Profile

Name	Jaeyoung Suh	Email	tjwodud04@gmail.com
GitHub	https://github.com/tjwodud04	About page	https://tjwodud04.github.io/intro/?lang=en
Hugging Face	https://huggingface.co/dddsaty	LinkedIn	https://www.linkedin.com/in/jaeyoungsuh/

Work Experience

Company	Employment Type	Position	Period
AIPIA	Contract	AI Engineer (Image Generation)	Oct 2025 – Jan 2026
AIO2O	Full-time	AI Engineer (LLM)	Sep 2023 – Mar 2025
BasgenBio	Full-time	AI Engineer (NLP, Computer Vision)	Mar 2022 – Sep 2023

Education

Degree	Institution	Period
M.S. in Artificial Intelligence (Data Science)	Hanyang University Graduate School	Mar 2020 – Feb 2022
B.S. in Computer Science	Korea National Institute for Lifelong Education (Bachelor's Degree by Examination)	Feb 2020
B.A. in History and English Language & Literature	Yonsei University, Wonju Campus	Mar 2014 – Feb 2018

Publications

Year	Title	PDF
2025	Jae Young Suh, Mingyu Jeon, "A Character-Based Korean Tourist Dialogue System with Proactive Recommendations and Live2D Embodiment", engrXiv preprint, 2025 Summary: Developed a web-based character-driven tourist guidance system by integrating the Korea Tourism Organization's TourAPI with a Live2D avatar. The system employs a proactive policy to actively suggest travel routes rather than responding passively to user queries. Evaluation results show a speech recognition accuracy of 6.4% WER, demonstrating the system's feasibility for immersive and visually guided tourist information services.	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 Summary: Conducted a quantitative study of social interaction and strategic decision-making in large language models using a buy-and-sell negotiation game. The analysis examined persona-driven strategy shifts, win rates, and Shapley values to assess how closely LLM behavior aligns with human negotiation patterns. The results highlight a clear gap between performance on knowledge-based benchmarks and behavior in realistic social contexts.	PDF
	Mingyu Jeon, Suwan Cho, and Jae Young Suh. "PPoGA: Predictive Plan-on-Graph with Action for Knowledge Graph Question Answering." Accepted to GMLLM 2025 (Frontiers in Graph Machine Learning for the Large Model Era), CIKM 2025 Workshop Summary: Proposed PPoGA, a planner–executor framework designed to improve flexibility in knowledge graph question answering. The method introduces predictive outcome simulation before action execution and a two-stage self-correction mechanism to support	PDF

	complex multi-step reasoning. Experimental results on benchmarks such as GrailQA show consistent improvements over prompt-based baselines.	
	Jae Young Suh and Mingyu Jeon. "A Modular Prototype of Emotion-Aware Proactive Voice Agent with Live2D Embodiment." Accepted to ProActLLM 2025 (Proactive Conversational Information Seeking with Large Language Models), CIKM 2025 Workshop Summary: Designed a modular voice agent architecture integrating Whisper-based speech recognition, emotion inference, and a proactive policy engine. The system controls intervention frequency through a cooldown mechanism and lightweight scoring model to maintain stability during interaction. Emotional reasoning based on the Eastern "Seven Emotions (Chiljeong)" framework is combined with a Live2D avatar to enable empathetic and context-aware dialogue.	PDF
2024	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. Summary: Applied domain-adaptive pretraining (DAPT) with LoRA to effectively inject Jeju Island–specific knowledge into a Llama 3 model. Built the JejuQA dataset consisting of over 47,000 Q&A pairs to improve regional information accuracy. The adapted model achieved higher F1 scores than the baseline and demonstrated reduced hallucination in domain-specific responses.	PDF
2021	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 Summary: Investigated data-driven approaches for building speech-based dialogue systems for social AI in cooperative game environments. Defined game states through analysis of human–human interactions and implemented an autonomous dialogue system connected to real-time game data. User studies identified six key factors, including responsiveness and naturalness, that contribute to more human-like AI behavior.	PDF

Certifications

Certification	Date Obtained
OPIc IH	Sep 2025
AICE(AI Certificate for Everyone)	Aug 2023
ADsP(Advanced Data Analytics semi-Professional)	Nov 2022
SQLD(SQL Developer)	Sep 2022
Engineer Information Processing (정보처리기사)	May 2019

Projects

Year	Affiliation	Project	Description	
2025		Live2D Character Voice Chatbot Demo	- Implemented a chatbot that responds to user speech by integrating OpenAI API with Live2D sample characters - Tech Stack: gpt-4o-audio-preview, Whisper, Flask, JavaScript, Vercel	Link
2024		LLM-based Q&A System Prototype	- Built a real-estate search prototype using Crawl4AI, PandasAI, and Streamlit Tech Stack: Python, OpenAI API, PandasAI, Streamlit	Link
2024	AIO2O	Jeju Tourism	- Conducted fine-tuning experiments using Llama 3 and	

		Data–based LLM Fine-tuning	Polyglot-ko models on Jeju tourism data - Organized the findings into a research paper and submitted it to a conference Tech Stack: Python, Llama3, Pytorch, Huggingface	
2022	BasgenBio	Biomedical Domain NER Experiment	- Performed Named Entity Recognition (NER) experiments on biomedical datasets from *Papers With Code*, classifying entities such as diseases, organizations, and genes - Trained and compared models including BioBERT and BioLinkBERT using Simple Transformers and Flair frameworks Tech Stack: Python, PyTorch, Simple Transformers, Flair	
2021	Hanyang University	HCI Project – Speech Dialogue System in Cooperative Game	Conducted experiments on human–AI interaction within the *Don't Starve Together* cooperative game environment as part of master's research - Implemented rule-based STT/TTS responses and collected internal feedback for improvement of the voice interface Tech Stack: pyttsx3, MS Azure STT/TTS, Lua (game mod integration)	Link

Supplementary Technical Document

<table>
<tr><th>Year</th><th>Description</th></tr>
<tr><td>2025</td><td>
Project
• Development of a Korean Voice Chatbot Prototype with Live2D Character Interaction
<table>
<tr><th>Version</th><th>Primary Goal</th><th>Technologies Used</th><th>Role & Key Contributions</th><th>Notes</th></tr>
<tr><td>v1</td><td>Implement a minimum viable version of a Korean voice-based Live2D character chatbot</td><td>whisper-1, gpt-4o audio-preview, Live2D, Flask, Vercel, HTML/CSS, JavaScript</td><td>• Designed a single-turn dialogue pipeline: STT → LLM → TTS
• Implemented basic lip-sync by linking Live2D mouth movements with audio output
• Connected Flask backend with browser-based frontend and deployed via Vercel</td><td>• Served as the foundation for later emotion expression and proactive dialogue experiments (v2)
• Focused on establishing a fully working minimal unit</td></tr>
<tr><td>v2</td><td>Enhance character personality through emotion/state awareness and proactive messaging</td><td>gpt-4o mini tts, gpt-4o mini search preview, whisper-1, Live2D, Flask, Redis, Vercel, HTML/CSS, JavaScript, SSE</td><td>• Designed emotion-aware prompts and response templates; adjusted speaking style and reaction tone
• Designed a tourism-domain variant with real-world information recommendation flow (v2-2)</td><td>• Extended the existing structure into a "stateful character" while preserving the core pipeline
• Used as a reference point for designing the real-time streaming architecture in v3</td></tr>
<tr><td>v3</td><td>Enable real-</td><td>gpt-realtime,</td><td>• Re-architected the batch pipeline (STT →</td><td></td></tr>
</table>
</td></tr>
</table>

		time conversational interaction	WebRTC, Live2D, Vercel, HTML/CSS, JavaScript	LLM → TTS) into a fully streaming structure • Synchronized Live2D motion and audio with realtime events; handled network latency and edge cases • Unified system messages and tone to support both Korean and English	

Project

• MCP Implementation for Claude Desktop using Korea Tourism Organization API

Technologies & Tools

• FastMCP, Claude Desktop App, Cursor

Overview

• Developed an MCP (Model Context Protocol) that enables Claude Desktop to access and respond to tourism information using public data provided by the Korea Tourism Organization. Implemented data communication between Claude Desktop and external APIs via the FastMCP library, enabling retrieval of tourism details based on user queries.

Roles & Contributions

• Integrated the Korea Tourism Organization's public API to retrieve regional tourism information.

• Developed an MCP module using FastMCP and Cursor to handle data communication between Claude Desktop and external APIs.

2024

Project

• Real Estate Search Proof-of-Concept (PoC)

Technologies & Tools

• Crawl4AI, PandasAI, OpenAI API, Pandas, Streamlit

Overview

• Built a prototype Q&A system capable of interpreting natural language queries (e.g., "Apartments in Gangnam under 1 billion KRW, 30 pyeong or larger") and returning filtered real-estate information in tabular form. Combined PandasAI and OpenAI API to generate contextual responses based on collected datasets.

Roles & Contributions

• Collected real-estate data from the web using Crawl4AI and processed it into structured tables with PandasAI.

• Implemented a Streamlit interface for interactive query input and result visualization.

Project

• Fine-Tuning Llama 3 with Small-Scale Jeju Tourism Data

Technologies & Tools

• Hugging Face Transformers, LoRA, OpenAI API

	Overview • Fine-tuned the Llama 3 model for Korean question-answering (QA) tasks using Jeju tourism data owned by the company. The project aimed to test whether domain-specific responses could be effectively generated from small datasets. Roles & Contributions • Performed fine-tuning of the Llama 3 model using the Hugging Face library. • Compiled and published the experimental results in an academic conference paper.
2021	Project • Development of a Speech Dialogue System in Cooperative Game Environments Technologies & Tools • OBS Studio, Zoom, MS Azure STT/TTS, pyttsx3, Loomie Virtual Avatar, Excel / Google Sheets Overview • Master's-level research project based on the *Don't Starve Together* game environment, where an AI character cooperated with human players via speech interaction. Designed a rule-based speech dialogue system to simulate AI-human cooperation and validated its functionality through gameplay experiments. Roles & Contributions • Designed a speech output structure that generated dialogue based on in-game contexts and integrated it with the Loomie virtual avatar for real-time rule-based interaction. • Conducted interaction tests with student participants and identified areas for improvement based on experimental feedback.