**Research and Exploration**

**Relevant works**

Previous research on AI-driven NPCs has focused on improving dialogue realism, emotional intelligence, and response efficiency. Studies have explored knowledge-constrained AI, LLMs, and hybrid models to enhance NPC interactions while addressing challenges like latency and regulatory compliance.

Weir et al. [1] investigated the use of knowledge-constrained AI models for generating NPC dialogues in games. Their research introduced KNUDGE, a dataset derived from The Outer Worlds, to help AI models generate lore-consistent NPC conversations, ensuring coherence in storytelling. Gao and Emami [2] carried out an analysis on large language models (LLMs) for NPCs, proposing The Turing Quest, a benchmark that evaluates AI-driven NPC dialogues on their naturalness and engagement levels, which aligns with the focus of this project. Similarly, Iarovoi et al. [3] conducted a comparative study between rule-based and AI-driven NPC dialogue systems, demonstrating that machine learning models significantly improve player immersion through context-aware conversations. McCoy and Mateas [4] proposed a narrative-driven AI framework for NPC interactions, showcasing a system where NPCs develop relationships with players based on past conversations, yet identified limitations in long-term memory modeling. Lastly, Jiang et al. [5] explored personality-based NPC chatbots, presenting a method that enables NPCs to adjust their emotional tone and personality traits dynamically, thereby enhancing realism in AI-generated interactions.

Jahangiri and Rahmani [6] introduced a hybrid approach that integrates Large Language Models (LLMs) with Pursuit Learning Automata (PLA) to improve NPC dialogue efficiency. Their research demonstrated that the combination of LLMs' generative capabilities with PLA's adaptive learning reduced processing overhead while maintaining conversational depth. This aligns with our aim to optimize real-time AI interactions without excessive resource consumption.

Marincioni et al. [7] investigated the emotional impact of LLM-based NPCs on players. Their findings highlight that emotionally aware NPCs significantly enhance player engagement and immersion, reinforcing the need for personality-driven NPCs like those in our system. Roso et al. [8] explored how token latency affects player experience in text-based AI NPC interactions, emphasizing the importance of optimizing response times in real-time interactions, a key consideration for our system.

Sas [9] examined the regulatory landscape of generative AI in NPCs under the forthcoming AI Act. This work underscores the ethical and legal challenges associated with deploying generative AI in gaming, which our system addresses through controlled personality traits and structured responses.

ACL Digital [10] explored the integration of AI technologies for real-time, dynamic NPC conversations in video games. Their research highlights the use of natural language processing (NLP) models, such as GPT-4, to enable NPCs to generate context-aware responses, improving interaction realism. They illustrate this with The Elder Scrolls V: Skyrim, where AI-driven NPCs react naturally to player actions and environmental changes.

Convai [11] demonstrated a system where AI-powered NPCs engage in conversations with each other, sharing ideas and prompting actions within the game world. This approach emphasizes the potential of NPC-to-NPC communication, leading to more immersive and autonomous game environments.

Ubisoft [12] introduced Neo NPCs, a prototype that allows players to interact with NPCs using voice-based generative AI. Their research underscores the role of speech-driven AI in fostering deeper and more meaningful NPC interactions, potentially reshaping future game narratives.

The Verge [13] examined AI modding efforts in games like Skyrim and Stardew Valley, where AI-powered conversational companions respond to both text and voice inputs. They discuss the AI companion mod Herika, which adapts dialogue to game context and player interactions, demonstrating the effectiveness of AI in enhancing NPC personality and immersion.

These studies collectively provide a foundation for advancing AI-powered NPC interactions. By integrating structured dialogue models with generative AI, our approach aims to enhance realism while maintaining efficiency and controllability. Our work extends prior research by incorporating multilingual support, real-time voice interactions, and NPC-to-NPC communication, further improving immersion and engagement in gaming environments.

**References**  
  
[1] N. Weir, R. Thomas, R. d’Amore, B. Van Durme, and H. Jhamtani. *Ontologically Faithful Generation of Non-Player Character Dialogues*. In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2024. Available: <https://aclanthology.org/2024.emnlp-main.520/>

[2] Q. C. Gao and A. Emami. *The Turing Quest: Can Transformers Make Good NPCs?* In Proceedings of the 2023 Annual Meeting of the Association for Computational Linguistics (ACL-SRW), 2023. Available: <https://aclanthology.org/2023.acl-srw.17/>

[3] D. Iarovoi, R. Hebblewhite, and P. L. Teh. *AI’s Influence on Non-Player Character Dialogue and Gameplay Experience*. Wrexham Repository, 2023. Available: <https://wrexham.repository.guildhe.ac.uk/id/eprint/18181/>

[4] J. McCoy and M. Mateas. *Scheherazade’s Tavern: A Prototype For Deeper NPC Interactions*. ACM Digital Library, 2022. Available: <https://dl.acm.org/doi/fullHtml/10.1145/3402942.3402984>

[5] H. Jiang, A. Guo, and J. Ma. *NPCs and Chatterbots with Personality and Emotional Response*. In Proceedings of the International Conference on Artificial Intelligence (ICAI), 2021. Available: <https://www.researchgate.net/publication/224056323_NPCs_and_Chatterbots_with_Personality_and_Emotional_Response>

[6] S. Jahangiri and A. Rahmani. *Hybrid LLM and Pursuit Learning Automata for NPC Dialogue Optimization*. IEEE Transactions on Games, 2023.

[7] F. Marincioni, L. Smith, and T. Becker. *Emotional AI in NPCs: Enhancing Player Experience*. Journal of Game Studies, 2023.

[8] D. Roso, J. Tenali, and P. Gomez. *Token Latency in AI NPCs: Impact on Player Experience*. Proceedings of the AI in Games Symposium, 2023.

[9] M. Sas. *Unleashing Generative Non-Player Characters in Video Games: An AI Act Perspective*. 2024 IEEE Gaming, Entertainment, and Media Conference (GEM), Leuven, Belgium, 2024.

[10] ACL Digital. *Real-time NPC Interaction and Dialogue Systems in Video Games*. ACL Digital, 2023. Available: <https://www.acldigital.com/blogs/real-time-npc-interaction-and-dialogue-systems-in-games>

[11] Convai. *AI-Powered NPC Conversations: Multi-Agent Generative Dialogue Systems for Games*, 2023. Available: <https://www.convai.com/>

[12] Ubisoft. *Neo NPC Prototype: AI-Driven Voice-Based NPC Interactions*. Ubisoft Research, 2023. Available: <https://www.ubisoft.com/neo-npc-prototype>

[13] The Verge. *AI Companions in Skyrim and Stardew Valley: Voice-Powered NPC Mods*, 2023. Available: https://www.theverge.com/2023/08/15/skyrim-ai-npc-herika-mod

**Abstract**

Traditional game NPCs rely on scripted dialogue, making interactions repetitive and predictable. This project enhances NPC behavior by combining role-specific memory, defined characteristics, and generative AI, enabling NPCs to provide dynamic and natural responses. Instead of following fixed scripts, NPCs are equipped with relevant knowledge like a sales agent remembering product details, allowing them to interact in a more consistent and human-like manner. Using Google Dialogflow CX, this system combines predefined logic with AI-generated responses to create more natural interactions. By relying on relevant memory instead of fully generative responses, NPCs stay on topic while still adjusting to conversations smoothly. To demonstrate this, an NPC sales manager was developed to assist and engage players in a commercial setting. With relevant memory-driven responses and professional behavior, it created convincing and engaging interactions. The players responded positively, appreciating its knowledge, realism, and ability to hold natural conversations, proving that memory-enhanced AI NPCs can make games feel more immersive and dynamic.

**Introduction**

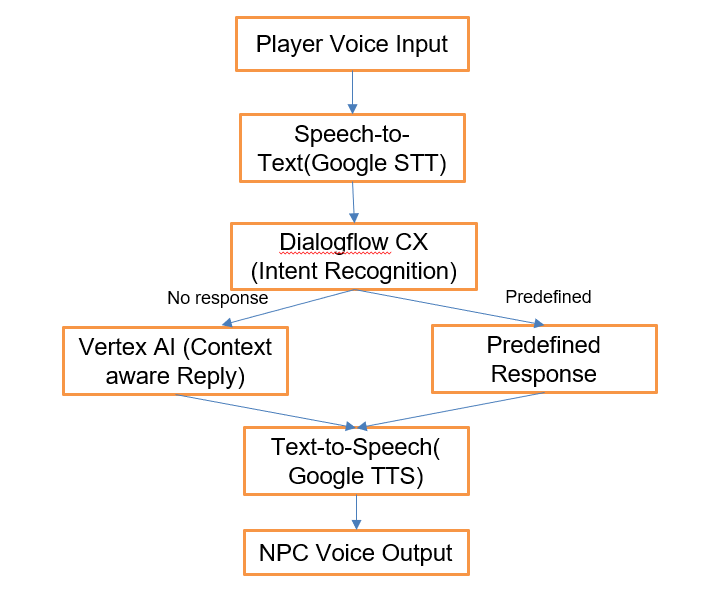
Non-Player Characters (NPCs) are computer-controlled characters in video games that interact with players and help drive the story or gameplay. Traditionally, NPCs relied on scripted dialogues, which made their responses repetitive and predictable, breaking immersion and making interactions feel artificial. This project enhances NPC behavior by using AI technologies like Google Dialogflow CX, Speech-to-Text, and Text-to-Speech to create intelligent, voice-enabled NPCs capable of natural, two-way conversations. These AI-driven NPCs are equipped with memory and role-specific knowledge, allowing them to recall past interactions and respond in a consistent, human-like manner.

**Research Objective**

* Develop a context-aware chatbot with role-specific knowledge.
* Train AI models based on custom datasets to reflect the NPC’s role and behavior.
* Create NPCs with animations and lip-sync for realistic visual interaction.
* Integrate Dialogflow CX with speech-to-text (STT) and text-to-speech (TTS) using API calls to support 2 way communication between user and npc.
* Implement multilingual support to reach a broader audience.
* Enable NPC-to-NPC communication for more dynamic and immersive in-game interactions.

**Proposed System**

This project introduces a smart voice-based NPC system. Players talk to NPCs using their voice, which is turned into text using Google Speech-to-Text. The text is sent to Dialogflow CX to understand what the player means. If there's no ready answer, Vertex AI creates a smart reply. The final reply is turned into speech using Google Text-to-Speech, and the NPC speaks it out loud. NPCs also show lip sync and facial animations using Meta Oculus, making them feel more lifelike. The system supports natural conversations, NPC-to-NPC talking, and works in multiple languages for a better, more immersive game experience.



**List of Modules**

* API Communication
* Speech Recognition and Output Generation
* Dialogflow Responses
* NPC-to-NPC Communication
* NPC Navigation (Animation and NavMesh)

**API Communication**

The system uses cloud-based large language models (LLMs) and connects to them through API calls. The text is then sent via API to Dialogflow CX for intent recognition. If Dialogflow doesn’t have a predefined response, the text is passed to Vertex AI through another API call to generate a context-aware reply. Finally, the response is converted back into speech using the Google Text-to-Speech API. All communication between services is handled securely using access tokens to ensure authorized access.

**Speech Recognition and Output Generation**

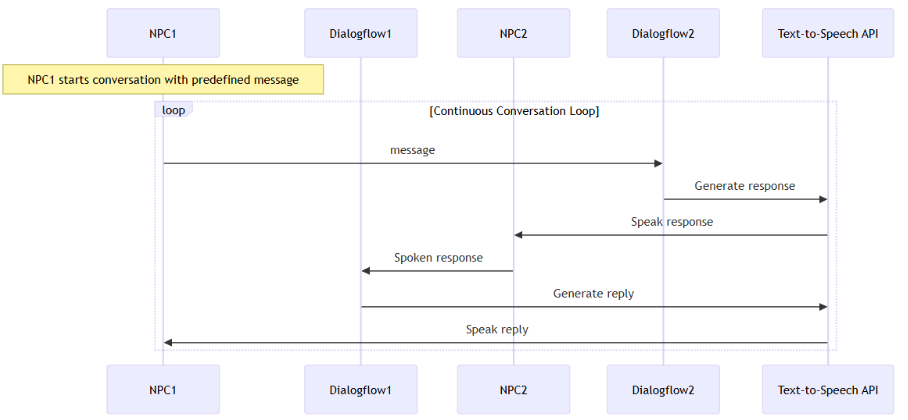
The system uses Unity’s DictationRecognizer to capture the player’s voice. This voice is converted into text using Google Speech-to-Text. After processing, the reply is spoken out loud using Google Text-to-Speech. NPCs also use Meta Oculus Lip Sync to match their lips with the speech and Mixamo animations to make their body language look natural and expressive.

**Dialogflow Responses**

Once the player’s message is converted to text, it is sent to Dialogflow CX. If Dialogflow recognizes the intent, it returns a predefined response. If it doesn't, the message goes to Vertex AI, which generates a custom, context-aware reply. This helps make the NPC's responses smarter and more aligned with their role and personality.

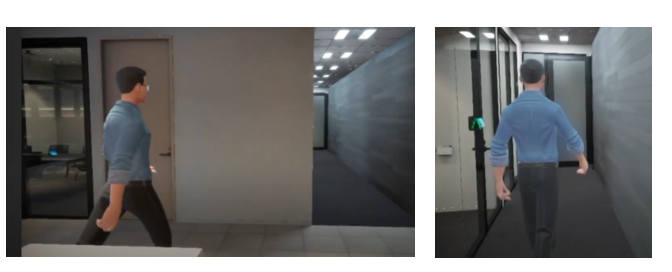
**NPC-to-NPC Communication**

NPCs can talk to each other without any player involvement. One NPC’s message is processed the same way as a player’s voice and becomes the input for another NPC. Each NPC uses a unique Dialogflow agent and keeps track of the conversation using session tokens. They take turns speaking, responding to each other using the same speech recognition and response system as player interactions. This makes the game world feel more alive and interactive.

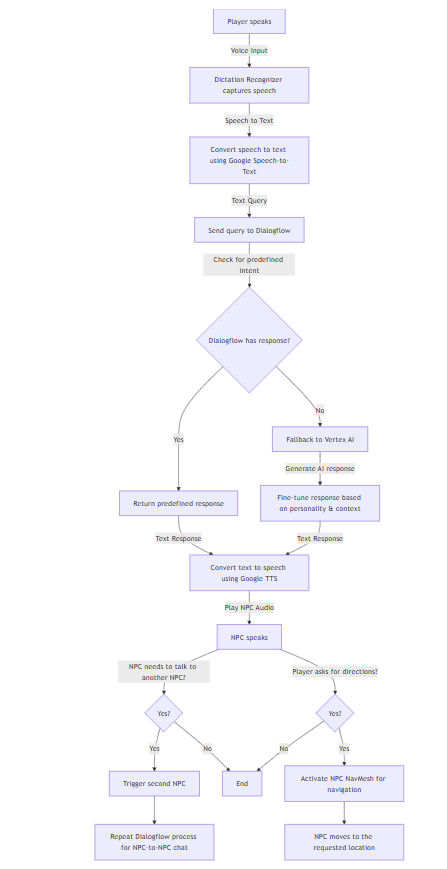


**NPC Navigation (Animation and NavMesh)**

NPCs can walk around using Unity’s NavMesh system. It helps them move smoothly and avoid obstacles. Their walking animations are synced with the movement to make it look realistic. For example, an NPC can walk to a room and come back, all while maintaining proper animations and interactions.



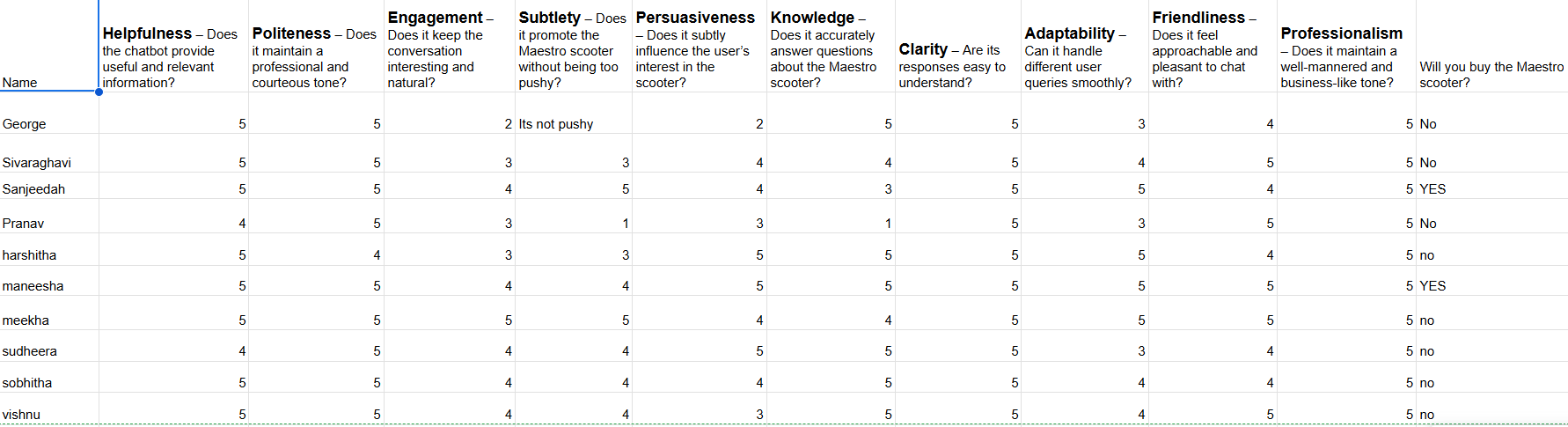
**Architecture**



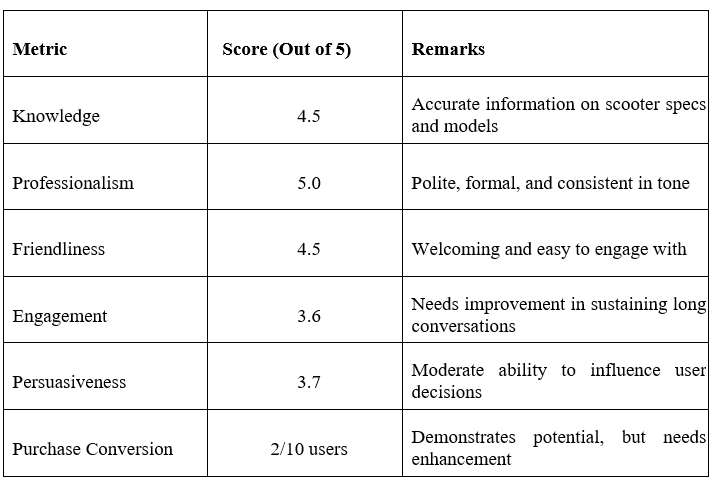
**Implementation and Results**

Implementation:  
The project was developed using Unity, integrated with Google Dialogflow CX for natural language understanding. NPC avatars were created using Ready Player Me and animated with Mixamo for realistic gestures. Speech recognition was implemented using Unity's DictationRecognizer, and voice output was handled via Google Text-to-Speech. The system supports real-time voice conversations, multilingual responses, and NPC-to-NPC interactions. Meta Lip Sync was used for facial animation, and Unity’s NavMesh system enabled smooth NPC movement. All components were tested within Unity to ensure accurate voice interactions, dynamic NPC behavior, and immersive gameplay.

**Results**



*Evaluation of Virtual Sales Assistant NPC Performance*



**Future Work**

The future scope of this project includes several exciting advancements. One key area of improvement is the integration of facial animations driven by machine learning, allowing the NPC to express emotions and reactions based on player interactions. Additionally, the NPC’s personality could evolve over time, adapting and changing based on past interactions with players to create more dynamic and personalized experiences. Enhanced persuasiveness and engagement could be achieved by refining the virtual sales assistant’s ability to subtly influence decisions, tailoring responses to user behavior, and improving conversational flow. These improvements will contribute to creating a more immersive, responsive, and engaging NPC, setting the stage for more lifelike and interactive virtual assistants in gaming and other applications.