

몰입도 극대화를 위한 공감 대화 연구 현황

한국전자기술연구원 ■ 신사임·정민영·장진예·김산·조병길

1. 서론

대화 (Dialogues)란, 인간이 삶에 필요한 것들을 주고받기 위해 관련된 대상들과 의미 및 정보를 메시지로 교류하는 행위를 의미한다. 대화가 원만하고 효과적으로 이루어지기 위해 반드시 필요한 공감 (共感, Empathy)이란, 듣는 사람이 말하는 사람과 같은 수준에서 느끼는 것이다 [1]. 즉, 공감 대화 (Empathetic dialogues)란, 대화 상대의 입장에서 생각하고 느끼면서, 이를 기반으로 대화를 표현하는 인간의 행위이다.

상용 서비스가 활성화 되어있는 대표적인 대화 기술들을 살펴보면, 대화를 통해 특정 임무를 수행하도록 지원하는 AI 스피커와 챗봇 서비스 등이 있다. 현재, 임무 기반 대화 (Task-oriented dialogue) 기술을 기반으로 하는 대화 서비스들은 사용자와의 교감 측면에서의 접근이 없는 특정 기능을 제공하며, 사용자의 상태 및 상황에는 무관하게 동일한 발화 (Utterances)와 상태 전이 (State transition)를 수행하는 방식이다. 즉, 임무 기반 대화 기술의 특정 임무의 단발적인 수행만을 목적으로 하여 사용자 만족도가 낮은 점을 서비스 확산의 한계로 지목하고 있다.

반면, 인공지능 기술을 활용한 공감 대화 기술은 다양한 인공지능 기술을 활용하여 실시간으로 사용자의 상황과 상태를 파악하여 상황에 적절하면서도 필요한 발화를 지원하는 대화 기술이다. 공감 대화 기술의 상용화는 대화처리를 통한 서비스의 영역을 감성 및 증강지적 서비스로의 확대를 가능하게 하고 있다.

인간과 공감을 형성하는 공감 대화 기술의 적용 분야는 기존의 대화 기반 서비스들의 확장을 포함하여 다음과 같은 광범위한 사회문제 해결 및 교감형 서비스에 적용이 가능하다.

(메디칼) 정신건강 관련 서비스: 정신적인 불안정으로 인해 상시 도움이 필요하고 의료적 관찰이 필요한 정신질환자를 대상으로 보호자가 잠깐 자리를 비우거나 의료진과의 면담 상황이 아닌 시점에도 환자를 관찰하고 모니터링하는 인공지능 기반 정신질환자 돌봄 서비스

(사회문제 해결) 독거노인 돌봄 서비스: 독거노인 주생활 공간에 공감 서비스 디바이스를 배포하고 생활 관리사가 방문하지 못하는 시간에도 AI 디바이스가 노인의 상태에 적합한 대화로 공감 서비스를 지원하고 안정감을 제공

(휴먼-컴퓨터 인터페이스) AI 말벗 서비스: 사용자의 미묘한 감정 상태와 심리 상황을 가장 정확하게 인지할 수 있는 대화기술을 기반으로 사용자 눈높이와 취향에 맞는 다양한 서비스¹⁾를 제공

최근 수년간 거대 언어모델 (Large language model) [2] 기반의 자연어처리 기법의 빠른 확산과 모델 규모의 확장으로 인공지능 기반 언어처리를 활용한 높은 성능의 서비스들의 등장으로, 고도화된 인공지능 대화기술의 확산이 기대를 모으고 있는 시점이다. 반면에, 수백억 규모의 초거대 자본이 투자된 소수의 서비스들²⁾을 제외하면 사용자들이 요구하는 대화 만족도를 충족시키지 못하고 있는 것도 현실이다.

본 논문에서는 현재 국내에서 연구개발이 진행되고 있는 공감 대화 기술의 현황을 살펴보고, 미래 공감 대화 기술의 고도화를 위한 연구개발 방향성에 대해 살펴보고자 한다. 이를 위한 논문의 구성은 다음과 같다. 2장에서는 현재 대표적인 공감 대화 연구개발 프로젝트를 소개하고 이들의 기술적 특성을 설명하고, 3장에서는 미래 공감 대화 기술 고도화를 위한 인공지능 언어처리 연구의 방향성을 도출하고 결론을 맺고자 한다.

† 지원문구 또는 감사의 글: 이 논문은 2022년도 정보(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. 2022-0-00320, 상황인지 및 사용자 이해를 통한 인공지능 기반 1:1 복합대화 기술 개발)

1) 검색, 추천, AI 비서 등
2) ChatGPT [3], Bard [4] 등

2. 공감 대화 기술 현황

인공지능 기술 관점으로 볼 때, 공감 대화를 위한 요소 기술은 다음의 그림처럼 멀티모달 인지, 의도 및 감정 이해, 공감 표현 기술을 들 수 있으며, 각각의 기술의 정의는 다음과 같다.

멀티모달 인지: 다양한 모달리티의 데이터를 인공지능 기술로 사용자 상태와 사용자 주변 환경을 분석하고, 공감 대화를 위한 상황인지에 도움이 되는 다양한 상태와 상황 정보를 수집하는 기술

의도 및 감정 이해: 이전 발화들과 현재의 사용자 입력 발화를 ‘멀티모달 인지’ 모듈에서 인식된 상황 정보와 함께 분석하여, 사용자 입력 발화의 의도 (Intent)와 감정 (Emotion)을 분석하는 기술

공감 표현 기술: 이전 발화와 현재 사용자 입력 발화를 ‘멀티모달 인지’와 ‘의도 및 감정 이해’ 모듈에서 분석한 정보들을 모두 입력으로 활용하여, 적절한 공감을 표현하는 시스템 발화를 생성하는 기술

공감 대화 서비스의 고도화를 위해서는 관련 인공지능 처리 기술들, 특히 대화처리와 직접적으로 연관 있는 자연어처리 기술의 고도화가 필수적이다. 이를 위해서는, 다양한 한국어 공감 대화 연구용 데이터셋 구축 및 공개 활성화가 전제되어야 하며, 공감 대화의 공감성과 몰입도를 높이기 위한 대표적인 연구 방향성 별 연구 현황을 소개하고자 한다.

2.1 사전학습 기반 대화 모델링

공감 대화 서비스의 성패는 대화의 높은 몰입도가 좌우하며, 서비스 사용자가 지속적으로 서비스 에이

전트와 대화를 주고받으면서 흥미가 떨어지지 않도록 사용자를 이해하고 배려하면서 사용자가 공감할 만한 시스템 발화를 제공하는 데 있다. 따라서, ‘공감 표현 기술’에서 다양한 상황에 자연스럽게 대응하는 다양한 발화의 생성이 가능한 자연어처리 생성 성능이 보장되어야 한다. 또한, 안정적이면서도 정확한 ‘멀티모달 인지’와 ‘의도 및 감정 이해’의 분석 모듈의 상황 정보 수집 성능도 ‘공감 표현 모듈’의 적절한 공감 발화 생성을 좌우하기 때문에, 이들의 성능 고도화도 공감 대화의 성능향상에 매우 중요하다.

따라서, 최근 자연어처리 성능을 혁신적으로 향상시키고 있는 사전학습모델 (Pre-trained model)의 적용을 통한 공감 대화 요소 엔진들의 성능 고도화 추이가 공감 대화 기반 서비스의 상용화를 빠르게 앞당기고 있다. 자연어처리를 위한 사전학습모델, 즉, 거대언어모델을 활용하는 기법으로는 자기 지도학습 (Self-supervised learning)을 기반으로 수백만 개 파라미터 규모의 사전학습모델을 특정 임무 (Downstream task)에 최적화 되도록 재학습하는 미세 조정 (Fine-tuning) 방식과 [5], 수천만 개 파라미터 규모의 초거대 모델에 임무를 설명하여 원하는 결과를 얻는 비지도 학습 기반의 프롬프트 학습 (Prompt learning) 기법이 있다. 미세 조정 방식을 통한 대화 모델링에 활용 가능하도록 무료로 공개된 한국어 기반 거대 언어모델로는 생성기 (Decoder) 기반의 Ko-GPT2 (Korean-Generative Pre-trained Transformer)³⁾와 이해기 (Encoder)와 생성기를 모두 포함하고 있는 KE-T5 (Korean-English Text-to-Text Transfer Transformer)⁴⁾ [6], ET5 (ETRI Text-



그림 1 공감 대화 기술 개요

3) <https://github.com/SKT-AI/KoGPT2>

4) <https://github.com/AIRC-KETI/ke-t5>

to-Text Trasfer Transformer)⁵⁾ 등이 있다.

2.2 부가 정보 주입 기반 대화 모델링

공감 대화에서 사용자와의 원활한 공감을 위해서는 단조로운 대화를 벗어나서 다양한 주제와 상황 정보를 활용하여, 흥미롭게 대화를 이어 나가야 한다. 특히, 대화 중인 사용자에게 특화된 발화는 사용자의 적극적인 공감을 촉진하여서, 대화에 몰입하여 지속적인 서비스 활용을 유도할 수 있게 된다. 공감 발화 생성을 위해 대화 모델의 입력으로 주입하기 위한 부가 정보들은 ‘멀티모달 인지’ 모듈에서 도출한 사용자 상태 및 상황 정보들을 최대한 활용할 수 있다.

[7]은 부가 정보를 대화 모델에 주입하여 적응적 한국어 대화를 생성하는 모델링 기법을 보여주는 초창기 연구 사례를 보여준다. 여섯 종의 날씨 및 시간 관련 상황 정보를 벡터 (Vector)로 임베딩화 (Embedding) 하여 대화 발화와 함께 학습한 결과, 상황 정보를 활용하지 않은 대화 모델보다 더 좋은 성능을 보이는 것을 실험으로 증명하였다.

[8]은 영상 및 사용자 음성과 발화를 실시간으로 분석하여 사용자의 감정 상태를 분석하여 감정 기반 공감 발화를 생성하여 대응하도록 학습한 공감 대화 모

델링 기법을 소개한다. 본 연구는 발화 단위로 인식한 사용자 감정과 사용자의 대화 전략을 공감 대화 생성 모델에 지난 발화들과 입력으로 주입하여, 공감 발화를 생성하도록 학습하는 모델링 기법을 제안하였다. 또한, 주 발화는 아니지만 사용자가 대화 중간에 대응하는 공감 호응 (Back-channel)을 분석하여, 사용자의 공감 정도와 대화의 감정과 의도 등을 분석하기 위해 활용하는 공감 호응 모델링 기법 연구를 최초로 제안하였다 [9, 10].

대화 과정에 사용자의 공감을 유도하는 전략 중 하나로 대화 과정에 사용자의 관심과 상황에 맞는 지식을 공유하며 인터랙션을 주고받는 지식 기반 대화가 있다. 아래 그림의 왼쪽 지식 기반 대화 모델링 기법에서 볼 수 있듯이, 지식 기반 대화를 위해서는 대화의 입력과 출력을 학습하는 표현 모듈 외에 ‘멀티모달 인지’ 기술의 하나로 대화에 활용되기 적합한 지식을 추출하는 추가적인 ‘지식 추출기 (Retriever)’를 네트워크에 포함하고 있다 [6]. 이와 같은 구조를 통해, 지식 기반 대화 모델은 이전 발화와 주제에 맞는 발화에 활용할 자연어 지식을 함께 입력으로 받아서 지식 정보가 포함된 정보성 발화를 상황에 맞게 생성하도록 학습된다.

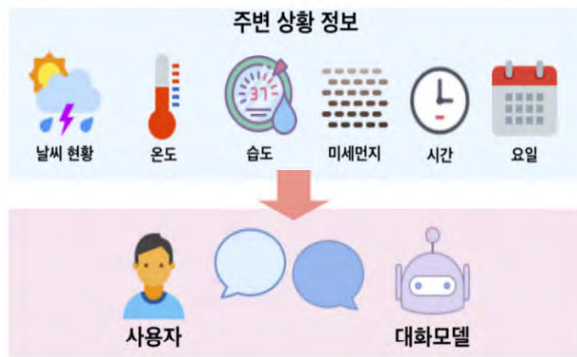


그림 2 상황 정보를 활용하는 대화 모델링 개요 [7]

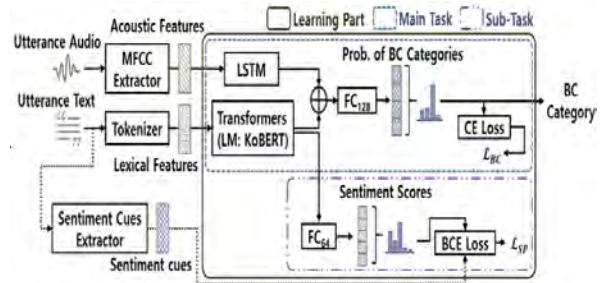
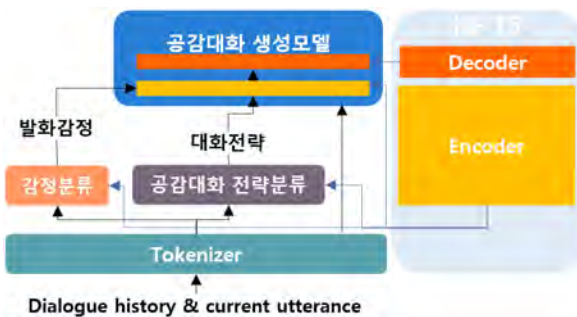
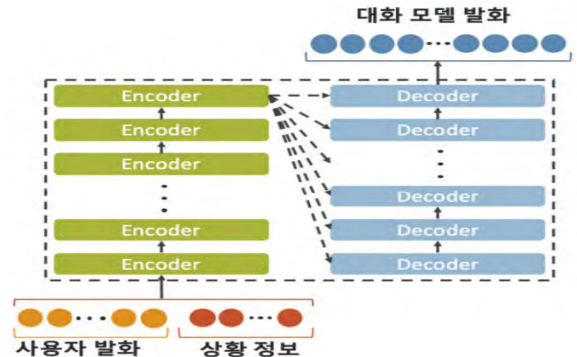


그림 3 감정 기반 공감 발화 엔진 및 공감 호응 인식 엔진 개요 [8]

5) <https://aiopen.etri.re.kr/et5Model>

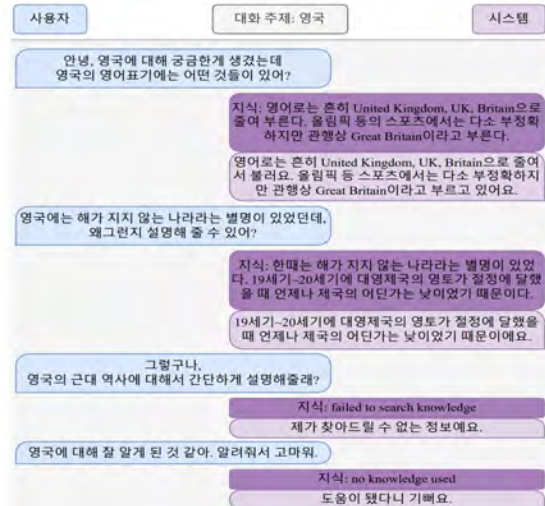
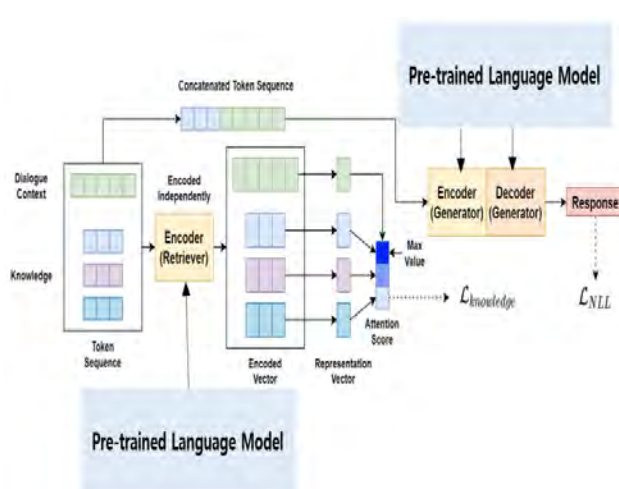


그림 4 지식 기반 대화 학습 기법 및 대화의 예

2.3 페르소나를 이해하는 대화 모델링

페르소나 (Persona)는 심리학에서 타인에게 비치는 외적 성격을 나타내는 용어이다 [11]. 인공지능이 인간 수준으로 자연스러운 대화를 생성하기 위해서는 대화 서비스 설계 시 서비스에 적절한 챗봇의 페르소나를 정교하게 설정하고, 대화 모델의 학습을 위한 데이터셋 구축 과정에서 해당 페르소나가 발화에 드러날 수 있도록 대화의 구성과 발화를 구축해야 한다.

공감 대화를 위한 대화 모델에서는, 인공지능에 이전트의 페르소나를 넘어서 대화 상대의 페르소나를 이해하고 이를 배려하기 위한 마음 이론 (Theory of Mind, ToM) 기반의 모델링 기법의 연구가 진행 중이다 [12]. 마음 이론은 타인이 자신과는 다른 신념, 욕구, 의도, 관점을 가지고 있다는 것을 이해하는 것으로 [12], 마음 이론을 반영한 대화 모델링 기법은 공감 대화의 과정에서 인공지능이 대화 사용자의 성격과 상태를 점진적으로 이해하게 되면서 사용자와의

공감도를 높이는 효과적인 공감 대화 서비스의 제공이 가능하다.

다중-성격 신뢰 추적기 (Multi-character belief tracker) 기반 대화 모델링 기법은 대화 과정에 고려해야 하는 모든 관점의 마음 이론 신뢰 그래프를 실시간으로 구성하여 관리하면서, 발화를 생성하는 거대언어모델의 프롬프트로 이전 발화들과 현재 상태의 신뢰 그래프들을 모두 제공하는 기법을 제안하였다. 이를 통해, 사용자와 인공지능의 마음 상태를 모두 반영한 발화들이 생성되도록 유도할 수 있어서, 기존 기법들보다 공감도 관점에서 높은 대화 만족도를 보여주었다 [12].

공감 대화를 위한 상대방과 인공지능의 페르소나를 이해하는 일관된 대화를 주고받기 위해서는, 인공지능이 특정 페르소나에서 연관하여 생각할 수 있는 대화의 주제를 다양하게 활용하여 발화를 생성할 수 있어야 한다. 이를 위한 페르소나 상식 지식베이스인 PEACOK (Persona commonsense knowledge for

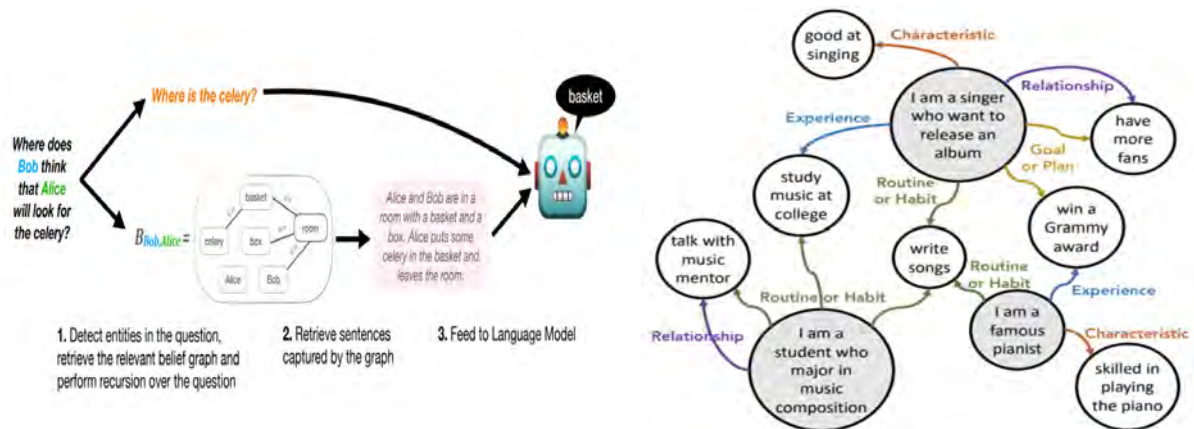


그림 5 신뢰 그래프 기반 대화 모델링 과정 (좌, [13]) 및 페르소나 대화를 위해 구축한 상식 그래프의 예 (우, [14])

consistent and engaging narratives)도 구축 및 공개되었다 [14]. PEACOK는 다양한 페르소나⁶⁾에 대한 특성, 습관, 취미, 행동, 관계에 대한 객체 (Entity) 간의 관계를 그래프 형태로 표현하여 구축하였다. PEACOK의 페르소나 관련 일반적인 지식들은 공감 대화의 발화 생성 과정에 사용자와 인공 지능의 페르소나와 연관된 그래프들을 이전 발화와 함께 발화 생성 과정에 함께 주입함으로써, 페르소나의 특성을 반영한 발화의 생성을 유도하는 데 활용할 수 있다.

3. 결 론

현재 공감 대화 모델링 기법은 돌봄, 상담 및 말벗 서비스 같은 감정 공유를 요하는 특정 서비스를 지원하기 위한 활용성을 보인다. 그러나, 궁극적으로 대화의 공감성은 대화의 목적성과 상관없이 인간을 대체할 수 있는 인공 지능 대화 기술의 확보를 위해 기본적으로 필요한 기술이라고 할 수 있다. 향후, 공감 대화 모델링 기법은 다양한 형태의 임무 지향 대화 (Task-oriented dialogues)를 요구하는 다양한 목적의 휴먼 에이전트 서비스⁷⁾에서도 기본적으로 제공되어야 하는 기술로 자리 잡을 것이다.

이 같은 공감 대화 기술의 빠른 고도화와 확장을 요구하는 흐름에 따라, 공감 대화 기술의 향후 연구의 방향성은 미래 인공 지능 기반 대화기술의 고도화와 발맞추어서 연구 방향성도 다양해질 것이다. 향후 인공 지능 대화 에이전트는 단발성 대화의 처리를 넘어서서 긴 (Multi-turn) 대화, 중장기적으로 산발적이면서도 연속적으로 진행되는 연속 (Multi-session) 대화를 처리하는 과정에서 지속적인 대화의 공감성을 유지하는 연구가 필요하다. 또한, 최근 빠르게 늘어나고 있는 무인 키오스크 기반의 업무처리 서비스에 활용하기 위한 비언어 (Nonverbal)와 준언어적 (Para-verbal) 표현과 언어적 (Verbal) 표현을 연계 활용하여 대화의 공감성을 높이는 복합대화 기반 공감 대화 모델링 기술의 연구도 필요할 것이다 [15].

참고문헌

- [1] 김나연, “말의 품격을 더하는 보이스 스타일링”, 도서출판 보일러, 2020.
- [2] Wikipedia, “Large language model”, https://en.wikipedia.org/wiki/Large_language_model.

- [3] N. Summers, “ChatGPT - Release Notes”, 2023.
- [4] Google Support, “Where you can use Bard”, 2023.
- [5] L. Bouchard, “What is Self-Supervised Learning? | Will machines ever be able to learn like humans?”. Medium, 2021.
- [6] S. Kim, J. Y. Jang, M. Jung, and S. Shin, “A Model of Cross-Lingual Knowledge-Grounded Response Generation for Open-Domain Dialogue Systems”, In proceeding of conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 352 - 365, 2021.
- [7] 장진예, 정민영, 박한무, 신사임, “외부 상황 정보를 활용하는 적응적 대화 모델의 구현”, 한글 및 한국어 정보처리 학술대회, pp. 456-459, 2019.
- [8] M. Jung, Y. Lim, S. Kim, J. Y. Jang, S. Shin, K. Lee, “An emotion-based Korean multimodal empathetic dialogue system”, In proceeding of workshop on When Creative AI Meets Conversational AI (CAI2), pp. 16 - 22, 2022.
- [9] J. Y. Jang, S. Kim, M. Jung, S. Shin, G. Gweon, “BPM_MT: Enhanced backchannel prediction model using multi-task learning”, In proceeding of conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 3447 - 3452, 2021.
- [10] 신동혁, 양재희, 장진예, 신사임, “상담 전략을 통합한 정서 교감형 챗봇 개발을 위한 데이터셋 구조 제안”, 한글 및 한국어 정보처리 학술대회, pp. 179-184, 2023.
- [11] Wikipedia, “페르소나”, <https://ko.wikipedia.org/wiki/페르소나>.
- [12] Wikipedia, “Theory of mind”, https://en.wikipedia.org/wiki/Theory_of_mind.
- [13] M. Sclar, S. Kumar, P. West, A. Suhr, Y. Choi, Y. Tsvekov, “Minding language models’ (Lack of) theory of mind: A plug-and-play multi-character belief tracker”, In proceeding of annual Meeting of the Association for Computational Linguistics (ACL), pp. 13960-13980, 2023.
- [14] S. Gao, B. Borges, S. Oh, D. Bayazit, S. Kanno, H. Wakaki, Y. Mitsufuji, A. Bosselut, “PEACOK: Persona commonsense knowledge for consistent and engaging narratives”, In proceeding of annual Meeting of the Association for Computational Linguistics (ACL), pp. 6569 - 6591, 2023.
- [15] 김재운, 장진예, 김산, 정민영, 강현욱, 신사임, “한국어 발화 문장에 대한 비언어 표현 정보를 자동으로 생성하는 모델”, 한글 및 한국어 정보처리 학술대회, pp. 91-94, 2023.

⁶⁾ 직업, 역할 등

⁷⁾ AI 튜터, AI 비서, 정보제공을 위한 AI 스피커 등

약 력



신 사 임

2000 숙명여대 전산학과 졸업 (학사)
 2002 한국과학기술원 전산학과 졸업 (석사)
 2018 서강대학교 컴퓨터공학과 졸업 (박사)
 2006~현재 한국전자기술연구원 인공지능연구센터 센터장
 관심분야: 인공지능, 자연어처리
 Email : sishin@keti.re.kr



정 민 영

2012 연세대학교 컴퓨터과학과 졸업 (학사)
 2019 연세대학교 컴퓨터과학과 졸업 (석박사통합)
 2019~현재 한국전자기술연구원 인공지능연구센터 선임연구원
 관심분야: 자연어처리, 인공지능 서비스 구축 운용
 Email : minyoung.jung@keti.re.kr



장 진 예

2005 경북대학교 전기전자 및 컴퓨터공학과 졸업 (학사)
 2007 경북대학교 전자공학과 졸업 (석사)
 2013 Penn State University 정보과학과 졸업 (석사)
 2017~현재 한국전자기술연구원 인공지능연구센터 선임연구원
 2021 서울대학교 융합대학원 수료 (박사)
 관심분야: 자연어처리, 대화 모델링
 Email : jinyea.jang@keti.re.kr



김 산

2017 경희대학교 전자전파공학과 졸업 (학사)
 2019 경희대학교 전자공학과 졸업 (석사)
 2019~현재 한국전자기술연구원 인공지능연구센터 선임연구원
 관심분야: 자연어처리, 지식처리, 생성 AI
 Email : kimsan0622@keti.re.kr



조 병 길

2015 한국과학기술원 전산학과 졸업 (학사)
 2017 한국과학기술원 전산학과 졸업 (석사)
 2023 한국과학기술원 전산학과 졸업 (박사)
 2023~현재 한국전자기술연구원 인공지능연구센터 선임연구원
 관심분야: 자연어처리, 복합추론, 인공지능 모델 보안
 Email : byunggil@keti.re.kr