

# Development of an Immersive VR Learning System with LLM-Integrated Avatar for Educational Assistance

## Thesis defense

Rezwanul Ashraf Ruddro  
Master of Science in Industrial Engineering  
Lamar University

Supervisor : Dr. Wenhao Yang

Research focus: immersive educational interaction, system integration, and comparative evaluation



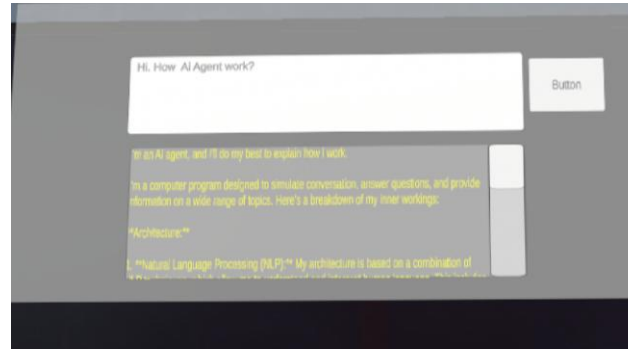
# Why this research matters

## Educational VR



- Supports presence, spatial context, and experiential learning in educational settings.
- Is particularly useful for simulation-based and embodied interaction.
- However, many educational VR systems still rely on scripted instruction or predefined interaction.

## Large language models



- Strong at open-ended question answering and explanation.
- Flexible natural-language interaction across varied prompts.
- Usually experienced through conventional screen-based interfaces.

## The unresolved gap

- Many VR learning systems lack open-ended conversational support.
- Large language models are typically used through screen-based interfaces rather than immersive environments.
- Avatar-based systems often remain scripted or decorative instead of functioning as intelligent educational assistants.
- Existing work rarely combines VR, embodied avatar interaction, and LLM-based support within one unified educational workflow.

### Thesis position

This thesis addresses these gaps by developing an immersive VR learning system in which an LLM-powered avatar provides real-time educational assistance within the virtual classroom.

## 1. Immersive Virtual Reality in Education

- Enhances engagement, presence, and experiential learning in higher education.
- Effective for simulation-based and spatial learning environments.

## 2. Large Language Models in Education

- Support intelligent tutoring, academic assistance, and open-ended question answering.
- Enable natural-language interaction across diverse educational contexts..

## 3. Pedagogical Agents and Avatars

- Improve learner motivation, social presence, and engagement in digital environments.
- Embodied agents facilitate more interactive and immersive learning experiences.

## 4. Research Gap

- Existing VR systems lack real-time conversational intelligence.
- LLM-based tools are typically confined to screen-based interfaces.
- Few studies integrate immersive VR, embodied avatars, and LLM-powered assistance into a unified educational framework.

# Problem statement

## What remains insufficient

- Many VR learning systems still lack open-ended conversational support.
- AI educational tools are usually delivered through screen-based interfaces.
- Avatar-based systems often remain scripted rather than dynamically instructional.
- Existing work rarely integrates VR, embodied avatars, LLM-based assistance, and usable in-environment input/output into one workflow.

## Existing educational systems

- VR learning systems support presence, simulation, and spatial engagement.
- AI educational assistants support open-ended question answering and flexible academic assistance.
- Avatar-based learning environments can improve interaction and social presence.

## Problem statement

There remains a need for an educational system that allows a learner to remain inside a virtual environment, submit questions through an in-world interface, and receive dynamically generated responses without leaving the immersive space.

### Scope of the thesis

This study focuses on prototype-level feasibility, interaction flow, and comparative system behavior rather than long-term learning-outcome measurement or a full intelligent tutoring architecture.

# Aim, objectives, and research questions

## Research aim

Design, develop, and evaluate an immersive VR learning system that uses an LLM-powered conversational avatar to provide real-time educational support.

## Core objectives

- Build the virtual learning environment in Unity.
- Integrate an avatar-based assistance mechanism.
- Enable in-world question entry through a virtual keyboard.
- Connect a locally hosted or open-source LLM.
- Display responses clearly inside VR.
- Evaluate behavior through a case study and PC baseline comparison.

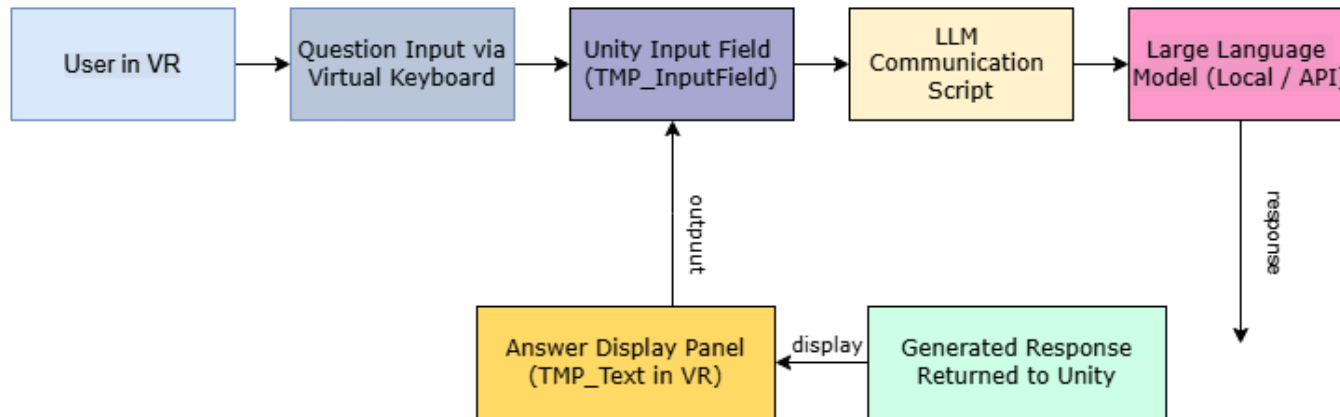
## Research questions

- 1 How can a VR learning environment be integrated with an LLM-powered avatar for educational assistance?
- 2 What architecture and interaction design are needed for real-time communication in VR?
- 3 How accurately and relevantly does the system respond compared with a PC-based AI baseline?
- 4 To what extent does the system improve interactivity in the VR learning environment?

# Proposed system overview

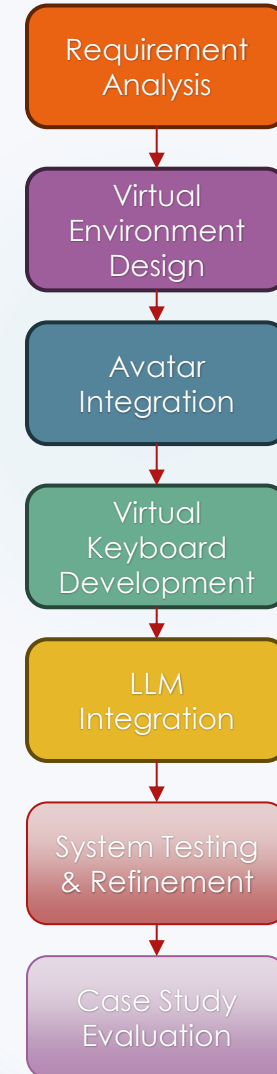
## LLM Communication and Response Generation Pipeline

Flow of user questions from the VR interface to the language model and back to the immersive environment



The user enters a question in VR, Unity forwards it through the communication script, the language model generates a response, and the answer is displayed back inside the immersive environment.

## Research and Development process



# Architecture and interaction workflow

## Interaction workflow

### Ask

The learner types a question using the VR virtual keyboard.

### Send

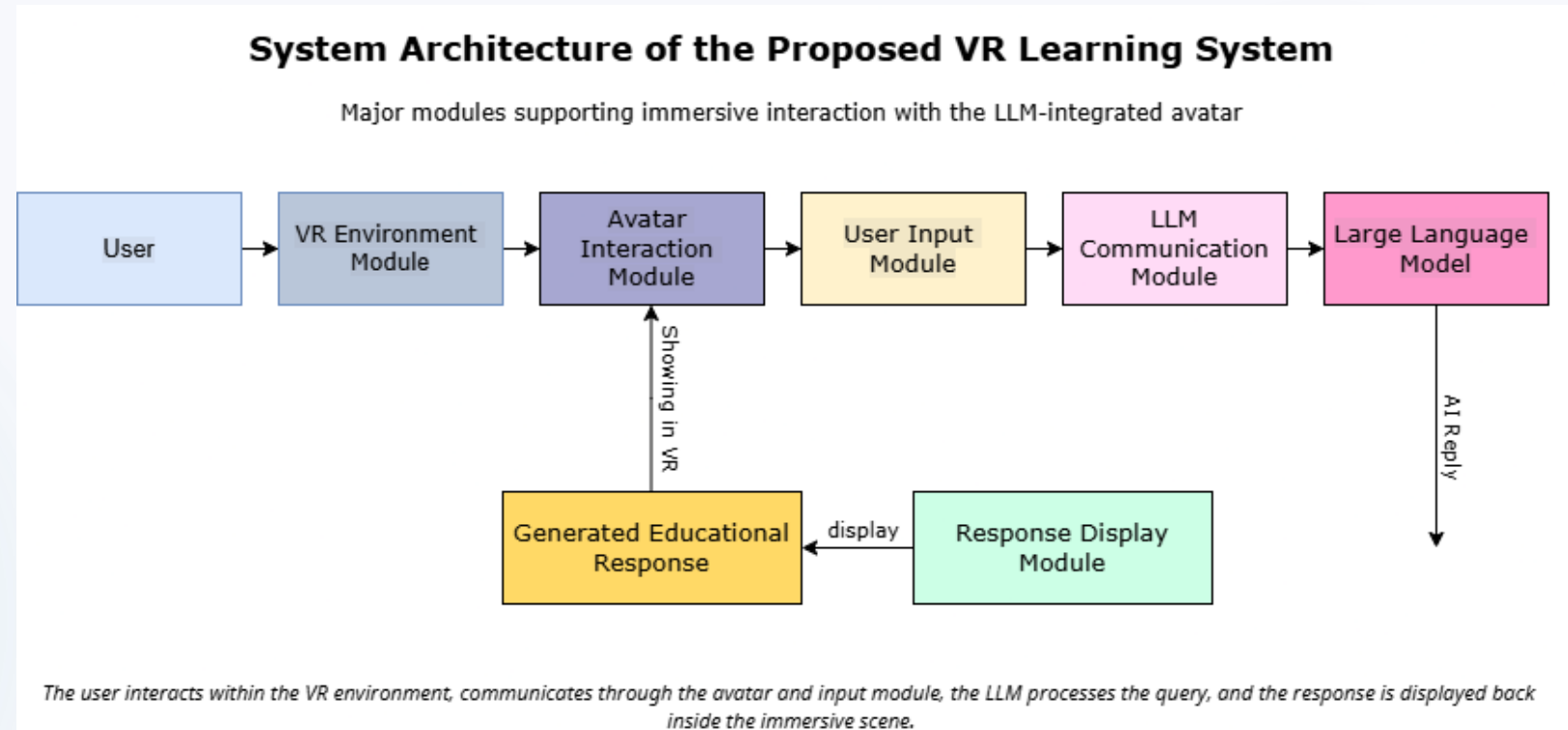
Unity captures the prompt and transmits it through the LLM communication module.

### Generate

The LLM processes the request and generates a response.

### Display

The answer is rendered inside the VR environment on the response panel.





# VR environment spotlight

A classroom-centered spatial layout designed for legibility, focus, and embodied interaction

Classroom context

Virtual keyboard

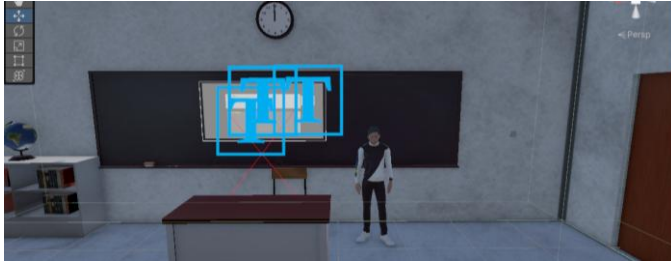
Avatar

Response panel

## Design rationale

- The classroom metaphor reduces orientation burden and supports a familiar educational context.
- The avatar serves as the visual and interactional anchor of the scene.
- Keyboard and response panel were placed within a comfortable forward-facing view.
- Readable typography and strong contrast were necessary because all output is consumed through a headset.

# Implemented components walkthrough



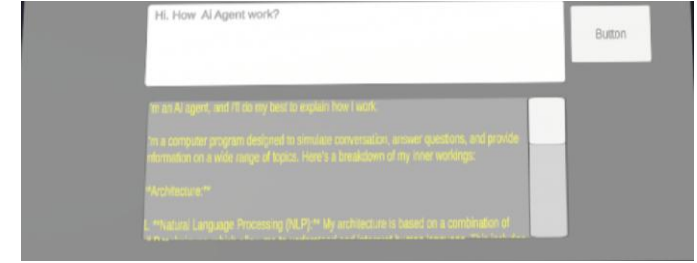
## VR classroom environment

Provides spatial context and a stable educational scene.



## Virtual keyboard input

Supports in-world text entry through controller-based selection.

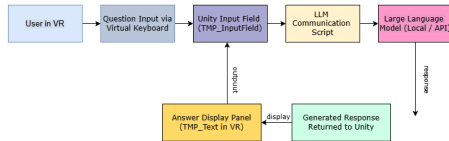


## Response display panel

Shows generated output in a readable in-environment format.

### LLM Communication and Response Generation Pipeline

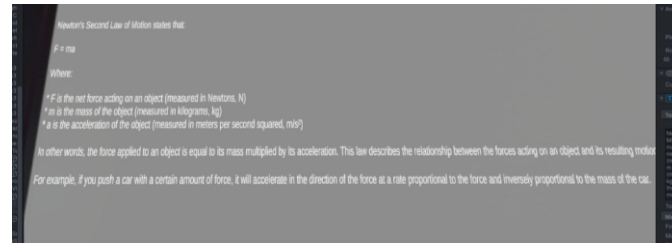
Flow of user questions from the VR interface to the language model and back to the Immersive environment



The user enters a question in VR, Unity forwards it through the communication script, the language model generates a response, and the answer is displayed back inside the immersive environment.

## Communication pipeline

Moves prompts from Unity to the model and returns generated text.



## End-to-end educational exchange

Question entry, model processing, and answer display occur in one loop.

All five core components were implemented and operated as one functional proof of concept. This is important because the thesis contribution depends on integration, not isolated modules.

# Technology Stack

## Unity Engine

Provided a real-time 3D platform for developing the immersive virtual classroom.

## XR Interaction Toolkit

Enabled controller-based interaction, raycasting, and intuitive VR navigation.

## C# Scripting

Implemented interaction logic and enabled asynchronous communication with the LLM.

## Ollama (Llama 3)

Delivered local, low-latency AI inference while ensuring data privacy.

## TextMeshPro & Unity UI

Ensured high-quality text rendering and readable in-world responses.

## Meta Quest 3

Supported standalone deployment and real-time immersive evaluation.

# Key Technical Challenges

## Real-Time LLM Integration in VR

- Challenge:** Establishing reliable communication between Unity and the LLM.
- Solution:** Implemented HTTP-based asynchronous requests using C# scripts.
- Outcome:** Enabled seamless, real-time AI assistance within the immersive environment.

## In-World Text Input and Response Display

- Challenge:** Allowing users to enter queries without leaving VR.
- Solution:** Developed a virtual keyboard using the XR Interaction Toolkit and TextMeshPro.
- Outcome:** Ensured intuitive interaction and improved usability.

## Performance and Latency Optimization

- Challenge:** Maintaining smooth performance during AI inference.
- Solution:** Deployed a local Llama 3 model via Ollama and **optimized response handling.**
- Outcome:** Achieved responsive and stable system performance.

*Prototype developed in Unity and evaluated using the Meta Quest 3.*

# Demo Video



IMMERSIVE VR  
LEARNING SYSTEM  
WITH LLM-  
INTEGRATED AVATAR

# Educational Case Studies in the VR Learning System

# CASE STUDY 1: MATHEMATICS – ACCELERATION

## TYPE OF QUESTION: CONCEPTUAL

▶ **Question:** What is acceleration?

▶ **AI Response:** Acceleration is the rate of change of velocity of an object with respect to time.

▶  $a = \Delta v / t$ .

▶ **Educational Value:** Demonstrates conceptual understanding and mathematical reasoning.

*Acceleration is the rate of change of velocity of an object with respect to time. It is a measure of*

*Mathematically, acceleration is represented by the symbol 'a' and is calculated as:*

$$a = \Delta v / \Delta t$$

*where:*

- \* a is the acceleration*
- \*  $\Delta v$  is the change in velocity (the final velocity minus the initial velocity)*
- \*  $\Delta t$  is the time over which the change occurs*

*...erates from 0 m/s to 10 m/s in 2 seconds, its acceleration would be:*

# CASE STUDY 2: PHYSICS – FREE FALL

## TYPE OF QUESTION: CONCEPTUAL

to is known as acceleration due to gravity, or  $g$ . It's a fundamental concept in physics that has been extensively experimentally verified. In a vacuum, where there is no air resistance, all objects fall at the same rate regardless of mass.

Universal gravitation, every object with mass attracts every other object with mass, and the force of attraction between them depends on the square of the distance between them. Under the sole influence of gravity, the relationship between mass and gravitational acceleration is much simpler.

rate regardless of mass:

is\*\*: Regardless of an object's mass, gravity always pulls with the same strength. This means that every object falls toward Earth at the same rate. In a vacuum or in situations where air resistance and friction are negligible, objects fall at the same rate because there is no net force opposing the force of gravity. The energy an object has (such as its potential energy due to its height) is converted into kinetic energy as it falls. Since all objects are converting their energy into kinetic energy equally, they all fall at the same rate.

gravitational field like on Earth, gravity warps spacetime around massive objects. This curvature causes objects with different masses to follow the same path.

comes from Sir Isaac Newton's law of universal gravitation and his subsequent development of the concept of acceleration.

▶ **Question:** Why do objects fall at the same rate regardless of mass?

▶ **AI Response:** The phenomenon you're referring to is acceleration due to gravity, or  $g$ .

**Educational Value:** Encourages scientific reasoning and conceptual learning.

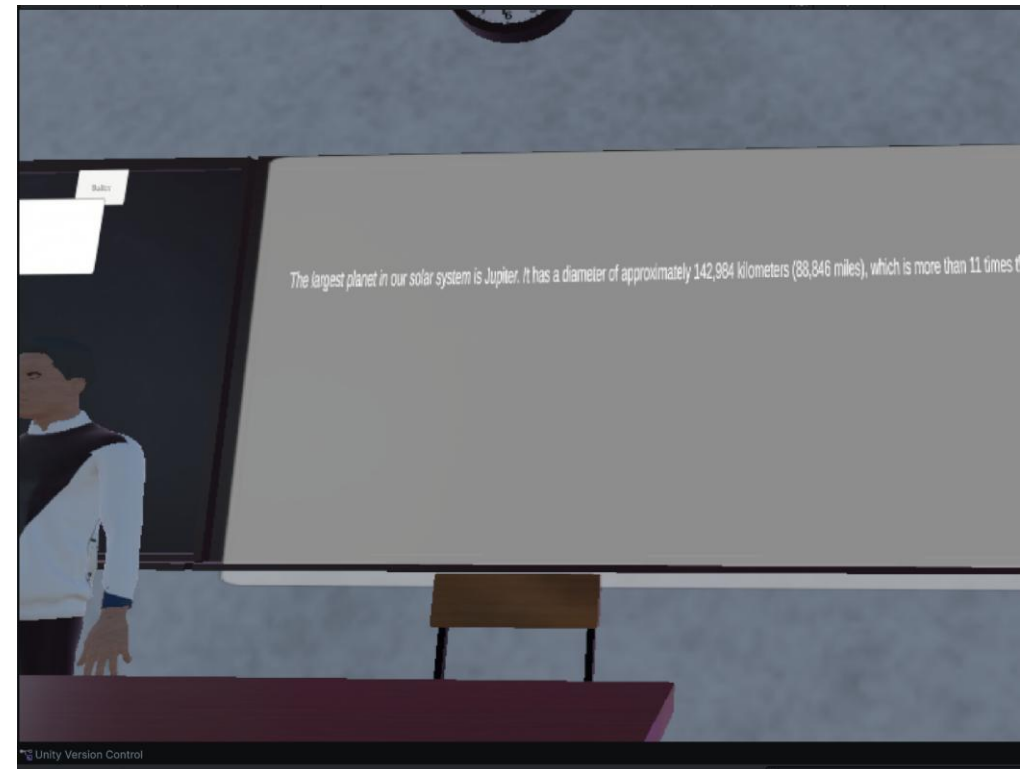
# CASE STUDY 3: ASTRONOMY – SOLAR SYSTEM

## TYPE OF QUESTION: FACTUAL

▶ **Question:** What is the largest planet in the solar system?

▶ **AI Response:** The largest planet in our solar system is Jupiter. It has a diameter of approximately 142,984 kilometers (88,846 miles), which is more than 11 times the diameter of the Earth.

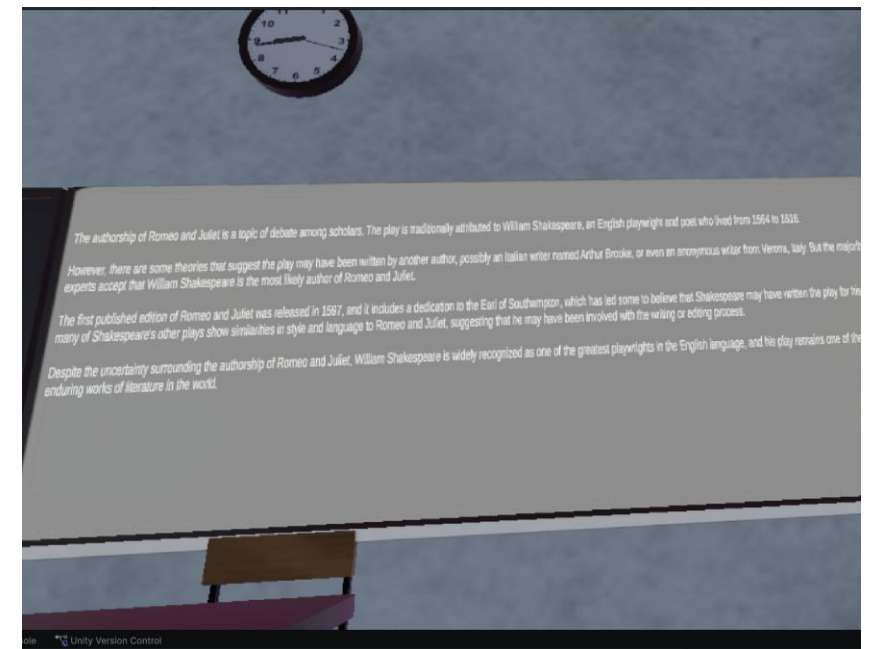
**Educational Value:** Supports factual recall and scientific literacy.



# CASE STUDY 4: LITERATURE – Romeo and Juliet

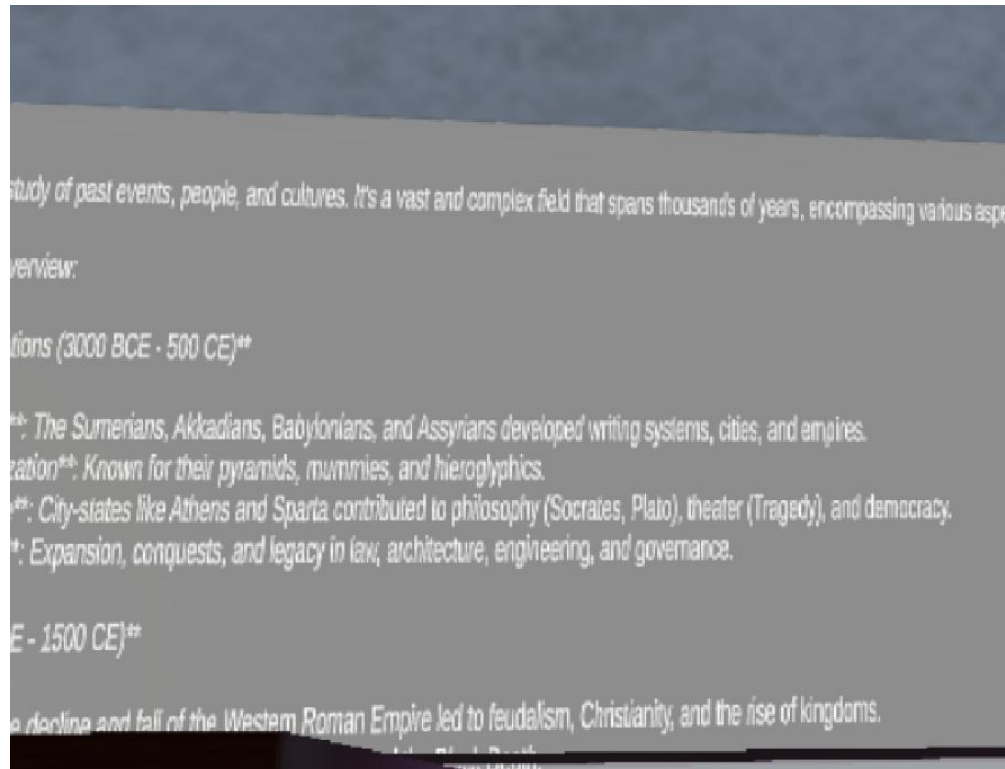
## TYPE OF QUESTION: FACTUAL

- ▶ **Question:** Who wrote Romeo and Juliet?
- ▶ **AI Response:** The authorship of *Romeo and Juliet* is a topic debate among scholars. The play is traditionally attributed to William Shakespeare, an English playwright and poet who lived from 1564 to 1616.
- ▶ **Educational Value:** Supports literary knowledge and factual recall in English literature.



# CASE STUDY 5: AMBIGUOUS QUERY – HISTORY

## TYPE OF QUESTION: AMBIGUOUS



- ▶ **Question:** Tell me about history?
- ▶ **AI Response:** History is the study of past events, people, and cultures. It is a vast and complex field that spans thousands of years, encompassing various aspects such as politics, social structures, economy, culture, science, and technology.
- ▶ **Educational Value:** Supports interpretation of ambiguous prompts and broad exploratory learning.

# Case Study – Questions

15 prompts spanning multiple cognitive styles were used to test correctness, relevance, clarity, and latency.

## **Factual**

- What is Newton's Second Law?
- What is the boiling point of water?
- What is photosynthesis?
- What is the formula for velocity?
- When did World War 2 end?
- Who wrote Romeo and Juliet?
- What is the capital of France?
- What is the chemical formula of carbon dioxide?
- What is the largest planet in the solar system?

## **Conceptual**

- What is the function of DNA?
- What is acceleration?
- Explain why objects fall at the same rate regardless of mass.

## **Open-ended**

- What are the main causes of World War 1?

## **Procedural**

- How do I solve a quadratic equation?

## **Ambiguous**

- Tell me about history?

# Evaluation Overview

Evaluation Method: Structured Case Study

Purpose: Assess performance and feasibility of the proposed VR learning system

System Evaluated: VR-based LLM-integrated educational assistant

Comparison Baseline: PC-based AI system

Focus: Accuracy, responsiveness, and educational effectiveness

# Evaluation Objectives

- ▶ Assess the accuracy and relevance of AI-generated responses
- ▶ Evaluate clarity and educational usefulness of system outputs
- ▶ Measure response latency in the immersive VR environment
- ▶ Compare performance with a PC-based AI baseline
- ▶ Validate the feasibility of integrating VR and Large Language Models
- ▶ Examine the effectiveness of avatar-mediated interaction

# Case Study Methodology

- ▶ Method: Structured Case Study Approach
- ▶ Total Questions: 15 educational queries
- ▶ Categories: Factual, Conceptual, Open-ended, Procedural, and Ambiguous
- ▶ Evaluation Strategy: Question–response analysis
- ▶ Comparative analysis with a PC-based AI baseline
- ▶ Suitable for prototype-stage system validation

# Evaluation Metrics



Correctness:  
Measures factual  
accuracy of  
responses



Relevance: Evaluates  
alignment with user  
queries




Clarity: Assesses  
readability and  
coherence of  
responses



Response Latency:  
Measures time taken  
to generate answers



Ensures a  
comprehensive and  
objective system  
evaluation



# Case Study Evaluation Framework

- ▶ Correctness: 1 = Correct, 0.5 = Partially Correct, 0 = Incorrect
- ▶ Relevance: Rated as High, Medium, or Low
- ▶ Clarity: Rated as High, Medium, or Low
- ▶ Response Latency (VR): Time recorded in seconds
- ▶ Response Latency (PC): Time recorded in seconds
- ▶ Provides standardized and consistent performance assessment

# Experimental Setup



Development  
Platform: Unity



VR  
Framework:  
XR Interaction  
Toolkit



Programming  
Language:  
C#



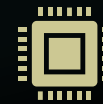
Conversation  
AI: LLaMA 3  
via Ollama



User Interface:  
TextMeshPro  
and Unity UI



Hardware:  
Meta Quest 3



Baseline  
System: PC-  
based  
ChatGPT

# Summary of Evaluation Design

Methodology:  
Structured Case  
Study

Sample Size: 15  
Educational  
Questions

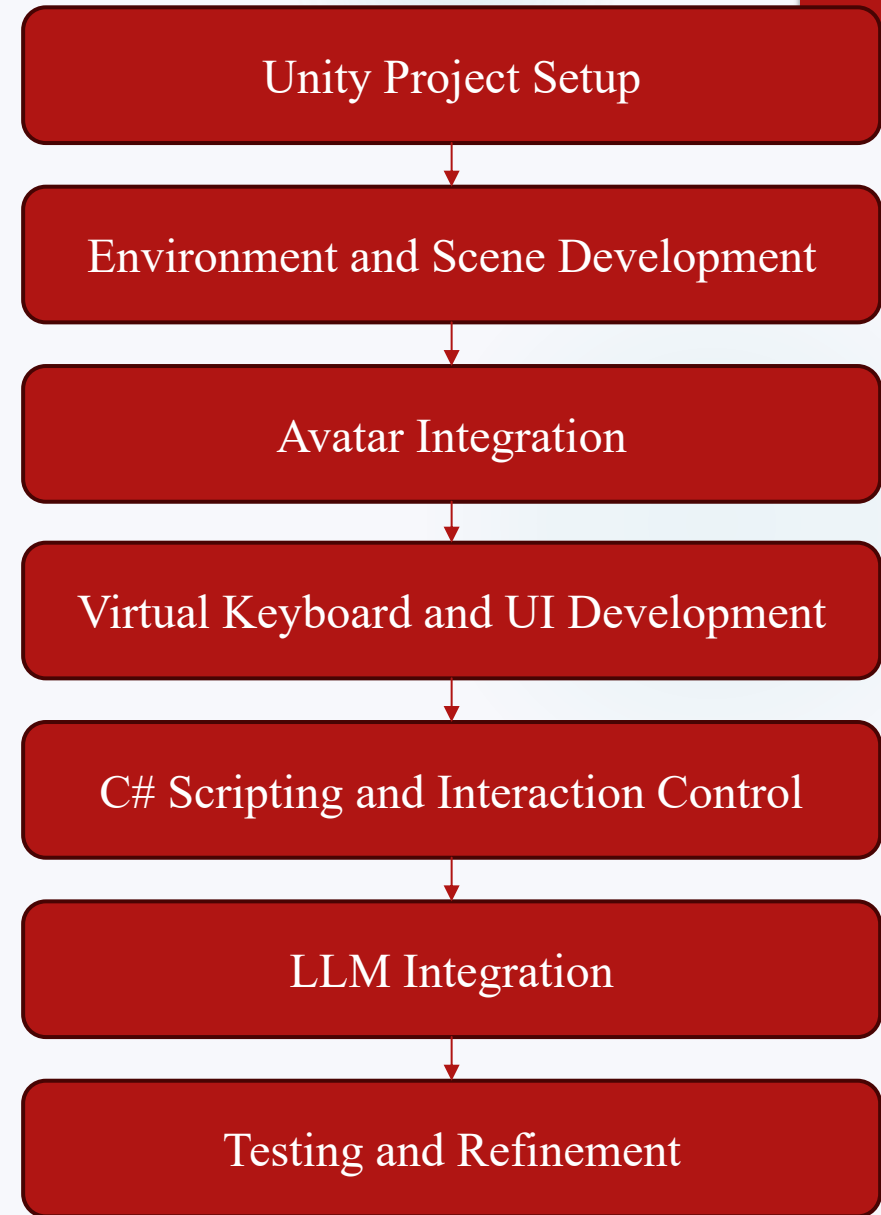
Evaluation Metrics:  
Correctness,  
Relevance, Clarity,  
and Latency

Comparison: VR  
System vs. PC-  
Based AI

Outcome:  
Validation of an  
immersive AI-  
assisted learning  
environment

Demonstrates the  
feasibility of VR-  
LLM integration in  
education

# Implementation Process



# Implementation results

5

Implemented components

15

Case-study prompts

1.00

Average correctness

15 / 15

High relevance cases

8.37 s

Average VR latency

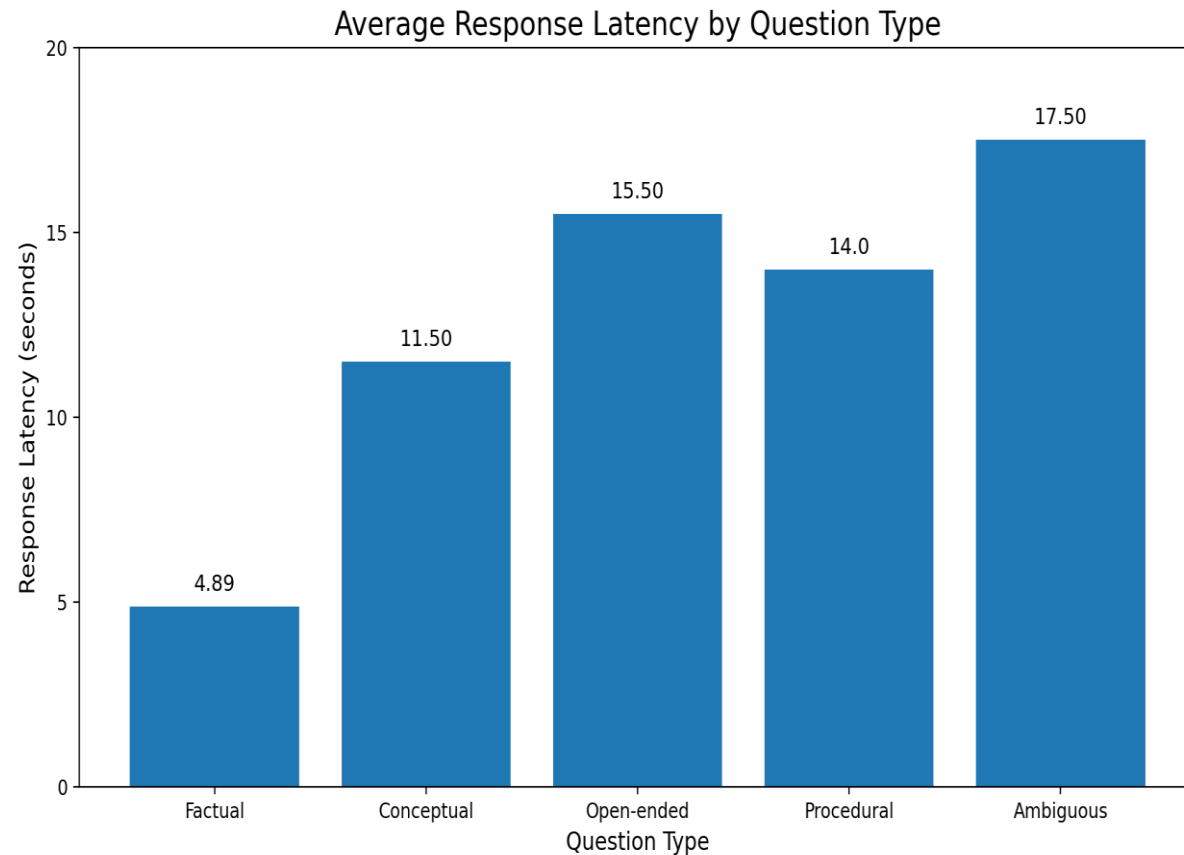
## What worked technically

- The classroom scene provided a stable immersive context for interaction.
- The avatar remained the focal point of the educational exchange.
- The virtual keyboard supported full in-environment question entry.
- The communication module successfully connected Unity to a local LLM.
- The response panel rendered readable output inside the same VR scene.

## Interpretation

The prototype is not merely a visual mockup. It demonstrates that VR interaction, embodied avatar presence, prompt transmission, local LLM processing, and in-scene response display can operate as one integrated educational workflow.

# Case study results: response quality and latency



## Main findings

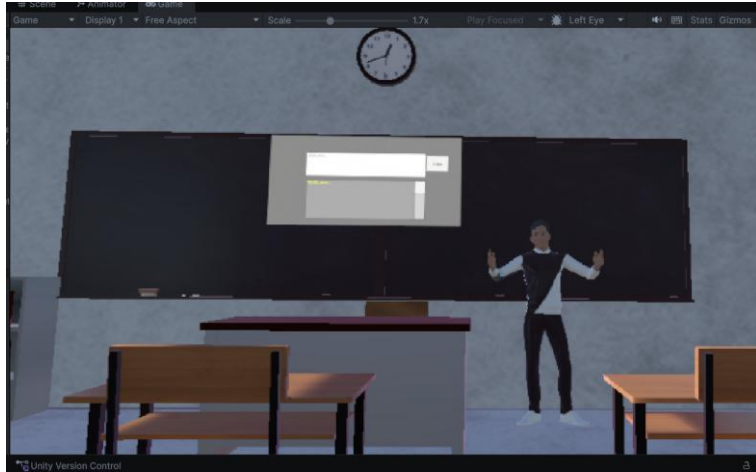
- All 15 VR responses were rated correct, highly relevant, and clear within this case study.
- Factual prompts were answered most quickly.
- Conceptual prompts required more time because the system generated explanatory answers.
- Open-ended, procedural, and ambiguous prompts produced the highest latency values.

## What this means

The prototype preserved answer quality across the sample, but generation time and interaction cost increased as prompts became more cognitively demanding.

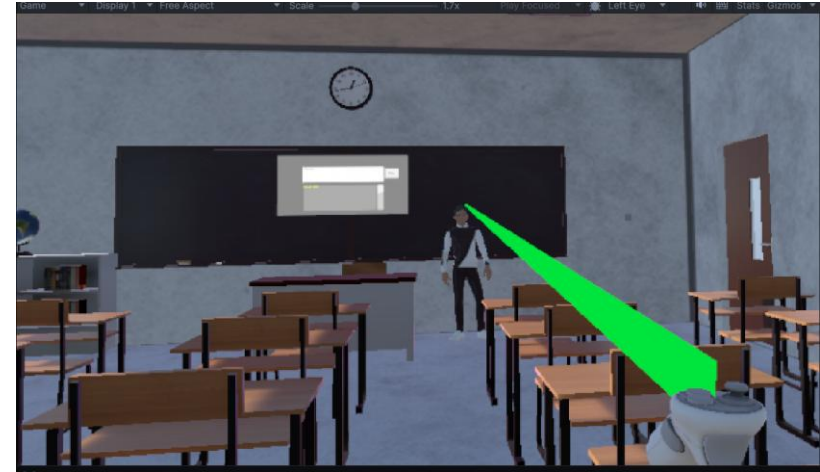
# Representative response examples

## Example 1: conceptual explanation



- The response was not just a short definition; it included formula structure and explanation.
- This matters because conceptual usefulness is more relevant educationally than simple fact recall.

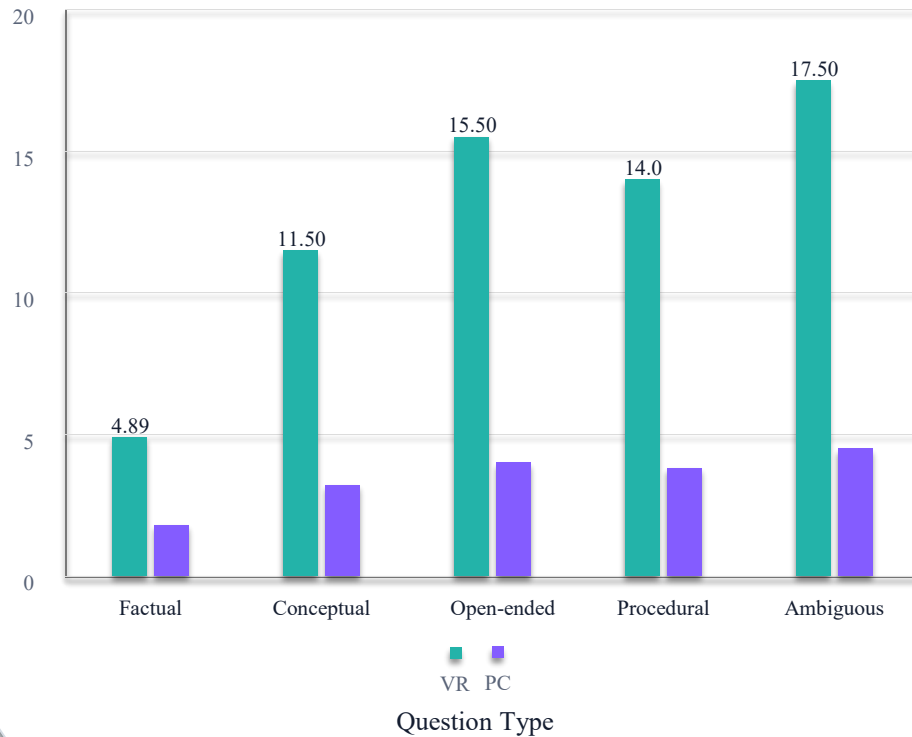
## Example 2: longer, more complex response



- The longer latency reflects increased explanation depth and rendering delay.
- This example shows why VR interaction should be judged as a trade-off between immersion and efficiency.

# VR system vs. PC-based AI baseline

Average Response Latency by Question Type



## How to present this fairly

- The same 15 prompts were used in both conditions.
- Both conditions achieved an average correctness score of 1.00 on this bounded test set.
- The PC baseline responded faster: 2.20 s average versus 8.37 s in VR.
- The PC condition used ChatGPT through a standard desktop interface.
- The VR condition benchmark. used Ollama/Llama 3 integrated through Unity, including controller input, communication overhead, and in-environment rendering.
- Therefore, this comparison should be understood as a practical baseline comparison, not a same-model
- The main trade-off is speed versus immersive, embodied educational interaction.

1.00

VR correctness

1.00

PC correctness

2.20 s

PC avg.

Defense message: PC wins on speed; VR adds presence, context, and embodied interaction.

# Discussion and main contributions

## Technical contribution

Implemented a five-component VR learning architecture consisting of the virtual environment, avatar interaction interface, virtual keyboard input system, LLM communication module, and response display module. Built a Unity-to-Ollama communication pipeline in C# using asynchronous HTTP requests so prompts entered in VR could be processed by a local Llama 3 model. Displayed AI-generated responses inside the same immersive scene through TextMeshPro, keeping the full question-response process within VR.

## Interaction contribution

Designed the avatar as the main point of educational interaction rather than as a decorative character. Implemented an in-world virtual keyboard with XR controller raycasting so users could enter questions without leaving the immersive environment. Organized the workflow so that question entry, model processing, and response review all occurred within one continuous interaction space.

## Educational contribution

Demonstrated that open-ended educational assistance can be embedded inside a VR classroom rather than delivered only through screen-based AI tools. Showed that the system could support factual, conceptual, procedural, open-ended, and ambiguous questions within a single prototype workflow. Established a practical baseline comparison with PC-based AI, showing preserved answer quality but higher latency in the immersive condition.

# Limitations and future work

## Current limitations

- The main limitations of this thesis are response latency in the local VR setup, less natural text input through the virtual keyboard, a limited case-study-based evaluation, and the fact that the prototype is not yet a full adaptive intelligent tutoring system.

## Future work priorities

- Replace controller-based text entry with voice interaction.
- Use streaming responses to reduce perceived delay.
- Improve grounding through prompt engineering or retrieval augmentation.
- Expand comparative testing and eventually move toward controlled user studies.

## Most impactful near-term improvement

Voice input would reduce friction, lower cognitive load, and make the educational interaction feel substantially more natural in VR.

# Reference

- Radianti, J., T. A. Majchrzak, J. Fromm, and I. Wohlgenannt. 2020. “A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda.” *Computers & Education* 147: 103778.
- Makransky, G., and G. B. Petersen. 2021. “The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality.” *Educational Psychology Review* 33 (3): 937–958.
- Kasneci, E., K. Sessler, S. Küchemann, M. Bannert, D. Dementieva, F. Fischer, U. Gasser, et al. 2023. “ChatGPT for good? On opportunities and challenges of large language models for education.” *Learning and Individual Differences* 103: 102274.
- Schroeder, N. L., O. O. Adesope, and R. B. Gilbert. 2013. “How effective are pedagogical agents for learning? A meta-analytic review.” *Journal of Educational Computing Research* 49 (1): 1–39.

# Conclusion

**Immersive VR + LLM assistance is technically feasible and educationally promising—especially when the goal is to keep the learner inside an interactive, embodied learning environment.**

**Thank you  
Questions?**