

프로필

이름	서재영	이메일	tjwodud04@gmail.com
GitHub	https://github.com/tjwodud04	소개 페이지	https://tjwodud04.github.io/intro/?lang=ko
Hugging Face	https://huggingface.co/dddsaty	LinkedIn	https://www.linkedin.com/in/jaeyoungsuh/

경력

회사명	근무형태	직무	기간
(주)아이피아	계약직	AI 엔지니어 (이미지 생성)	2025.10 ~ 2026.01
(주)에이아이오토오 부설연구소	정규직	AI 엔지니어 (LLM)	2023.09 ~ 2025.03
(주)바스젠바이오 기업부설연구소	정규직	AI 엔지니어 (NLP, CV)	2022.03 ~ 2023.09

학력

학위	학교명	기간	학위/전공명
석사	한양대학교 일반대학원	2020.03 ~ 2022.02	인공지능석사 (데이터사이언스전공)
학사	독학학위제	2020.02	공학사 (컴퓨터과학)
학사	연세대학교 원주캠퍼스	2014.03 ~ 2018.02	문학사 (역사문화학, 영어영문학)

논문

연도	논문명	PDF
2025	Jae Young Suh, Mingyu Jeon, "A Character-Based Korean Tourist Dialogue System with Proactive Recommendations and Live2D Embodiment", engrXiv preprint, 2025 개요: 한국 관광공사 TourAPI와 Live2D 아바타를 결합한 웹 기반 캐릭터 관광 안내 시스템을 구축했습니다. 사전 예방적 (Proactive) 정책을 적용하여 사용자에게 능동적으로 여행 코스를 추천하는 기능을 구현했습니다. 평가 결과 WER 6.4%의 음성 인식 정확도를 기록했으며, 시각적 피드백을 통한 몰입감 있는 관광 정보 제공 가능성을 입증했습니다.	PDF
	Mingyu Jeon, Jaeyoung Suh, Suwan Cho, and Dohyeon Kim, "How Far Can LLMs Emulate Human Behavior?: A Strategic Analysis via the Buy-and-Sell Negotiation Game", arXiv preprint, 2025 개요: '구매 및 판매 협상 게임' 시뮬레이션을 통해 거대언어모델(LLM)의 사회적 상호작용 및 전략적 의사결정 능력을 정량적으로 평가했습니다. 모델별로 부여된 페르소나에 따른 협상 전략의 변화와 승률, Shap Value를 분석하여 인간 행동 모사 능력을 측정했습니다. 실험을 통해 기존 지식 기반 벤치마크와 실제 사회적 맥락에서의 모델 수행 능력 간 차이를 규명했습니다.	PDF
	Mingyu Jeon, Suwan Cho, and Jae Young Suh. "PPoGA: Predictive Plan-on-Graph with Action for Knowledge Graph Question Answering." Accepted to GMILLM 2025 (Frontiers in Graph Machine Learning for the Large Model Era), CIKM 2025 Workshop 개요: 지식 그래프 질의응답(KGQA)의 유연성을 높이기 위해 플래너-실행자 구조를 가진 'PPoGA' 프레임워크를 제안했습니다. 행동 전 결과를 시뮬레이션하는 예측 처리와 2단계 자체 수정 메커니즘을 통해 복잡한 추론 문제 해결 능력을 강화했습니다. GrailQA 등 주요 벤치마크에서 기존 Prompting 방식을 능가하는 성능을 달성했습니다.	PDF
	Jae Young Suh and Mingyu Jeon. "A Modular Prototype of Emotion-Aware Proactive Voice Agent with Live2D"	PDF

	<p>Embodiment." Accepted to ProActLLM 2025 (Proactive Conversational Information Seeking with Large Language Models), CIKM 2025 Workshop</p> <p>개요: Whisper STT, 감정 인식, 선제적 정책 엔진을 통합한 모듈형 음성 에이전트 아키텍처를 설계했습니다. 쿨다운 메커니즘과 경량 스코어링 모델을 통해 상황에 맞는 능동적 개입 빈도를 제어하며 시스템 안정성을 확보했습니다. 동양의 '칠정(七情)' 분류법 기반 감정 추론과 Live2D 외형을 결합한 공감 대화 시스템을 구현했습니다.</p>	
2024	<p>Jae Young Suh, Eunchan Lee, Yohan Jeong, Donggil Park, and Sungmin Ahn. "Teaching Large Language Models to Understand Jeju Island with Domain-Adaptive Pretraining." 2nd International Conference on Foundation and Large Language Models (FLLM), pp. 21–28, 2024.</p> <p>개요: 도메인 적응형 사전 학습(DAPT)과 LoRA 기법을 활용해 Llama 3 모델에 제주도 특화 지식을 효과적으로 학습시켰습니다. 47,000여 개의 Q&A 쌍으로 구성된 JejuQA 데이터 세트를 구축하여 지역 정보 제공의 정확도를 높였습니다. 학습 결과 기준 모델 대비 F1 점수가 향상되었으며, 환각 현상을 억제한 정확한 도메인 지식 전달이 가능함을 입증했습니다.</p>	PDF
2021	<p>Jae Young Suh, Casey C. Bennett, Benjamin Weiss, Eunseo Yoon, Jihong Jeong, and Yejin Chae. "Development of Speech Dialogue Systems for Social AI in Cooperative Game Environments." IEEE Region 10 Symposium (TENSYMP 2021), pp. 1–4, 2021</p> <p>개요: 협동 게임 환경 내 소셜 AI를 구축하기 위해 데이터 기반의 음성 대화 시스템 개발 방법론을 연구했습니다. 사람 간 상호 작용 분석을 통해 게임 상태를 정의하고, 실시간 게임 데이터와 연동되어 자율적으로 발화하는 시스템을 구현했습니다. 사용자 테스트를 통해 AI가 더 사람과 닮을 수 있게끔 하는 반응성, 자연스러움 등 6가지 핵심 구성 요소를 도출했습니다.</p>	PDF

자격증

자격증명	취득일
OPIc IH	2025.09
AICE(AI Certificate for Everyone)	2023.08
ADsP(Advanced Data Analytics semi-Professional)	2022.11
SQLD(SQL Developer)	2022.09
정보처리기사	2019.05

프로젝트

연도	소속	프로젝트명	내용	
2025		Live2D 캐릭터 음성 챗봇 데모 구현	- OpenAI API와 Live2D 샘플 캐릭터를 연동하여 사용자 음성에 반응하는 챗봇 구현	링크
2024		LLM 기반 질의응답 시스템 프로토타입	- Crawl4AI, PandasAI 및 Streamlit을 사용한 부동산 검색 프로토타입 구현 기술스택: Python, OpenAI API, PandasAI, Streamlit	링크
2024	에이아이오투오	제주 관광 데이터 기반 LLM 튜닝	- Llama 3, Polyglot-ko 모델을 활용하여 제주도 관광지 데이터를 기반으로 fine-tuning 실험 수행 - 논문으로 정리하여 학회에 투고 기술스택: Python, Llama3, Pytorch, Huggingface	
2022	바스젠바이오	바이오 도메인	- Papers With Code 기반 공개 바이오 의학 데이터를	

		NER 테스트	<p>활용하여 질병, 기관, 유전자 등 개체를 분류하는 NER(Named Entity Recognition) 모델 실험을 수행</p> <ul style="list-style-type: none"> - Simple Transformers 및 Flair 프레임워크 기반으로 BioBERT, BioLinkBERT 등 모델 학습 및 결과 비교 <p>기술스택: Python, PyTorch, Simple Transformers, Flair</p>	
2021	한양대학교	HCI 프로젝트	<p>석사과정 중 "Don't Starve Together" 협동 게임 환경에서 사람과 AI의 상호작용에 대한 실험을 수행</p> <ul style="list-style-type: none"> - 규칙 기반 STT/TTS 응답을 구성하고, 음성 인터페이스에 대한 내부 간이 피드백 및 개선점 수집 <p>기술스택: pyttsx3, MS Azure STT/TTS, Lua (게임 모드 연동)</p>	링크

추가 기술서

연도	내용																								
	프로젝트명 <ul style="list-style-type: none"> • Live2D 캐릭터와 음성 인터랙션을 활용한 한국어 챗봇 프로토타입 개발 내용																								
	<table border="1"> <thead> <tr> <th>버전</th><th>주요 목표</th><th>사용 기술</th><th>역할 및 내용</th><th>비고</th></tr> </thead> <tbody> <tr> <td>v1</td><td>한국어 음성 기반 Live2D 캐릭터 챗봇의 최소 동작 버전 구현</td><td>whisper-1, gpt-4o audio-preview, Live2D, Flask, Vercel, HTML/CSS, JS</td><td> <ul style="list-style-type: none"> STT → LLM → TTS로 이어지는 단일 대화 파이프라인 설계 • Live2D 입 모양과 음성 출력 연동, 기본 립싱크 구현 • Flask 백엔드와 브라우저 프론트엔드 연결, Vercel 배포 </td><td> <ul style="list-style-type: none"> • 이후 감정 표현·선제 대화 실험(v2)의 기반 • 동작하는 최소 단위 확보에 초점 </td></tr> <tr> <td>v2</td><td>감정·상태 반영 및 선제 메시지 지원으로 캐릭터성 강화</td><td>gpt-4o mini tts, gpt-4o mini search preview, whisper-1, Live2D, Flask, Vercel, HTML/CSS, JS, SSE</td><td> <ul style="list-style-type: none"> • 감정 인식 프롬프트·응답 템플릿 설계, 말투·반응 톤 조정 • 관광 도메인 버전에서 실제 관광 정보 추천 흐름 설계 (v2-2) </td><td> <ul style="list-style-type: none"> • 이전 구조를 유지하면서 "상태를 가진 캐릭터"로 확장 • 이후 실시간 스트리밍 구조(v3) 설계 시 기준점으로 활용 </td></tr> <tr> <td>v3</td><td>실시간 대화 구현</td><td>gpt-realtime, WebRTC, Live2D, Vercel, HTML/CSS, JS</td><td> <ul style="list-style-type: none"> • 일괄 처리(STT→LLM→TTS)를 스트리밍 구조로 재설계 • Realtime 이벤트에 맞춰 Live2D 모션·음성 동기화, 네트워크 자연·예외 처리 정리 • 한·영 모두 지원하도록 시스템 메시지·톤을 정리 </td><td></td></tr> </tbody> </table>					버전	주요 목표	사용 기술	역할 및 내용	비고	v1	한국어 음성 기반 Live2D 캐릭터 챗봇의 최소 동작 버전 구현	whisper-1, gpt-4o audio-preview, Live2D, Flask, Vercel, HTML/CSS, JS	<ul style="list-style-type: none"> STT → LLM → TTS로 이어지는 단일 대화 파이프라인 설계 • Live2D 입 모양과 음성 출력 연동, 기본 립싱크 구현 • Flask 백엔드와 브라우저 프론트엔드 연결, Vercel 배포 	<ul style="list-style-type: none"> • 이후 감정 표현·선제 대화 실험(v2)의 기반 • 동작하는 최소 단위 확보에 초점 	v2	감정·상태 반영 및 선제 메시지 지원으로 캐릭터성 강화	gpt-4o mini tts, gpt-4o mini search preview, whisper-1, Live2D, Flask, Vercel, HTML/CSS, JS, SSE	<ul style="list-style-type: none"> • 감정 인식 프롬프트·응답 템플릿 설계, 말투·반응 톤 조정 • 관광 도메인 버전에서 실제 관광 정보 추천 흐름 설계 (v2-2) 	<ul style="list-style-type: none"> • 이전 구조를 유지하면서 "상태를 가진 캐릭터"로 확장 • 이후 실시간 스트리밍 구조(v3) 설계 시 기준점으로 활용 	v3	실시간 대화 구현	gpt-realtime, WebRTC, Live2D, Vercel, HTML/CSS, JS	<ul style="list-style-type: none"> • 일괄 처리(STT→LLM→TTS)를 스트리밍 구조로 재설계 • Realtime 이벤트에 맞춰 Live2D 모션·음성 동기화, 네트워크 자연·예외 처리 정리 • 한·영 모두 지원하도록 시스템 메시지·톤을 정리 	
버전	주요 목표	사용 기술	역할 및 내용	비고																					
v1	한국어 음성 기반 Live2D 캐릭터 챗봇의 최소 동작 버전 구현	whisper-1, gpt-4o audio-preview, Live2D, Flask, Vercel, HTML/CSS, JS	<ul style="list-style-type: none"> STT → LLM → TTS로 이어지는 단일 대화 파이프라인 설계 • Live2D 입 모양과 음성 출력 연동, 기본 립싱크 구현 • Flask 백엔드와 브라우저 프론트엔드 연결, Vercel 배포 	<ul style="list-style-type: none"> • 이후 감정 표현·선제 대화 실험(v2)의 기반 • 동작하는 최소 단위 확보에 초점 																					
v2	감정·상태 반영 및 선제 메시지 지원으로 캐릭터성 강화	gpt-4o mini tts, gpt-4o mini search preview, whisper-1, Live2D, Flask, Vercel, HTML/CSS, JS, SSE	<ul style="list-style-type: none"> • 감정 인식 프롬프트·응답 템플릿 설계, 말투·반응 톤 조정 • 관광 도메인 버전에서 실제 관광 정보 추천 흐름 설계 (v2-2) 	<ul style="list-style-type: none"> • 이전 구조를 유지하면서 "상태를 가진 캐릭터"로 확장 • 이후 실시간 스트리밍 구조(v3) 설계 시 기준점으로 활용 																					
v3	실시간 대화 구현	gpt-realtime, WebRTC, Live2D, Vercel, HTML/CSS, JS	<ul style="list-style-type: none"> • 일괄 처리(STT→LLM→TTS)를 스트리밍 구조로 재설계 • Realtime 이벤트에 맞춰 Live2D 모션·음성 동기화, 네트워크 자연·예외 처리 정리 • 한·영 모두 지원하도록 시스템 메시지·톤을 정리 																						
	프로젝트명 <ul style="list-style-type: none"> • 한국관광공사 API 기반 Claude Desktop용 MCP 구현 																								

	<p>사용 기술 및 도구</p> <ul style="list-style-type: none"> FastMCP, Claude Desktop App, Cursor <p>프로젝트 개요</p> <ul style="list-style-type: none"> 한국관광공사에서 제공하는 공공 데이터를 활용하여, Claude Desktop 앱 내에서 관광지 정보를 검색하고 응답할 수 있는 MCP를 개발하였습니다. FastMCP 라이브러리를 통해 Claude Desktop과 외부 API 간의 통신을 구현하였으며, 사용자 질의에 따라 관광 정보를 실시간으로 제공하는 구조를 설계하였습니다. <p>주요 역할 및 수행 내용</p> <ul style="list-style-type: none"> 한국관광공사의 공공 API를 연동하여 관광지 정보를 조회하는 기능을 구현하였습니다. FastMCP 라이브러리와 Cursor를 활용해 Claude Desktop과 외부 API 간의 데이터 통신을 처리하는 MCP 모듈을 개발하였습니다.
2024	<p>프로젝트명</p> <ul style="list-style-type: none"> 부동산 검색 PoC <p>사용 기술 및 도구</p> <ul style="list-style-type: none"> Crawl4AI, PandasAI, OpenAI API, Pandas, Streamlit <p>프로젝트 개요</p> <ul style="list-style-type: none"> 사용자의 자연어 질의(예: "강남구 10억 이하 30평 이상 아파트")를 분석하여 조건에 맞는 부동산 정보를 필터링하고, 결과를 표 형태로 제공하는 질의응답 시스템의 프로토타입을 구현하였습니다. 수집한 데이터를 기반으로 PandasAI와 OpenAI API를 연동해 응답을 하도록 구성하였습니다. <p>주요 역할 및 수행 내용</p> <ul style="list-style-type: none"> Crawl4AI를 활용해 웹에서 부동산 데이터를 수집하고, PandasAI로 표 형태로 가공하였습니다. Streamlit을 이용해 질의 입력과 결과 출력이 가능한 사용자 인터페이스를 구현하였습니다. <p>프로젝트명</p> <ul style="list-style-type: none"> 소량의 제주도 지역 관광 데이터 기반 Llama 3 Fine-Tuning 프로젝트 <p>사용 기술 및 도구</p> <ul style="list-style-type: none"> Hugging Face Transformers, LoRA, OpenAI API <p>프로젝트 개요</p> <ul style="list-style-type: none"> 사내에서 보유한 제주도 관광 관련 데이터를 활용하여, Llama 3 모델을 한국어 질의응답(QA) 태스크에 특화된 형태로 미세조정(Fine-Tuning)하였습니다. 소규모 데이터로도 도메인에 적합한 응답을 생성할 수 있는지 검증하는 것을 목표로 하였습니다. <p>주요 역할 및 수행 내용</p> <ul style="list-style-type: none"> Hugging Face 라이브러리를 활용해 Llama 3 모델의 Fine-Tuning을 수행하였습니다.

	<ul style="list-style-type: none"> • 실험 결과를 정리하고, 해당 내용을 학회 논문으로 게재하였습니다.
2021	<p>프로젝트명</p> <ul style="list-style-type: none"> • 협동 게임 환경 내 사회적 AI를 위한 음성 대화 시스템 개발 <p>사용 기술 및 도구</p> <ul style="list-style-type: none"> • OBS Studio, Zoom, MS Azure STT/TTS, pyttsx3, Loomie Virtual Avatar, Excel / Google Sheets <p>프로젝트 개요</p> <ul style="list-style-type: none"> • "Don't Starve Together" 게임 환경을 기반으로, 인간과 협동하는 AI 캐릭터가 음성으로 상호작용하도록 구현한 석사과정 연구입니다. 사용자와 AI 간의 상호작용 과정을 규칙 기반 음성 대화 시스템으로 설계하고 실험을 통해 기능을 검증하였습니다. <p>주요 역할 및 수행 내용</p> <ul style="list-style-type: none"> • 게임 내 상황에 따라 발화를 생성하는 음성 출력 구조를 설계하고, Loomie 가상 아바타와 연동하여 실시간 규칙 기반 상호작용을 구현하였습니다. • 실험 참여 학생들과 상호작용 테스트를 진행하고, 결과를 바탕으로 개선점을 도출하였습니다.