Yichuan Zhang

Yichuan.Zhang22@student.xjtlu.edu.cn | https://github.com/jadegreened | https://zeachone.com/

Education

University of Liverpool   2022 – 2026/06(Expected)
BSc in Information and Computing Science   Liverpool, England
- GPA: 3.84
- University Academic Achievement Scholarship Winner   2023 – 2024

Work Experience

Openinterx|Engineering Department   2024/07 – 2024/10
Advisor: shawn shen
Software Engineer internship

University of Bristol|BristollG Lab   2024/02 – 2024/06
Advisor: shawn shen
Remote Research Assistant

Xi'an Jiaotong-Liverpool University|X-CHI Lab   2023/10 – 2024/01
Advisor: Prof. Hai-Ning Liang
Research Assistant

Xi'an Jiaotong-Liverpool University|IOT Lab   2023/04 – 2023/09
Advisor: Prof. Yu-ji Dong
Research Assistant

Publications

[AAMAS'2024] Shiyao Zhang, Yuji Dong, **Yichuan Zhang**, Terry R Payne & Jie Zhang "Large Language Model Assisted Multi-Agent Dialogue for Ontology Alignment." 2024 International Conference on Autonomous Agents and Multiagent Systems. DOI: https://dl.acm.org/doi/abs/10.5555/3635637.3663238

Projects

AI Semantic Alignment System   2023/04 – 2023/7
Advisor: Prof. Yuji Dong
- Developed an AI-driven ontology alignment dialog system using Java and MELT framework. The system explored the semantic understanding ability of Large Language Model in the context of ontology alignment in IOT.
- With the continuous development of large language models, the advanced GPT-4 has demonstrated remarkable semantic understanding capabilities. In the ontology of the Internet of Things (IoT) domain, human-expert-created ontologies are filled with abundant semantic information, which plays a crucial role in ontology alignment. Previously, alignment work required human experts to spend weeks completing it. This has inspired us to explore the potential of large language models in ontology alignment.
- In this project, we utilized the dialogue model "Correspondence Inclusion Dialogue" to develop a program within the ontology alignment framework MELT and conducted rigorous agent modeling. This allowed different agents to possess distinct capabilities, ensuring high system accuracy. On the Anatomy dataset provided by OAEI, our system achieved a final score of 54.4% precision, 59.9% recall, and an F-measure of 66.6%.

Context Awareness Text Augmented System   2024/01 – 2024/06
Advisor: Prof. Shawn Shen
- Real-time multimodal perception of the surrounding environment was achieved using HoloLens devices, employing state-of-the-art models from various fields to convert visual data to text, audio to text, IMU data to text, and user speech to text.
- Experimental scenes such as a keyboard were built using Unity, and tools within MRTK were utilized to collect video, audio, IMU, and microphone data. Real-time data communication was achieved using sockets. To overcome issues with HoloLens local network transmission, a public server was deployed using the Spring Boot framework to handle data reception and distribution.

- Inspired by the paper **KWickChat: A Multi-Turn Dialogue System for AAC Using Context-Aware Sentence Generation by Bag-of-Keywords(IUI 27th)**, the recommendation system in this work does not use conventional NLP methods but rather employs a fine-tuned LLaMa-7B model and carefully designed prompts for prediction. With each user input, the system calls a remote server, where a vector database retrieves semantically closest information from multimodal data stored in text form, and then provides this data to the large model for prediction.

Memory Augmented Video Agent 2024/06 − 2024/09
Advisor: Prof. Shawn Shen
- Contributed code to the upcoming open-source project **Memory-Augmented-Video-Agent (MAVA)**. The MAVA system primarily handles exceptionally long videos. My responsibilities included splitting the project code by logic into separate containers, deploying the project on the cloud, and rewriting the core code within the RAG (Retrieval-Augmented Generation) module for long video processing.
- Used FFmpeg for streaming uploaded videos, keeping video frame data in memory manageable and greatly improving request handling concurrency. During video processing, it applies semantic-based compression to store only minimally relevant image data. This approach preserves the video's original information while significantly lowering storage costs and reducing system load.
- Successfully completed alpha version testing of the entire project, achieving concurrent request handling throughout the system.

Skills

Programming Languages: C#, Java, SQL, Python, JavaScript, HTML, CSS
Tools and Frameworks: Unity, Oculus SDK, Hololens, Flask, React, Springboot

Research Interest

HCI/AR/VR/AI Agent