

AVR2025-UncannyNPC: OpenAI Realtime NPC Integration

Interaktive VR-NPCs mit KI-gesteuerter Sprachverarbeitung

Projektinformationen

Projekttitel:	AVR2025-UncannyNPC
Untertitel:	Interactive NPCs with Real-time Voice AI in Unity VR
Team:	crack666 & maiossa
Repository:	https://github.com/crack666/AVR2025-UncannyNPC
Technologie:	Unity 2022.3 + OpenAI Realtime API + Meta Quest (PC-Streaming)

Projektziel

Hauptziel: Entwicklung einer immersiven VR-Umgebung mit intelligenten NPCs, die über natürliche Sprache in Echtzeit interagieren können.

Kernfunktionen:

- **Echtzeitsprachverarbeitung:** Integration der OpenAI Realtime API für nahtlose Sprachkommunikation
- **VR-Optimierung:** Entwickelt für Meta Quest 2/3 mit PC-Streaming, Standalone in Entwicklung
- **Audiovisueller Realismus:** Synchronisierte Lippenbewegungen und natürliche Animationen
- **Skalierbare Architektur:** Erweiterbar für multiple NPCs und komplexe Szenarien

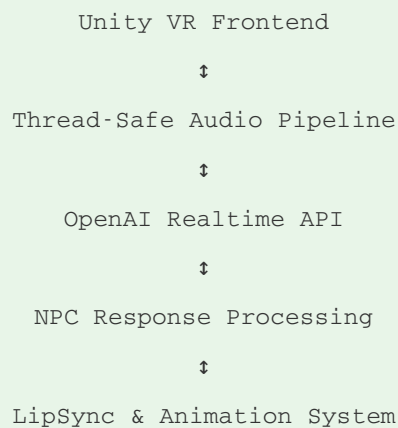
Technische Spezifikationen

Technologie-Stack:

- **Unity 2022.3 LTS:** Hauptentwicklungsumgebung mit XR-Support

- **XR Interaction Toolkit:** VR-Interaktionssystem für Meta Quest
- **OpenAI Realtime API:** KI-gesteuerte Sprachverarbeitung
- **ReadyPlayerMe:** Professionelles Avatar-System
- **Mixamo:** Charakteranimationen und Rigging
- **uLipSync + Custom Fallback:** Präzise Lippensynchronisation

System-Architektur:



Lösungsansätze & Innovationen

1. Thread-sichere Audio-Architektur

Herausforderung: Unity's Hauptthread kann bei intensiver VR-Verarbeitung blockieren

Lösung: Implementierung einer separaten Audio-Pipeline mit Background-Threading für unterbrechungsfreie Sprachverarbeitung

2. Latenz-Optimierung

Herausforderung: Netzwerklatenz beeinträchtigt VR-Immersion

Lösung: Intelligente Buffering-Strategie mit Predictive Loading und lokaler Fallback-Verarbeitung

3. Hybrid LipSync-System

Herausforderung: Standard-LipSync unzureichend für realistische NPCs

Lösung: Kombination aus uLipSync-Präzision und Custom Fallback-Algorithmus für robuste Synchronisation

Implementierungsstatus

Grundarchitektur (Abgeschlossen)

- Unity VR-Projekt Setup mit XR Toolkit
- OpenAI Realtime API Integration
- Basis-NPC System und ReadyPlayerMe Integration

Audio-Pipeline (Abgeschlossen)

- Thread-sichere Audio-Verarbeitung implementiert
- Echtzeitsprachaufnahme und -übertragung
- Streaming Audio Playback mit Latenz-Optimierung

Animation & LipSync (Aktuell)

- uLipSync Integration in Unity-Pipeline
- Custom Fallback-System für robuste Synchronisation
- Avatar-Animation Pipeline mit Mixamo

VR-Optimierung (Geplant)

- Performance-Tuning für Quest Hardware
- UI/UX Verfeinerung für VR-Umgebung
- Multi-NPC Support und Szenario-Erweiterung

Erreichte Meilensteine

Erreichte Meilensteine:

- OpenAI Realtime API Integration in Unity
- Thread-sichere Audio-Pipeline Grundarchitektur
- VR-Interaktion mit NPCs (PC-Streaming)
- ReadyPlayerMe Avatar-System Integration
- Basis LipSync-System implementiert

- Automatisiertes Setup-System für Entwickler

Performance-Ziele:

- **Audio-Latenz:** < 1000ms End-to-End (Zielwert)
- **VR Frame Rate:** 72/90 FPS (Quest 2/3 Zielwerte)
- **Platform:** PC-VR Streaming (Quest Standalone in Entwicklung)
- **Entwicklungsstatus:** Core-Funktionen implementiert
- **Test-Platform:** Primär PC-basierte VR-Entwicklung

Ausblick & Erweiterungsmöglichkeiten

Zukünftige Entwicklungen:

- **Emotion Recognition:** KI-basierte Emotionserkennung in Sprache und Gestik
- **Persistent Memory:** NPCs mit langfristigem Gesprächsgedächtnis
- **Multi-Language Support:** Mehrsprachige NPC-Interaktionen
- **Advanced Facial Animation:** Expression Mapping und Micro-Expressions
- **Cloud Scaling:** Serverbasierte NPC-Intelligenz für komplexe Szenarien

Projektfazit

Das AVR2025-UncannyNpc Projekt demonstriert eine funktionsfähige Integration der OpenAI Realtime API in Unity VR-Umgebungen. Die entwickelte Lösung ermöglicht grundlegende Sprachinteraktionen mit NPCs in VR, wobei der Fokus auf der technischen Machbarkeit und Entwicklungserfahrung liegt.

Die implementierten Systeme für Thread-Management, Audio-Streaming und Avatar-Integration bilden eine solide Grundlage für weitere Entwicklungen in der VR-KI-Interaktion. Das Projekt zeigt sowohl die Möglichkeiten als auch die aktuellen technischen Herausforderungen bei der Echtzeitsprachverarbeitung in VR auf.