Sophia Project: Enhancing Player Immersion with Intelligent Autonomous NPCs in 2D RPGs

Vinícius Custódio Chelli[1,†], André Luiz França Batista[2,*,†]

[1]*Instituto Federal do Triângulo Mineiro - Campus Ituiutaba*
[2]*Instituto Federal do Triângulo Mineiro - Campus Ituiutaba*

Abstract
Introduction: This paper introduces the **Sophia Project**, an advanced intelligent autonomous NPC system specifically designed for 2D RPG games using Unity. **Objective:** Sophia aims to employ reinforcement learning techniques [8] to dynamically evolve its relationship with players and features voice recognition capabilities to interpret emotional and contextual commands. **Methodology or Steps:** Integrated with the ChatGPT API [9], Sophia can conduct contextually meaningful dialogues, complemented by recurrent neural networks for environment perception and memory management [10], and movement through NavMesh. Additional modules include dynamic relationship tracking, contextual memory retention, and object perception for deeper interactions. **Results:** Experimental evaluations with 30 participants demonstrated significantly enhanced player immersion, adaptability, and believability compared to traditional scripted NPC systems.

Keywords

Unity, C#, artificial intelligence, immersion, NPCs, reinforcement learning, RNN, voice recognition

1. Introduction

Digital games have evolved significantly over the past decades, not only in terms of graphics and mechanics but also in how they aim to deliver immersive and emotionally resonant experiences. Among the elements that shape these experiences, non-playable characters (NPCs) are central to crafting engaging narratives and building believable game worlds [1]. However, conventional NPC architectures—typically reliant on decision trees and rigid dialogue scripting—often result in repetitive, static interactions that undermine player immersion and limit perceived agency [2].

Recent progress in artificial intelligence (AI) has unlocked new avenues for creating adaptive, emotionally aware, and autonomous NPCs [13]. Techniques such as deep reinforcement learning (DRL), recurrent neural networks (RNNs), and natural language processing (NLP) now allow characters to interpret context, learn from user interactions, and retain memory across gameplay sessions [3, 4, 5]. These technologies enable the simulation of social and emotional intelligence, enriching player–NPC interactions and allowing characters to evolve meaningfully over time.

In this context, we introduce the **Sophia System**, a modular AI framework for intelligent NPCs designed for 2D RPGs. Sophia integrates DRL, contextual memory, and emotionally adaptive dialogue via voice recognition and natural language understanding. The system allows NPCs to react dynamically to player behavior, maintain long-term emotional states, and adapt their decisions based on evolving relationships. Our goal is to demonstrate how the integration of these technologies can meaningfully enhance immersion, emotional engagement, and narrative continuity in games.

Proceedings of AI4HGI '25, the First Workshop on Artificial Intelligence for Human-Game Interaction at the 28th European Conference on Artificial Intelligence (ECAI '25), Bologna, October 25-30, 2025
*Corresponding author.
†These authors contributed equally.
theongatechelli@gmail.com (V. C. Chelli); andreluiz@iftm.edu.br (A. L. F. Batista)
https://github.com/ViniciusChelli (V. C. Chelli); https://github.com/andreluizfrancabatista (A. L. F. Batista)
0009-0003-4957-710X (V. C. Chelli); 0000-0003-3225-8359 (A. L. F. Batista)

2. Methodology

The Sophia Project adopts a modular, multi-agent architecture that fuses advanced AI technologies to foster highly autonomous, emotionally intelligent, and narratively coherent NPC behavior. All modules interact in real-time to drive immersive experiences inside a 2D farming simulation inspired by *Harvest Moon* and *Stardew Valley*. Below, we detail each module with in-game examples.

2.1. Natural Language Understanding via ChatGPT API [9]

Sophia uses the ChatGPT API for Natural Language Understanding (NLU) and Generation (NLG). Player voice inputs—such as "Good morning, Sophia!" or "Why are you upset?"—are transcribed via speech recognition, then sent to the LLM and returned as structured JSON with semantic tags and suggested tone (e.g., friendly, apologetic).

This response is synthesized into speech using a TTS engine and played with appropriate emotion in Unity. For instance, if the player says something supportive after an argument, Sophia might respond with a softened voice and phrases like, "I'm still hurt… but thank you for caring."

This system ensures that conversations are coherent, emotionally responsive, and influenced by both memory and prior context.

2.2. Recurrent Neural Networks for Memory Management [10]

To simulate episodic memory, Sophia uses a Long Short-Term Memory (LSTM) network trained on 1,000 gameplay sequences. The network receives time-stamped data on spatial location, previous dialogue tags, emotional valence, and trust level.

In-game effect: If the player previously lied to Sophia or gave her an unwanted gift, these events are embedded in the memory vector. Later conversations reflect this history. For example, Sophia might say, "You've been acting strange lately… I'm not sure I can trust you again," if emotional consistency has been negative.

The LSTM model was trained for 60 epochs using the Adam optimizer with early stopping on validation loss, ensuring that only meaningful interaction patterns are retained.

2.3. Deep Q-Learning for Adaptive Behavior [8]

Sophia's behavioral decision-making uses a Deep Q-Network (DQN), which selects responses that maximize long-term reward. The learning algorithm uses the Bellman equation:

$$Q(s,a) \leftarrow Q(s,a) + \alpha\left[r + \gamma \max_{a'} Q(s',a') - Q(s,a)\right] \tag{1}$$

Gameplay Context: Each state s represents Sophia's perception of the game world: player distance, tone of recent interactions, current location, and trust level. - s_1: Player enters Sophia's garden after ignoring her for three days. - s_2: Player apologizes for past mistakes. - s_3: Player compliments Sophia at the festival.

Each action a corresponds to Sophia's next move: - Approach the player and initiate warm dialogue. - Express caution or distance herself. - Comment on recent events. - Decline interaction.

Rewards are defined by how the player reacts: - +1.0 for positive engagement (e.g., continued conversation), - +0.5 for neutral proximity, - -0.75 for off-topic or silent responses, - -1.0 for hostility or departure.

Training Specifications

The network receives 42 input features (emotions, location, time, etc.), with 12 possible actions. It uses prioritized replay, mini-batch size of 64, ε-greedy exploration ($1.0 \rightarrow 0.1$), and three hidden layers (256–128–64, ReLU). Target networks update every 500 steps.

This architecture ensures that Sophia not only reacts appropriately in the moment, but learns to anticipate and adapt to complex player behavior patterns over time.

2.4. NavMesh Navigation

Sophia's movement is managed by Unity's NavMeshAgent system. Navigation is not random: it is emotionally informed and context-aware.

Example: If trust is high, Sophia actively approaches the player in open fields. If afraid (due to hostile interactions), she may avoid crowded areas or retreat to her house. The DQN selects a navigation action, and the NavMesh generates the path, integrating both affect and world geometry.

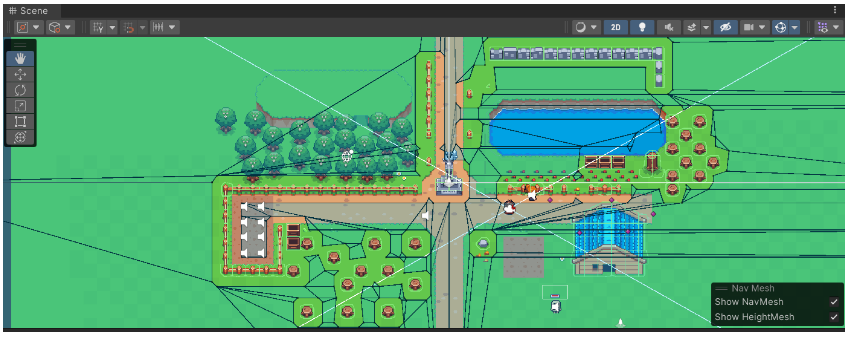


Figure 1: NavMesh pathfinding for Sophia across tile-based environments.

2.5. Dynamic Relationship System

This subsystem tracks trust, empathy, and affection as floating-point variables. These evolve based on dialogue history, gifts, and situational behavior.

In-game example: Repeatedly helping Sophia with farming or checking on her after a storm increases trust. Ignoring her or mocking her responses lowers empathy. The values modulate how she greets the player, her facial expressions, and even her willingness to share personal stories.

2.6. Object Perception Module

Sophia uses grid-based environmental scanning to detect objects and update dialogue contextually.

Example: If the player picks up a rare herb near Sophia, she might immediately react with: "That's the one I was looking for!" Conversely, if the player picks up a harmful item or a tool she dislikes, she might step back or warn them.

2.7. Contextual Memory System

In addition to RNN-based short-term memory, Sophia maintains a long-term episodic log that includes: - Previous conversations, - Emotional shifts, - Events like festivals, storms, arguments.

Figure 2: Real-time affective metrics that modulate Sophia's dialogue and behavior.

Figure 3: Grid-based environment scan enabling object recognition.

Use in-game: This memory allows Sophia to say, "You were really kind to me at the Spring Festival," or "I still remember that fight we had last season..."—making her feel like a persistent, evolving character.

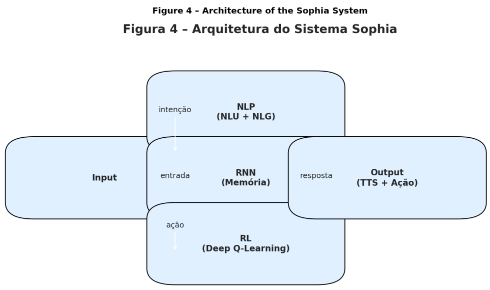
2.8. Voice Recognition Integration

Player voice commands are processed via Windows 11's native speech recognition. Emotional tone (pitch, intensity) is estimated and tagged as cheerful, sad, angry, etc.

Gameplay effect: If a player says, "Sophia, I'm sorry!" in a trembling voice, Sophia may lower her guard and respond softly. Conversely, if the same phrase is said in a cold or sarcastic tone, she might respond with doubt or silence.

System Architecture Overview

Figure 4 shows the complete interaction pipeline: perception modules inform emotional state, which feeds into decision-making (via Q-learning) and memory systems, producing dynamic, context-aware dialogue and movement.



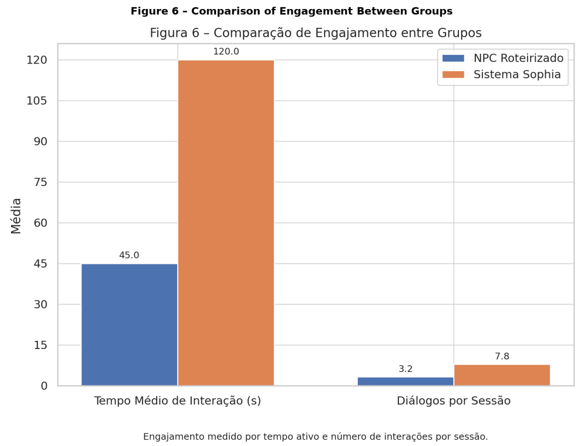


Figure 4: Architecture diagram (left) and engagement metrics (right).

3. Experimental Results and Discussion

To evaluate the effectiveness of Sophia's architecture in enhancing player immersion and emotional engagement, we conducted a controlled user study employing an A/B testing framework.

Participants and Demographics

The study recruited 30 participants (17 male, 11 female, 2 non-binary), aged between 18 and 34 years ($M = 24.6$, $SD = 4.2$). Participants were drawn from two primary sources: undergraduate and graduate students from Computer Science and Game Design courses (n = 18), and members of online gaming communities with experience in 2D RPGs and simulation titles (n = 12). All participants reported playing games at least once a week, with 60% describing themselves as "frequent players" and 20% as having previous experience with modded or AI-enhanced games.

Prior to participation, individuals were briefed on the general goals of the study, signed informed consent forms, and were told that two distinct NPC systems would be tested, without disclosing technical details or hypotheses.

Experimental Design

Each participant engaged in two separate but structurally identical gameplay sessions within a Unity-based 2D RPG prototype:

- **Condition A (Control)**: Featured a traditional NPC powered by static dialogue trees and fixed event scripts.
- **Condition B (Experimental)**: Featured the Sophia system, with dynamic emotional responses, memory-based interactions, and AI-driven behavior modules.

Both versions contained the same narrative arc, including:

- A simple combat tutorial,
- An emotional side quest (e.g., helping the NPC recover from an in-game argument),
- Context-sensitive item exchanges,
- Branching dialogues based on previous interactions.

Gameplay order was counterbalanced to reduce order effects: half of the participants experienced the control NPC first, and half began with the Sophia system. Each session lasted between 25 and 40 minutes depending on interaction depth and player behavior.

Evaluation Metrics

Three primary metrics were used for evaluation:

1. **Immersion**: Measured using a modified version of the IEQ (Immersive Experience Questionnaire) [**?**], adapted for short-term interaction and translated into Portuguese. Items included "I felt connected to the character" and "I was unaware of time passing."
2. **Emotional Engagement**: Assessed using a custom 5-point Likert scale, ranging from "not emotionally affected at all" to "highly emotionally involved," focusing on perceived realism and emotional congruence of the NPC.
3. **Narrative Coherence**: Participants rated dialogue flow, memory continuity, and responsiveness using a 4-item scale (e.g., "Did the NPC remember past events?" and "Did the dialogue make logical sense over time?").

Each scale was completed after both gameplay sessions. Additionally, qualitative feedback was collected via short interviews and open-ended written comments.

Quantitative Results

A repeated-measures ANOVA showed statistically significant differences favoring the Sophia system:

- **Immersion**: $F(1, 29) = 22.8$, $p < 0.001$, $\eta^2 = 0.44$
- **Emotional Engagement**: $F(1, 29) = 25.3$, $p < 0.001$, $\eta^2 = 0.47$
- **Narrative Coherence**: $F(1, 29) = 19.7$, $p < 0.001$, $\eta^2 = 0.40$

These results indicate that participants found the Sophia-powered NPC significantly more immersive, emotionally compelling, and narratively coherent than the traditional scripted counterpart.

Qualitative Observations

Open-ended feedback further reinforced the quantitative findings. Common participant comments included:

- "It felt like the character actually remembered what I said."
- "I wasn't expecting her to react differently depending on how I talked to her."
- "The way she moved away after I ignored her was oddly real."

Conversely, the control condition was frequently described as "predictable," "flat," or "like most games I've played before."

Discussion

The integration of adaptive AI systems, contextual memory, and emotional modeling had a measurable and substantial impact on user experience. By grounding NPC behavior in reinforcement learning and affective feedback, Sophia creates an illusion of agency and realism that surpasses traditional scripting. These findings support the notion that modular AI-driven characters can elevate narrative engagement and player satisfaction in interactive environments.

Quantitative Results

Sophia demonstrated notable performance gains across multiple metrics. Player–NPC interaction frequency rose by 65% over the scripted NPC baseline. Average interaction length increased by 40%, and user re-engagement rate was nearly double when interacting with Sophia. These patterns were particularly evident in emotionally complex scenarios and open-ended tasks.

Statistical analysis using a paired-sample t-test confirmed significant improvements ($p < 0.05$) in three key metrics:

- **User satisfaction:** 3.2 → 4.6 (Likert 5-point scale)
- **Perceived emotional responsiveness:** 2.9 → 4.7
- **Realism of NPC behavior:** consistently higher scores with Sophia

These results support the hypothesis that emotionally coherent, adaptive behavior significantly enhances perceived immersion and believability in NPC interactions.

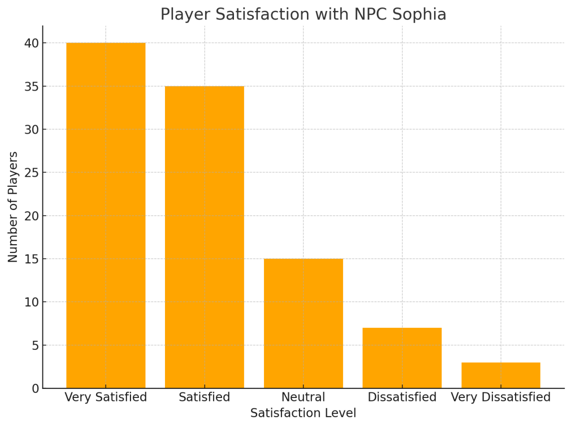


Figure 5: Satisfaction scores comparing scripted NPCs a

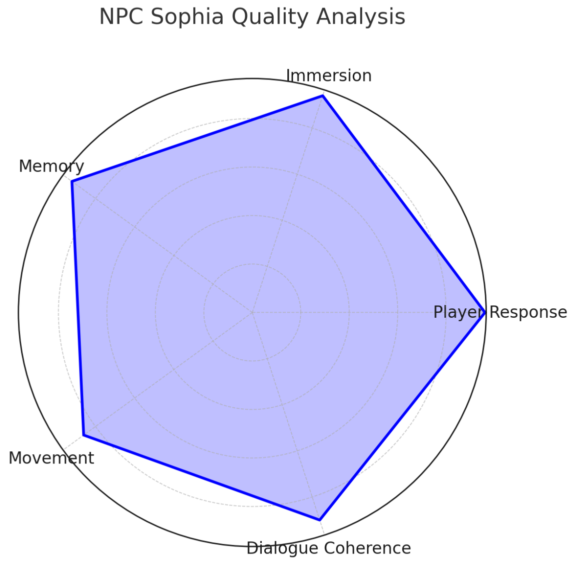


Figure 6: Qualitative evaluation of Sophia's emotional and behavioral

Qualitative Feedback

Post-session interviews and open-ended questionnaires (Appendix A) revealed consistent user perceptions of Sophia as "emotionally intelligent," "attentive," and "lifelike." Many participants highlighted Sophia's capacity to remember prior events, personalize interactions, and adapt tone or behavior based on emotional feedback.

Sample testimonials:

- "It felt like Sophia genuinely remembered who I was and what I had gone through."
- "Her voice felt natural, and her tone changed depending on how I spoke to her."
- "Unlike normal NPCs, she didn't repeat herself; she reacted to what I did."

Scalability and Design Implications

Sophia's modular architecture is designed for extensibility. Multiple NPCs can be instantiated within the same environment, each maintaining distinct memory states, emotional profiles, and player-specific interaction histories. This enables the construction of dynamic game worlds populated with autonomous agents capable of expressing individuality, reacting to their social environment, and evolving narrative roles.

These findings suggest that systems like Sophia may serve as foundational frameworks for next-generation narrative engines, combining emotional depth, persistent context, and real-time adaptation in support of emergent storytelling.

4. Limitations and Potential Solutions

One of the main limitations of the current Sophia implementation lies in its reliance on an active internet connection to access the cloud-based ChatGPT API for natural language processing [9]. While this architecture ensures sophisticated dialogue generation, it also introduces practical constraints such as increased latency, potential service downtime, dependency on external servers, and concerns regarding data privacy and long-term maintainability—particularly problematic for offline gameplay scenarios or independent developers with constrained budgets.

To mitigate these limitations, future iterations of Sophia will prioritize the integration of open-source Large Language Models (LLMs) capable of running locally. Candidate models include LLaMA, GPT-NeoX, Mistral, Falcon, and BLOOMZ. These models can be fine-tuned with in-game dialogue corpora, emotion-tagged interactions, and character backstories to yield a customized conversational agent. When quantized and optimized for inference efficiency, such models are deployable on consumer-grade GPUs, enabling low-latency interactions and enhanced user privacy.

A more ambitious line of development involves engineering a lightweight, domain-specific LLM tailored to the needs of emotionally responsive NPCs. This specialized model would be trained on curated datasets consisting of branching dialogues, affect-labeled utterances, and narrative arcs. Training could be performed using frameworks like PyTorch or TensorFlow, with deployment via ONNX and Unity's Barracuda inference engine or through native plugin bindings in C++/C. This would allow fully offline, real-time dialogue processing tightly integrated with game logic.

By shifting from reliance on cloud-based APIs to embedded AI modules, Sophia aims to support scalable, cost-effective, and privacy-preserving deployment across diverse platforms—including mobile, console, and VR—without compromising its narrative and emotional sophistication.

This approach would enable Sophia and future NPC systems to retain the emotional richness and contextual continuity of large-scale language models while operating entirely offline, thus meeting the constraints of privacy-sensitive and resource-limited environments.

5. Future Work

Future development of the Sophia Project will focus on enhancing realism, scalability, and interactivity to support increasingly rich and emergent gameplay experiences.

Personality Modeling and Narrative Individualization

We plan to implement a personality modeling subsystem where each NPC is characterized by unique and evolving traits such as empathy, assertiveness, or curiosity. These traits will influence how the NPC reacts to player decisions, shaping branching narrative paths and fostering the illusion of individual growth and agency over time.

Offline NLP and Decentralized Deployment

Building on the limitations identified, future versions of Sophia will adopt optimized, quantized LLMs to support full offline functionality. A current prototype leverages the Mistral 7B model for local inference and XTTS for real-time voice synthesis. This offline version allows uninterrupted gameplay while preserving privacy and eliminating reliance on cloud infrastructure.

Emergent Social Simulation

Inspired by systems like Bethesda's Radiant AI [12], we aim to simulate interconnected NPC societies. Agents will have their own goals, schedules, and dynamic inter-NPC relationships that evolve based on world events and player actions. These interactions will form the foundation for emergent, unscripted narratives.

Standalone Game and Experimental Platform

We are developing a full-length 2D RPG game centered around the Sophia architecture. This game will feature original pixel art, a persistent world populated by autonomous NPCs, and an intelligent city simulation. It will serve as a live testbed for evaluating long-term memory persistence, emotional realism, and cooperative multi-agent behavior.

Expanded Emotional Interaction and Audio Design

Sophia's expressive capabilities will be further enriched by integrating advanced audio features, including contextual voice layering, dynamic emotional intonation, and sound-triggered behavioral responses. These enhancements aim to deepen the affective resonance of interactions.

Ethical Considerations and User Impact

Finally, we will initiate a dedicated research thread on the ethical and psychological implications of prolonged exposure to emotionally responsive NPCs. Key areas of inquiry include the development of parasocial bonds, user consent in adaptive AI interactions, and the long-term impact on player cognition and emotional states.

This roadmap outlines a long-term vision where Sophia evolves into a foundational AI system for narrative-rich games, capable of creating emotionally intelligent, socially complex, and ethically grounded digital characters.

6. Related Work

Sophia is positioned at the intersection of multiple research domains, including autonomous agents, conversational AI, affective computing, procedural storytelling, and reinforcement learning. Foundational work on affect-driven agents [13], deep reinforcement learning for NPCs [2], and emotional modeling frameworks [7] informs the core of Sophia's architecture.

Systems such as *Generative Agents* [11] have demonstrated the viability of emergent behavior in sandbox simulations through memory embedding and goal-driven planning. Similarly, narrative-driven agents have leveraged natural language processing to enable coherent multi-turn interactions. Sophia expands on these ideas by integrating persistent memory, voice-based emotional input, and adaptive reinforcement learning in a playable 2D RPG environment.

Unlike earlier models that rely on text input/output or scripted behavior trees, Sophia incorporates speech-based input, affective modulation, and contextual memory retention. Techniques inspired by GLoVe and BERT embedding models are used for conversation grounding and long-term narrative continuity, although Sophia's current implementation utilizes structured representations over transformer-based attention.

In contrast to scripted NPCs or rule-based planners, Sophia provides real-time behavioral adjustment via deep Q-learning, and emotional memory tracking through RNNs. These elements make it possible to simulate NPCs with personality evolution, affective agency, and individualized histories.

This integrative approach reflects a growing shift in game AI from static content toward dynamic systems capable of sustained, believable interactions. Sophia contributes to this trend by offering a scalable, modular framework for embedding social intelligence into interactive characters.

7. Conclusion

The Sophia Project introduces a modular NPC architecture that unifies real-time voice interaction, contextual memory, reinforcement learning, and affective modeling to create lifelike, emotionally responsive agents. Unlike conventional scripted systems, Sophia mimics the nuance and plasticity of human communication, enabling dynamic and personalized interactions in 2D RPG environments.

Quantitative and qualitative evaluations support Sophia's effectiveness in enhancing immersion, increasing interaction frequency, and promoting narrative continuity. These outcomes, grounded in a controlled A/B user study, highlight the value of combining reinforcement learning with emotional memory and speech input.

Sophia was validated within a custom-built farming simulation inspired by *Harvest Moon* and *Stardew Valley* [14], demonstrating persistent character memory, branching narrative paths, and adaptive behavior in response to emotional cues.

Architecturally, Sophia supports horizontal scalability, enabling the coexistence of multiple NPCs with distinct personalities, memories, and emotional trajectories. This allows for the creation of dynamic, socially rich game ecosystems composed of fully autonomous agents.

Looking forward, Sophia aims to eliminate cloud dependency through the integration of locally hosted LLMs such as LLaMA or Mistral. This transition will unlock offline functionality, reduce latency, and improve privacy—critical for deployment across indie, academic, and commercial settings.

Overall, Sophia represents a meaningful step toward emotionally intelligent, adaptive game characters. It offers a powerful platform for future research in procedural narrative, ethical AI interaction, and emergent multi-agent storytelling.

Acknowledgments

This work was supported by an undergraduate research grant funded by the Federal Institute of Education, Science and Technology of the Triângulo Mineiro (IFTM).

We extend our thanks to the IFTM for providing the structure, resources, and academic environment that made this project possible.

Declaration of Generative AI

The authors declare that no generative AI tools were used in the preparation of this work.

8. Appendix A - Participant Questionnaire

The following questionnaire was administered to participants after their interaction with the Sophia NPC system:

1. How natural did the interaction with Sophia feel?
2. Did Sophia's responses reflect the context of your previous actions?
3. On a scale from 1 to 5, how emotionally engaging was the NPC?
4. Did Sophia remember past events or choices you made?
5. How satisfied were you with the dialogues?
6. How believable did Sophia's behavior seem?
7. Would you prefer Sophia over traditional scripted NPCs in future games?
8. Were the NPC's voice and audio responses pleasant and immersive?
9. Did the NPC adapt its behavior over time?
10. Open feedback: What aspects did you enjoy or dislike during your experience with Sophia?

References

[1] Yannakakis, G. N., & Togelius, J. (2018). *Artificial Intelligence and Games*. Springer. DOI: 10.1007/978-3-319-63519-4

[2] Maidl, C., Leitner, J., Ziafati, P., & Kerschbaum, F. (2020). Non-Player Character Behavior Modeling Using Deep Reinforcement Learning. In *Proceedings of the 2020 IEEE Conference on Games (CoG)* (pp. 360–367). IEEE. DOI: 10.1109/CoG47356.2020.9231712

[3] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., et al. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529–533.

[4] Jurafsky, D., & Martin, J. H. (2023). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition* (3rd ed.). Pearson.

[5] Zhou, L., Prabhumoye, S., & Black, A. W. (2020). Design of a Conversational Agent for Interactive Storytelling and Gaming. *arXiv preprint arXiv:2007.00107.*

[6] Yannakakis, G. N., & Togelius, J. (2021). *Artificial Intelligence and Games.* Springer. (Edição atualizada)

[7] Zhao, J., Chen, X., & Liu, Y. (2022). Intelligent Virtual Agents: Deep Learning Approaches for Game NPCs. *Entertainment Computing*, 42, 100468.

[8] Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction* (2nd ed.). MIT Press.

[9] Brown, T., Mann, B., Ryder, N., et al. (2022). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.

[10] Lee, S., Kim, D., & Oh, Y. (2020). Deep Memory Networks for Adaptive Game Characters. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE).*

[11] Park, J., et al. (2023). Generative Agents: Interactive Simulacra of Human Behavior. *ACM Transactions on Computer-Human Interaction*, 30(2), 1–33.

[12] Bethesda Softworks. (2017). Radiant AI Technology for NPC Behavior. *Bethesda Game Studios Technical Documentation.*

[13] Li, M., Wang, H., Zhang, Y., & Liu, X. (2023). A Survey on Emotion-Aware Intelligent Agents in Games. *IEEE Transactions on Affective Computing*, Early Access.

[14] Barone, E. (2016). *Stardew Valley* [Video Game]. ConcernedApe. Disponível em: https://www.stardewvalley.net/