Developing a Cyber-Physical-Social Metaverse System for Interactive Cultural Heritage Experiences

Louis Nisiotis
School of Sciences
UCLan Cyprus
LNisiotis@uclan.ac.uk

Charalampos Nikolaou
School of Sciences
UCLan Cyprus
CNikolaou@uclan.ac.uk

Nikita Markov
School of Sciences
UCLan Cyprus
NMarkov@uclan.ac.uk

Aimilios Hadjiliasi
School of Sciences
UCLan Cyprus
AHadjiliasi1@uclan.ac.uk

***Abstract*—The Metaverse is significantly impacting the field of Digital Cultural Heritage, providing new opportunities for preserving, presenting, engaging, and interacting with historical and cultural artefacts through immersive virtual environments. This paper presents the development and implementation of an Intelligent Reality Virtual Museum prototype system for interactive digital cultural heritage experiences, within the context of a Cyber-Physical-Social System (CPSS) architectural framework for Metaverse applications development. The prototype integrates advanced Artificial Intelligence and Large Language Models in fusion with Digital Twins, Robotic and Gaming technologies to deliver immersive cultural heritage experiences that blend the real with the digital worlds. The paper presents the system architecture, system design, and development process, and the results of a user evaluation study assessing the usability and user engagement with the system during an interactive cultural heritage experience. The results revealed a usable system capable of offering an engaging and entertaining interactive experience to its users, highlighting the potential to be used for supporting and enhancing cultural heritage accessibility and engagement in our current digital age. System limitations and plans for future work are presented, focusing on refining and expanding the system capabilities towards the development of a new type of a Cyber-Physical-Social Eco-Society of systems for Metaverse applications to support and enhance the way users interact with and experience cultural heritage.**

***Index Terms*—Metaverse, Large Language Models, Artificial Intelligence, Digital Twins, Virtual Museums, Robots, Cultural Heritage.**

I. Introduction

Preservation of cultural heritage has always been of great importance to protect historical and artistic knowledge and achievements, and the evolution of human societies in general. The use of emerging digital technologies we have at our disposal today provide significant opportunities to disrupt how cultural heritage is curated, preserved, and experienced. With the introduction of the Metaverse in our lives, opportunities and possibilities arise for engaging with cultural heritage in virtual persistent spaces that blend the real and digital worlds interchangeably. However, the development of Metaverse environments and applications require advanced technological implementations. Among a plethora of technologies to support the development of Metaverse applications, Cyber-Physical-Social Systems (CPSS) are used, integrating physical elements with computational processes and social inputs to enable real-time and advanced human-machine interactions. Especially when combined with emerging technologies such as Artificial Intelligence (AI), eXtended Reality (XR), digital twins, robots, and gaming technologies in fusion, these systems can seamlessly blend the real with the virtual worlds, demonstrating the concept of the Metaverse, and are especially relevant in cultural heritage where immersive and interactive experiences can improve engagement and accessibility.

The work presented in this paper builds on previous work of the lead author on a CPSS-based architectural framework for Metaverse applications development [1], attempting to realise parts of its implementation and evaluating the system's usability and user engagement. Specifically, the paper focuses on the development and implementation of a prototype system that leverages Large Language Model (LLM) and symbolic AI techniques, digital twins, computer graphics and gaming technologies to deliver immersive cultural heritage experiences through the concept of an interconnected Intelligent Reality Virtual Museum Metaverse application. The prototype aims for a seamless integration of human, physical, and computational components, demonstrating how CPSS can be used for cultural heritage in engaging and interactive ways, and contribute towards the development of a Cyber-Physical-Social Eco-Society of systems powering the Metaverse.

II. Background and Context

A. Technology and the Metaverse in Digital Cultural Heritage

Cultural heritage plays an important role in shaping our social values, beliefs, and identities, and connects our societies and selves with our past. Organizations and stakeholders responsible for preserving cultural heritage aim to ensure that historical sites, exhibits, tangible and intangible artifacts remain accessible and informative to the public, but at the same time, they seek new ways to enhance and improve visitor engagement [2]. The latest advancements of digital technologies and software in AI, XR, computer graphics, Internet of Things (IoT), robotics, sensors, drones, digital twins, gaming technologies and other, have significantly impacted the domain of cultural heritage, providing opportunities for accessibility, transferability and preservation of historical artifacts, and are being key drivers for the field of Digital Cultural Heritage (DCH) [2], [3]. DCH focusses on the reconstruction, digitisation, and preservation of tangible and intangible historical artefacts, locations and buildings, cultural traditions and

2836-3795/25/$31.00 ©2025 IEEE
DOI 10.1109/COMPSAC65507.2025.00088

practices, and other relevant historical phenomena [4]. DCH practices require sophisticated data capture tools, specialised software and hardware, which are now more accessible due to technological hardware and software advancements and cost reductions [2]. Over the past decade for example, XR and AI have significantly impacted the DCH domain. XR technologies and software provide tools for capture, recreation, digitization, and access to immersive digital experiences in interactive forms, supporting user engagement and cultural understanding, while AI has been increasingly used in heritage conservation, translation, authenticity recognition, personalisation and other applications, significantly impacting how cultural content is interacted and preserved [5]. Furthermore, the recent advancements in Generative AI (GenAI) have had a profound impact on DCH. Models such as ChatGPT can extract information from large documents to create concise exhibition texts, audio guide scripts, exhibit labels, catalog entries, produce personalized storytelling, automatic artifact descriptions, and enable the development of real-time conversational agents [6]. Virtual agents for instance, converging XR and AI for mimicking human intelligence, are been increasingly used for enabling sophisticated social simulations and autonomous decision-making, with important applicability in DCH [2]. Moreover, digital twins is another significant technology that drew interest and is impacting DCH, providing the capability of digitization and deployment of dynamic virtual models of real-world entities updated in real-time, applied in visualization, preservation, and risk management among other DCH-related domains [7].

The introduction of emerging disruptive technologies into our lives has given rise to the concept of the Metaverse. The Metaverse is a significant evolution in digital experiences, offering persistent virtual environments that seamlessly integrate elements from the real and the digital worlds [8]. For cultural heritage, the Metaverse offers opportunities for preserving and presenting historical artifacts, landmarks, stories and narratives in immersive, interactive, and entertaining ways. The field sees the Metaverse as a vehicle for expression and exploration through human-centred approaches that provide users with new ways to interact with and experience digital worlds [9]. Cultural heritage stories evolve over time, reflecting changes in characters, location, scientific achievements, religious values and beliefs and many other tangible and intangible elements. The Metaverse can document and preserve these stories dynamically, capturing their progression across digital timelines and making them accessible and relevant for future generations [7]. Such method of cultural heritage presentation can be particularly appealing to younger generations who are influenced by modern technology, such as video games, high-speed internet, and realistic computer graphics, and who expect interactive and immersive experiences during their museum visits [2]. The Metaverse can help to preserve and present stories and knowledge that communities identify as part of their heritage through unique opportunities to document, digitize, and demonstrate them to support their longevity and continued significance. The advancements in technologies powering the Metaverse and attitudes towards acceptance [10] have enabled the development of innovative systems capable of realising and supporting multi-user, and human-to-machine interactions in immersive and interconnected blended spaces [1]. Particular types of complex computing systems that are used to realise and deploy Metaverse systems and applications are Cyber-Physical, and Cyber-Physical-Social Systems.

B. Cyber-Physical and Cyber-Physical-Social Systems

Cyber Physical Systems (CPS) integrate computational algorithms with physical components through interconnected infrastructures, smart devices, and physical environments capable of sensing, actuating and communication capabilities, to combine physical operations with digital information systems [11]–[14]. CPS have been successfully used in a plethora of domains such as automotive, manufacturing and other industries, and the field has a significant trend towards the integration and convergence of CPS with emerging technologies for advancing their capabilities [14]. However, in CPS humans are commonly assumed to be external entities interacting and providing inputs to the system [15]. CPS mostly focus on connecting the physical world with computational processes and the cyberspace, lacking mechanisms to incorporate human interaction and influence within the system [16]. They are inadequate of supporting Human-in-the-Loop systems [17], and especially for realizing Metaverse applications, human interaction and influence are key components [1].

Cyber-Physical-Social Systems (CPSS) are extending CPS by introducing the human and social dimensions into systems, making them the focus of significant academic and industrial interest with applicability in a plethora of domains [15], [18]. A CPSS is defined as a complex computing system *"comprising cyber, physical and social components, which exists or emerges through the interactions between those components. A CPSS comprises at least one physical component responsible for sensing and actuation, one cyber component for computations and one social component for actuating social functions"* [15]. CPSS enable to establish social connections and support human communication and exchange of experiences, incorporating the concept of social computing into complex computing systems, and facilitating social collaboration, interaction, and knowledge to develop social dynamics [19], [20]. A true CPSS emerges when smart devices interact with humans, effectively integrating social aspects into the system [15].

CPS and CPSS have been adopted to support cultural heritage in multiple ways with promising results over the past few years. Previous work on CPS and CPSS has demonstrated functional examples of how to develop Intelligent Reality Virtual Museums capable of seamlessly blending the real with the virtual worlds, agents and elements through the convergence of XR, AI and Robots [2], [10], [21], heritage building management and to support conservation [22], personalisation though modeling and orchestrating the interactions between a museum and its visitors [23], and to handle human behaviour complexity to establish Human-Machine Interaction [24] among other.

The convergence of technologies within CPS and CPSS have led to the emergence of a Cyber-Physical-Social Eco-Society of Systems (CPSeS) [2], [25], towards the realisation of interconnected network of systems, enabling seamless interactions between real-world physical components, digital environments, and social elements, bringing us closer to the vision of the Metaverse. However, despite the potential of CPSS to develop Metaverse systems and applications, designing such systems presents challenges due to the complexity of integrating heterogeneous networks, hardware, and software [1]. Additionally, incorporating human and social influences into such systems remains a significant obstacle, with limited methodologies available to address these issues effectively [26]. From an engineering perspective, the Metaverse employs key features of CPSS and the concept is fundamentally relevant, since CPSS are integrating physical, artificial, and mental worlds, operating in physical and cyber domains, and influenced by users and their behaviour [17].

C. LLMs in Digital Cultural Heritage

With the massive attention drawn to the field of Generative AI as a result of the significant recent advancements (i.e. ChatGPT, DALL-E etc.), one of the specific research domains presenting important innovations focusses on integrating Large Language Models (LLMs) into chatbots and virtual agents for real-time response generation and human-AI conversations in virtual environments [27]–[30]. Especially within the field of DCH, LLM-powered chatbots demonstrate the ability to create immersive and educational experiences, offering significantly improved usability and engagement compared to traditional chatbot systems which their conversations and interactions are predefined and scripted [31]. Furthermore, recent research sees the successful development and utilisation of LLM-enabled CPS [32]. Examples include the use of conversational agents within CPS [33], systems designed for natural language interaction with users [34], and systems leveraging logic and reasoning capabilities to aid decision making [35], among others [32]. LLM-powered CPS are categorized as: i) *Assistant*, supporting CPS by processing data, grounding context, and facilitating interaction with the external world through natural language, images, and other inputs, enabling perception and communication capabilities; and ii) *Brain*, acting as decision-makers, analyzing and organizing information to plan and reason based on pre-trained knowledge, to enable CPS to perform complex planning and reasoning tasks [32].

Even though there are many examples of successful implementation on LLMs in DCH and other fields, there is a notable gap in the literature and practical applications where LLMs are leveraged beyond conversational and decision-making tasks, and especially within CPSS. In particular, limited research explores the use of LLMs for real-time information generation that dynamically influences system behaviour, environment content design, and driving the overall virtual experience. Most implementations focus on integrating LLMs with Non-Playing Characters (NPC), virtual agents, and chatbots, overlooking the potential of using LLMs as a technology to drive intelligent systems that adapt to user interactions and environmental factors. The prototype presented in this paper aims to address this gap by presenting a work in progress towards a practical example of integrating LLMs into a CPSS, contributing to the vision of integrating AI-driven systems into the Metaverse.

III. System Architecture Framework

A system architecture framework aiming to provide the foundation for deploying such advanced capabilities, has been previously published by the lead author [1]. This framework incorporates AI-driven decision-making, CPSS integration, symbolic and generative AI to support seamless interactions between physical artifacts, digital agents and elements, and user inputs to create a new type of a Cyber-Physical-Social Eco Society of systems for Metaverse applications (Fig. 1).

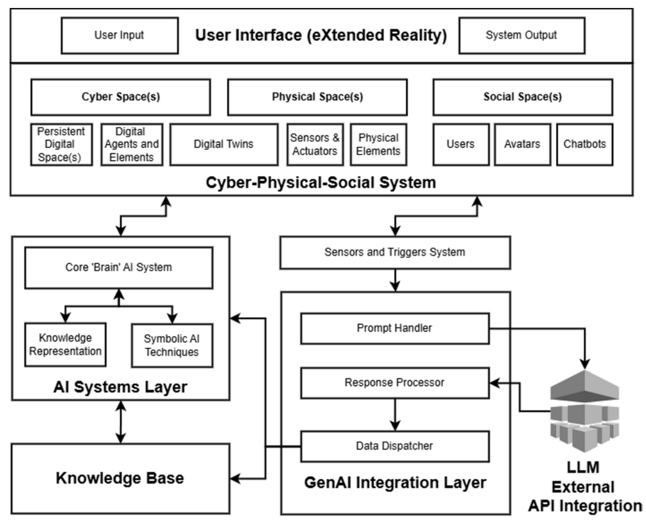

Fig. 1. AI-Driven CPSS Architecture [1].

The proposed framework is built on a CPSS architecture with multiple layers to enable immersive, real-time interactions in Metaverse applications. The *User Interface Layer* provides immersive environments through XR technologies, allowing real-time user input. The *CPSS Integration Layer* connects physical, digital, and social components, integrating sensors, actuators, and digital twins for cyber-physical synchronization and incorporating users and avatars for social interactions. Orchestration is managed by the *AI Systems Layer*, with a central AI 'brain' driving decision-making and dynamically adjusting system behaviour based on user inputs and environmental data. The *GenAI Integration Layer* uses APIs to connect high-performance AI tools, generating contextual content triggered by a *Sensors and Triggers* subsystem that exists in the real and digital spaces. Finally, the *Knowledge Base Layer* serves as a central repository for system data, supporting information sharing across components. The prototype presented in this paper demonstrates the framework's potential for delivering adaptive cultural heritage experiences within the context of an Intelligent Reality Virtual Museum.

IV. Prototype System Development

The prototype is a work in progress system based on previous work [1] aiming to create an interactive and immersive cultural heritage Metaverse experience. It is developed using Unity3D to design the digital environment and implement its key features (Fig. 2). Currently, the main emphasis is to develop, test the system functionalities, and initially evaluate user experience, therefore, it runs on Windows PCs for now, with future plans to introduce XR hardware support and functionalities. The system supports multi-user interaction through the free version of Photon Engine Software-as-a-Service that allows up to 20 users to interact in the virtual world simultaneously (photonengine.com). The Photon's Voice over IP (VoIP) service and a custom instant messaging system were also implemented to support social interaction.

Fig. 2. Example screenshots of the virtual museum environment.

A. Robotic Digital Twin Integration

The environment features the digital twin of the TurtleBot2 (See Fig. 3). This is a mobile research robot placed within the physical environment, performing actuations in the real world which are simulated and shadowed in the digital space in real time. The role of the robot in the system is to facilitate real time information exchange between the real and digital spaces, demonstrating how such types of agents can be placed either in open, inaccessible, or restricted physical locations, serving as a mobile guide that autonomously navigates to different exhibits and delivering contextual educational information to users. TurtleBot2 is equipped with a high-resolution camera, sensors, a laser scanner for navigation and obstacle avoidance, and a small mechanical arm for object handling. TurtleBot2 runs on the Robotic Operating System (ROS) open-source framework that offers tools, libraries, and visualization features for advanced robotic functionalities [36].

To enable Cyber-Physical communication and synchronization between Unity and the Robot, the ROS-Unity integration plug-in available through the Unity Robotics Hub was utilised [37]. The plug-in provides the ROS-TCP Connector Unity component responsible for sending and receiving messages from ROS. This connector is paired with the ROS-TCP Endpoint, which serves as the communication node within ROS. Both components use the TCP/IP protocol for data transmission. The plug-in also supports importing robot models using Universal Robot Description Format (URDF) files, allowing the creation of 3D digital twins of physical robots in Unity, enabling synchronous updates by exchanging messages for key data, such as position updates, object detection, and interaction commands. For example, as the TurtleBot2 moves in the real world, its digital twin in Unity replicates its movement in real time, and also live streams the camera feed from the eyes of the robot to a plane attached on the digital twin within the digital space, connecting the two realities. The implementation is currently established through a local network connection, and we aim to expand the system to support remote user access through 5G wireless technology.

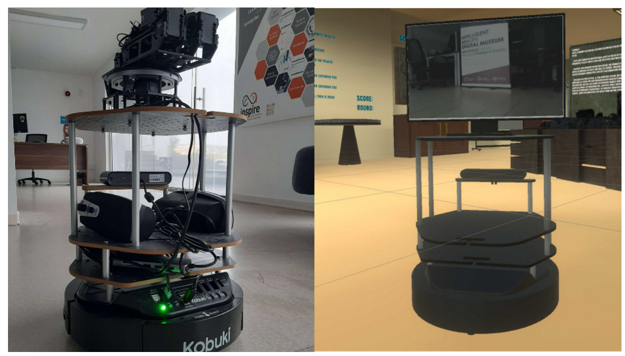

Fig. 3. The real and Digital Twin of the TurtleBot2

B. Generative AI Integration

To deploy the GenAI component of the system, several factors had to be taken into consideration. Modern LLMs have shown great capabilities in conversational tasks and chatbot applications [38], but their effectiveness is limited by several key factors such as reliance on internet-based training data, outdated knowledge, and accuracy limitations, among others [39]. Using commercial services such as ChatGPT incurs costs [40], which depending on the usage can become significantly high. Furthermore, running fully fetched locally hosted LLMs instances has expensive requirements in terms of cost, hardware, and infrastructure, as well as advanced technical knowledge [41]. However, the costs and powerful infrastructure requirements barriers have been significantly lowered with smaller-scale custom-built solutions. Several open-access engines and models are available for use, where developers can deploy locally and experiment with. Notably, NVIDIA also introduced the 'Chat with RTX' (ChatRTX) platform, which enables users to develop custom LLMs and personalized AI chatbots on local machines without the need for advanced technical expertise or expensive hardware [42]. For the needs of this prototype, a performant solution was required that would be managed and controlled internally to avoid excessive costs and uncontrolled actions, data sharing, and external service reliance.

1) Deploying a LLM-Powered Server: To deploy a local LLM model, we have conducted several tests with different engines and models. For our project prototype, it was essential to develop a Client-Server architecture to ensure that all current and future devices running the CPSS virtual museum could connect to the system and enable dynamic communication and interaction between components, as we plan to extend compatibility to include XR technology. Therefore, it was necessary to establish data exchange between Unity and the LLM, to generate personalized responses and provide adaptive content, through personalised prompts formulated based on users interactions within the virtual world. The aim was to support functionalities such as real-time exhibit descriptions, virtual guide dialogues, and context-specific knowledge generation to support the user experience.

Initially, we experimented with ChatRTX version 0.4 (installer size - 11 GB), downloaded from the NVIDIA website [43]. By default, user interactions with the ChatRTX system are provided through a web interface [44], and to bypass the standard interface and enable direct requests from Unity, we have customised a publicly available 'Python API for Chat With RTX' implementation [45]. While the ChatRTX was performing well, we encountered technical difficulties in extending the system with custom components for additional capabilities, and we decided to proceed with open-source LLM engines and models instead. Among different runtime engine environments, *Ollama* engine (ollama.com/) was selected for deploying and executing our locally hosted LLM, providing flexibility and control on the model's deployment without relying on external APIs or cloud-based processing. We have tested different Large and Small Language Models to identify a model that would offer appropriate and performant responses. After several tests, we decided to adopt *Mistral 7B*, which demonstrated high processing speed and relatively low memory consumption.

The LLM was running on a PC operating Microsoft Windows 10, equipped with an Intel i5-12650H CPU, NVIDIA RTX 4060 GPU (8 GB GDDR6 VRAM), 16 GB DDR5 RAM, and 2 TB SSD storage. We opted for a medium-spec server to initially evaluate the system's performance under limited resources, providing a baseline for optimization and scalability assessments before deploying the system on higher-capacity infrastructure. Currently, the system is running in a local network configuration, but future development aims to support remote connections.

2) Establishing the LLM-Unity Communication: To establish the communication interface between the LLM server and Unity, the *LangChain* was used to set up the client-server infrastructure for establishing incoming and outgoing communication requests. LangChain (langchain.com/) is a framework for building LLM-based applications that organizes the process into chains linking the steps of generating prompts, processing text, and managing context (e.g., message history). A Flask-based Python server was deployed, responsible for: i) receiving user requests from Unity over the network through POST requests in JSON format using the UnityWebRequest class; ii) processing the input and forwarding it to the LLM; iii) retrieve the responses from the LLM; iv) and return the generated output to Unity where the *Response Processor* component (See Fig. 1) is responsible for handling the information that is fed back into the system.

3) Enabling Conversation Memory: One of the key requirements of the LLM-powered system was to remember previous user exchanges to provide personalised and relevant responses to the system. To achieve this, a cloud-based *Upstash Redis* serverless database service (upstash.com/) was used to conduct initial testing. A session-based memory storage system was implemented, allowing the model to retain past messages and generate responses that consider previous interactions. On our immediate future plans, we aim to replace this cloud-based solution with our own local database system. However, given the sensitivity around storing conversational data and the potential implications for user privacy, special attention has been given to data protection. We currently store conversation memory temporarily during active sessions, deleting them once the session ends, ensuring no long-term storage of personal user data. In the future, users will be provided the option to delete their conversation history at any time during their session or after. Our immediate future plans include replacing the cloud-based solution with a locally hosted database system to give us full control over data handling.

C. Environment Design

1) Digital Twin of a Real Setting: The main scene of the virtual environment replicates a 100m^2 computer lab at the University of Central Lancashire, Cyprus (UCLan Cyprus), serving as the demonstrator test-bed for the initial deployment and evaluation of the virtual museum prototype in a controlled but realistic and manageable setting. In the virtual environment, digital exhibits represent historical artifacts and sites for users to explore and interact with. Dummy artifacts are positioned within the physical space to simulate real-world exhibits, to demonstrate how a real physical museum may layout and display its exhibits. To accurately model the physical space, the TurtleBot2 robot was used to generate a 2D map using the gmapping laser-based method. The map was subsequently used in Unity to reconstruct the space in 3D (Fig 4). Unity's built-in level design tools were then employed to refine the digital environment further for a near-accurate representation of the physical space.

2) Cultural Heritage Exhibits Development: The environment features digital reconstructions of indicative significant historical locations and artifacts from Cypriot cultural heritage including:

- The exterior building layout of the church of Panayia Aggeloktisti (Fig 5a), its famous mosaic located inside the church (Fig 7), its iconostasi, and two framed paintings (Fig 6);
- Prophet Elias Church on the top of Profitis Elias mountain (Fig 5b);

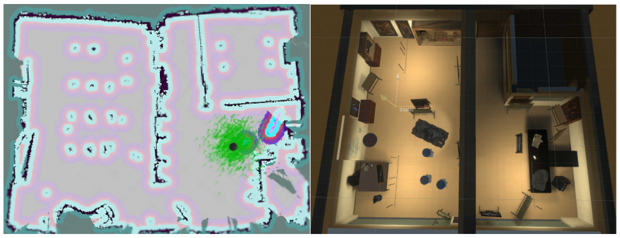

Fig. 4. Digital twin process of the physical lab space

- Ayios Mamas Church ruins in the deserted village of Ayios Sozomenos (Fig 5c);
- The exterior building layout of the UNESCO protected heritage site of Panayia Asinou (Fig. 5d).

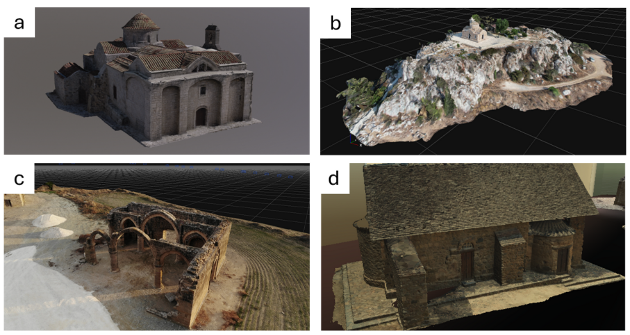

Fig. 5. Digital Reconstructions: a) Church of Panayia Aggeloktisti, b) Prophet Elias Chapel, c) Ayios Mamas Ruins, d) Church of Panayia Asinou.

Fig. 6. Digital reconstructions of Panayia Aggeloktisti's mosaic, iconostasi, and paintings.

To develop the 3D reconstructions, a team of undergraduate Computer Games Development students from UCLan Cyprus under the supervision of the lead author, conducted multiple visits to the various locations to capture detailed images of their exterior and key interior artifacts. High-resolution photographic data was collected using drones with HD cameras. The collected photographs were processed using 3DF Zephyr photogrammetry software (3dflow.net) to generate detailed 3D models. These models are initially generated with high polygon counts and texture inconsistencies, and they were then optimized using a chain of various CAD tools (see example process in Fig. 7). These optimization steps improved loading times, rendering efficiency, and helped to maintain consistent frame rates. Additional open-access 3D assets were used to further improve the visual design of the environment.

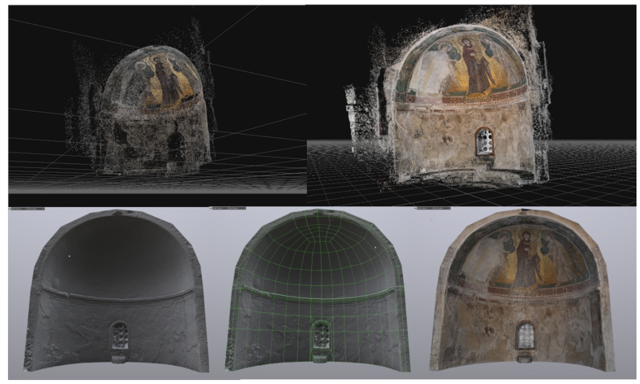

Fig. 7. Examples of the process of reconstructing the Mosaic of the Panayia Aggeloktisti Church starting from generating a sparse point cloud, followed by the generation of the dense point cloud and 3D mesh, and later processed for geometry decimation, retopology and texturing.

3) Environment AI Functionalities: The functionality and behaviour of the Unity environment, its encompassing agents, and gameplay elements are driven by a number of AI subsystems. The environment is structured around a central 'brain' system component, which oversees the system's decision-making and coordination. It manages several Blackboard AI subsystems used as shared data structures to store and organize information about users, environment components, and system states. The environment employs a grid-based approach to divide the virtual space into different sections, facilitating tracking of user activities and their interactions.

The environment has specific trigger areas that are linked to the 'Brain' AI, and in particular the *Prompt Handler* specific subcomponent responsible for dynamic prompt formulation based on user-specific information retrieved from the blackboards to request personalized responses from the GenAI. An example of a formulated prompt is: *"User [UserID] is now in [AreaID], which features exhibits about [ExhibitTopic]. They previously explored [PreviousAreas] and interacted with [SpecificExhibits]. Generate a brief description of the current exhibits, including notable historical facts and connections to their earlier exploration in the following format: [Context], [Fact], [Suggested Action]"*. The responses are processed by the *Response Processor* component and parsed onto the *Data Dispatcher* responsible for feeding them back into the system. Currently, the information is shown in informational panels within the environment and stored within the central *Knowledge Base* of the system, and we aim to develop AI-driven conversational agents in the future.

4) Educational Content and Interactive Activities: The virtual museum is designed to deliver an engaging educational experience through a combination of informational content and interactive activities. Informational signs are placed around in the environment near historical exhibits to provide educational information and exhibit descriptions (Fig. 2). The panels are dynamically populated with real-time educational content generated by the LLM adapted based on user interactions for personalized learning. Additionally, the environment features several entertainment-focused mini-games designed to enhance user engagement. Upon successfully completing these mini-games, users unlock access to further educational materials, aiming to blend entertainment and learning to create an immersive and rewarding experience (Fig. 8).

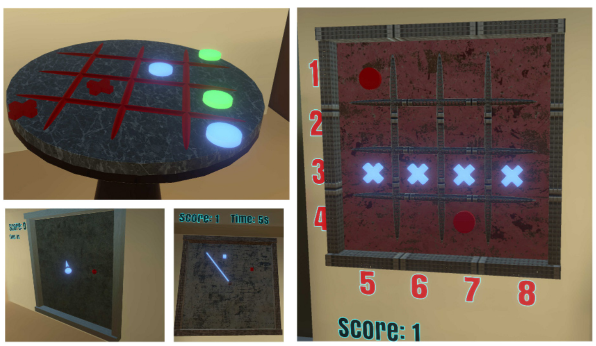

Fig. 8. Examples of mini-games in the environment.

V. Usability and User Engagement Evaluation

A. Study Design and Experimental Procedures

We evaluated the usability and user experience of the system through the System Usability Scale (SUS) [46], [47] and the User Engagement Scale (UES) [48]. SUS is a validated and straightforward tool that assesses usability based on user interaction and satisfaction. It consists of 10 statements, alternating between positive and negative wording, that measure effectiveness (can users achieve their objectives), efficiency (how much effort is required), and satisfaction (overall experience). Users rate each statement on a 1 to 5 scale (1 = Strongly Disagree, 7 = Strongly Agree), with scores adjusted for positive or negative phrasing. The SUS score is calculated by summing the adjusted scores and multiplying by 2.5, yielding a score between 0 and 100, with an average score of 68 indicating acceptable usability. The UES is rated in 1-7 point Likert scale (1 = Strongly Disagree, 7=Strongly Agree) and evaluates the user engagement in the context of interacting with a digital system based on four key attributes: i) Focused Attention (FA), measuring how absorbed individuals are interacting with a system, ii) Aesthetic Appeal (AA), measuring the attractiveness and visual appeal of the system's interface, iii) Perceived Usability (PU), measuring the degree of control and effort expanded in the system, and iv) Perceived Reward (PR) exploring how satisfying and rewarding is the experience provided by the system. The scale also provides the Overall Engagement score (OE), which is the result of aggregating the results of all items, and dividing the sum by the number of the scale items [48].

An open call for participation was announced at UCLan Cyprus, and 26 participants (22 Male, 4 Female) between 19 and 37 years old (M = 23, SD = 3.87) agreed to join the system evaluation experience. Participants were highly familiar with desktop computers (M = 4.5, SD = 0.65, out of 5) with diverse gaming habits. Over half (53.8%) played regularly (>= 4 times per week), with 26.9% gaming daily. However, 30.8% played very rarely, and 15.4% only once per week. Most participants (38.5%) played for 1-3 hours per day, while 3.8% spent more than 10 hours gaming daily. Weekly gaming frequency varied, with 15.4% playing more than 10 times per week, while 3.8% did not play at all.

Prior to engaging with any of the study materials, they were briefed on the study's purpose and its objectives, and their informed consent was sought. Participants were each allocated a desktop PC in a computer lab, and experienced the environment in groups of 3 for each session, in the presence of a researcher. Participants were requested to join the virtual world, navigate the environment and interact with the exhibits, and read educational information provided by the system, either static or dynamically through the LLM server based on their interactions. They were requested to engage and complete the mini games, exchange messages with each other, and also interact with the Digital Twin of the TurtleBot2, seeing it moving and providing live video feed in the virtual world, while at the same time it was moving around in the physical space. The sessions lasted 25-35 minutes for each group. After their experience with the environment, they were asked to complete the SUS, and UES questionnaires.

B. Results

Before conducting any statistical analyses, data normality was examined using the Shapiro-Wilk test, which indicated that the distribution of SUS and the overall UES score were approximately normally distributed. Therefore, parametric descriptive statistics using Means (M) and Standard Deviations (SD) and correlational tests using Pearson's correlation were employed. The SUS score revealed a mean score of 74.33 (SD=14.98), which is higher than the 68 threshold score, indicating satisfying usability of the overall system. The UES results (Table I) revealed that participants have positively perceived their Overall Engagement experience with the system (OE: M = 5.29, SD = .94), reporting high level of ease, control and usability (PU: M = 5.75, SD = 1.19) corroborating the SUS results. They expressed relatively good level of absorption in the experience (FA: M = 4.64, SD = 1.46), reported high levels of interface attractiveness and visual appeal (AA: M = 5.26, SD = 1.07), and high level of reward and satisfaction (RW: M = 5.52, SD = .99). The findings suggest that the experience provided by the system was well-perceived, with users finding it engaging, absorbing, usable, and visually appealing. The results were further explored for correlations between the

SUS and UES, using Pearson's correlation test. The results indicated strong positive correlations between SUS scores and Overall Experience (OE: $r(24) = .782$, $p < .001$), Usability (PU: $r(24) = .736$, $p < .001$), Reward (PR: $r(24) = .747$, $p < .001$) and Aesthetic Appeal (AA: $r(24) = .637$, $p < .001$). Focused Attention displayed a moderately positive correlation with SUS score (FA: $r(24) = .438$, $p = .025$). The findings suggest significant links between the user engagement factors and perceived system usability.

TABLE I
DESCRIPTIVE STATISTICS FOR USER ENGAGEMENT SCALE

Factor	Mean	SD	Min	Max
Focused Attention	4.64	1.46	1.00	7.00
Perceived Usability	5.75	1.19	3.33	7.00
Aesthetic Appeal	5.26	1.07	2.67	7.00
Reward	5.52	0.99	2.00	7.00
Overall Engagement	5.29	0.94	2.58	6.75

Following the completion of the post-experience questionnaire, participants were informally interviewed about their experience with the system. Most of the participants found value in using the system for learning, with comments such as: *"I believe I could use this technology to learn, if the content is interesting.."*, *"...it is a fun way to learn by playing games and playing with a robot which I can control over a game"*, and *"I think this is a different method to learn new information"*. Furthermore, the majority expressed positive perceptions towards the idea of playing mini-games to unlock new information. Some indicated that the digital entertainment aspect of the environment played a key role in their engagement with the system, making the experience more interactive and rewarding. Participants noted that the gaming approach helped maintaining their interest and explored the virtual museum more actively. Earning access to content through mini-games provided a sense of accomplishment, and they also indicated that the inclusion of entertainment elements differentiated the experience from traditional museum visits, making it an enjoyable way to engage with cultural heritage.

However, some concerns were expressed considering the context presented in the environment i.e. *"I don't know if the context and the mechanics provided are matching"*, indicating that a closer alignment with the mini games and the educational content should be considered. Regarding the use of the robot in the environment, participants appreciated its functionalities but expected more meaningful involvement in the overall experience: *"I like the idea that I can interact with the robot, but its usage should be more clear"*, indicative of the current stage of system development, where the robot is currently only used for live video feed and pathfinding movement in the physical space.

During the initial evaluation of the system, occasional instances of information inaccuracies and hallucinations were observed in the LLM-generated responses. While the system successfully delivered personalized and context-aware content in most cases, there were moments when the LLM provided responses that were either partially incorrect or plausible-sounding but inaccurate. These occurrences were more noticeable when users interacted with areas requiring highly specific cultural knowledge. Although not all users immediately identified these inaccuracies, their presence highlights a key limitation of the current system, and it is in our immediate future plans to explore this further.

VI. DISCUSSION

The introduction of the Metaverse has impacted multiple industrial, educational and social domains as well as our professional and daily lives. The field of cultural heritage in particular has benefited, especially by leveraging Metaverse applications for experiencing immersive interactive DCH experiences. The deployment of complex computing systems to power such applications requires extensive technical knowledge, expensive hardware requirements and detailed guidelines, and their widespread deployment is still at its infancy.

The prototype system presented in this paper represents an example of a CPSS Metaverse application, demonstrating the convergence of multiple emerging technologies to enhance the way cultural heritage is experienced, preserved, and interacted in the Metaverse. The convergence of LLMs, digital twins, gaming technologies, and AI-driven interaction mechanisms in fusion as demonstrated in this system, provides capabilities for developing an intelligent and immersive cultural hub that actively engages users in immersive and interactive ways. The integration of a real robot in the physical space and its virtual counterpart in the digital world demonstrates how cyber-physical agents can create new interactive possibilities for museums and educational spaces, blending real-world presence with digital storytelling, and enabling opportunities for users to access remote, inaccessible or restricted locations through virtual worlds. The convergence of technologies and software components work together to deliver an engaging and dynamic experience that responds and adapts to user interactions, to provide a more customized cultural heritage experience than traditional museum visits with pre-defined museum content. The system leverages multi-user interaction capabilities, where users can explore the virtual museum individually or collaboratively, interact with robotic digital twins, virtual agents and with each other, establishing the social computing aspect of CPSS. The system also leverages the affordances of mini-games to support engagement and digital entertainment and make the experience interactive and stimulating.

The system usability and user engagement evaluation results indicated that the prototype successfully engages users and delivers a usable experience. Users revealed positive perceptions towards the system usability, the reward mechanisms, and its aesthetic appeal, and of the system's ability to engage and sustain their interest throughout the experience. They also emphasised that the digital entertainment aspect of the experience played a key role in maintaining their engagement, highlighting its importance in sustaining engagement in virtual museums, particularly for younger audiences who are familiar with interactive and game-like experiences. Furthermore, the integration of robots and their digital twins in this system creates the premises for developing innovative and highly contextualized interactive experiences that blend physical and virtual environments. The robots can be placed to operate in physically restrictive areas like fragile archaeological sites, inaccessible locations, or sensitive museum zones, and perform different tasks, and real-time data mirroring and bidirectional interaction to be facilitated by their digital twins within a virtual setting. This type of Intelligent Reality Virtual Museum can further support existing DCH practices by creating new ways of interacting with cultural heritage, offering opportunities for immersive, interactive and entertaining digital experiences, with further applicability in a plethora of domains. The prototype demonstrator and the framework implementation presented in this paper provide a blueprint and practical example for further Metaverse systems application development.

VII. Conclusions, Limitations and Future Work

The CPSS-based Metaverse application prototype presented in this paper and its system usability and user engagement evaluation highlight the potential of converging AI, digital twins, and gaming technologies to develop engaging interactive cultural heritage experiences. The prototype offers adaptive, interactive, and multi-user interactivity, demonstrating a practical example of how CPSS Metaverse systems move beyond static representations of cultural heritage to create an Intelligent Reality Virtual Museum.

One key limitation is that the system is currently restricted to a local network environment, with both the LLM server and the robot operating within the same local network. This limits accessibility for remote users and prevents further testing. The *GenAI Integration Layer* currently under development is hindered by technical challenges impacting its functioning. The *Prompt Handler* component requires further refinement to improve prompt formulation, and the *Response Processor* to improve processing relevance. The *Data Dispatcher* also needs refinements to support the efficient distribution of generated content to agents and other interactive elements within the environment. Another limitation is that the LLM server occasionally experiences response delays and in some occasions it generates inaccurate information. The effect of hallucination has been observed, where the model generates plausible but factually incorrect or misleading information that negatively affects the user experience and the trust on the systems contextual knowledge. Furthermore, while the initial usability and engagement evaluations were promising, the small sample size and lack of representation from diverse user groups with varying experiences highlight the need for further validation with a wider audience. Additionally, there is a need to better align the gaming and educational components within the environment. Even though the developed mini-games have improved engagement, it is important that the interactive experiences and gameplay support the learning outcomes more directly. Lastly, while the current prototype demonstrates the feasibility of fusing real and virtual environments, it currently operates at a limited scale and is hindered by challenges related to scalability, including potential latency, computational overhead, and networking issues that should be explored further under higher loads and in remote multi-user conditions.

To address some of these challenges and further enhance the system, future developments will focus on expanding system accessibility by deploying the system on an open network with 5G connectivity, allowing remote users to engage with the virtual museum, the robot and the LLM server. Stress-testing under various conditions is planned to evaluate system and network performance to assess scalability. Additionally, exploration of alternative LLMs, Small Language Models (SLMs), and long-term memory storage solutions will be explored to improve response accuracy, efficiency, and contextual awareness. We plan to conduct a detailed evaluation of different LLM models and their generated responses, focusing on their reliability, factual correctness, and system performance through expert review and user feedback for a systematic assessment, as well as how users perceive and react to instances of AI-generated inaccuracies. Future work also aims at incorporating direct user chat interactions with the LLM server for real-time conversational exchanges. The system will also be extended in the near future to include XR integration, enabling immersive interactive experiences, and evaluated using a larger and more diverse sample. More artifacts, digital exhibits and heritage stories will be incorporated, along with the introduction of intelligent virtual agents capable of dynamic interactions, driven by the LLM server. Future work will also focus on expanding game-based elements, such as quests, AI-driven adaptive challenges and multiplayer collaboration to further support the entertainment and educational values of the system. Further evaluations will be conducted to assess user experience, system usability in XR settings, user acceptance, overall system performance, LLM-generated content reliability, and security aspects of the system. The intended improvements and future research plans aim to bring the system development closer to realizing a fully intelligent, immersive, and interconnected cultural heritage Metaverse application, to support and enhance the way users interact with and experience cultural heritage.

Acknowledgment

In memory of our dear friend, mentor, and collaborator Dr Lyuba Alboul.

References

[1] L. Nisiotis, "A system architecture for AI-driven interactive systems within a CPSS framework for metaverse applications," in *2024 IEEE International Symposium on Emerging Metaverse (ISEMV)*, pp. 9–12, 2024.

[2] L. Nisiotis, A. Hadjiliasi, F. Alexandrou, and L. Alboul, "Interwoven Spaces with XR, AI, and Robots: Merging Realities in Space and Time," in *Museums and Technologies of Presence*, pp. 243–261, Routledge, 2023.

[3] E. Markopoulos, P. Markopoulos, M. Liumila, Y. Almufti, and C. Romano, "Digital cultural strategies within the context of digital humanities economics," in *Advances in Human Factors in Wearable Technologies and Game Design* (T. Ahram, ed.), (Cham), pp. 283–295, Springer International Publishing, 2020.

[4] L. Pujol and E. Champion, "Evaluating presence in cultural heritage projects," *International Journal of Heritage Studies*, vol. 18, no. 1, pp. 83–102, 2012.

[5] Z. Stojanović and E. Čajić, "Application of artificial intelligence in cultural heritage," *Thematic Proceedings*, p. 333, 2024.

[6] D. H. Spennemann, "Generative artificial intelligence, human agency and the future of cultural heritage," *Heritage*, vol. 7, no. 7, p. 3597, 2024.

[7] X. Zhang, D. Yang, C. H. Yow, L. Huang, X. Wu, X. Huang, J. Guo, S. Zhou, and Y. Cai, "Metaverse for cultural heritages," *Electronics*, vol. 11, no. 22, 2022.

[8] M. Y. Almoqbel, A. Naderi, D. Y. Wohn, and N. Goyal, "The metaverse: A systematic literature review to map scholarly definitions," in *Companion Publication of the 2022 Conference on Computer Supported Cooperative Work and Social Computing*, CSCW'22 Companion, (New York, NY, USA), p. 80–84, Association for Computing Machinery, 2022.

[9] C. Innocente, F. Nonis, A. Lo Faro, R. Ruggieri, L. Ulrich, and E. Vezzetti, "A metaverse platform for preserving and promoting intangible cultural heritage," *Applied Sciences*, vol. 14, no. 8, 2024.

[10] L. Nisiotis and L. Alboul, "Evaluating presence and technology acceptance of an intelligent reality virtual museum prototype," in *2022 IEEE 2nd International Conference on Intelligent Reality (ICIR)*, pp. 1–6, IEEE, 2022.

[11] Q. Chang, R. Gao, Y. Lei, L. Wang, C. Wu, *et al.*, "Cyber-physical systems in manufacturing and service systems," *Mathematical Problems in Engineering*, vol. 2015, pp. 1–2, 2015.

[12] A. Ghosh, D. Chakraborty, and A. Law, "Artificial intelligence in internet of things," *CAAI Transactions on Intelligence Technology*, vol. 3, no. 4, pp. 208–218, 2018.

[13] J. Lee, B. Bagheri, and H.-A. Kao, "Recent advances and trends of cyber-physical systems and big data analytics in industrial informatics," in *Proceedings of international conference on industrial informatics (INDIN)*, pp. 1–6, 2014.

[14] N. Singh, P. K. Panigrahi, Z. Zhang, and S. M. Jasimuddin, "Cyber-physical systems: a bibliometric analysis of literature," *Journal of Intelligent Manufacturing*, pp. 1–37, 2024.

[15] B. A. Yilma, H. Panetto, and Y. Naudet, "Systemic formalisation of cyber-physical-social system (CPSS): A systematic literature review," *Computers in Industry*, vol. 129, p. 103458, 2021.

[16] L. Monostori, "Cyber-physical production systems: Roots, expectations and r&d challenges," *Procedia CIRP*, vol. 17, pp. 9–13, 2014.

[17] X. Wang, J. Yang, J. Han, W. Wang, and F.-Y. Wang, "Metaverses and demetaverses: From digital twins in CPS to parallel intelligence in CPSS," *IEEE Intelligent Systems*, vol. 37, no. 4, pp. 97–102, 2022.

[18] J. J. Zhang, F.-Y. Wang, X. Wang, G. Xiong, F. Zhu, Y. Lv, J. Hou, S. Han, Y. Yuan, Q. Lu, and Y. Lee, "Cyber-physical-social systems: The state of the art and perspectives," *IEEE Transactions on Computational Social Systems*, vol. 5, pp. 829–840, Sep. 2018.

[19] S. Pasandideh, P. Pereira, and L. Gomes, "Cyber-physical-social systems: taxonomy, challenges, and opportunities," *IEEE Access*, vol. 10, pp. 42404–42419, 2022.

[20] H. Ning, H. Liu, J. Ma, L. T. Yang, and R. Huang, "Cybermatics: Cyber–physical–social–thinking hyperspace based science and technology," *Future Generation Computer Systems*, vol. 56, pp. 504–522, 2016.

[21] L. Nisiotis and L. Alboul, "Initial evaluation of an intelligent virtual museum prototype powered by AI, XR and robots," in *8th International Conference on Augmented Reality, Virtual Reality, and Computer Graphics*, pp. 290–305, Springer, 2021.

[22] G. Nota and G. Petraglia, "Heritage buildings management: the role of situational awareness and cyber-physical systems," *Journal of Ambient Intelligence and Humanized Computing*, vol. 15, no. 4, 2024.

[23] N. Patrizi, S. K. LaTouf, E. E. Tsiropoulou, and S. Papavassiliou, "Museum and visitor interaction and feedback orchestration enabled by labor economics," *IEEE Transactions on Computational Social Systems*, vol. 10, no. 5, pp. 2870–2881, 2023.

[24] Y. Naudet, B. A. Yilma, and H. Panetto, "Personalisation in cyber physical and social systems: the case of recommendations in cultural heritage spaces," in *2018 13th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)*, pp. 75–79, IEEE, 2018.

[25] L. Nisiotis, L. Alboul, and M. Beer, "A prototype that fuses virtual reality, robots, and social networks to create a new cyber–physical–social eco-society system for cultural heritage," *Sustainability*, vol. 12, no. 2, p. 645, 2020.

[26] J. Zeng, L. T. Yang, M. Lin, H. Ning, and J. Ma, "A survey: Cyber-physical-social systems and their system-level design methodology," *Future Generation Computer Systems*, vol. 105, pp. 1028–1042, 2020.

[27] J. Li, X. Zheng, I. Watanabe, and Y. Ochiai, "A systematic review of digital transformation technologies in museum exhibition," *Computers in Human Behavior*, vol. 161, p. 108407, 2024.

[28] S. Varitimiadis, K. Kotis, A. Skamagis, A. Tzortzakakis, G. Tsekouras, and D. Spiliotopoulos, "Towards implementing an AI chatbot platform for museums," in *International conference on cultural informatics, communication & media studies*, vol. 1, 2020.

[29] J. Jeon, S. Lee, and H. Choe, "Beyond chatgpt: A conceptual framework and systematic review of speech-recognition chatbots for language learning," *Computers & Education*, p. 104898, 2023.

[30] G. Yenduri, M. Ramalingam, G. C. Selvi, Y. Supriya, G. Srivastava, P. K. R. Maddikunta, G. D. Raj, R. H. Jhaveri, B. Prabadevi, W. Wang, *et al.*, "Gpt (generative pre-trained transformer)–a comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions," *IEEE Access*, 2024.

[31] K. H. C. Lau, E. Bozkir, H. Gao, and E. Kasneci, "Evaluating usability and engagement of large language models in virtual reality for traditional scottish curling," *arXiv preprint arXiv:2408.09285*, 2024.

[32] W. Xu, M. Liu, O. Sokolsky, I. Lee, and F. Kong, "LLM-enabled cyber-physical systems: Survey, research opportunities, and challenges," in *2024 IEEE International Workshop on Foundation Models for Cyber-Physical Systems & Internet of Things (FMSys)*, pp. 50–55, IEEE, 2024.

[33] D. Branco, A. Amato, S. Venticinque, and R. Aversa, "Agents based cyber-physical diffused museums over web interoperability standards," *IEEE Access*, vol. 11, pp. 44107–44122, 2023.

[34] P. A. Jansen, "Visually-grounded planning without vision: Language models infer detailed plans from high-level instructions," *arXiv preprint arXiv:2009.14259*, 2020.

[35] K. Lin, C. Agia, T. Migimatsu, M. Pavone, and J. Bohg, "Text2motion: From natural language instructions to feasible plans," *Autonomous Robots*, vol. 47, no. 8, pp. 1345–1365, 2023.

[36] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, A. Y. Ng, *et al.*, "ROS: an open-source robot operating system," in *ICRA workshop on open source software*, vol. 3, p. 5, 2009.

[37] Unity Technologies, "ROS–Unity integration repository." Available from: https://github.com/Unity-Technologies/Unity-Robotics-Hub, Accessed on: 17/01/2025.

[38] G. Team, P. Georgiev, V. I. Lei, R. Burnell, L. Bai, A. Gulati, G. Tanzer, D. Vincent, Z. Pan, S. Wang, *et al.*, "Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context," *arXiv preprint arXiv:2403.05530*, 2024.

[39] H. Chaudhari, G. Severi, J. Abascal, M. Jagielski, C. A. Choquette-Choo, M. Nasr, C. Nita-Rotaru, and A. Oprea, "Phantom: General trigger attacks on retrieval augmented language generation," *arXiv preprint arXiv:2405.20485*, 2024.

[40] J. Yang, H. Jin, R. Tang, X. Han, Q. Feng, H. Jiang, S. Zhong, B. Yin, and X. Hu, "Harnessing the power of LLMs in practice: A survey on ChatGPT and beyond," *ACM Trans. Knowl. Discov. Data*, vol. 18, no. 6, 2024.

[41] NVIDIA, "NVIDIA HGX AI supercomputer," 2024. Available From: https://www.nvidia.com/en-gb/data-center/hgx/, Accessed on: 17/01/2025.

[42] NVIDIA, "Say what? Chat with RTX brings custom chatbot to NVIDIA RTX AI PCs," 2025. Available From: https://blogs.nvidia.com/blog/chat-with-rtx-available-now/, Accessed on: 17/01/2025.

[43] NVIDIA, "NVIDIA ChatRTX: Your Personalized AI Chatbot." Available From: https://www.nvidia.com/en-us/ai-on-rtx/chatrtx/, Accessed on: 17/01/2025.

[44] NVIDIA, "Demo: NVIDIA Chat with RTX. Your Personalized AI Chatbot," 2025. Available From: https://www.nvidia.com/en-in/ai-on-rtx/chat-with-rtx-generative-ai/, Accessed on: 22/01/2025.

[45] Vossr, "Python API for Chat With RTX," 2024. Available From: https://github.com/vossr/Chat-With-RTX-python-api, Accessed on: 17/01/2025.

[46] J. Brooke, "SUS: a quick and dirty usability scale," *Usability Evaluation in Industry*, 1996.

[47] J. Brooke, "SUS: a retrospective.," *Journal of usability studies*, vol. 8, no. 2, 2013.

[48] H. O'Brien, *Theoretical Perspectives on User Engagement*, pp. 1–26. Cham: Springer International Publishing, 2016.