성 분석에서 토큰 단위와 불용어 처리 기준에 따른 성능 비교**

정영석, 허탁성, 김유섭

한국소프트웨어종합학술대회 (KSC), 2019 [한국정보과학회]

**[KSC '19] LDA와 선호도 전파 알고리즘을 이용한 아동 발화의 화제수 측정**

오세은, 허탁성, 이윤경, 김유섭

한국소프트웨어종합학술대회 (KSC), 2019 [한국정보과학회]

**[HCLT '19] Sent2Vec을 이용한 아동 발화의 화제 변화 디텍팅**

허탁성, 이윤경, 김유섭

제 31회 한글 및 한국어 정보처리 학술대회 (HCLT), 2019 [한국정보과학회]

**[HCLT '18] 합성곱 신경망 모델과 극단 모델에 기반한 발화자 연령 예측**

허탁성, 김지수, 오병두, 김유섭

제 30회 한글 및 한국어 정보처리 학술대회 (HCLT), 2018 [한국정보과학회]

**[HCLT '18] 연령별 한국어 발화 자동 분석 서비스**

최지은, 오병두, 허탁성, 김유섭

제 30회 한글 및 한국어 정보처리 학술대회 (HCLT), 2018 [한국정보과학회]

**저널 논문**

**[(KCI), 한국독일언어문학회 '21] Die Applikabilität der automatischen Evaluation von Humanüberse tungen (인간번역 자동평가에서 정답자와 평가자가 다르다면)**

정혜연, 명혜정\*, 최혜림\*, 허탁성\*

\*공동저자

Zfdsl, Aug.2021 [KDSSL]

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

유용민\*, 허탁성\*, 박영준\*, 김경선

\*공동저자

Symmetry, Aug.2021 [MDPI] (Impact Factor: 2.713)

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

허탁성, 김종대, 박찬영, 김유섭

Applied Sciences, Mar.2021 [MDPI] (Impact Factor: 2.474)

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

허탁성, 김철호, 김종대, 박찬영, 김유섭

Sensors and Materials, Jan.2021 [MYU] (Impact Factor: 0.560)

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

변석수, 허탁성, 최정명, 정영석, 김유섭, 이원기, 김철호  
Scientific Reports, Jan.2021 [Nature] (Impact Factor: 3.998)

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

허탁성, 김종대, 박찬영, 김유섭  
Journal of Intelligent and Fuzzy Systems, Jan.2021 [IOS Press] (Impact Factor: 1.851)

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

허탁성, 김유섭, 최정명, 정영석, 서수영, 이준호, 전진평, 김철호  
Journal of Personalized Medicine, Dec.2020 [MDPI] (Impact Factor: 4.433)

[(KCI), 번역학연구 '20] [번역자동평가에서 풀리지 않은 과제](#)

김보영, 김연주, 서승희, 송신애, 이진현, 전경아, 최지수, 홍승빈, 정혜연, 허탁성  
Journal of Translation Studies, Mar.2020 [KATS]

## 보류 논문

[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](#)

조병철\*, 허탁성\*, 박영준, 유용민, 조원익, 김경선

\*공동저자

## 간단 소개

저는 자연어처리를 전문으로 하는 연구개발자입니다. 자연어처리 기술 중 특히 자연어 이해를 통해 다양한 문제를 자동화하는 데 관심이 많으며, 자연어처리 모델 개발 및 고도화를 하고 싶습니다. 텍스트를 이용하여 텍스트 분류, 대화 상태 추적, 질병 예측, 대화 능력 발달 정도 예측, 문장 유사도 평가, 데이터 증강 등 여러 가지 분야의 연구를 진행해왔습니다. 이러한 많은 연구 경험을 통해, 저는 새로운 문제를 해결하고자 할 때 중요한 점을 캐치할 수 있으며, 이를 해결할 수 있는 방법들을 빠르게 탐색 및 결정할 수 있는 등 많은 이점을 가지고 있다고 생각합니다.