Enhancing Game Development with Generative Artificial Intelligence Technologies: A Case Study of UNDO

Eduardo Barroso Nunes[1], **Maria Andréia Formico Rodrigues**[2]

[1]Universidade de Fortaleza (UNIFOR) – Secretaria de Segurança Pública do Ceará
Av. Washington Soares, 1321 – 60811-905 – Fortaleza – CE – Brazil

[2]PPGIA - Universidade de Fortaleza (UNIFOR) – GIRA Lab
Av. Washington Soares, 1321 – 60811-905 – Fortaleza – CE – Brazil

eduardo@edbn.dev, andreia.formico@gmail.com

Abstract. *This paper explores the use of Generative AI technologies in game development through the case study of the serious game UNDO, which integrates time-travel and environmental education. Players engage in recycling tasks to correct past environmental mistakes. By utilizing advanced AI technologies like ChatGPT-3.5, ChatGPT-4, GitHub Copilot, DALL-E 2, DALL-E 3, Midjourney, AudioLM, AIVA, SOUNDRAW, and ElevenLabs within Unity3D, UNDO demonstrates the potential of Generative AIs to enhance narrative, visuals, and audio. These tools facilitate content creation, enhance production efficiency, and reduce repetitive tasks, while still requiring human oversight for refinement. This study demonstrates Generative AI's effectiveness in enriching the game development process.*
Keywords *Generative AIs, Serious Game, Environmental Preservation, Recycling.*

1. Introduction

Generative AI technologies have revolutionized game development by automating content creation, enhancing innovation, and accelerating workflows [Taye 2023]. These technologies streamline the production process, enabling developers to focus on creativity and strategy. AI-driven tools generate diverse assets and simulate complex environments, enriching gaming experiences and facilitating the exploration of more challenging themes.

The integration of Generative AI in game development enhances productivity and creative freedom, making it an essential tool in the rapidly evolving industry [Schell 2008]. AI systems contribute to content creation, supporting gameplay that involves collaborative construction [Treanor et al. 2015]. While human oversight remains fundamental, the combination of human creativity, data, and AI capabilities continues to expand the possibilities in game design [Lin et al. 2023, Carvalho e Rodrigues 2023, Serpa e Rodrigues 2022]. AI tools for generating images, sounds, and text have significantly impacted the gaming and other media-rich industries. Techniques like *Stable Diffusion* are used to quickly create detailed visuals, enhancing efficiency [Rombach et al. 2022]. Projects like *Spiritual Voice Chat with a ChatGPT-driven Monk* [Tamulur 2022] and *Rhymates* [Bogomolny 2022] demonstrate AI's role in creating personalized, interactive experiences.

This work examines generative AI techniques through the prototype game UNDO, a serious game focusing on time travel and recycling. Generative AIs were used to create interactive narratives and enhance audiovisual elements, demonstrating the technology's potential to enrich and streamline the game development process.

2. Comparative Insights: AI Models in Language, Image, and Audio

AI is extensively used in the gaming industry for character, environment, physics, and sound development, as well as analyzing player behaviors [Huang e Huang 2023]. Advances in deep learning and neural networks have enabled Generative AI to generate personalized content, reducing manual programming [Takahashi 2021, Serpa e Rodrigues 2022, Taye 2023].

Generative natural language AI models like *ChatGPT-3*, *ChatGPT-3.5*, *ChatGPT-4*, *ChatGPT-4o*, Bing AI, and Open Assistant facilitate tasks such as text synthesis and programming. These models use machine learning to generate coherent and contextually relevant text, with *GitHub Copilot* assisting in code generation [LAION 2022, Mehdi 2023, OpenAI 2024].

For image creation, tools like *DALL-E 2*, *DALL-E 3*, and *Midjourney* generate high-quality visuals from text descriptions. These tools, including *Stable Diffusion*, offer diverse quality, realism, and style, and are widely used for artistic and educational projects [Rombach et al. 2022, Holz 2022]. Additionally, the *NVIDIA Canvas App* [NVIDIA Corporation 2021] uses Generative AI to turn simple brushstrokes into realistic landscapes. *CrAIyon*, with the VQGAN network, generates colorful images from text inputs, ideal for conceptual illustrations [Hai-Jew 2023].

In audio and music, Generative AIs like *AIVA*, *ElevenLabs*, *SOUNDRAW*, *AudioLM*, and *MusicLM* create and manipulate sound, enhancing creative processes in content creation. These technologies enable the production of original music, realistic synthetic voices, and comprehensive audio processing [AIVA Technologies 2020, SOUNDRAW, Inc. 2024, ElevenLabs 2022, Borsos et al. 2023, Agostinelli et al. 2023].

3. UNDO Overview

This section details the design and development of the UNDO game using Generative AIs, guided by a compact team of two professionals specializing in both game design and development. UNDO is a serious 3D prototype game that explores historical periods impacted by a time-travel accident, leading to environmental degradation. The original narrative revolves around a technical mishap with a time machine, placing the player in control to save Earth from the trash accumulated over various eras. The game's story is conveyed through dynamic cutscenes and animations. The player, embodied by the humanoid robot hero character, Jammo, is guided by the NPC Robot Sphere (Figure 1), who provides recycling knowledge and navigational assistance throughout the game [Silva e Rodrigues 2009, Barbosa e Rodrigues 2006]. Robot Sphere's role is important in explaining the plot and guiding the player through different phases, enhancing the game's educational and narrative depth.

3.1. Gameplay

The game UNDO starts at the Main Menu, where players can interact with the NPC Robot Sphere, review collected trash, access unlocked levels, and view cutscenes. It

includes three levels: Tutorial, Desert, and Forest. Upon selecting a level, an introductory cutscene sets the scene, followed by the player's task of constructing an antenna to enable trash collection. Players must sort collected trash into specific bins, each designated for different types of waste (e.g., organic, plastic, metal). The Robot Sphere, with voice narration enhanced by *ElevenLabs* text-to-speech technology, guides players by explaining the recycling process for each type of trash. Efficient trash management and time optimization are critical, with points awarded for correct actions. The game features a three-jump mechanic inspired by Super Mario 64, and the difficulty increases with each level, providing a challenging learning curve. UNDO's development followed a structured flow, using Unity's *Input System* for enhanced 3D navigation and gamepad support, focusing on platforming and map exploration. Synthetically, the development process of UNDO followed the structured flow diagram shown in Figure 1.

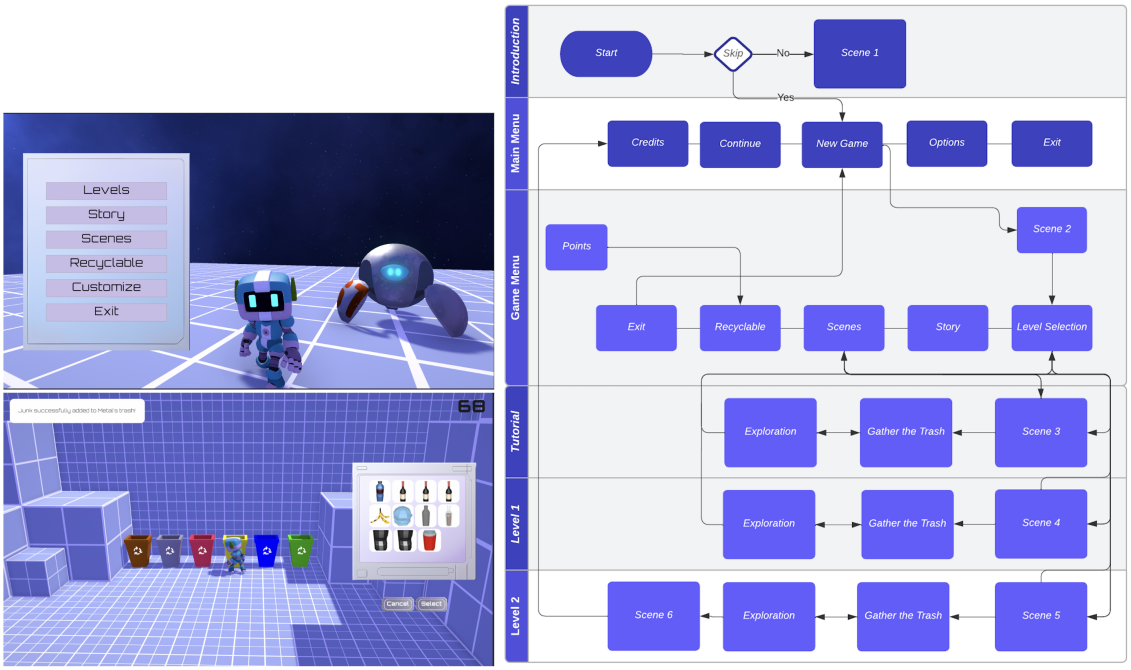

Figure 1. Main game menu with Jammo (left) and NPC Robot Sphere (right) (top left); Jammo by recycling bins with disposal notifications (bottom left); UNDO game flow diagram (right).

3.2. Generative AIs Used in UNDO

The motivation for integrating Generative AIs in UNDO extends beyond enhancing art, narrative, sound, and gameplay; it also explores their potential in animation, coding, and sound design. We utilized student versions of these technologies, supplemented with paid versions when necessary. These tools were iteratively applied during development cycles of testing, revision, and code refinement to ensure high-quality outcomes. The development was conducted on a DELL G15 laptop with a 12th Gen Intel i5-12500H CPU and Intel Iris Xe Graphics GPU. The technologies employed included:

- ***Midjourney* and *DALL-E***: These models were essential for creating visual art and user interfaces (UIs), enhancing the game's aesthetics and user experience. *Midjourney* was used for generating thematic elements across various settings

(indoor, desert, forest), while *DALL-E 2* and *DALL-E 3* were utilized for creating backgrounds for game menus, including the Configuration Menu and credits screen [Holz 2022, OpenAI 2023].

- ***ChatGPT* and *GitHub Copilot***: Employed for narrative generation and source code development, *ChatGPT-3.5* and *ChatGPT-4* crafted complex narratives and dialogues, and resolved gameplay development issues. *GitHub Copilot*, integrated into the *JetBrains Rider IDE*, facilitated C# game development, producing functional and optimized code efficiently [OpenAI 2024, GitHub 2021].
- ***ElevenLabs***: This text-to-speech technology brought the protagonist's dialogues to life, enhancing user engagement and providing a realistic auditory experience [ElevenLabs 2022].
- ***AIVA* and *SOUNDRAW***: Used for musical and sound composition, these models created the game's atmosphere and cohesive soundtrack. *AIVA* composed music [AIVA Technologies 2020], while *SOUNDRAW* customized soundtracks [SOUNDRAW, Inc. 2024].
- ***AudioLM***: This model was used for creating specific sound effects, such as footsteps and jumping sounds, executed via Python, providing precision and adaptability to enhance the game's auditory experience [Borsos et al. 2023].

3.3. Functionalities and Solutions

In this subsection, we explore the main functionalities and solutions implemented in UNDO, utilizing generative AI technologies to enhance game development and user experience.

Natural Language, Scripting, and Post-Processing: The Unity UI Toolkit, guided by ChatGPT, enabled the creation of UXML layouts and Unity Style Sheets (USS) for game scenarios like save management and level selection. ChatGPT provided structured UXML and USS examples, simplifying UI design. It also assisted in scripting C# for game state management and 3D scene navigation. Challenges in 3D character control, such as input precision and state management, were resolved by transitioning to Unity's new Input System and incorporating State Machines, as recommended by ChatGPT, enhancing control accuracy and responsiveness [Rodrigues et al. 2015, Serpa et al. 2020]. Post-processing complexities were addressed with ChatGPT-4's guidance, recommending effects like Bloom, Color Grading, Depth of Field, Vignette, and Ambient Occlusion, which improved visual appeal and accessibility, particularly for a young audience.

UI Elements and Background Design: For creating user interface (UI) elements like buttons and panels, the *Midjourney* tool was utilized. By using Midjourney's bot on Discord, unique visual designs were generated, aligning with the UNDO game's aesthetic. Detailed prompts allowed exploration of various styles, resulting in multiple design options. *Midjourney* facilitated experimentation with diverse styles, textures, and colors, enhancing the game's inventory system with futuristic, science fiction elements. The designs ensured visual cohesion across the game, even with distinct elements used for menus and inventory. This maintained a harmonious aesthetic, providing a clear and interactive user interface. For menu and cutscene backgrounds, *DALL-E* and *DALL-E 3* were employed. These tools effectively generated images that matched the game's theme and atmosphere, integrating seamlessly into the design framework.

Soundtrack and Sound Effects: For the UNDO game's menu and level soundtracks, *AIVA* and *SoundRaw* were employed. *AIVA* allowed for detailed music composition adjustments, including scales, styles, BPM, and instruments, akin to tools like *Fruit Loops* [Image-Line Software 2023]. *SOUNDRAW* facilitated quick music generation based on abstract themes like "sad" or "exciting," though with less fine-tuning. Together, these tools created a rich, diverse soundtrack that captured specific emotions and atmospheres. *AudioLM* was used for generating sound effects and ambient sounds for menus, item selections, character movements, and explosions, refined through iterative testing. *ElevenLabs* provided narrative and dialogue voiceovers in cutscenes, enhancing characters' and stories' realism. It also offered educational content in the "About" menu, enriching the auditory and informative experience for users.

4. Tests and Results

The primary developer initially created UNDO without Generative AIs to understand traditional development methods. Later, Generative AIs were integrated to explore modern alternatives and compare outcomes. In the first version, all game logic was developed internally, without external libraries or tools. The second version, enhanced with Generative AIs, included significant improvements in art, sound, and gameplay. New cutscenes, menus, and control mechanics were introduced, enriching the overall gaming experience. For character animation and cutscene creation, the *Timelines* tool from Unity was used, enabling synchronization of audio and visuals for a cohesive narrative. Additionally, *ChatGPT* was consulted for programming advice, such as implementing getters and setters in C#, to enhance coding practices. The prototype developed with Generative AIs demonstrated substantial improvements in the game's main phases—Tutorial, Desert, and Forest—particularly in visual quality, animation smoothness, and audio, while also accelerating the game development process.

5. Conclusion and Future Work

This study demonstrated the transformative impact of generative AI tools on game development, as exemplified by the UNDO game. Utilizing tools like *ChatGPT*, *DALL-E*, *Midjourney*, *AudioLM*, *AIVA*, *SOUNDRAW*, and *ElevenLabs*, we significantly enhanced narrative, visual, and audio elements, accelerating the development process. The integration with Unity3D enabled rapid content generation and improved creative output, despite challenges requiring multiple iterations and precise prompt crafting. Future work involves expanding UNDO with new levels, mini-games, and a customized virtual assistant. Further exploration of generative AIs can streamline repetitive tasks, allowing developers to focus on complex design and emotional integration, ultimately enhancing the gaming experience.

Acknowledgements

The authors thank the University of Fortaleza (UNIFOR) and CNPq (Ed. CNPq 09/2023) for the financial support.

References

Agostinelli, A., Denk, T. I., Borsos, Z., Engel, J., Verzetti, M., Caillon, A., Huang, Q., Jansen, A., Roberts, A., Tagliasacchi, M., Sharifi, M., Zeghidour, N., e Frank,

C. (2023). MusicLM: Generating music from text. https://arxiv.org/abs/2301.11325. Last Visited: 2024-01-10.

AIVA Technologies (2020). AIVA – The AI composing emotional soundtrack music. https://aiva.ai/. Last Visited: 2024-01-10.

Barbosa, R. G. e Rodrigues, M. A. F. (2006). Supporting guided navigation in mobile virtual environments. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, pages 220–226.

Bogomolny, E. (2022). Rhymates. https://bogomolnyelad.substack.com/p/board-game. Accessed: 2024-05-28.

Borsos, Z., Marinier, R., Vincent, D., Kharitonov, E., Pietquin, O., Sharifi, M., Roblek, D., Teboul, O., Grangier, D., Tagliasacchi, M., et al. (2023). AudioLM: A language modeling approach to audio generation. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 31:2523–2533.

Carvalho, V. M. e Rodrigues, M. A. F. (2023). Investigating and comparing the perceptions of voice interaction in digital games: Opportunities for health and wellness applications. In *2023 IEEE 11th International Conference on Serious Games and Applications for Health (SeGAH)*, pages 1–8. IEEE.

ElevenLabs (2022). Prime Voice AI. https://beta.elevenlabs.io/. Last Visited: 23/04/2024.

GitHub (2021). GitHub Copilot: Your AI pair programmer. https://github.com/features/copilot. Last Visited: 2024-05-20.

Hai-Jew, S. (2023). CrAIyon: Putting an art-making AI through its paces. *C2C Digital Magazine*, 1(18):10. Last Visited: 2023-04-20.

Holz, D. (2022). Midjourney. https://www.midjourney.com/home/. Last Visited: 2024-04-20.

Huang, J. e Huang, K. (2023). ChatGPT in Gaming Industry. In *Beyond AI: ChatGPT, Web3, and the Business Landscape of Tomorrow*, pages 243–269. Springer.

Image-Line Software (2023). Fruit Loops Studio. https://www.image-line.com/fl-studio/. Last Visited: 2024-04-10.

LAION (2022). Open Assistant. https://open-assistant.io/. Last Visited: 2024-03-10.

Lin, C.-H., Gao, J., Tang, L., Takikawa, T., Zeng, X., Huang, X., Kreis, K., Fidler, S., Liu, M.-Y., e Lin, T.-Y. (2023). Magic3D: High-resolution text-to-3D content creation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 300–309.

Mehdi, Y. (2023). Reinventing search with a new AI-powered Microsoft Bing and Edge: Your copilot for the web. https://blogs.microsoft.com/blog/2023/02/07/. Last Visited: 2024-03-10.

NVIDIA Corporation (2021). NVIDIA Canvas. Computer Software. https://www.nvidia.com/en-us/studio/canvas/. Last Visited: 2024-05-15.

OpenAI (2023). DALL-E 3. Computer Software. https://openai.com/dall-e-3. Last Visited: 2024-02-15.

OpenAI (2024). ChatGPT – Conversation with an AI language model. https://openai.com/chatgpt. Last Visited: 2024-07-10.

Rodrigues, M. A. F., Macedo, D. V., Serpa, Y. R., e Serpa, Y. R. (2015). Beyond fun: an interactive and educational 3D traffic rules game controlled by non-traditional devices. In *Proceedings of the 30th Annual ACM Symposium on Applied Computing*, pages 239–246.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., e Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10684–10695.

Schell, J. (2008). *The Art of Game Design: A book of lenses*. CRC press.

Serpa, Y. R., Nogueira, M. B., Rocha, H., Macedo, D. V., e Rodrigues, M. A. F. (2020). An interactive simulation-based game of a manufacturing process in heavy industry. *Entertainment Computing*, 34:100343.

Serpa, Y. R. e Rodrigues, M. A. F. (2022). Human and machine collaboration for painting game assets with Deep Learning. *Entertainment Computing*, 43:100497.

Silva, W. B. e Rodrigues, M. A. F. (2009). A lightweight 3D visualization and navigation system on handheld devices. In *Proceedings of the 24th ACM Symposium on Applied Computing*, pages 162–166.

SOUNDRAW, Inc. (2024). SOUNDRAW – AI Music Generator. Computer Software. https://soundraw.io/. Last Visited: 20/05/2024.

Takahashi, K. (2021). AICommand: ChatGPT integration with Unity Editor. https://github.com/keijiro/AICommand. Last Visited: 2024-05-22.

Tamulur (2022). Spiritual Voice Chat with a ChatGPT-driven Monk. https://tamulur.itch.io/spirtual-voice-chat-with-a-chatgpt-driven-monk. Last Visited: 2024-03-10.

Taye, M. M. (2023). Understanding of Machine Learning with Deep Learning: Architectures, workflow, applications and future directions. *Computers*, 12(5):91.

Treanor, M., Zook, A., Eladhari, M. P., Togelius, J., Smith, G., Cook, M., Thompson, T., Magerko, B., Levine, J., e Smith, A. (2015). AI-based game design patterns.