Milestone 3: Interim Report

Progress and Implementation Status

We have made significant progress in various aspects of the game, building its layers systematically to provide an immersive experience for the players.

Our **Functional Minimum** was "*turn-based combat working. Basic interaction with the LLM driven characters should all be working. Karmascore based on the conversations should be working.*" We are ready with the first two but the Karma System is currently missing. We had to implement other systems first as a base for the Karma system.

For the **Low Target** our goal was "*The same as the functional minimum, plus all of the interaction should be done with Voice to text and text to Voice, and there should be a very nice hands only interface inside the robot.*" Here we are finished with the Text to Voice and Voice to text system, but at the moment the Control is done with Buttons. We also have the exterior of the robots and vr character movement working. we prioritized it over having a beautiful interface

Additionally we have to say that we deviated a little bit from our initial Layer Plan because some things had to be done first or were prioritized by us in order to get a working version first. Other than that we realized that some elements were missing in our layer plan that's why we cannot mark the two above layers done. In our perspective we are good in time with our progress. The details follow in the next section.

In terms of game assets, we have completed the development of 3D characters, a battle arena, and giant robots. These were crafted using Maya and supplemented with AI-assisted tools. All assets have been saved as `.fbx` files, ready for seamless integration into Unity. For animations, we utilized Mixamo to conform humanoid models with skeletons, enabling a diverse library of animations including idle, walk, fall, and aiming sequences. Alongside visual elements, sound design has been a focus, involving the curation and editing of effects such as footsteps, gunfire, explosions, and the mechanical rumbling of robots. These efforts ensure that both visual and audio elements enrich the gameplay experience.

On the AI and LLM front, we conducted extensive testing to integrate conversational AI into the game for NPC interactions. ChatGPT provided initial dialogue capabilities but proved to be less engaging in tone and speed compared to Llama 3.1. While Llama delivers a more dynamic conversational style, it runs slower on local hardware, presenting a challenge for seamless implementation. A text-only prototype of the game was developed to explore LLM capabilities, including scenarios like waiting rooms, shops, combat sequences, and NPC dialogues. While functional, this prototype highlighted issues with maintaining game state accuracy, such as tracking turns and inventory, which necessitated a more robust approach to game state management.

Regarding game mechanics, we made strides in programming a state system to track player and NPC attributes, such as health, strength, inventory, and active combat scenarios. This hybrid approach combines AI interactions with an external system to maintain consistency. We also programmed realistic shooting mechanics with calculated projectile trajectories and implemented a 3D game space menu for in-game controls, specifically for robot management.

The player interaction systems were crafted with meticulous attention to detail, emphasizing a seamless and engaging experience. Joystick-based controls were implemented to handle both movement and rotation, ensuring precise responsiveness during gameplay. While the implementation process posed certain challenges, Unity's robust toolset facilitated the integration of these features effectively. The combat system was designed to provide players with diverse tactical choices, featuring a button-activated shooting mechanism for offense and a dynamic shield system for defense. These elements work in harmony to add complexity and strategy to combat scenarios, enhancing the overall depth and replayability of the gameplay experience.

However, the development process has not been without challenges. One unexpected difficulty was ensuring the accuracy of LLM-driven game logic. The AI occasionally misinterpreted the game state, leading to errors in aspects like turn tracking and inventory management.The animation pipelines introduced significant challenges, especially when adapting Mixamo's skeletons to align with the unique models designed for our game. Ensuring compatibility between the predefined skeletons and our custom rigging required extensive tweaking and experimentation to maintain proper animations. Beyond this, creating smooth transitions between animations demanded additional effort to achieve a polished and cohesive player experience.

On the positive side, some elements were easier than expected. For example, the joystick system for controlling rotation and movement came together quickly, thanks to Unity’s built-in frameworks. These successes have helped maintain momentum in the face of other challenges.

Another relatively straightforward yet impactful feature we implemented was the voice-to-text and text-to-voice functionality. This system leverages Meta’s Voice SDK, which comes equipped with built-in Wit.AI components designed to seamlessly integrate into Unity. The SDK’s out-of-the-box compatibility significantly streamlined the initial setup, allowing us to focus on fine-tuning the experience rather than reinventing core voice processing mechanics. These capabilities enable players to interact with the game environment through natural language commands, adding a layer of accessibility and immersion.

However, while the core functionality was simple to implement, refining its usability presented more nuanced challenges. A key difficulty lies in determining the optimal method to initiate and terminate voice commands and voice detection. Currently, these actions are triggered by a physical button press, which provides reliability but deviates from our ultimate goal of enabling a completely hands-free experience. For users playing on the Quest, our vision is to rely solely on voice commands and hand gestures for control, creating an intuitive and futuristic interface.

To achieve this, several areas require further exploration. We need to implement a robust voice activity detection system to automatically recognize when a player starts speaking while minimizing false positives from ambient noise or overlapping dialogue. Additionally, incorporating contextual awareness, such as interpreting the intent behind commands based on gameplay scenarios, could significantly enhance the system's responsiveness and fluidity. Moreover, integrating fallback mechanisms, like visual indicators or subtle haptic feedback, would provide users with clear confirmation of when the system is actively listening or has recognized a command. These enhancements aim to elevate the hands-free interaction to a level that feels natural and frictionless, aligning with the immersive potential of VR gaming.

Our experiences during development have led to numerous design revisions, each aimed at improving functionality, immersion, and player engagement. A major insight came from recognizing the inherent limitations of relying exclusively on the large language model (LLM) for state management and decision-making. To address these challenges, we implemented a robust hybrid state management system that seamlessly combines the LLM's dynamic AI capabilities with external tracking mechanisms. This approach not only improves the accuracy and consistency of in-game responses but also allows for more precise monitoring of gameplay events and player actions. By blending AI-driven adaptability with structured tracking, we achieved a balance that significantly enhances both performance and player experience.

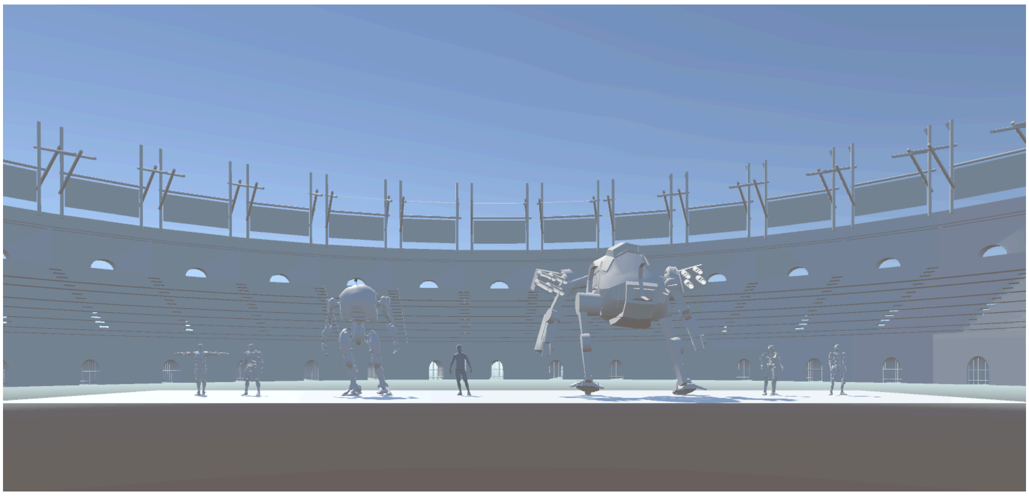
Combat mechanics underwent substantial refinement as well. To inject more variety and challenge into gameplay, we moved beyond rigid, pre-scripted behaviors for opponents and integrated basic AI logic that allows them to adapt to player strategies. This shift introduces an element of unpredictability to encounters, requiring players to think critically and adapt in real time. The opponents can now exhibit more nuanced behaviors, such as responding to player movement patterns, targeting vulnerabilities.

In addition to gameplay systems, we prioritized significant visual and audio upgrades to deepen the sense of immersion. A new day-night cycle was implemented with shaders that dynamically adjust lighting and environmental details based on the time of day. This feature not only enhances the realism of the game world but also creates new tactical opportunities. Combat moves were enriched with visual effects that reflect the energy and impact of each action, ensuring that every strike, block, or special ability feels distinct and satisfying.

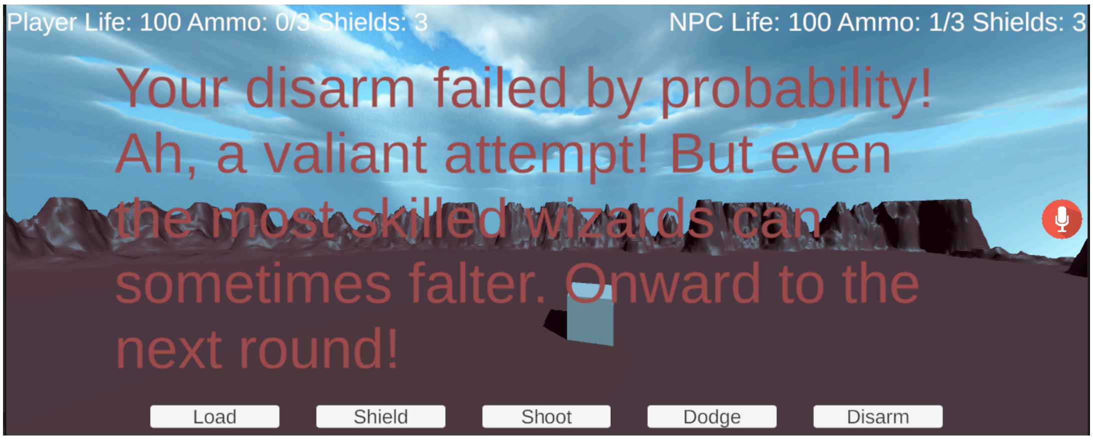
Complementing these visual enhancements, the audio design received equal attention. Action-specific sound effects were selected to provide auditory feedback that matches the intensity and context of the player’s choices, from the sharp clash of projectiles to the low hum of a charged shield.

These revisions collectively reflect our commitment to creating a game that not only engages players on a mechanical level but also immerses them in a vibrant and responsive world. Each enhancement was informed by player feedback and our own iterative testing, ensuring that every element—from state management to sensory details—works in harmony to deliver a seamless and captivating experience.

To support this development, we have prepared screenshots of our 3D assets, including characters, the arena, and robots as they appear in Unity. A video demonstration highlights the joystick’s functionality, showcasing movement and shooting mechanics. Screenshots of the day/night shader effects and combat illustrate the game's visual appearance.



Original characters in the arena (Unity)



Player Life: 100 Ammo: 0/3 Shields: 3

NPC Life: 100 Ammo: 1/3 Shields: 3

Your disarm failed by probability! Ah, a valiant attempt! But even the most skilled wizards can sometimes falter. Onward to the next round!

The fight when playing on PC, controlled by buttons and/or voice already includes basic animations for each choice

Joystick and button in vr world as controllers