香港教育大學
The Education University of Hong Kong

法政大学
HOSEI University

Going Beyond Physical Classrooms: Experiencing Immersive and Interactive Learning in the Metaverse with a Digital Human Teacher

Professor Yanjie SONG

Department of Mathematics and Information Technology

The Education University of Hong Kong

25 Feb. 2025, Hosei University,Japan

Department of Mathematics and Information Technology

LEARNING VERSE

Outline

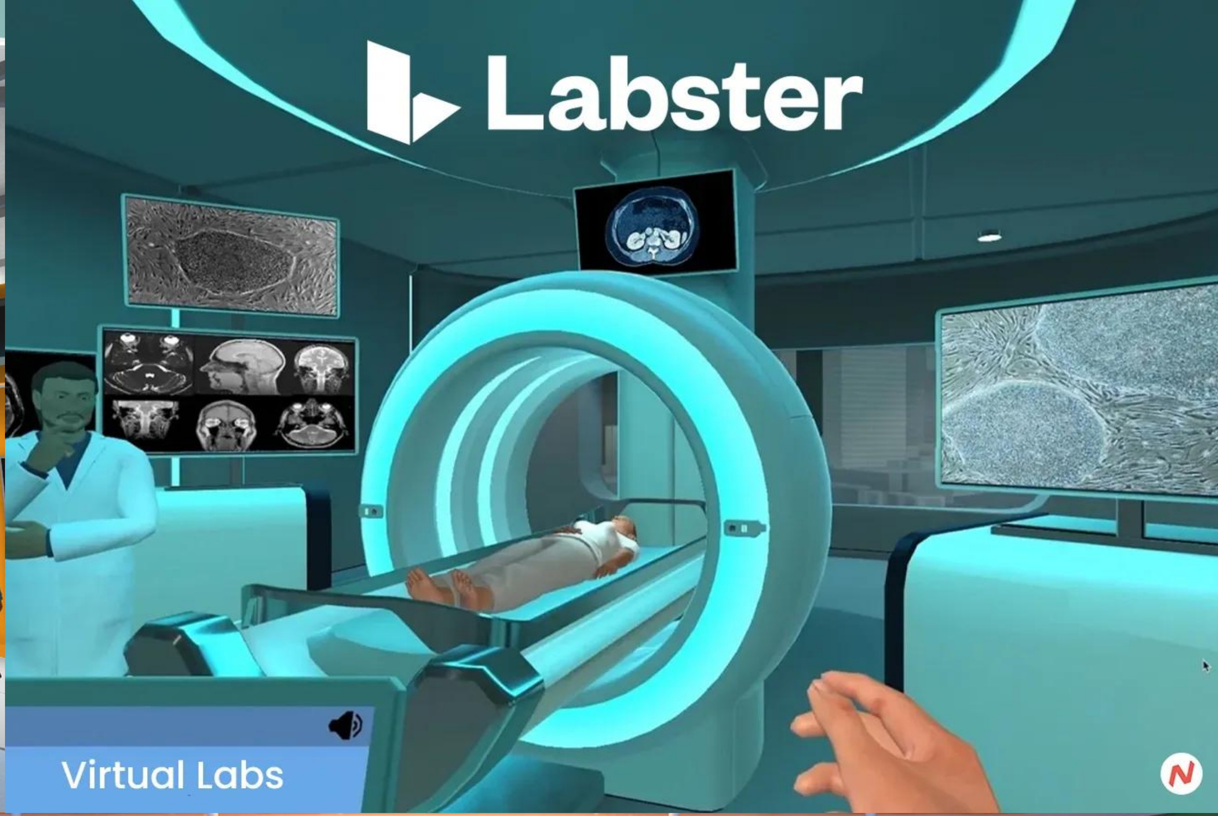
1. Introduction
2. The metaverse platforms and digital human
3. Demos of learning scenarios
4. Future work

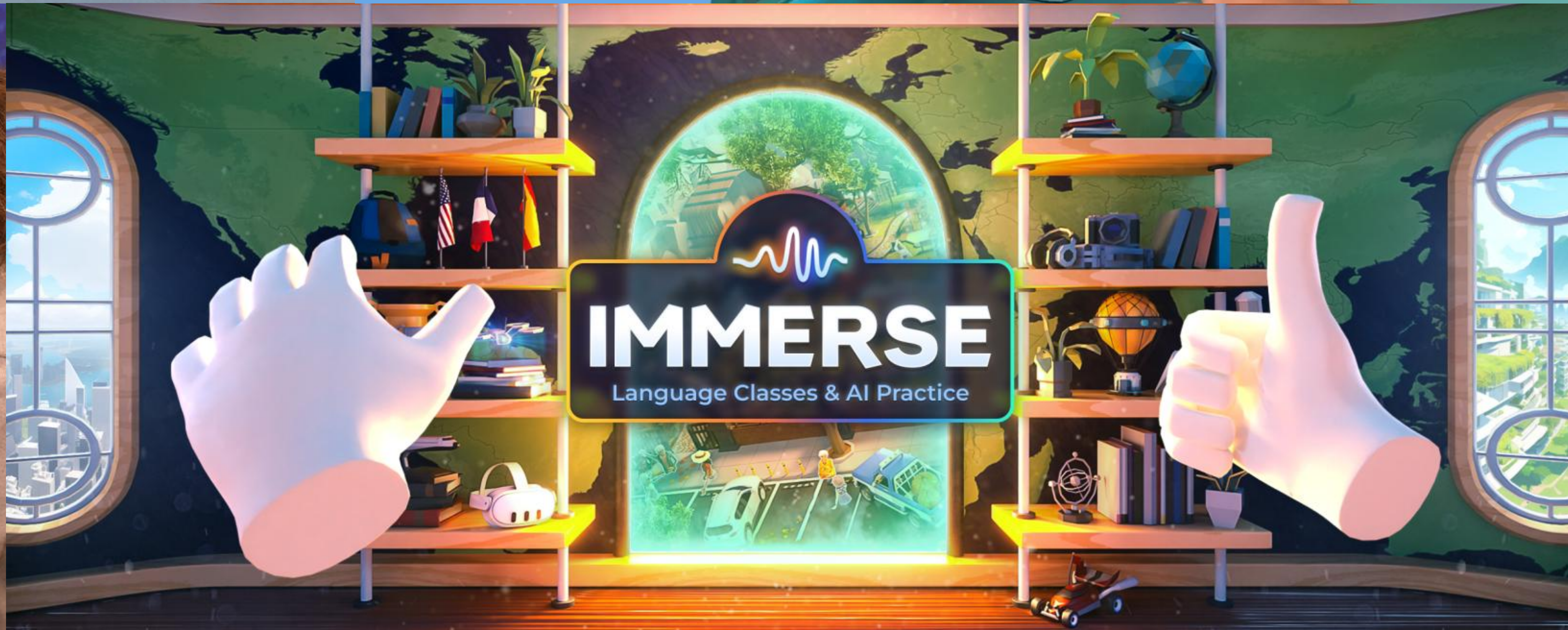
1. Introduction

1.1 Metaverse

Metaverses

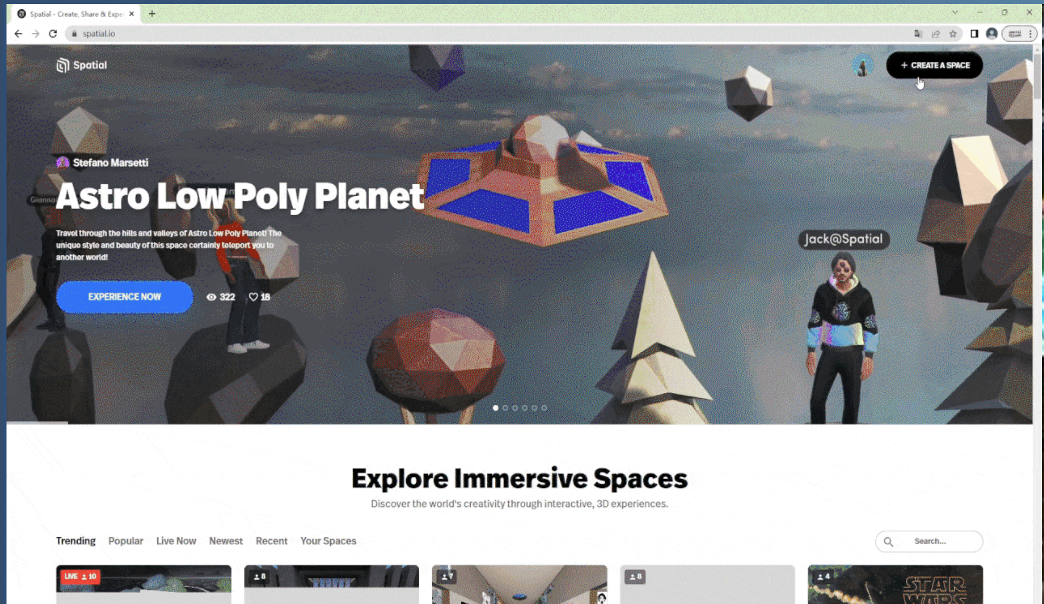
Evolvement of metaverses with generative AI

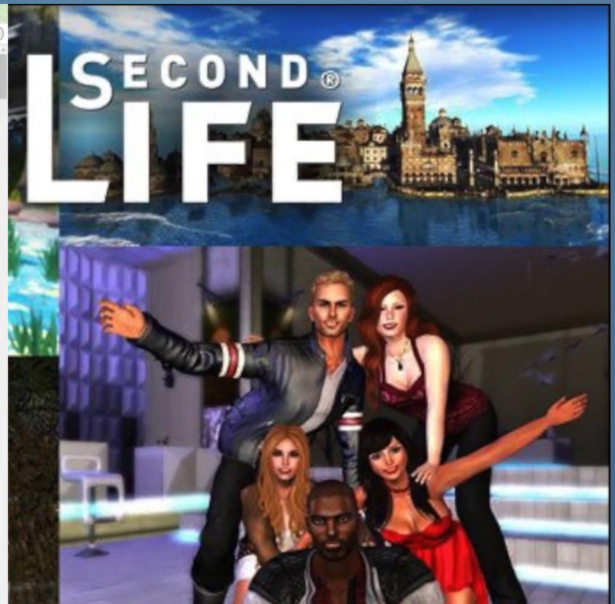




Issue 1: Excessive gamification

These platforms overemphasize on gaming which may distract students, affecting their learning outcomes.

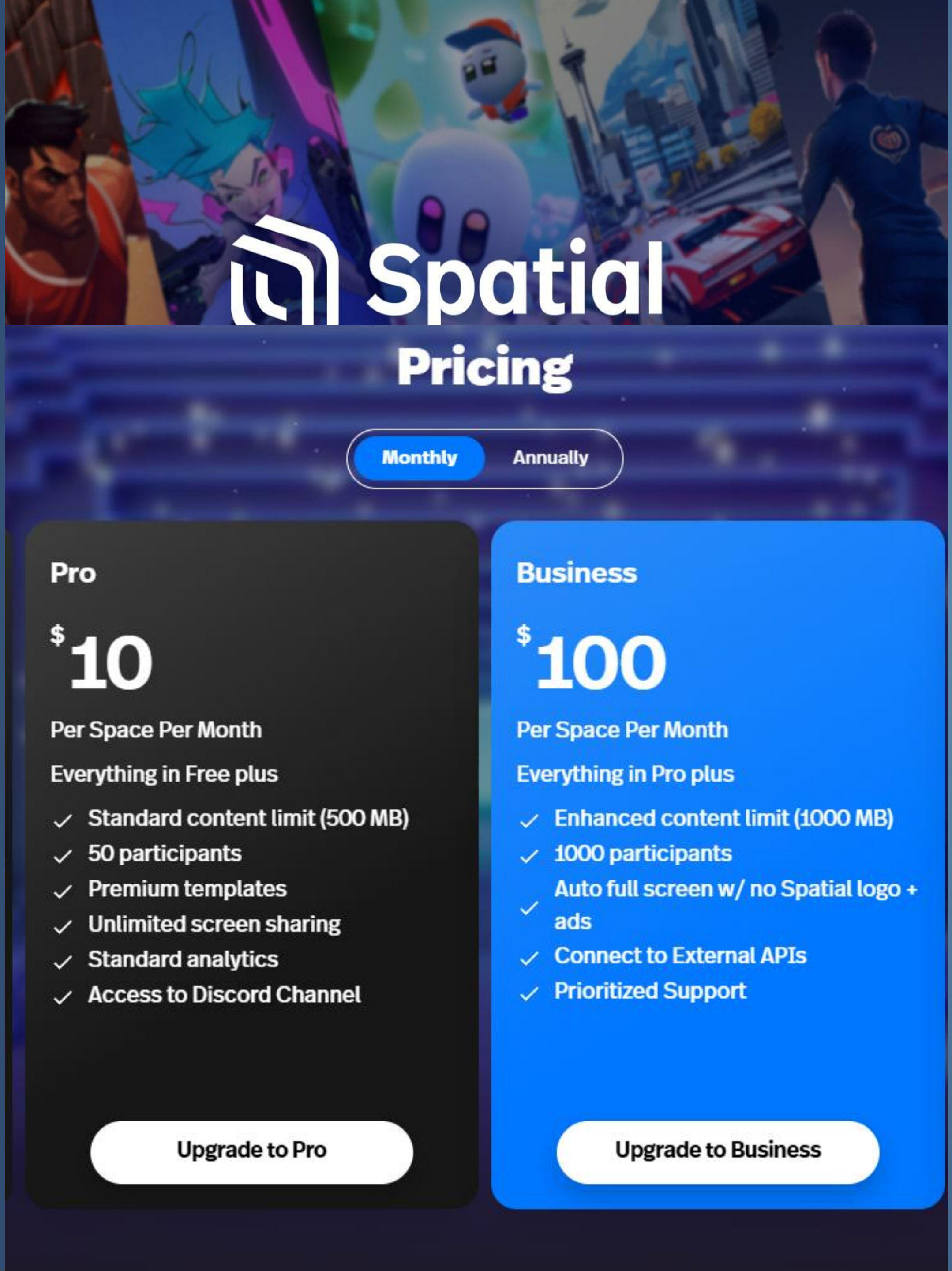




Issue 2:

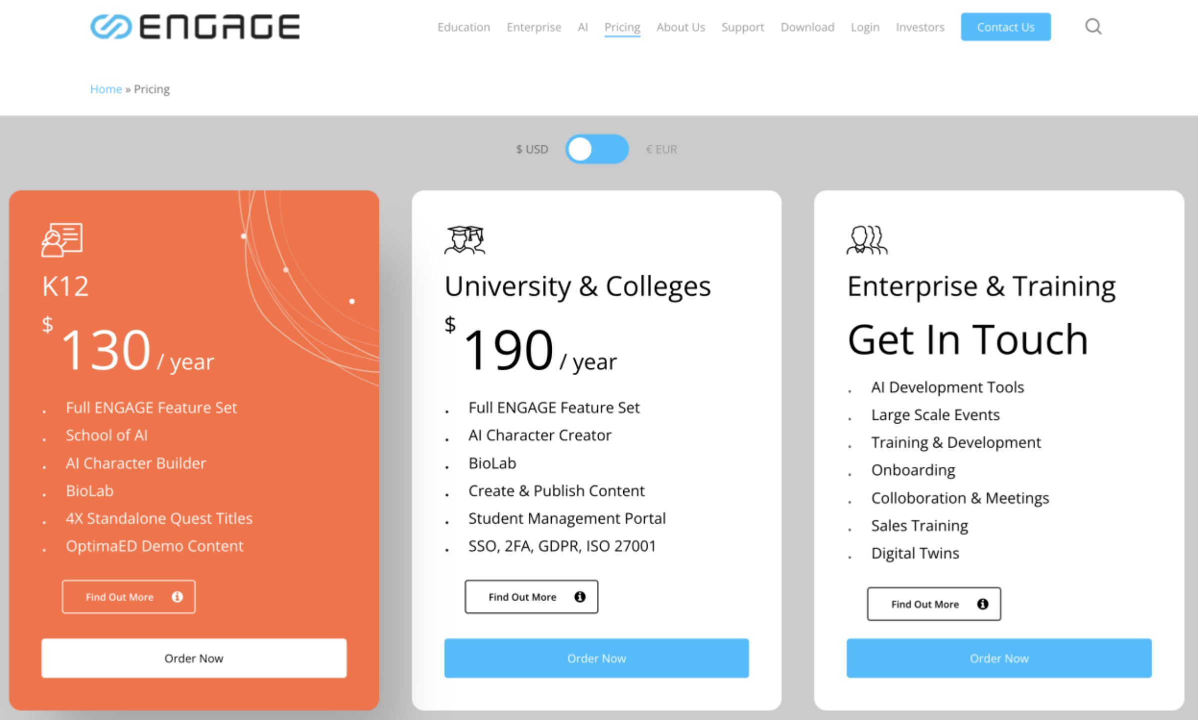
Lack of instructional design

These platforms' lack of instructional design complicates course implementation and management for teachers, and can lead to students losing track of their learning goals; and the AI technology adopted is very basic.



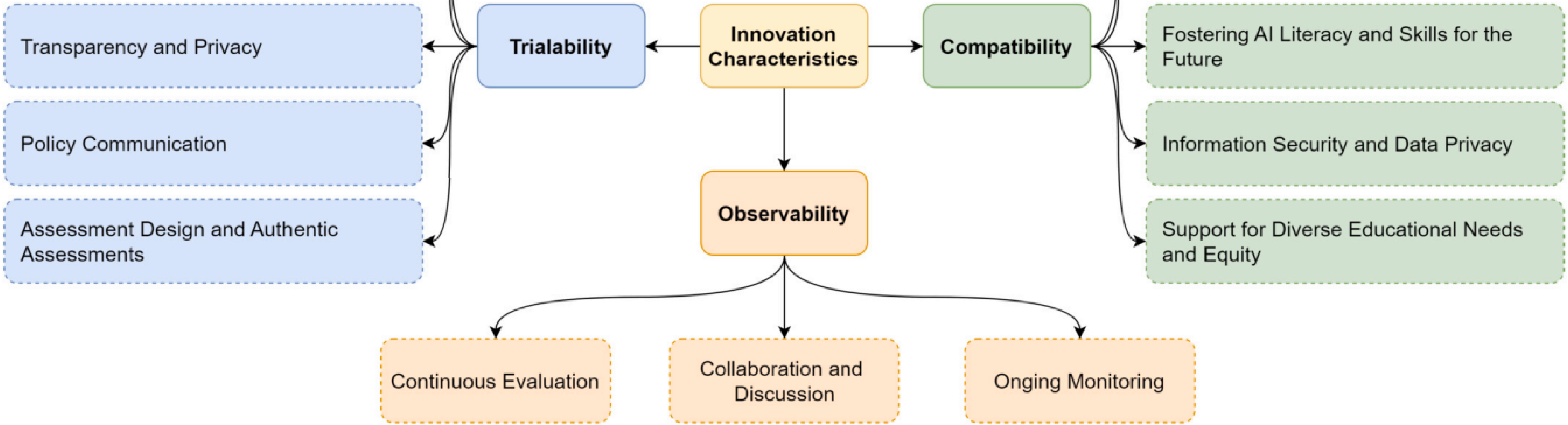
Issue 3: Expensive fees

Popular metaverse platforms now charge fees and restrict data collection, hindering student usage data collection for learning analytics.



1.2 Generative AI

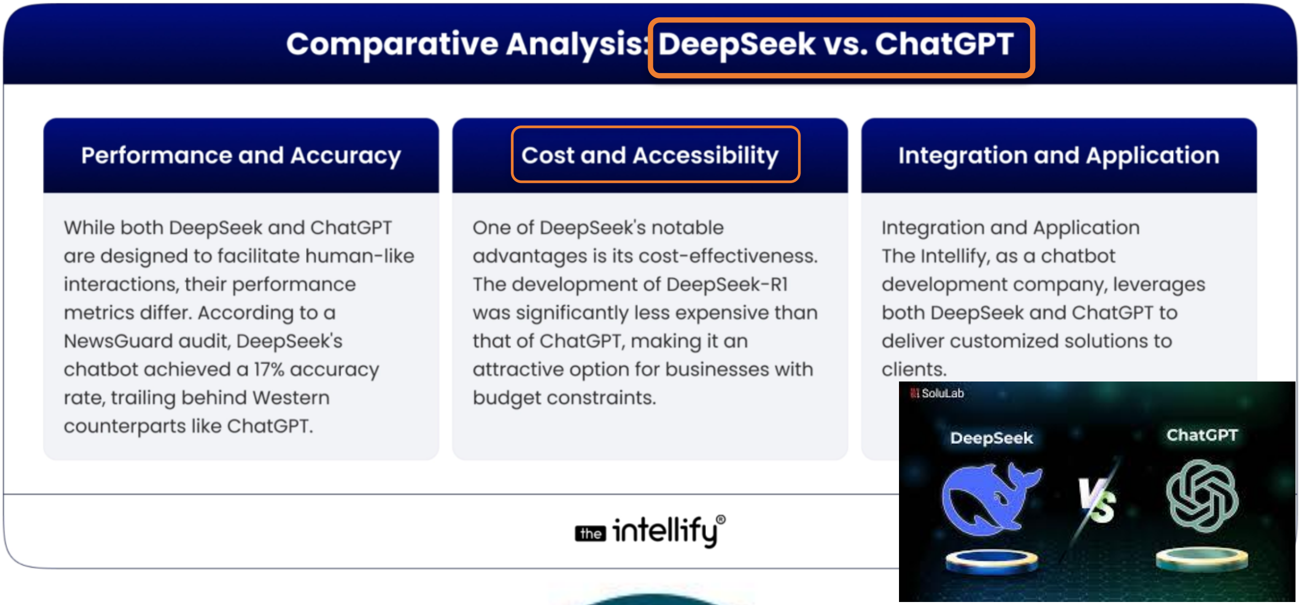
Generative AI refers to a category of AI algorithms that generate new outputs based on the data they have been trained on. Generative AI creates new content in the form of images, text, audio, and more.



Key themes related the compatibility, trialability, observability of generative AI integration which emerged from the analysed universities' policies and guidelines (Jin et al., 2025)

Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence*, 8, 100348

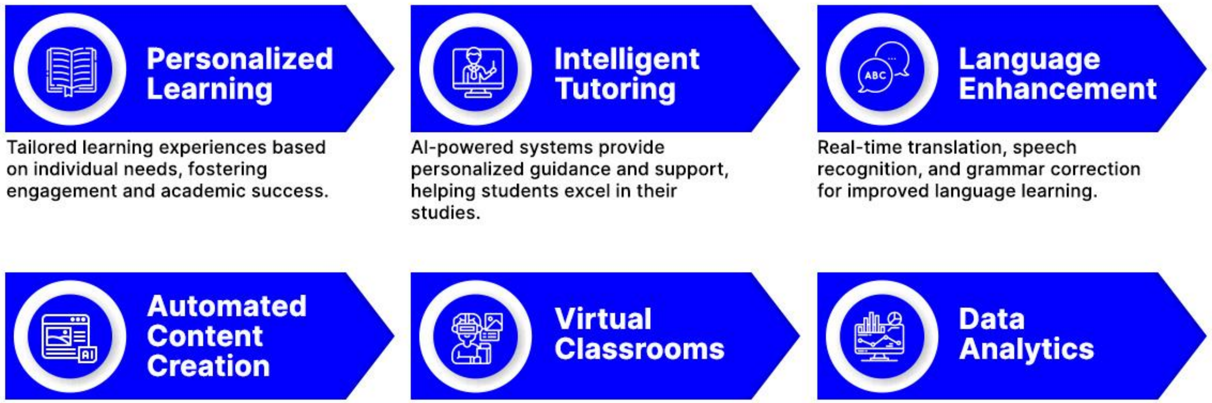
Comparative Analysis: DeepSeek vs. ChatGPT

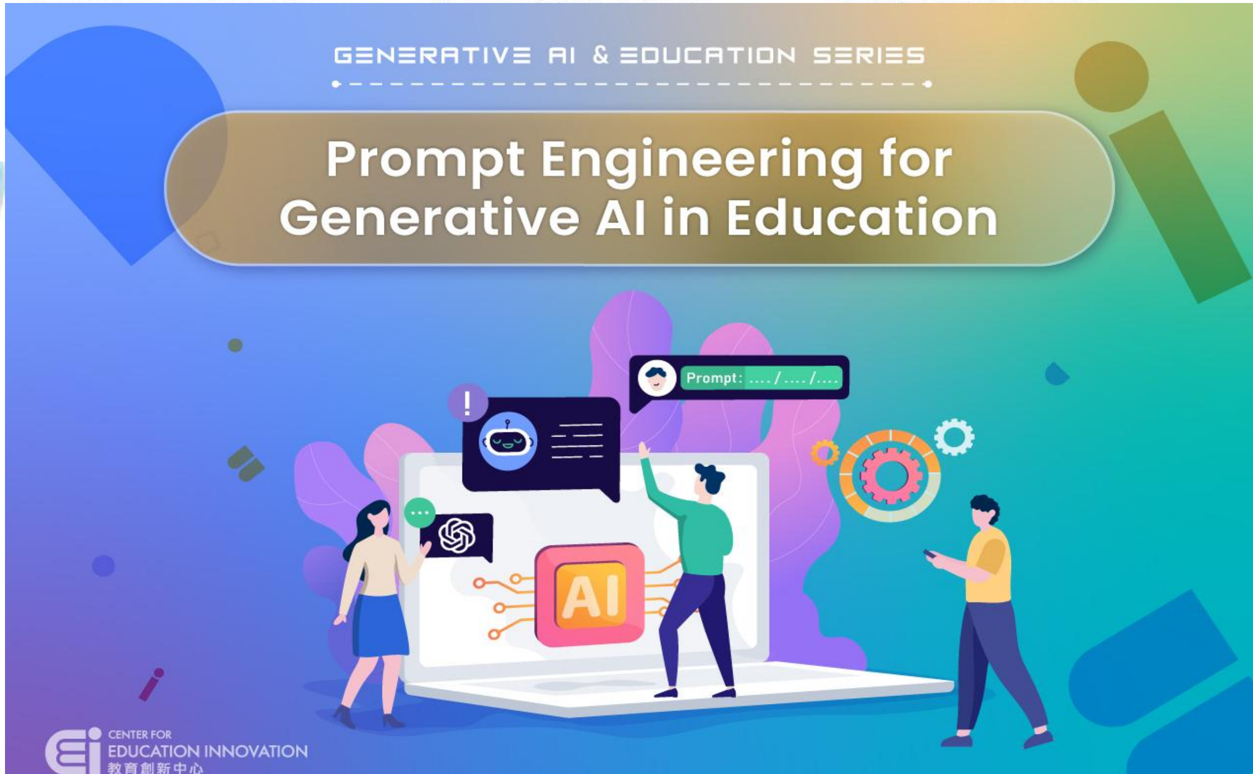


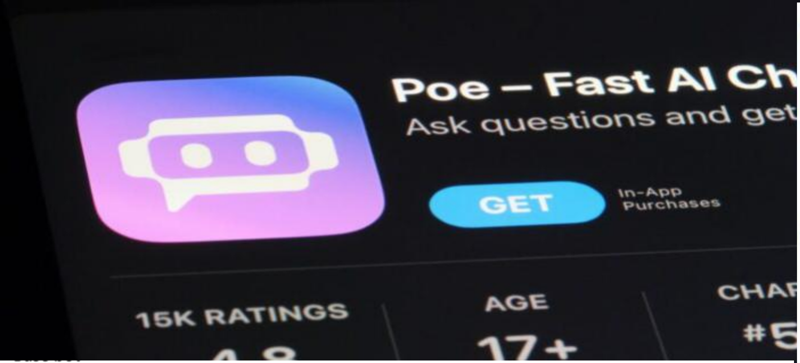
webclues

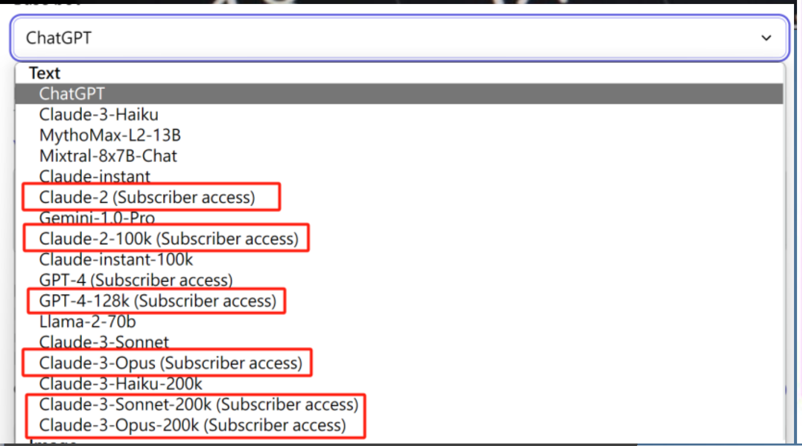
APPLICATIONS OF GENERATIVE AI IN EDUCATION

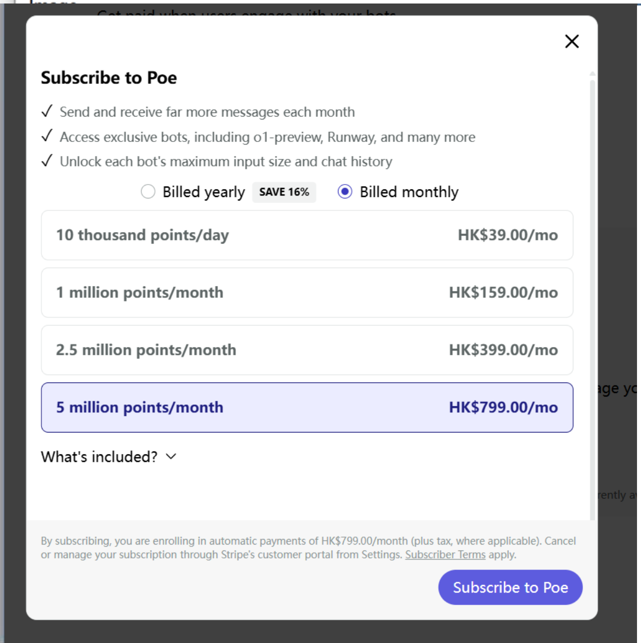
https://www.webcluesinfotech.com/the-role-of-generative-ai-in-shaping-the-future-of-education/











Issue 1:

Expensive fees

Issue 2:

Focus on AI tool applications instead of customising them for instructional design and implementation

Issue 3:

Personal privacy issues

Issue 4:

Fairness and equity

Our solution

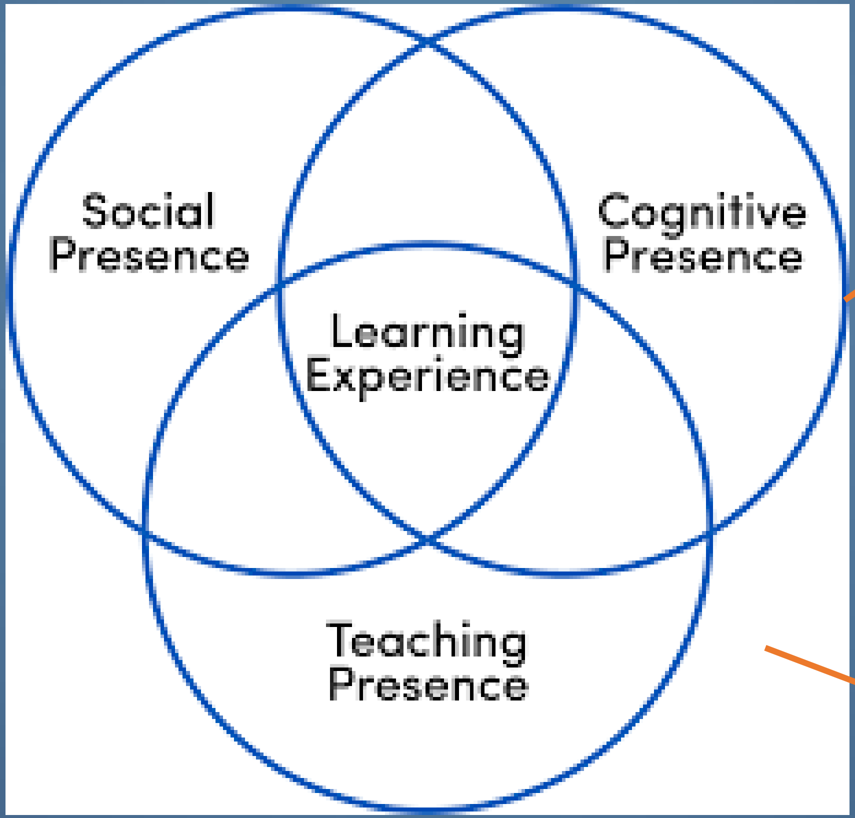
To integrate generative AI into the metaverse

2. The metaverse platforms - Learningverse and LearngverseVR and digital human

2.1 Learningverse

- Learningverse is a 3D metaverse platform, integrating a range of AI technologies for learners to conduct immersive and interactive learning activities in online or blended contexts.

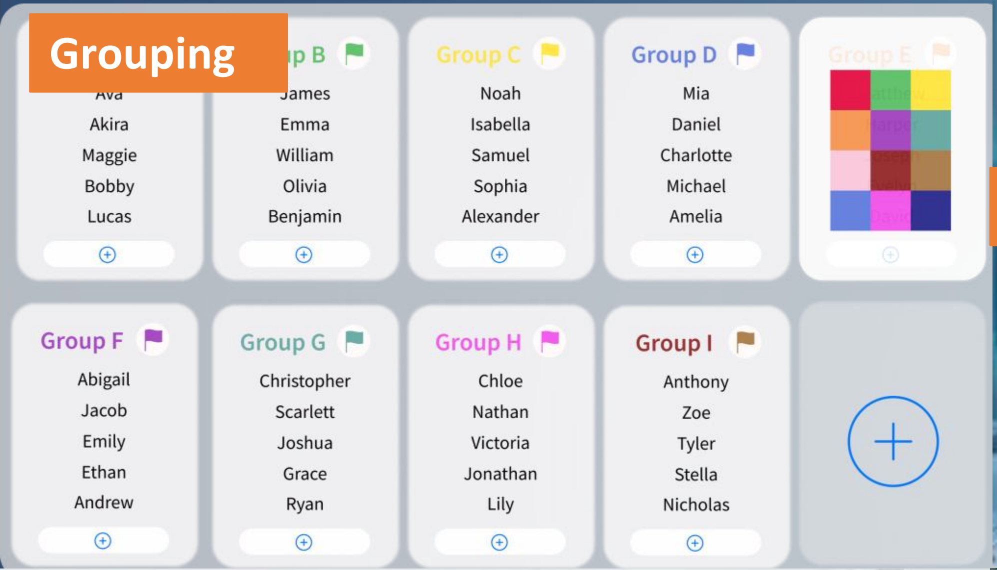
Community of inquiry as a theoretical underpinning for setting up the platform

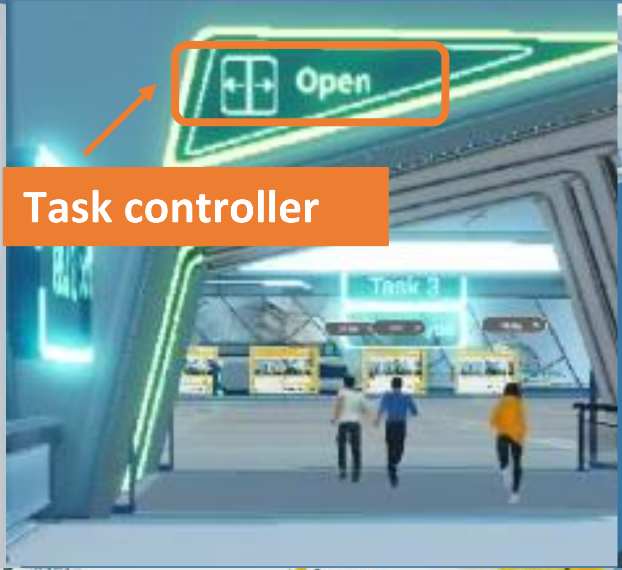


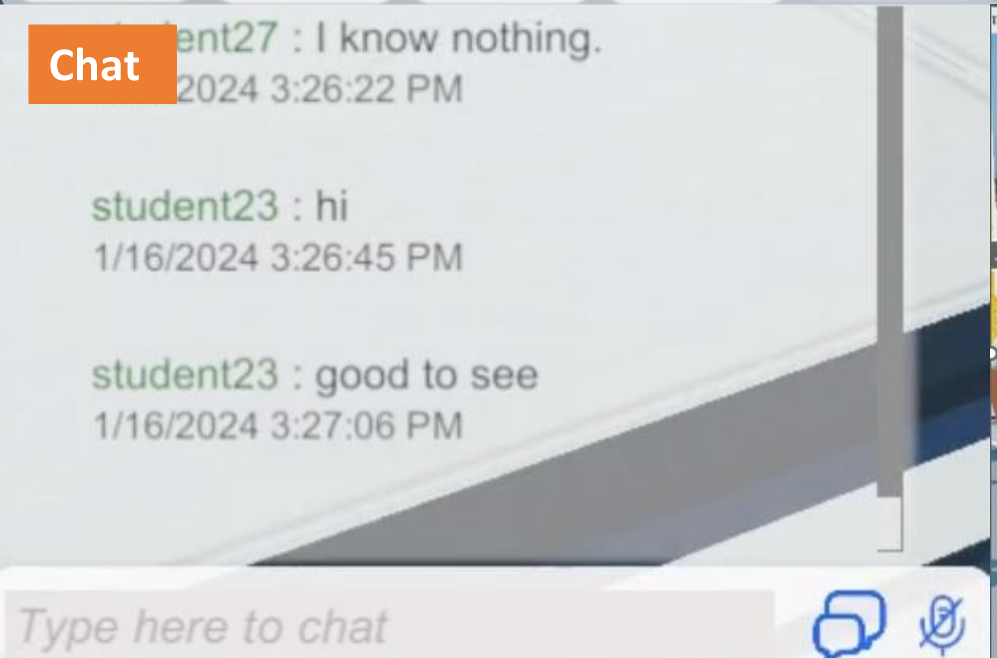
(Garrison et al.,2010)

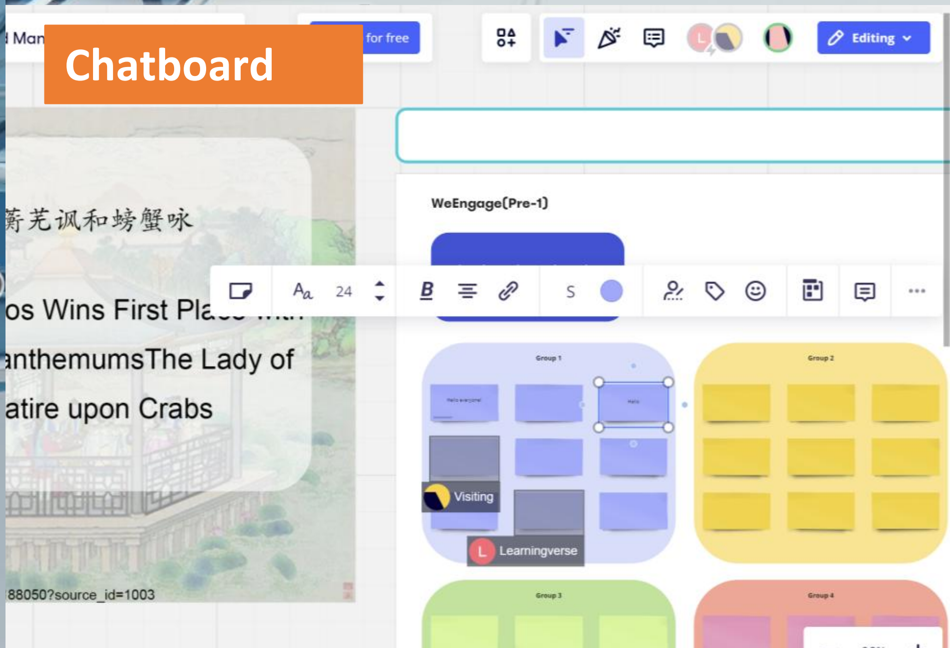
Song, Y., Cao, J., Wu, K., Yu, P. L. H., & Lee, J. C. K. (2023). Developing "Learningverse" - A 3-D Metaverse Platform to Support Teaching, Social, and Cognitive Presences. *IEEE Transactions on Learning Technologies*, 16(6), 1165-1178. https://doi.org/10.1109/TLT.2023.3276574

Specialised features for educational use

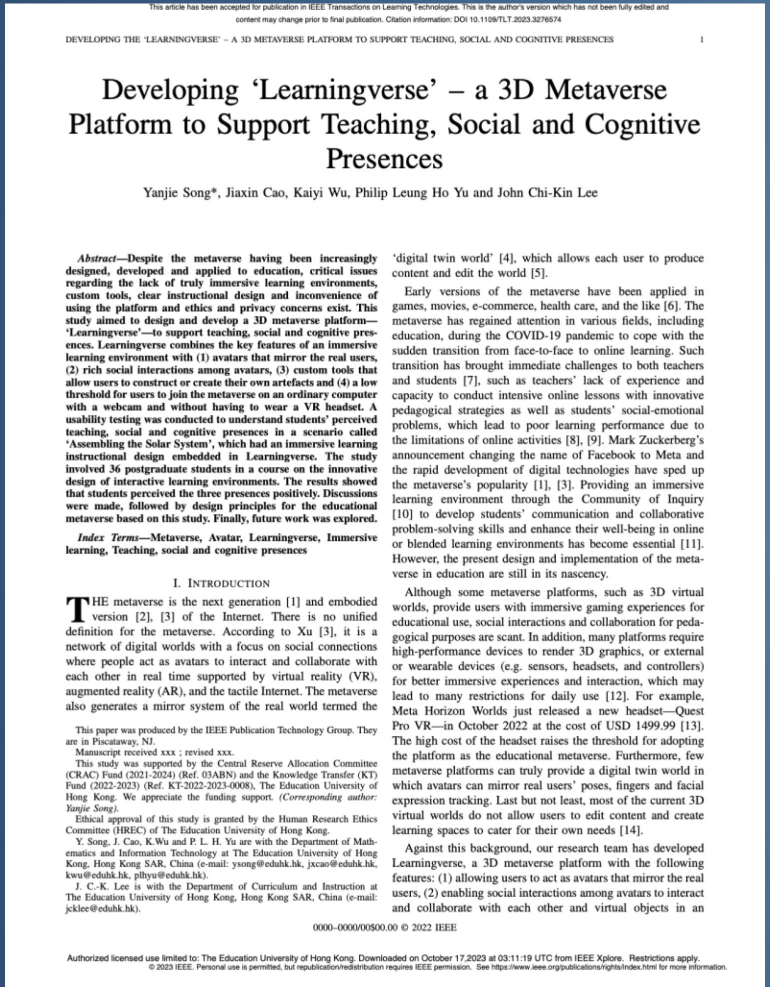








Publication and awards


This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2023.3276574

Developing 'Learningverse' – a 3D Metaverse Platform to Support Teaching, Social and Cognitive Presences

Yanjie Song*, Jiaxin Cao, Kaiyi Wu, Philip Leung Ho Yu and John Chi-Kin Lee

***Abstract*—Despite the metaverse having been increasingly designed, developed and applied to education, critical issues regarding the lack of truly immersive learning environments, custom tools, clear instructional design and inconvenience of using the platform and ethics and privacy concerns exist. This study aimed to design and develop a 3D metaverse platform—'Learningverse'—to support teaching, social and cognitive presences. Learningverse combines the key features of an immersive learning environment with (1) avatars that mirror the real users, (2) rich social interactions among avatars, (3) custom tools that allow users to construct or create their own artefacts and (4) a low threshold for users to join the metaverse on an ordinary computer with a webcam and without having to wear a VR headset. A usability testing was conducted to understand students' perceived teaching, social and cognitive presences in a scenario called 'Assembling the Solar System', which had an immersive learning instructional design embedded in Learningverse. The study involved 36 postgraduate students in a course on the innovative design of interactive learning environments. The results showed that students perceived the three presences positively. Discussions were made, followed by design principles for the educational metaverse based on this study. Finally, future work was explored.**

***Index Terms*—Metaverse, Avatar, Learningverse, Immersive learning, Teaching, social and cognitive presences**

I. Introduction

THE metaverse is the next generation [1] and embodied version [2], [3] of the Internet. There is no unified definition for the metaverse. According to Xu [3], it is a network of digital worlds with a focus on social connections where people act as avatars to interact and collaborate with each other in real time supported by virtual reality (VR), augmented reality (AR), and the tactile Internet. The metaverse also generates a mirror system of the real world termed the 'digital twin world' [4], which allows each user to produce content and edit the world [5].

Early versions of the metaverse have been applied in games, movies, e-commerce, health care, and the like [6]. The metaverse has regained attention in various fields, including education, during the COVID-19 pandemic to cope with the sudden transition from face-to-face to online learning. Such transition has brought immediate challenges to both teachers and students [7], such as teachers' lack of experience and capacity to conduct intensive online lessons with innovative pedagogical strategies as well as students' social-emotional problems, which lead to poor learning performance due to the limitations of online activities [8], [9]. Mark Zuckerberg's announcement changing the name of Facebook to Meta and the rapid development of digital technologies have sped up the metaverse's popularity [1], [3]. Providing an immersive learning environment through the Community of Inquiry [10] to develop students' communication and collaborative problem-solving skills and enhance their well-being in online or blended learning environments has become essential [11]. However, the present design and implementation of the metaverse in education are still in its nascency.

Although some metaverse platforms, such as 3D virtual worlds, provide users with immersive gaming experiences for educational use, social interactions and collaboration for pedagogical purposes are scant. In addition, many platforms require high-performance devices to render 3D graphics, or external or wearable devices (e.g. sensors, headsets, and controllers) for better immersive experiences and interaction, which may lead to many restrictions for daily use [12]. For example, Meta Horizon Worlds just released a new headset—Quest Pro VR—in October 2022 at the cost of USD 1499.99 [13]. The high cost of the headset raises the threshold for adopting the platform as the educational metaverse. Furthermore, few metaverse platforms can truly provide a digital twin world in which avatars can mirror real users' poses, fingers and facial expression tracking. Last but not least, most of the current 3D virtual worlds do not allow users to edit content and create learning spaces to cater for their own needs [14].

Against this background, our research team has developed Learningverse, a 3D metaverse platform with the following features: (1) allowing users to act as avatars that mirror the real users, (2) enabling social interactions among avatars to interact and collaborate with each other and virtual objects in an

This paper was produced by the IEEE Publication Technology Group. They are in Piscataway, NJ.

Manuscript received xxx ; revised xxx.

This study was supported by the Central Reserve Allocation Committee (CRAC) Fund (2021-2024) (Ref. 03ABN) and the Knowledge Transfer (KT) Fund (2022-2023) (Ref. KT-2022-2023-0008), The Education University of Hong Kong. We appreciate the funding support. *(Corresponding author: Yanjie Song).*

Ethical approval of this study is granted by the Human Research Ethics Committee (HREC) of The Education University of Hong Kong.

Y. Song, J. Cao, K.Wu and P. L. H. Yu are with the Department of Mathematics and Information Technology at The Education University of Hong Kong, Hong Kong SAR, China (e-mail: ysong@eduhk.hk, jxcao@eduhk.hk, kwu@eduhk.hk, plhyu@eduhk.hk).

J. C.-K. Lee is with the Department of Curriculum and Instruction at The Education University of Hong Kong, Hong Kong SAR, China (e-mail: jcklee@eduhk.hk).

"Learningverse" won "Bronze Medal" at the 48th International Exhibition of Inventions in Geneva, Switzerland 2023; and "Silver Medal" and "Special Award" at The International Invention Innovation Competition in Canada (iCAN) 2023

Song, Y., Cao, J., Wu, K., Yu, P. L. H., & Lee, J. C. K. (2023). Developing "Learningverse"—A 3-D Metaverse Platform to Support Teaching, Social, and Cognitive Presences. *IEEE Transactions on Learning Technologies*, *16*(6), 1165-1178.

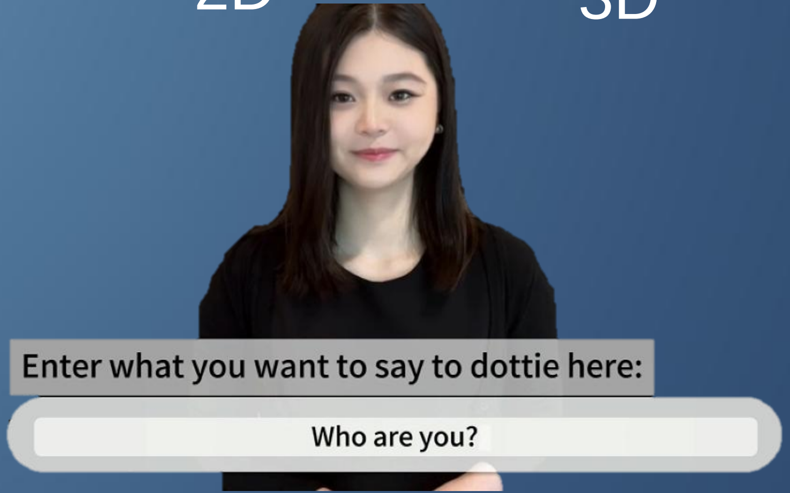
2.2 Digital human teacher

Integrating generative AI into the metaverse platform:

- Human-like interaction:
    - GenAI and LLMs represent cutting-edge technologies capable of mimicking human-like cognitive processes (Fan et al., 2023)
- Adaptive learning:
    - AI can adapt instructional strategies based on real-time student responses and learning progress (Rane, Choudhary, & Rane, 2023)
- Personalised feedback and guidance:
    - Individualised support (Lim et al., 2023)
    - Retrieval-augmented generation(RAG) for improving accuracy of feedback

2D

3D



Real shot

2D Digital human teacher

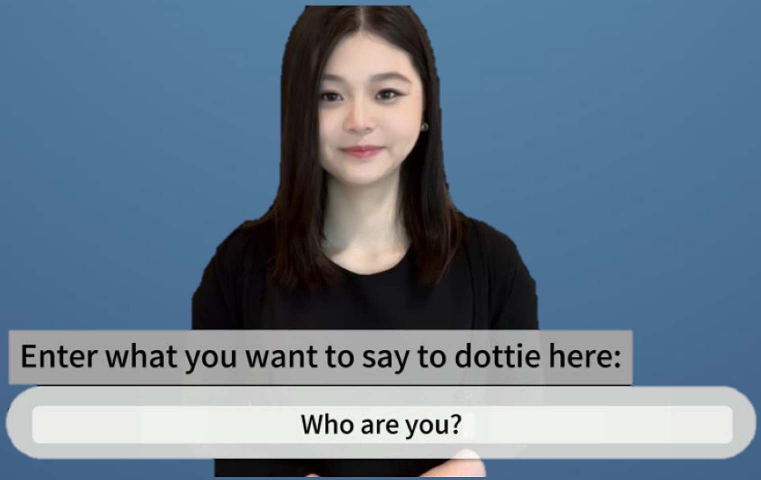
2D digital human in Learningverse

3D digital human – a nurse in a hospital

Real shot video synthesis of digital human

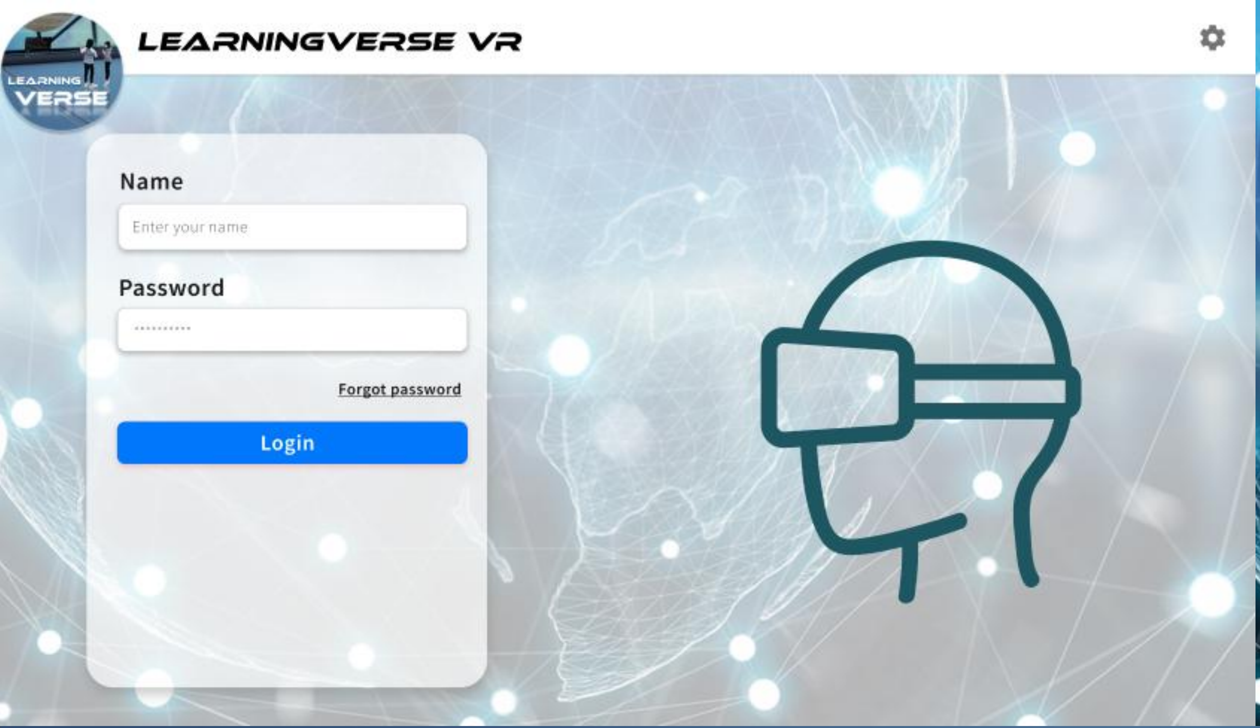
Digital human teacher (basic)

- Driven by Large Language Models (LLMs)
    - Domain-specific knowledge embedded in knowledge base
    - Using Retrieval Augmented Generation (RAG)
    - Leveraging prompt engineering to guide the interaction
- 2D plane avatar with real shot photos
- Text-to-text input, output and interaction



2.3 LearningverseVR

An immersive game-based learning platform based on Learningverse that uses generative AI and virtual reality to improve learners' immersive and interactive experiences and offer novel views on digital game-based learning.



Song, Y., Wu, K., & Ding, J. (2024). Developing an immersive game-based learning platform with generative artificial intelligence and virtual reality technologies–"LearningverseVR". *Computers & Education: X Reality*, *4*, 100069.

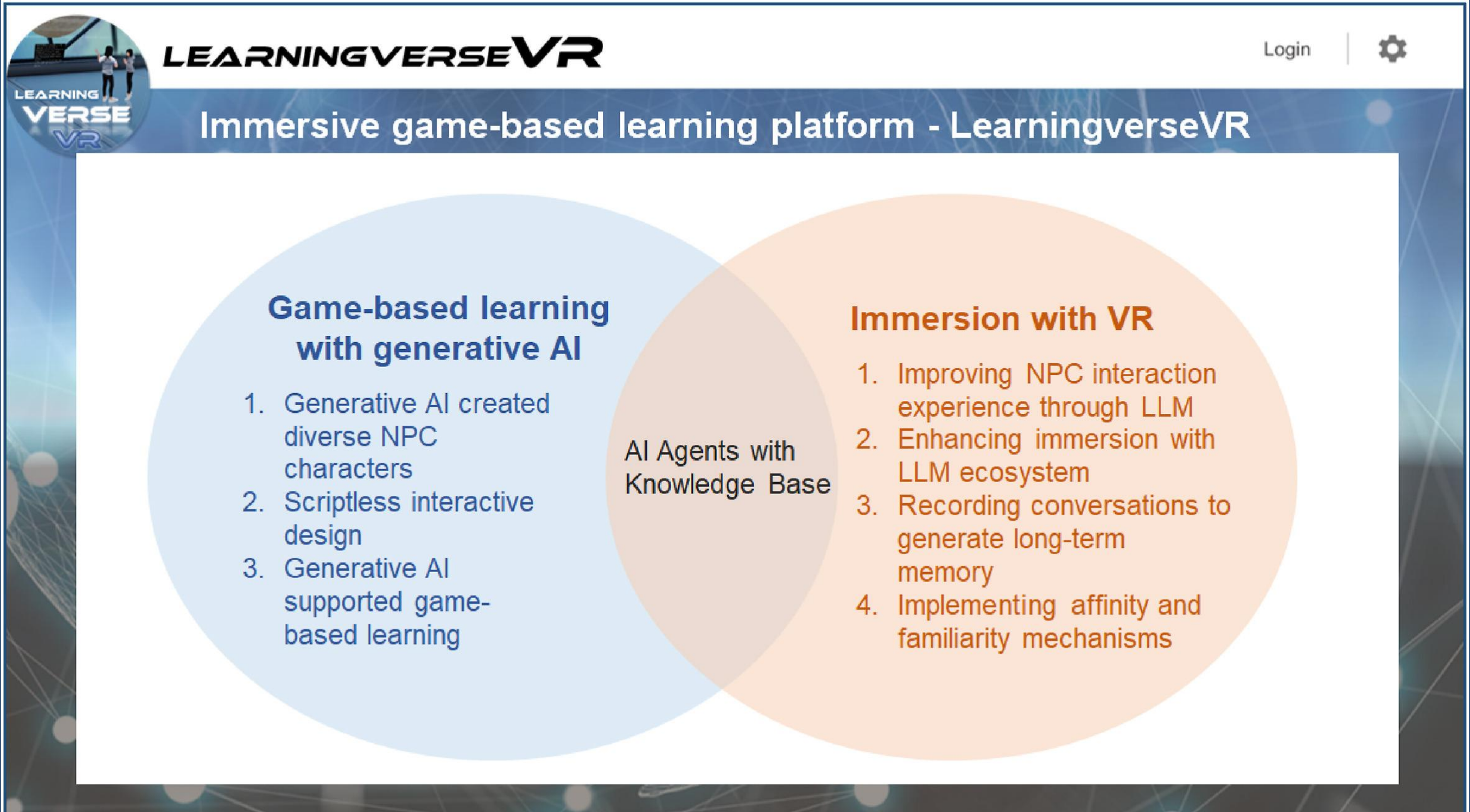
Differences between LearningverseVR and other platforms on the market

Unique features in LearningverseVR.

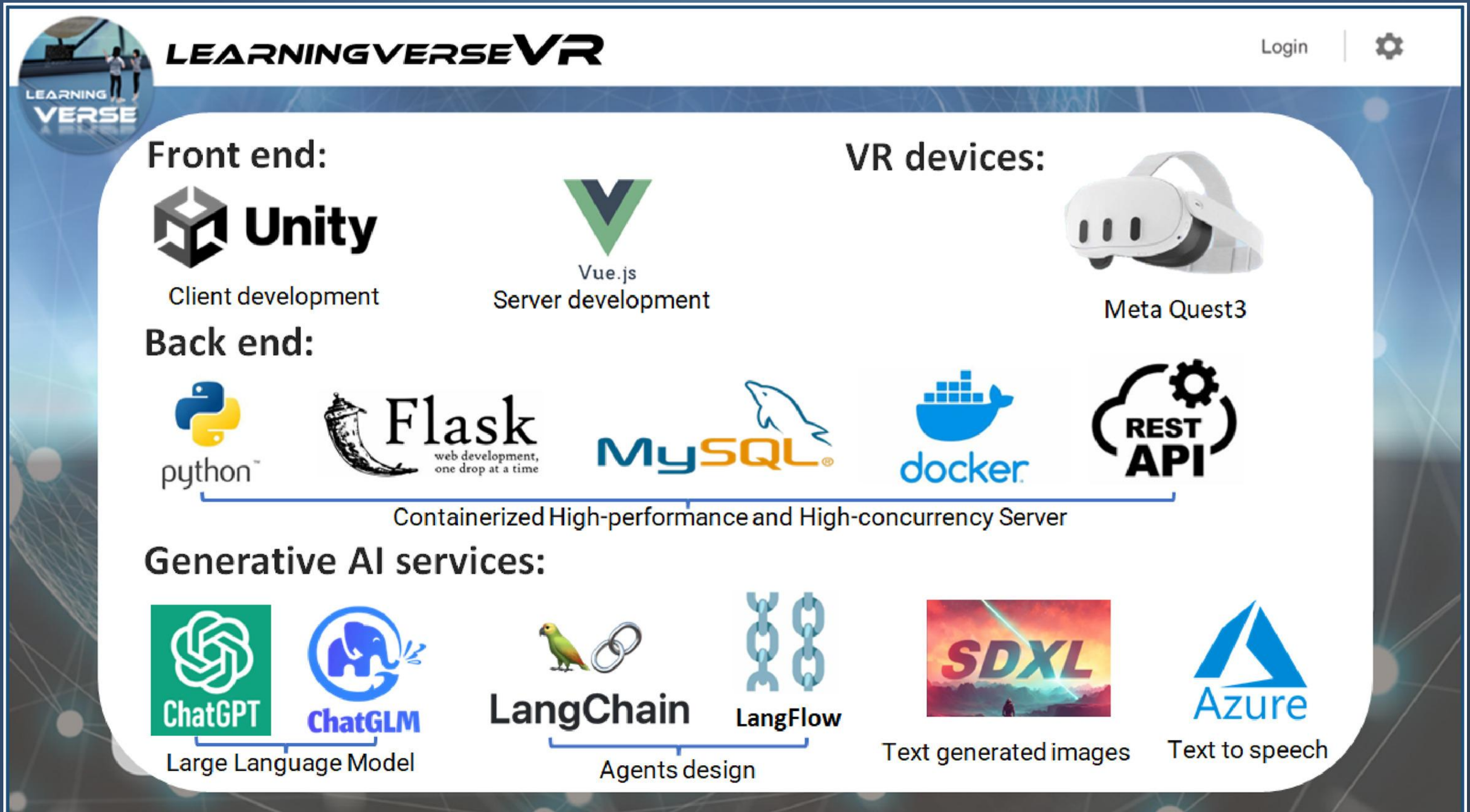
	LearningverseVR	Learningverse	Spatial	Minecraft	Rec room	Second Life	Roblox	Meta Horizon
Theoretical framework (Garrison et al., 1999)	✓	✓	×	×	×	×	×	×
Collaborative interactions (Zhao et al., 2022)	✓	✓	✓	✓	✓	✓	✓	✓
Behaviour & facial expression recognition	×	✓	×	×	×	×	×	×
Grouping mechanism (Blatchford & Russell, 2019)	✓	✓	×	×	×	×	×	×
Navigator & map (Chapman & Rich, 2018)	✓	✓	×	✓	×	×	×	×
Learning process controller	✓	✓	×	×	×	✓	×	×
Real-time communication (Blackmon, 2012)	✓	✓	✓	✓	✓	✓	✓	✓
Chatboards (Kirkup & Kirkwood, 2005)	✓	✓	✓	✓	✓	✓	×	×
Interactive whiteboard (Armstrong et al., 2005)	✓	✓	✓	×	×	×	×	×
Custom tool (Kye et al., 2021)	✓	✓	✓	✓	×	✓	×	✓
Game-based learning elements (Plass et al., 2015)	✓	×	×	✓	✓	✓	✓	✓
Generative AI (Hwang & Chien, 2022)	✓	×	×	×	×	×	×	×
VR function (Freina & Ott, 2015)	✓	×	×	✓	✓	×	✓	✓

LearningverseVR keeps most of the unique features of Learningverse, but adds generative AI-driven NPC interactions, game-based learning design, and the first-person VR experience.

Overview of LearningverseVR design framework



Technology infrastructure of LearningverseVR



Design of generative AI-driven NPCs /digital humans

Large Language Models (LLMs)
Text generation & Action

Knowledge base access
Context-awareness

Prompt engineering
NPC objective & behaviour

Generative visual model
Image & diagram generation

Audio encoder & decoder
Text-to-speech

API communication
Integration of various capabilities

3D reactions
Face emotion & body gestures

Retrieval and prompt engineering based on the LLMs

Generative visual model and text-to-speech converter creation of an agent

Agent-client communication

Multi-Agent ecosystem

AI-driven NPC

Agent workflow

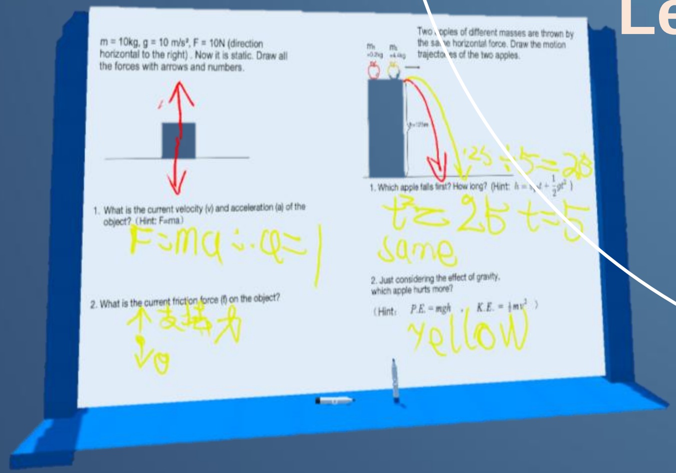
RAG

LLM

Generative AI

Unity

Game-based learning interaction



LearningverseVR Multi-Agent ecosystem

AI feedback

First-person view

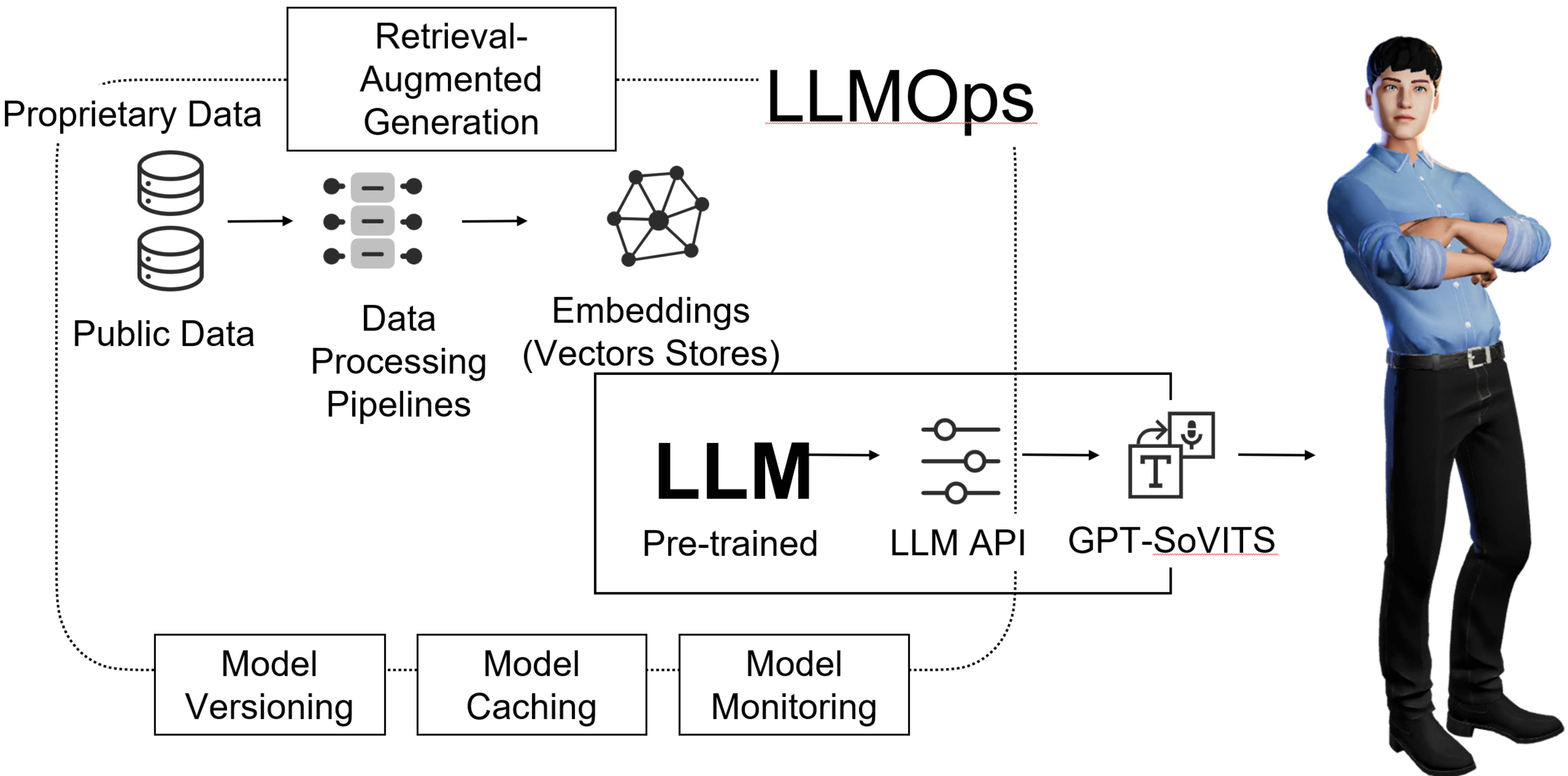
Song, Y., Wu, K., & Ding, J. (2024). Developing an immersive game-based learning platform with generative artificial intelligence and virtual reality technologies–“LearningverseVR”. *Computers & Education: X Reality*, *4*, 100069.

Multi-Agent ecosystem

Multi-Agent ecosystem is an autonomous interactive system where multiple AI agents (including characters, objects, plants) can communicate and share information with each other. Each agent possesses its own cognitive abilities and decision-making logic, enabling unscripted natural dialogue and behaviour simulation to create a more authentic and dynamic interactive experience.

Features	Multi-Agent ecosystem	Traditional ecosystem
Unscripted dialogues	✓	✗
Autonomous decision making	✓	✗
AI objects	✓	✗
AI plants	✓	✗
Dynamic behavior simulation	✓	✗
Inter-agent communication	✓	✗
Information sharing	✓	✗
Preset script interactions	✗	✓

Development of the Large Language Model Operation System (LLMOps) based on Dify



3. Demos of learning scenarios

3.1 Collaborative science inquiry learning in Learningverse on the topic of "Double Ninth Festival (重陽節)"

WeEngage

WeExplore

WeAnalyse

Task 1 WeEngage: Students work in groups to make pre-reflections on what they know about Double Ninth Festival on the Whiteboard.

Task 2 WeExplore: Students work in groups to explore the Learningverse space based on A Dream of Red Mansions, exploring and experiencing the ancient Double Ninth Festival folk customs

Task 3 WeAnalyse: Students analyse the Intangible Cultural Heritage knowledge and information about Double Ninth Festival.

WeReflect

WeExplain

Task 5 WeReflect: Quiz and post reflection on Whiteboard

Task 4 WeExplain: Students perform in groups to show their understanding of the festival (both the key concepts learning in the novel and customs in real life).

Learning Chinese culture on "Double Ninth Festival (重陽節)" supported by a digital human teacher

- Simulating human thinking and behaviour
- Interactive learning experiences
- Personalised learning pathways
- Real-time problem solving
- AI-scaffolded learning

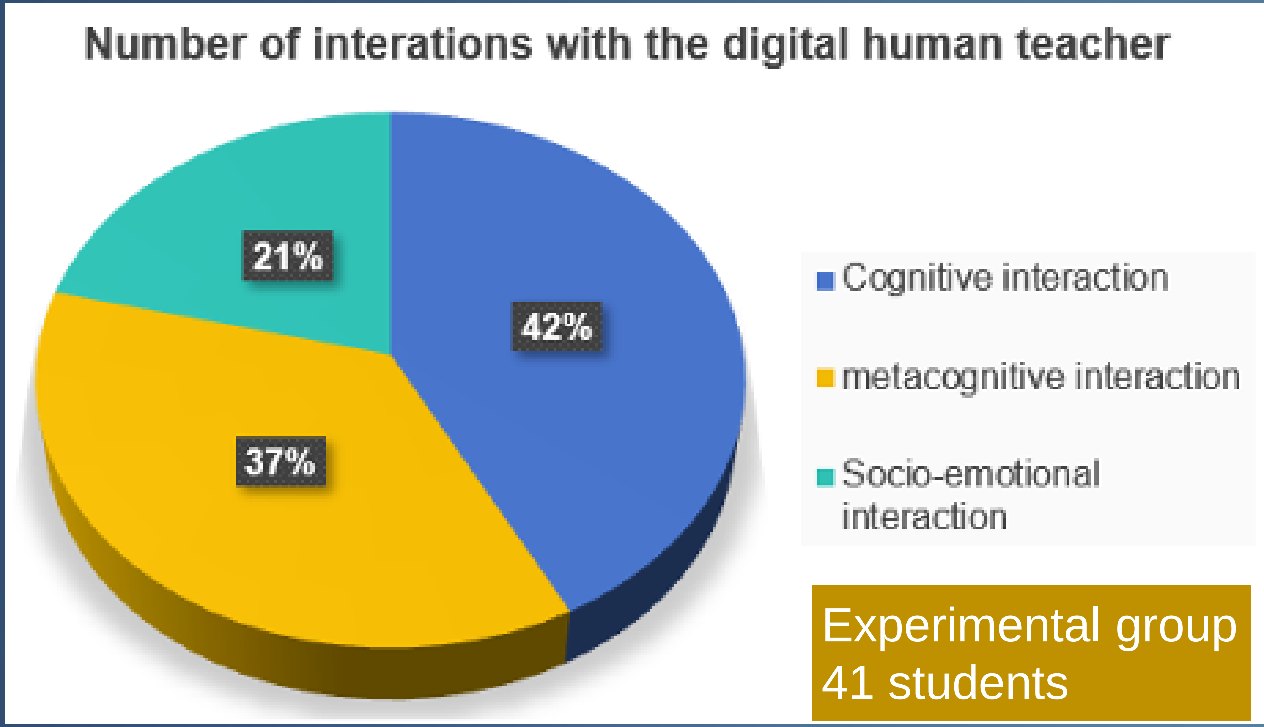
Digital human teacher

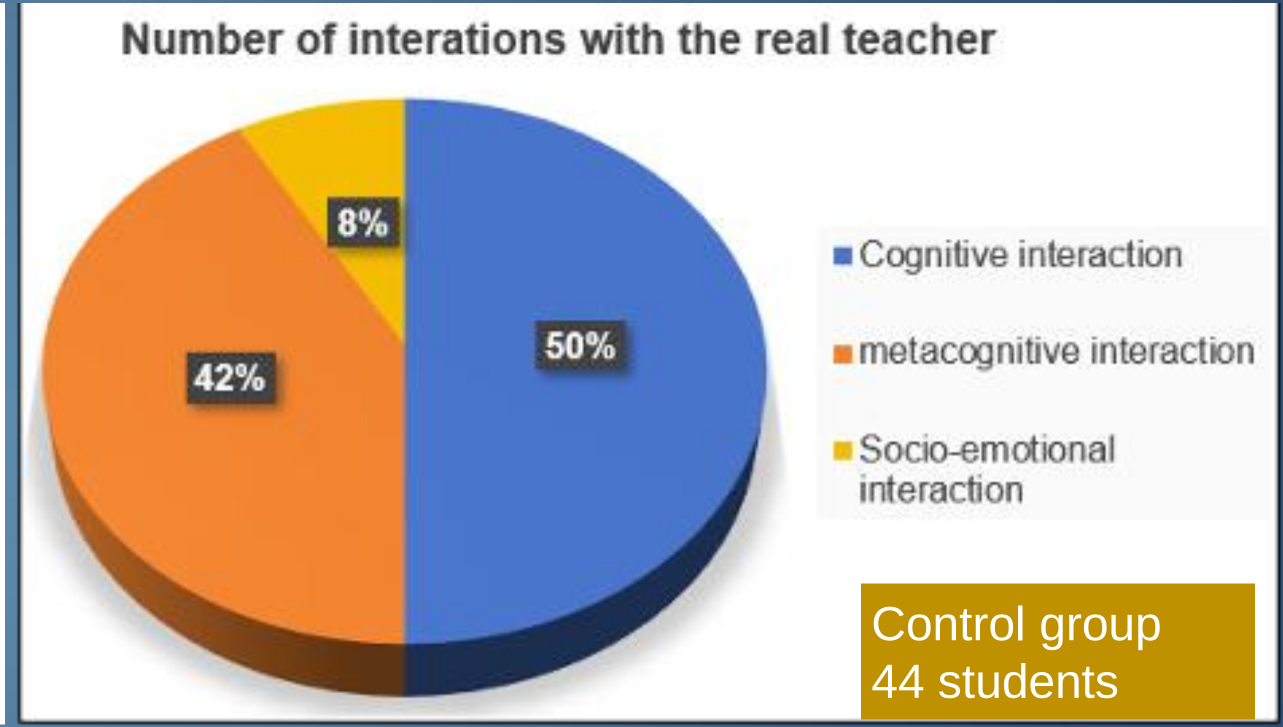
(Drigas et al., 2023; Fink et al., 2024; Johri et al., 2023; Wang et al., 2023)

Experimental group: 6 groups with 41 students
Control group: 7 groups with 44 students

Chinese culture (Double Ninth Festival 重陽節)

Results of interaction patterns between experimental and control groups





This Figure shows that in the experimental group, cognitive, metacognitive and socio-emotional interactions accounted for 42% (66 interactions), 37% (57 interactions) and 21% (33 interactions) between the digital human teacher and students respectively.

This Figure shows that in the control group, cognitive, metacognitive and socio-emotional interactions accounted for 50% (6 interactions), 42% (5 interactions) and 8% (1 interaction) between the real teacher and students respectively.

3.2 Collaborative inquiry-based learning on the topic of "Green energy" in Learningverse

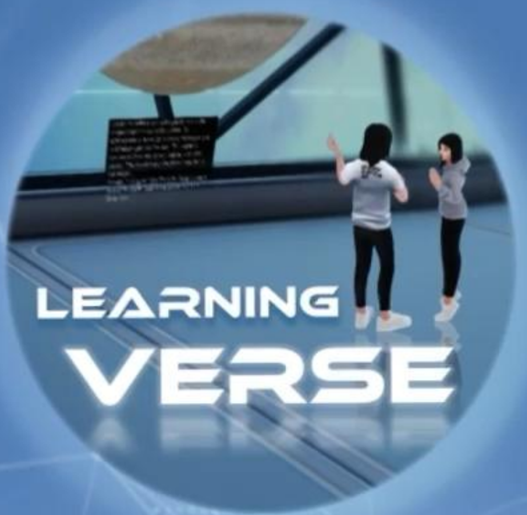
Task 1 WeEngage: Students interact with a steam engine model in the Learningverse, sharing their prior knowledge about energy.

Task 2 WeExplore: Students design power supply solutions for a virtual town using various energy sources and converters.

Task 3 WeAnalyse: Students analyse their power supply solutions based on cost, efficiency, sufficiency, and pollution.

Task 4 WeCreate: Students create their power supply solutions using interactive features in Learningverse.

Task 5 WeShare: Students share their power supply solutions with peers.



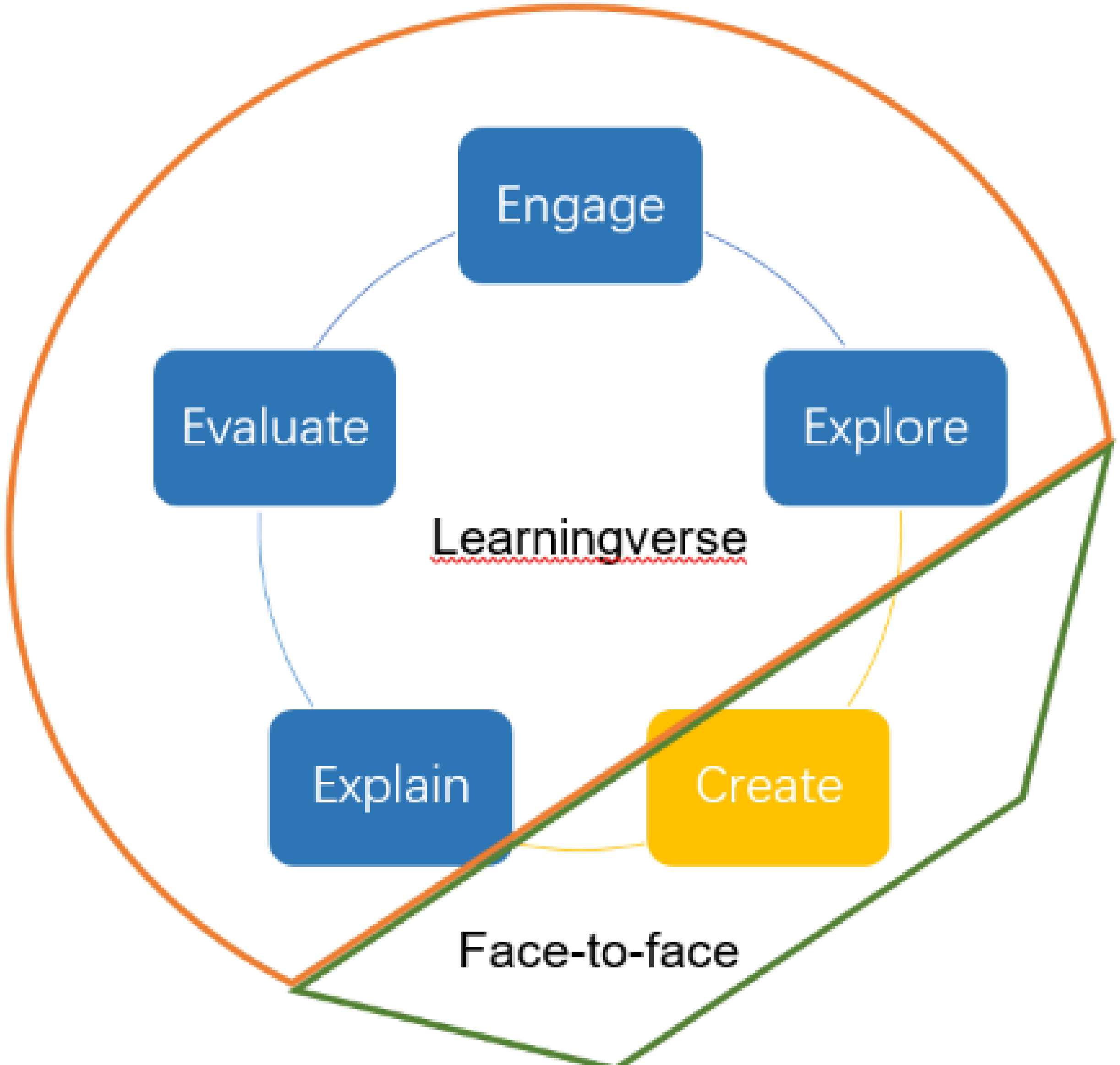
Learningverse

3.3: STEM training - Introduction to robotics "VEX GO" in a primary school in Learningverse

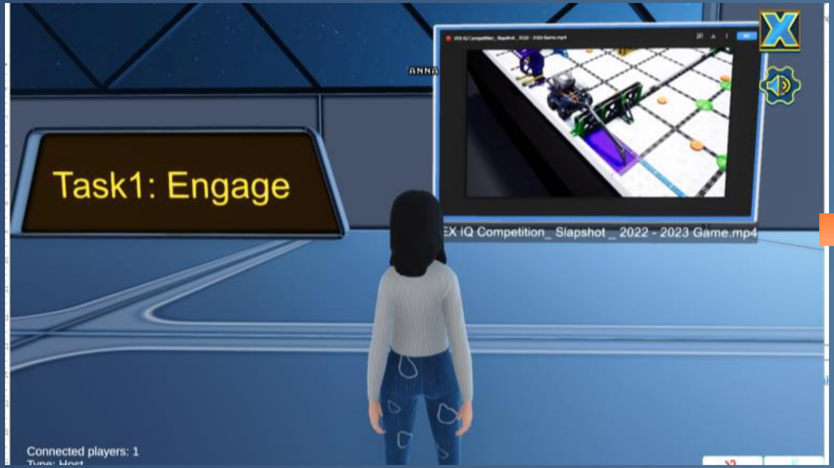
Learning objectives：

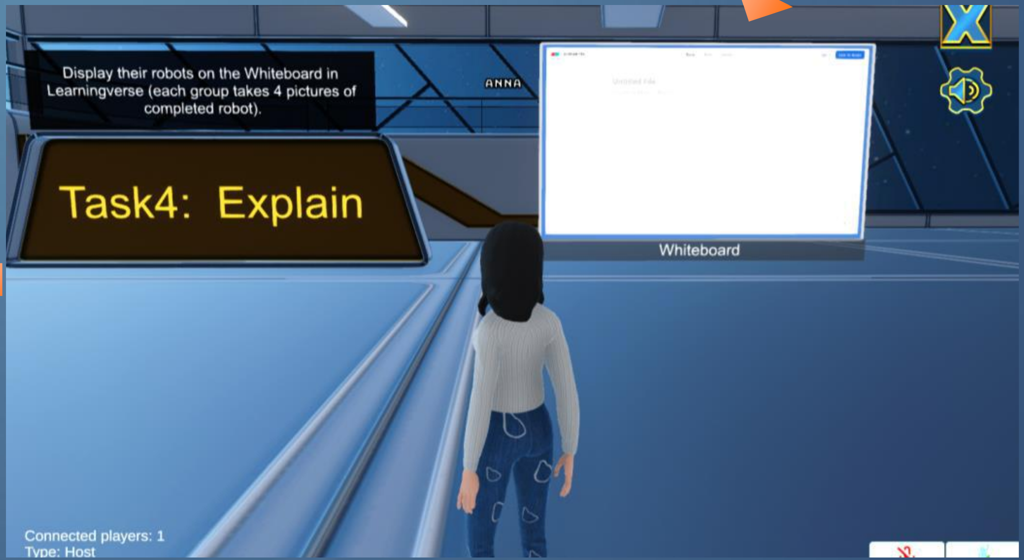
1. To know the basic components for assembling VEX GO
2. To learn how to use VEXcode GO and a Code Base to solve problems in groups

Inquiry based learning in a seamless immersive learning environment

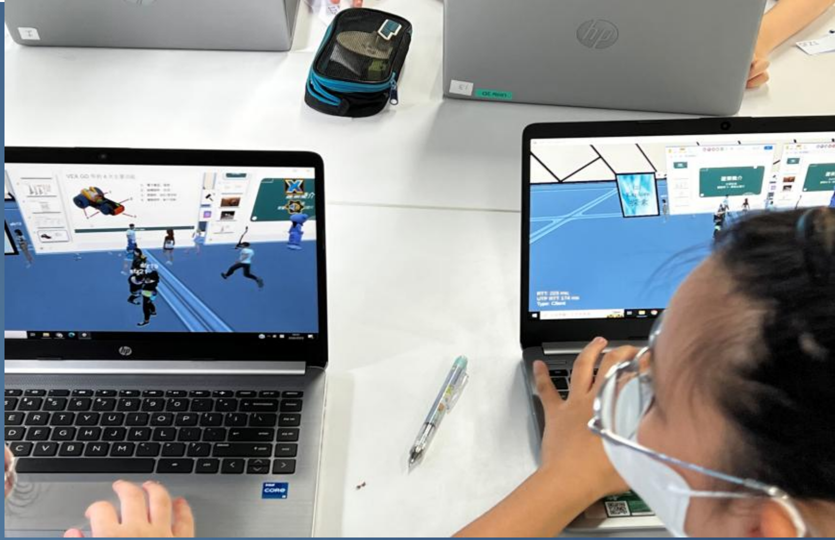


Inquiry based learning on assembling robotics in a seamless immersive learning environment





1. Engage

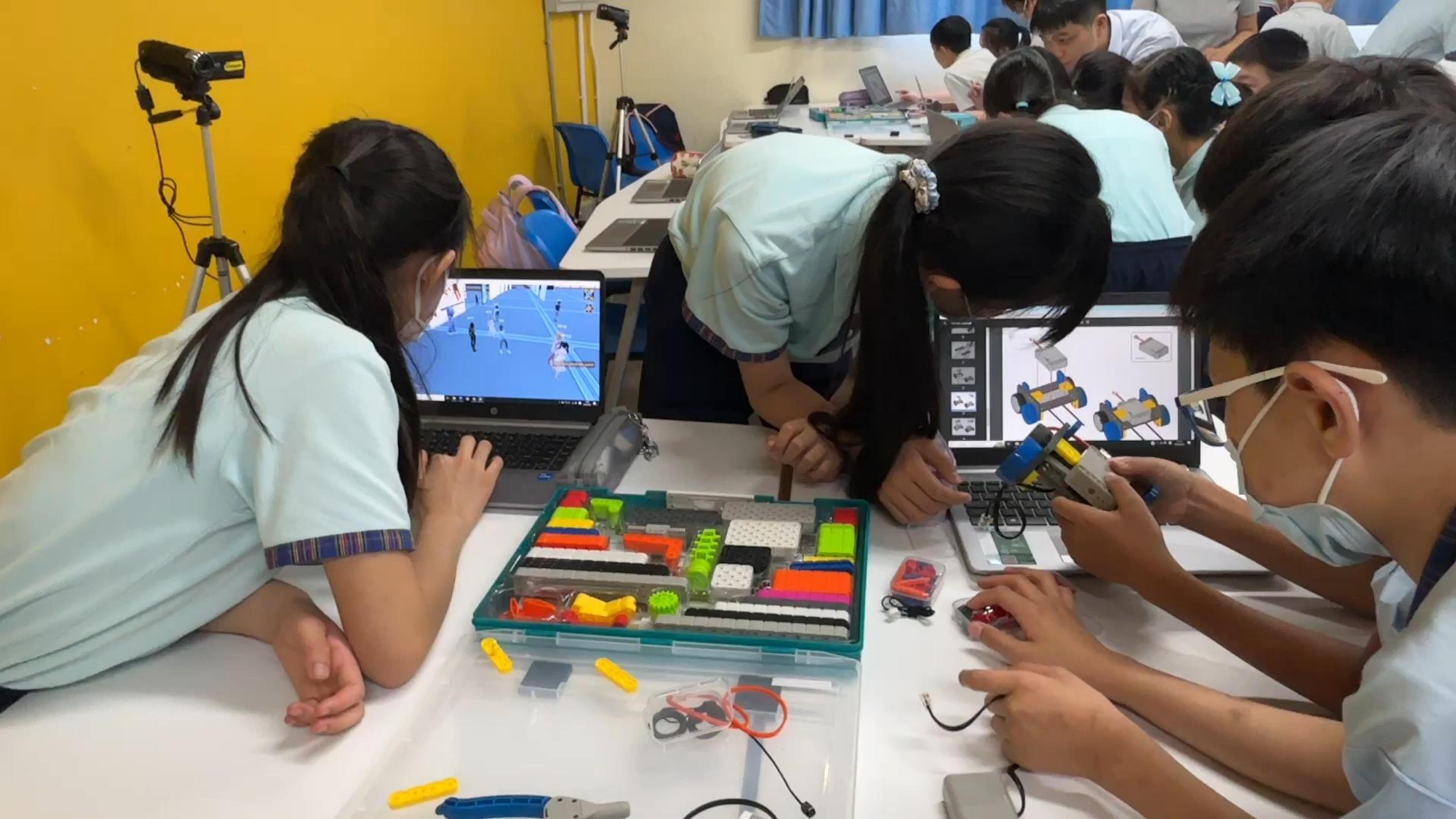


2. Explore information

5. Evaluate

4. Explain: Display their products in Learningverse

3. Create



3.4 STEM in LearningverseVR: Experiential learning to investigate "Newton's Laws of Motion"

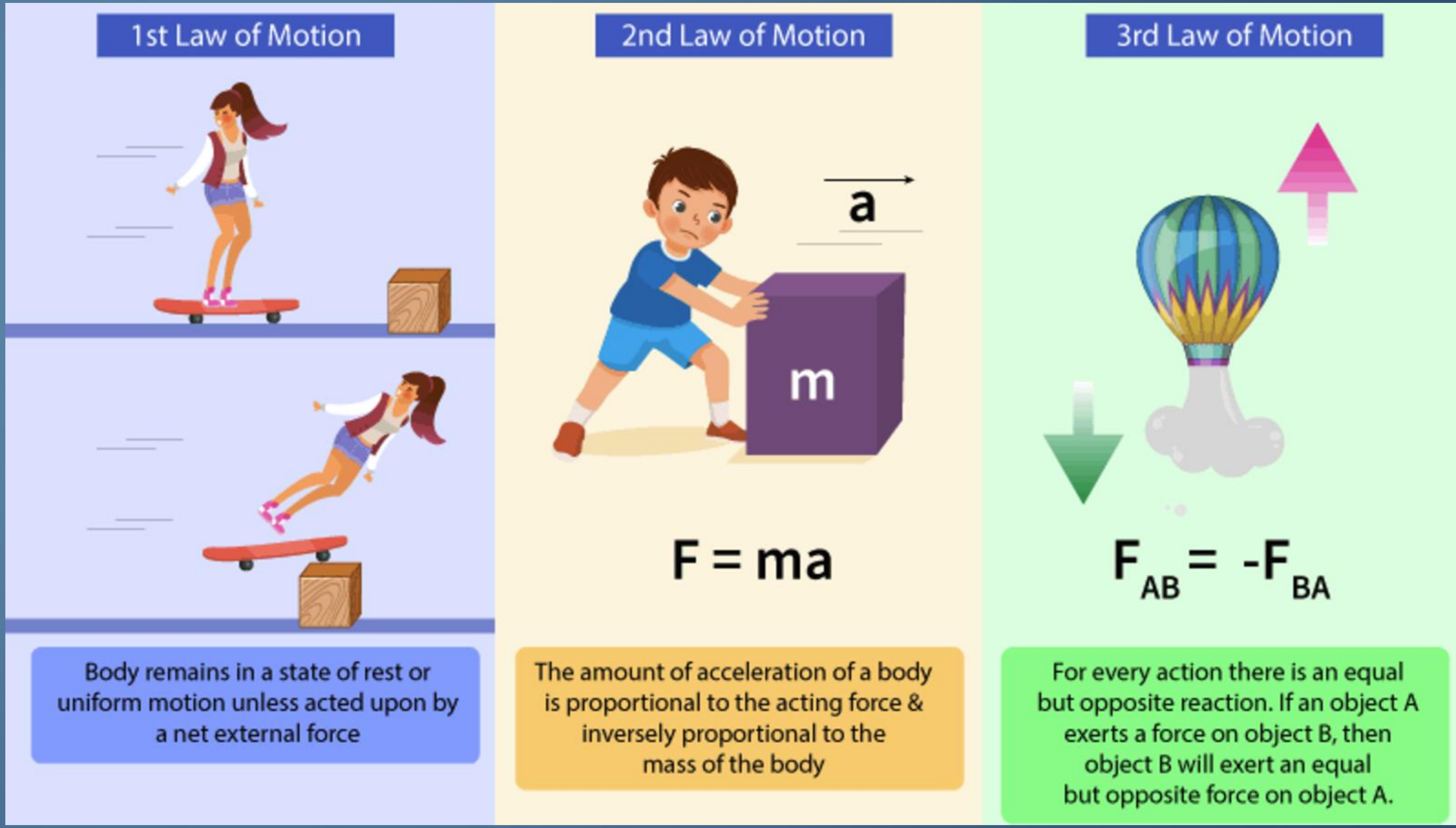
A pilot study

Topic
Magical VR adventure: Newton's laws of motion through apples and archery

Target students
Grade Level: Hong Kong Secondary 4 Students (Approximately 15-16 years old)

Prior Knowledge
- Basic understanding of gravity and its effects
- Familiarity with concepts of mass and weight

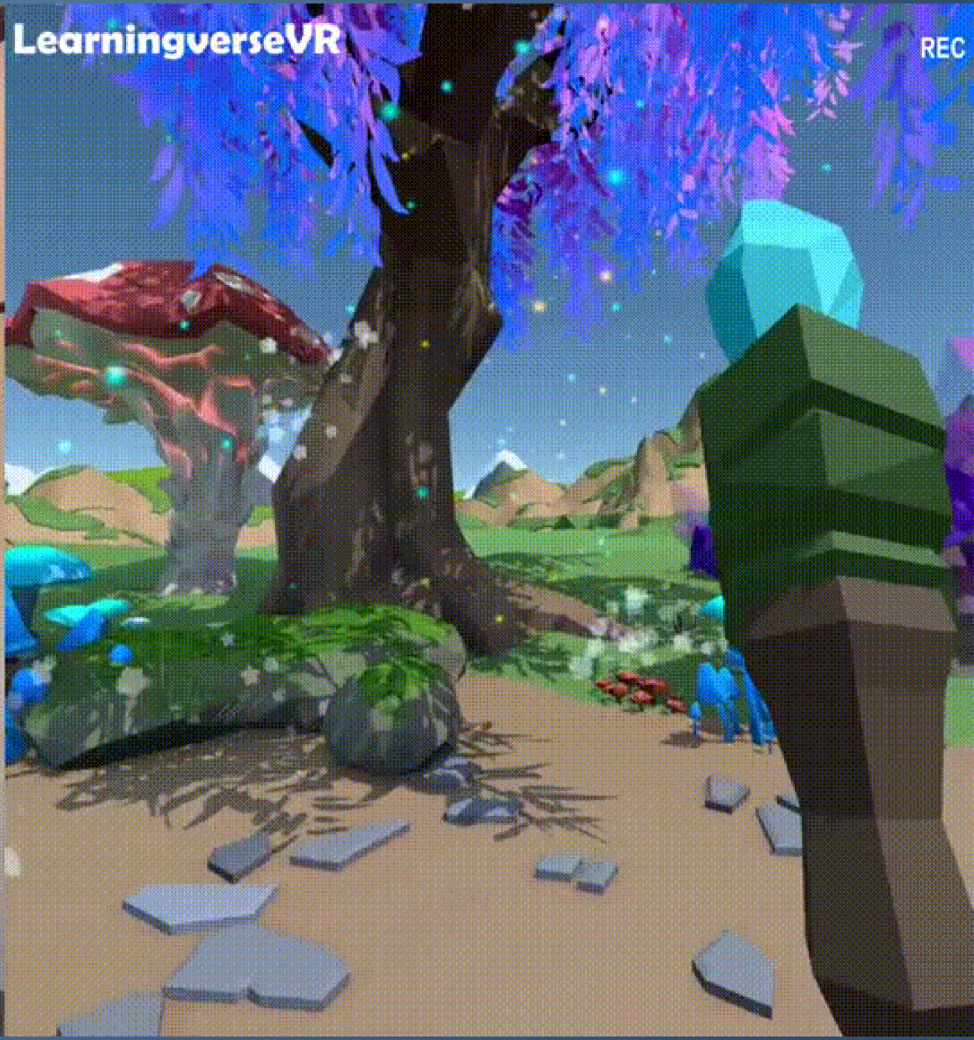
Learning objectives



Adapted from: EDB Science Education Curriculum Documents
https://www.edb.gov.hk/attachment/en/curriculum-development/kla/science-edu/Science(S1-3)_supp_e_2017.pdf

Game-based experiential learning in Learningvesrse

1) Concrete experience: Students engage in hands-on activities like catching falling apples, providing direct experience with the concepts of gravity and mass.

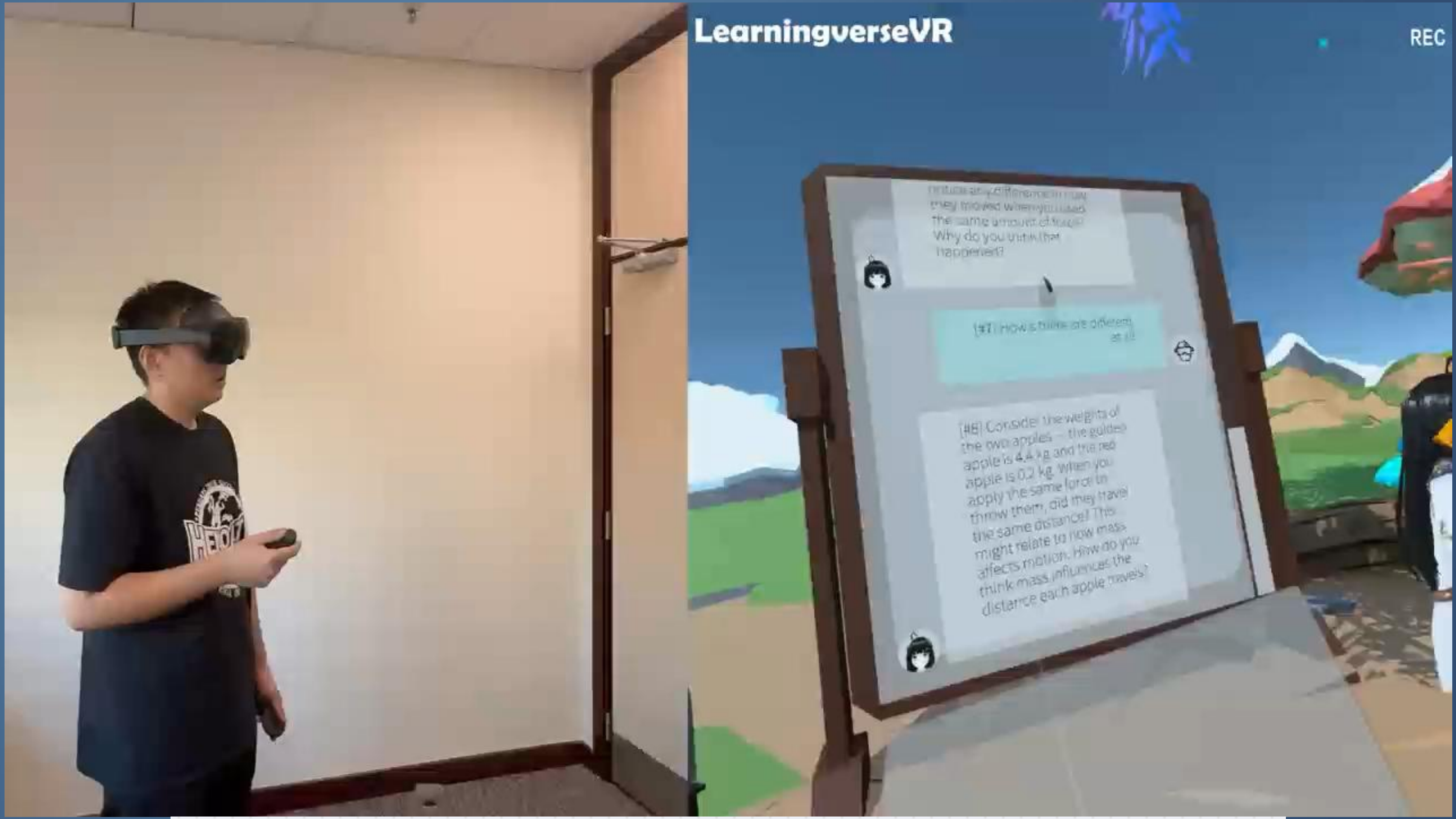


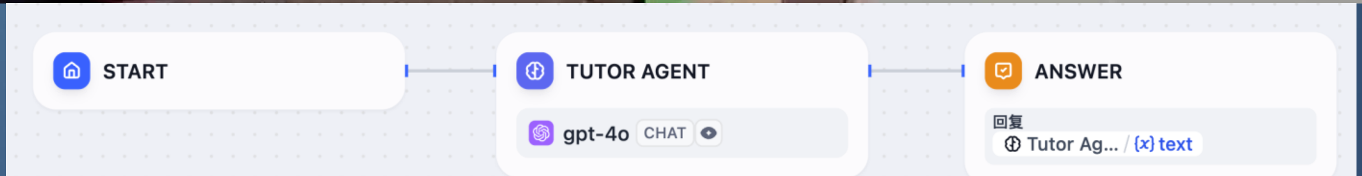
Game-based experiential learning in Learningvesrse

2) Reflective observation: Tasks like throwing apples into the storage box help students observe the effects of force and mass on motion.

Game-based experiential learning in Learningvesrse

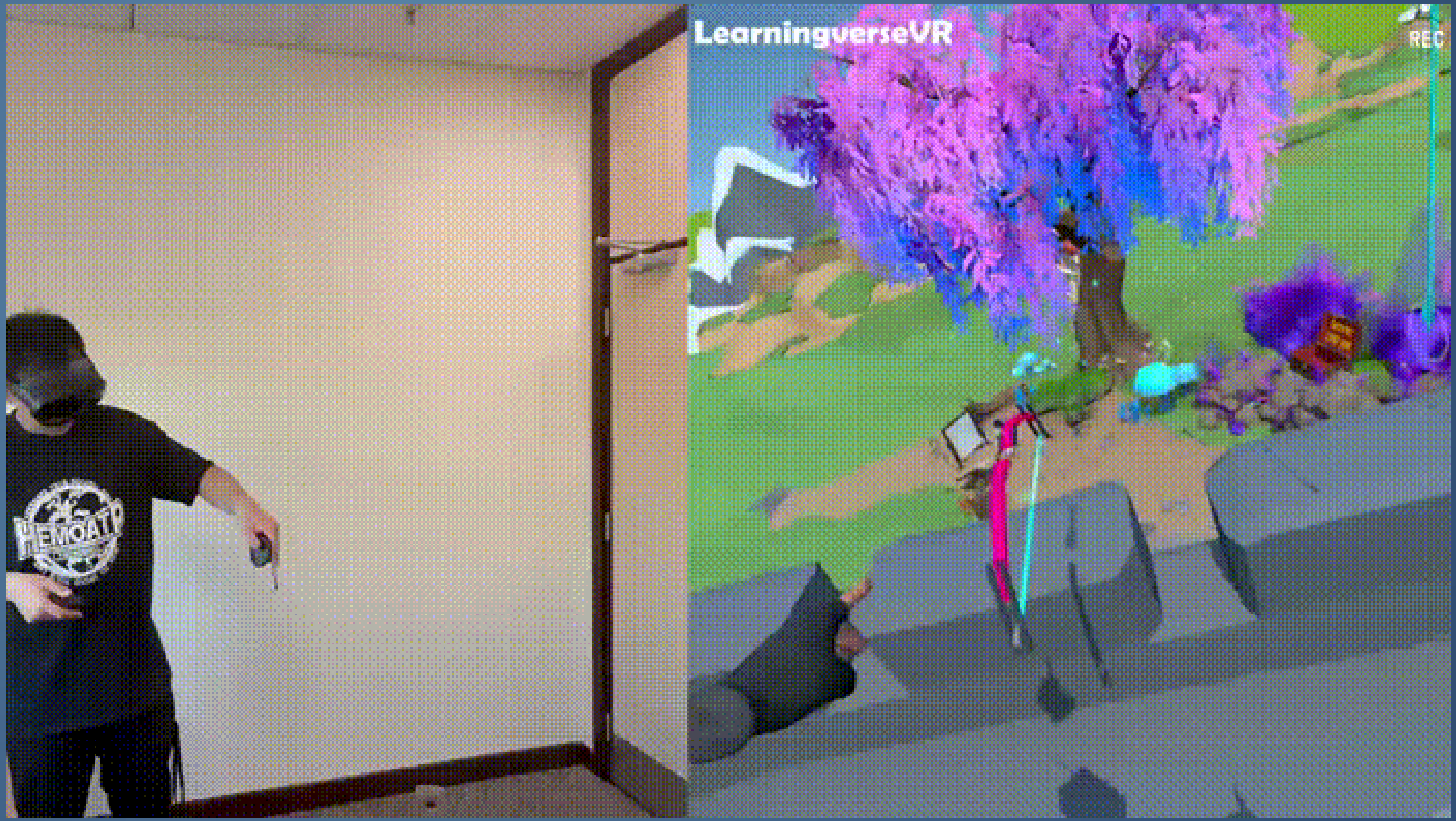
3) Abstract conceptualisation: The AI-driven NPC guides students in understanding of Newton's laws and related equations.





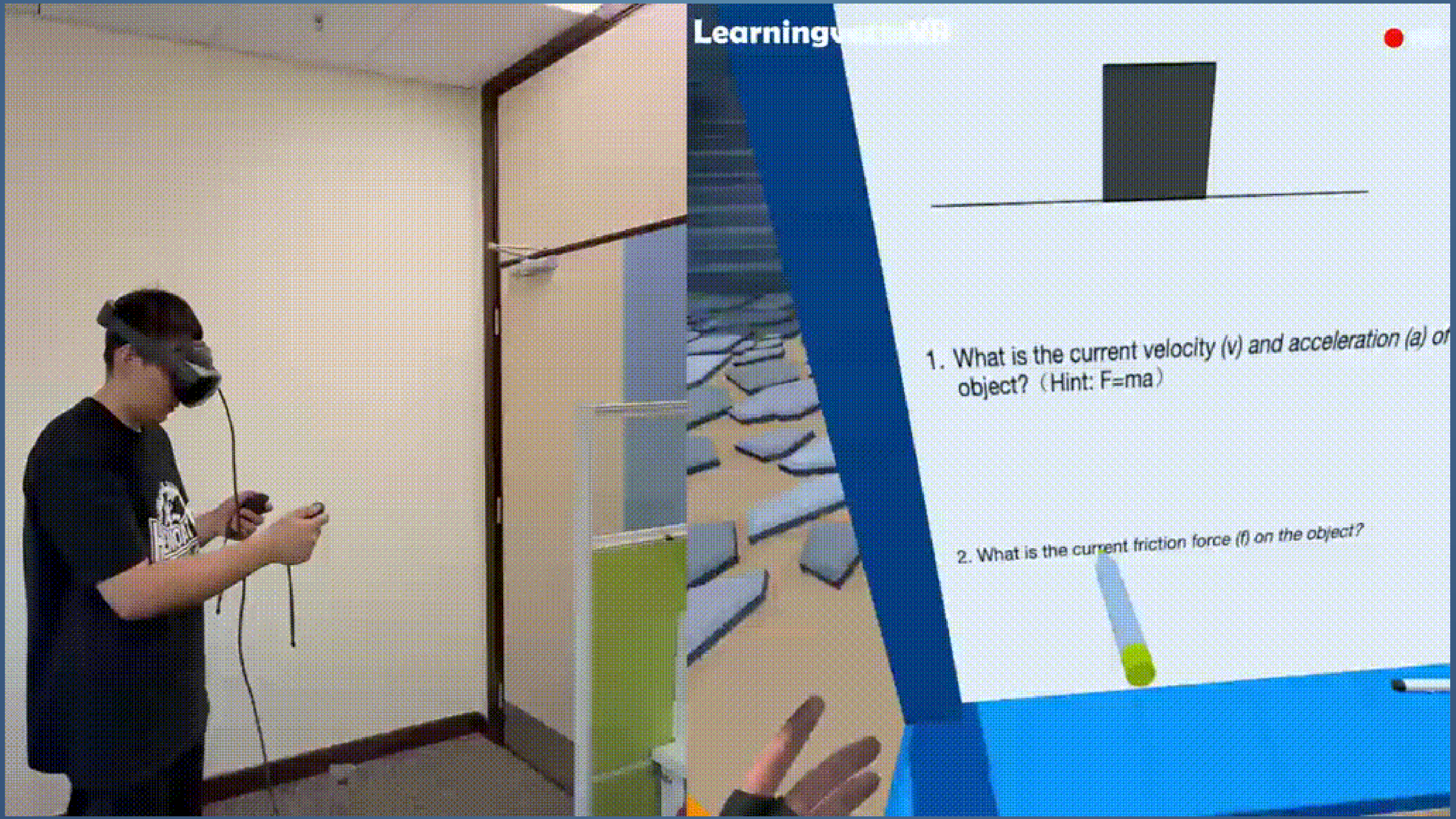
Game-based experiential learning in Learningvesrse

4) Active experimentation :
In the Archery challenge arena, students apply learned concepts to new challenges, experimenting with and observing arrows with different mass and weight (e.g. ice and wood), which results in different trails and air/flying speed.



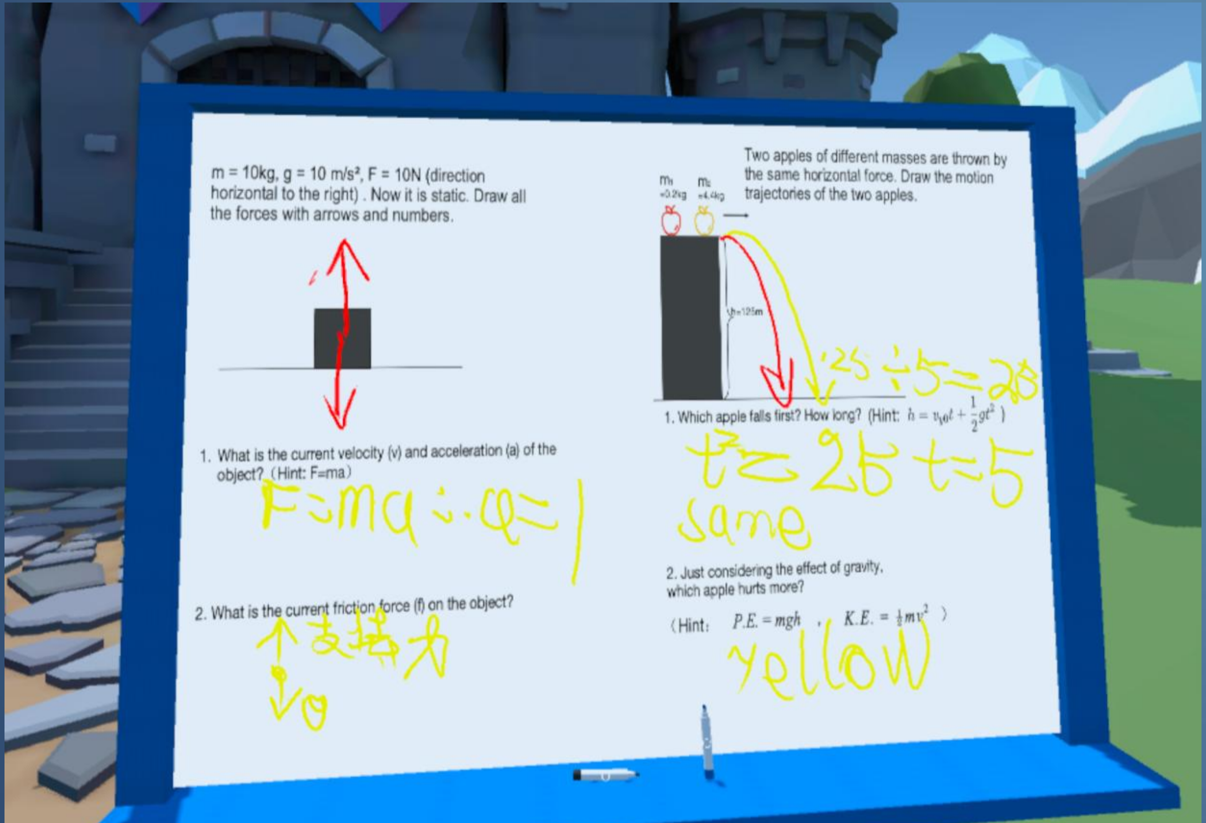
Game-based experiential learning in Learningvesrse

5) Immersive learning and AI feedback: students draw and write equations on a virtual whiteboard.



Game-based experiential learning in Learningvesrse

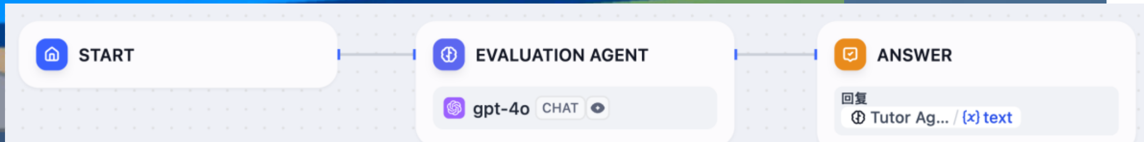
5) Immersive learning and AI feedback: students draw and write equations on a virtual whiteboard with AI feedback for learning support and review.



- **Missing Forces (Applied Force and Friction Force):** The student's diagram does not show the applied force of 10 N directed horizontally to the right, which is crucial for understanding the scenario. Additionally, there is no representation of the friction force, which should be equal and opposite to the applied force to maintain static equilibrium.

Correct Force Analysis Should Include:

- **Gravitational Force (G = mg):** This should be drawn vertically downward, representing the weight of the object, calculated as the product of mass and gravitational acceleration (10 kg × 9.8 m/s² = 98 N).
- **Normal Force (N):** Drawn vertically upward, equal in magnitude to the gravitational force, to maintain vertical equilibrium.
- **Applied Force (F = 10 N):** Horizontally to the right.
- **Friction Force (f):** Horizontally to the left, equal in magnitude to the applied force, to ensure no net horizontal movement.



Pilot study: Research questions, procedure and results

RQ1: What is the impact of LearningverseVR on students' understanding of Newtonian mechanics?

RQ2: What are students’ perceptions of the usability and effectiveness of LearningverseVR as a game-based learning platform?

Pre-test (5 mins) → Introduction LearningverseVR (5 mins) → Game-based learning experience (30 mins) → Post-test (5 mins) / Perception questionnaire (5 mins) → Semi-structured interview (10 mins)

Results:

RQ1: Students demonstrated improved conceptual understanding of newtonian mechanics

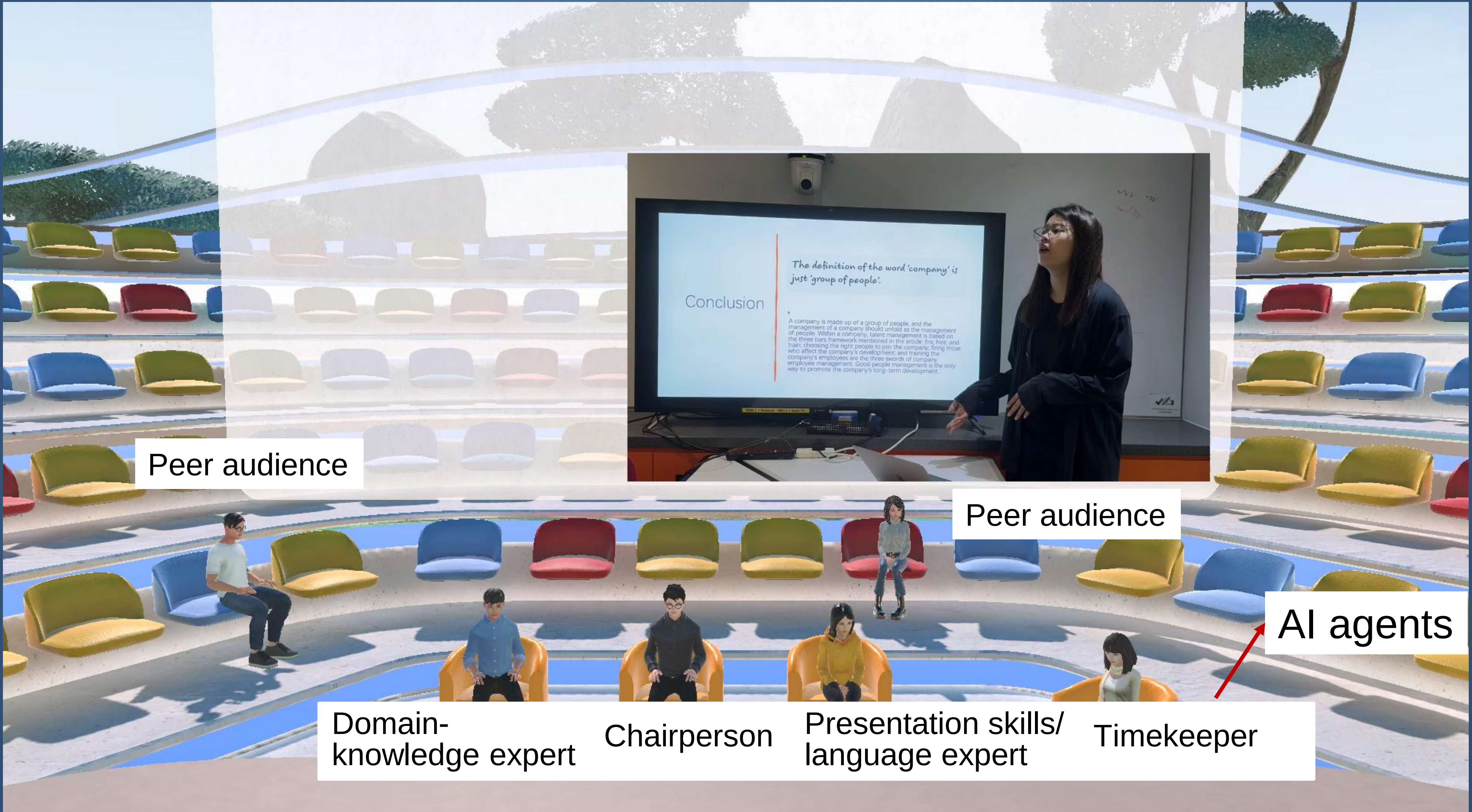
RQ2: Participants positively perceived learningverseVR’s usability and effectiveness, highlighting deeper knowledge understanding and enhanced learning interest, but some reported discomfort with the VR headsets.

3.5 “AGILE” (AI agents in the immersive learning environment) for academic presentation

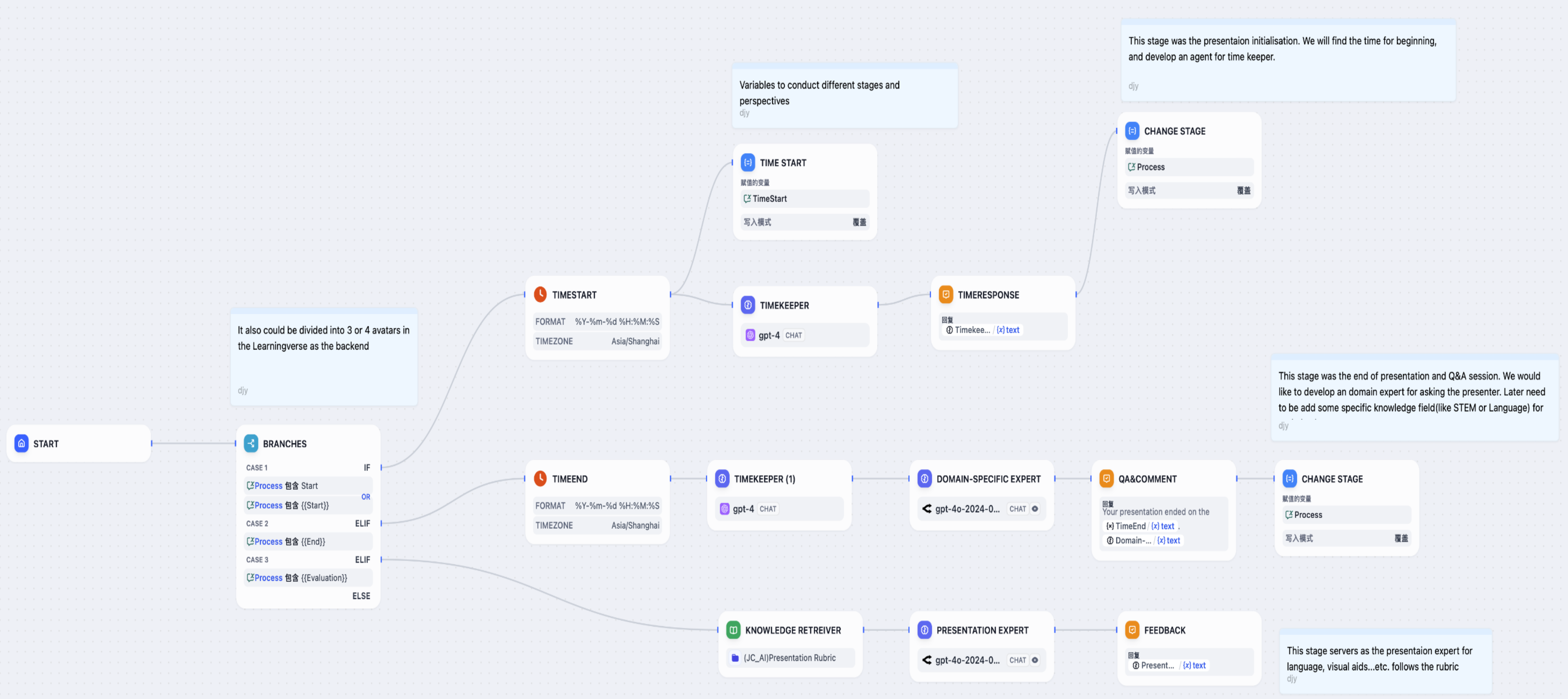
AI agents act as avatars in the immersive learning environment

Domain-knowledge Expert | Chairperson | Presentation skills/language expert | Timekeeper

“AGILE” for academic presentation in Learningverse



3.6 "AGILE" with the multimodal AI agent system in LearninverseVR using multi-LLMs in a workflow



Multimodal AI agent system prototype

Planning
- text

Preparing
- text
- image

Monitoring
- voice
- expression
- gaze
- gesture
- body language

Reflecting
- text
- image

Planning agent:
Assist learners in developing initial outlines, defining objectives, and organizing presentation topics.

Preparing feedback agent:
Analyse PPT slides and presentation scripts that learners upload, then provides a comprehensive feedback report.

Preparing assistant agent:
Assist learners in optimizing PPT slides and presentation scripts through iterative, multi-round dialogues.

Monitoring feedback agent:
Provide real-time insights on presentation by multimodal agents and NPCs agents.

Multimodel agents:
(e.g., voice agent, expression agent, gaze agent, gesture agent) Track and analyse non-verbal data.

NPCs agents (e.g., peer agent, knowledge expert agent, chairperson agent, presentation skills expert agent) Drive different roles in VR learning environments.

Reflecting feedback agent:
Summarise the overall presentation process data and produces a summary report.

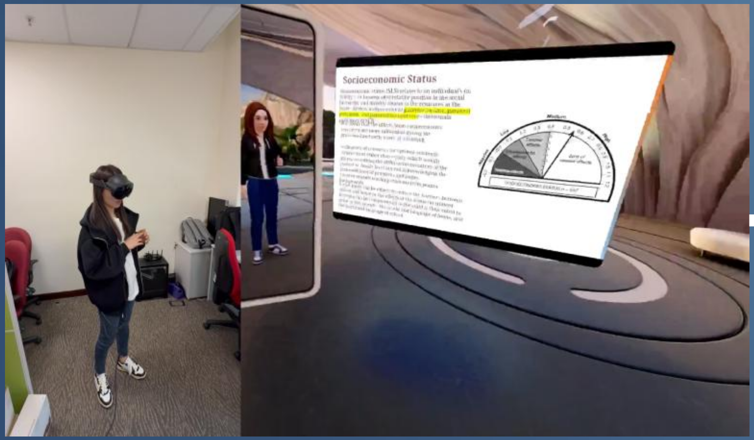
Reflecting assistant agent:
Generate three topic-related questions and communicates with learners to spur deeper reflection

Monitoring: Multimodal AI agent prototype of AGILE

Option academic achievement rather than equality which usually means some smoothing the different differential resources at student or family level but not acknowledging the increasing level of problems and issues.

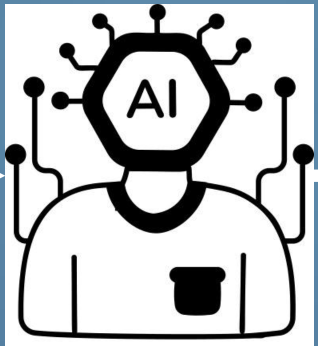
Text

Speech



Video

Multimodal data input at the same time



Multimodal data feedback through different agents.

Text

Feedback: "While focusing on academic achievement is crucial, it's essential to also tackle the underlying issues that contribute to resource differentials among students and families."

Speech

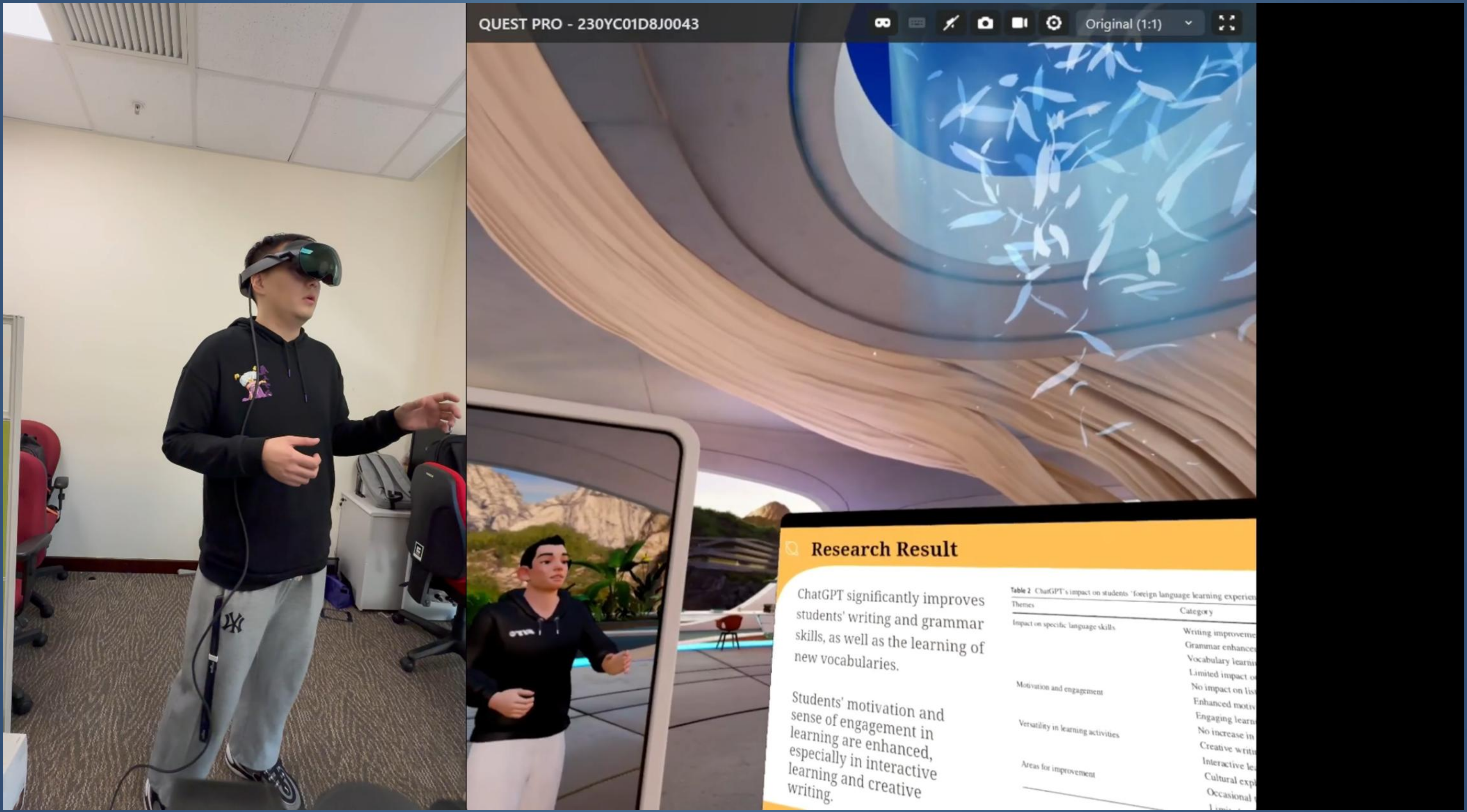
Feedback: Need to increase volume and improve fluency.

Video

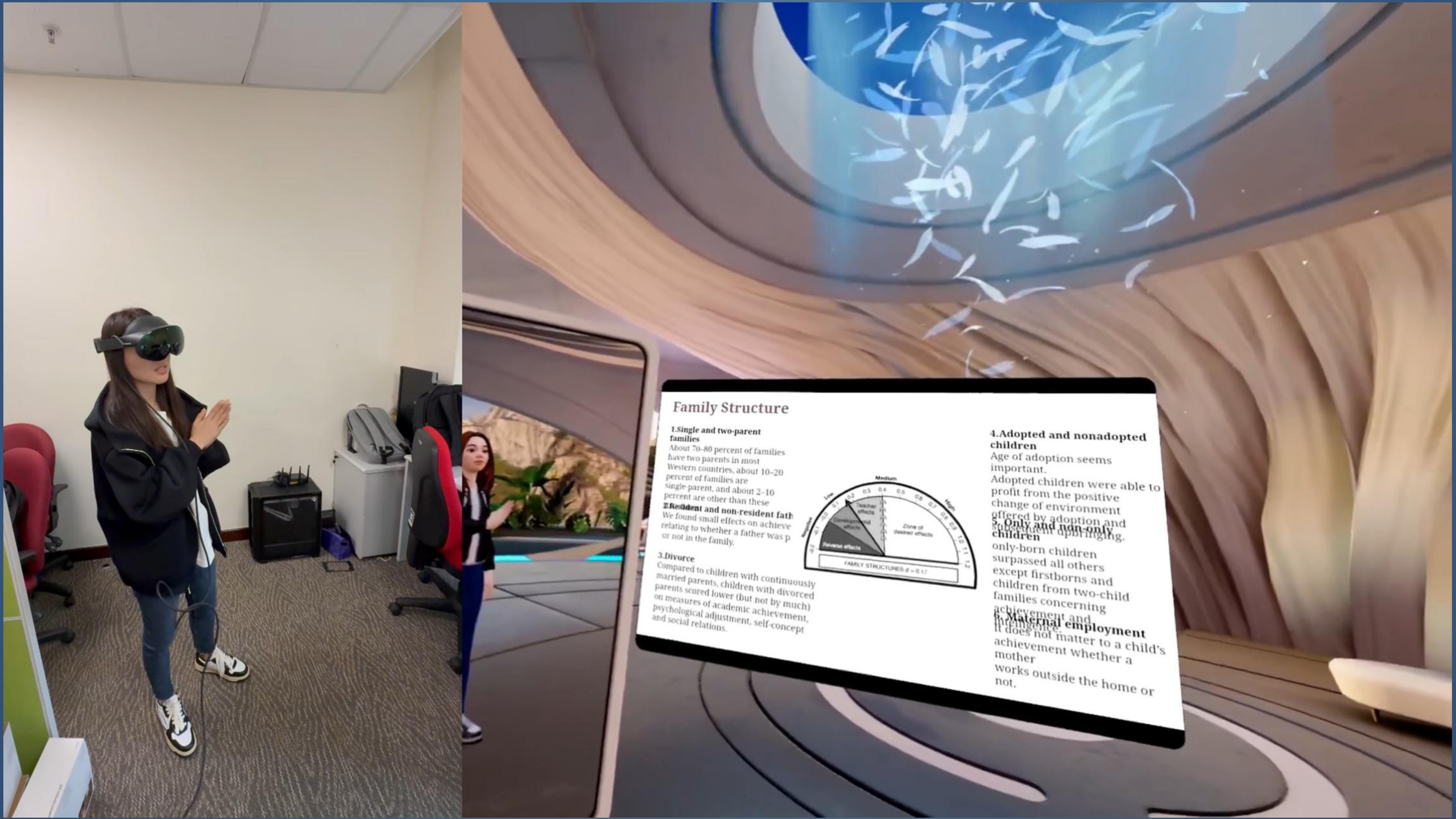
Feedback: The body is tense. You can try interacting with the virtual PPT.

Immersive VR presentation space

1st Academic presentation in “AGILE” with VR

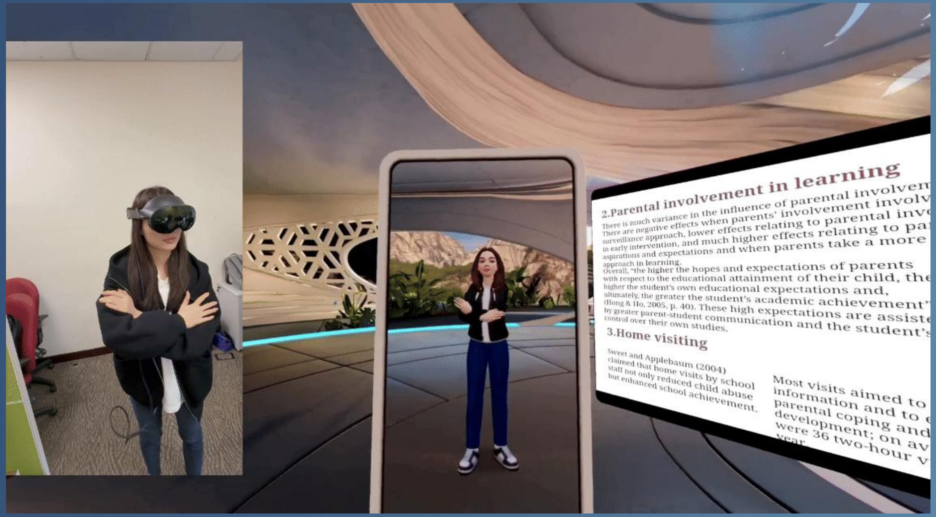


2nd Academic presentation in 'AGILE' with VR

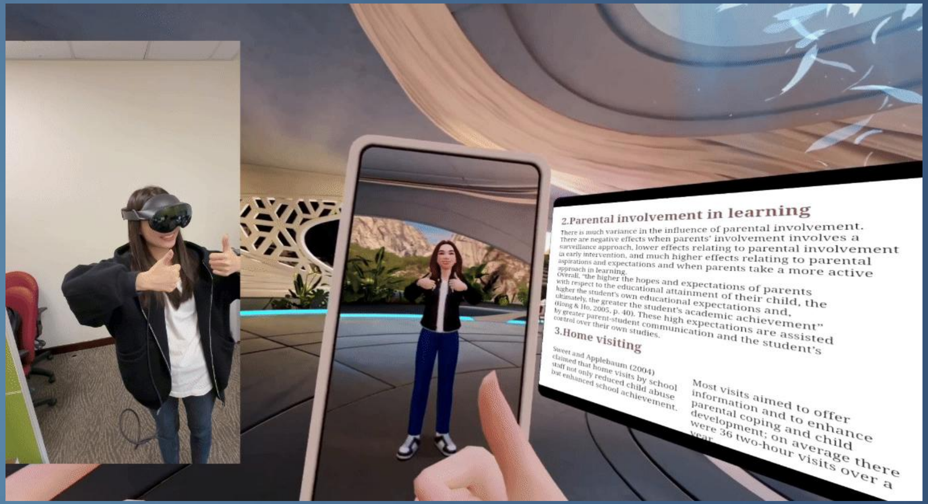


Academic presentation in “AGILE” with VR

Gaze

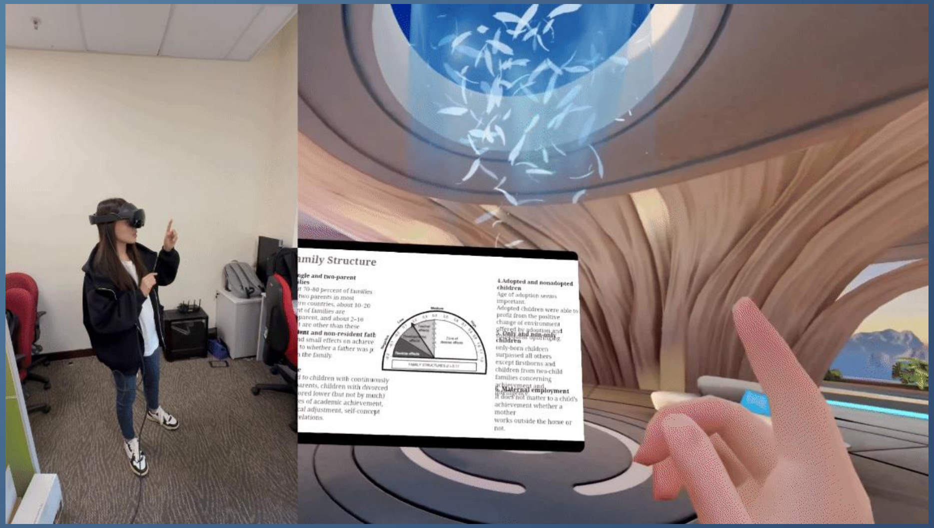


Gesture



Expression

Body language

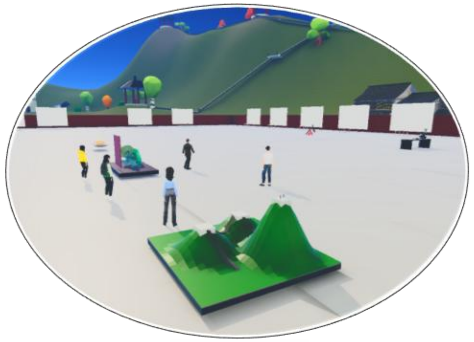


Wrap-up

1. Introduction to metaverses and Gen AI

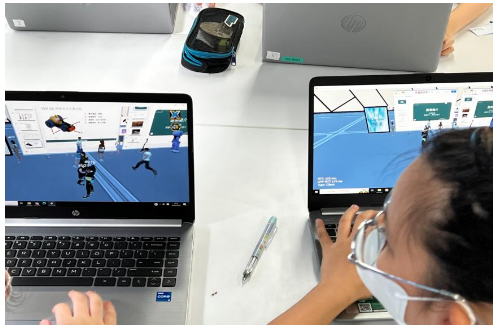
2. The metaverse platforms - Learningverse and LearngverseVR + digital human

3. Six learning scenarios in Learningverse or LearningverseVR

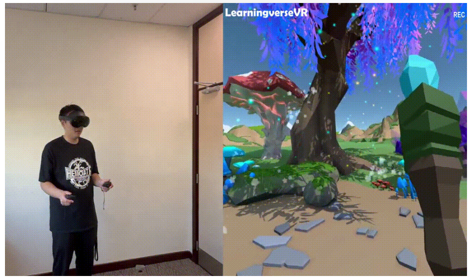


1) "Double Ninth Festival (重陽節)" in Learningverse

2) "Green Engergy" in Learningverse



3) "Robotics "VEX GO" in Learningverse



4) "Newton's Law of Motion" in Learningverse VR

5) Academic presentation in "AGILE" in Learningverse VR

6) Academic presentation with multimodal agent system in "AGILE" in LearningverseVR

4. Future work

Future work:

Intelligent digital human multimodal real-time interaction upgrade

1. Real-time speech-driven & affective response
   Establish a low-latency real-time communication framework enabling digital humans to converse with users while dynamically generating context-appropriate facial expressions and body movements.
2. Autonomous response system development
   Enable digital humans to achieve proactive dialogue initiation and autonomous interaction planning through dynamic decision-making logic architecture, transcending passive single-turn Q&A limitations.
3. Dynamic environmental adaptive mobility
   Enable digital humans to autonomously adjust behavioural logic based on virtual scene requirements (e.g., obstacle avoidance, social distancing).

Thank you!

Contact: Song Yanjie
Email: ysong@eduhk.hk