A System Architecture for AI-Driven Interactive Systems Within a CPSS Framework for Metaverse Applications

Louis Nisiotis
University of Central Lancashire, Cyprus
Larnaca, Cyprus
LNisiotis@uclan.ac.uk

***Abstract*—This paper proposes a system architecture for deploying AI-driven interactive systems within the context of a Cyber-Physical-Social System framework, specifically designed for applications in the emerging Metaverse, applicable to a plethora of domains. The proposed architecture integrates advanced technologies to facilitate immersive user experiences and intelligent system interactions based on user actions and real-time data inputs from the real and the digital world(s), digital twins, artificial agents and elements. The integration of the proposed architecture components aim to support sophisticated interactions between physical, digital entities, and humans within the Metaverse. The paper presents the theoretical underpinnings and system description of the proposed architecture, contributing to future developments in Metaverse systems and technologies.**

***Index Terms*—Metaverse, eXtended Reality, Artificial Intelligence, Cyber-Physical-Social Systems, Generative AI.**

I. Introduction

In recent years, significant advancements in technology have set the stage for the rise of the Metaverse, a collective virtual shared space created by the convergence of artificially enhanced physical reality and persistent virtual spaces, influenced by humans and their behaviour. While the Metaverse is rooted in science fiction, originating from Neal Stephenson's 'Snow Crash' book [1] describing a virtual environment parallel to reality where humans live and interact as avatars, it is now starting to become a reality. This shift is driven by rapid advancements across various technological domains, disrupting the way humans experience the world in unprecedented ways. While the rise of the Metaverse is drawing significant attention, it also presents numerous challenges spanning across several societal, ethical, and technological domains [2]. Especially in the technological landscape, there is a notable lack of universal standards, complicating interoperability, compatibility, and accessibility. The absence of protocols and guidelines for technology deployment means that applications can vary widely, and potentially creating fragmented Metaverse applications. This paper contributes to this field, and proposes a system architecture for deploying AI-driven interactive systems within the context of Cyber-Physical-Social Systems (CPSS) for the development of Metaverse applications, applicable to a plethora of domains.

II. Background and Context

A. eXtended Reality and Artificial Intelligence

The Metaverse is widely regarded as a forthcoming technological evolution. As the concept is still developing, there is lack of agreement on what the Metaverse really is, and how it will evolve. Metaverse is often used to describe persistent virtual environments that combine elements from the real and digital worlds, avatars, agents, and digital objects, blending the physical world within virtual spaces. Recent technological advancements in software and hardware such as Artificial Intelligence (AI), eXtended Reality, the Internet of Things (IoT), computer graphics and others, have fostered the development of intelligent reality systems that seamlessly merge the real with the digital worlds, disrupting human-machine interactions and enabling more intuitive immersive experiences across various sectors [3]. eXtended Reality (XR) in particular is one of the key contributing technologies in realising the Metaverse. XR is an umbrella term that encapsulates Virtual (VR), Augmented (AR), and Mixed Reality (MR) software and hardware technologies [4], providing access to immersive experiences broadening the scope of digital possibilities in Metaverse applications. The mainstream adoption of XR is significantly increasing, driven by decrease in costs of ownership, improvements in portability and comfort, and expansion of application development and community support [5]. Today, XR is widely utilized across governments, corporate, industry, academic and other settings, and in our personal lives, impacting the digital transformation of the society in general.

When XR is used in conjunction with other emerging disruptive technologies, opportunities to create a powerful fusion for the development of complex computing systems and intelligent reality environments facilitating the visual aspect of the Metaverse arise [3]. For example, the convergence of AI tools such as machine learning, natural language processing, neural networks, deep learning, intelligent agents, and more recently generative AI into XR introduces new functionalities and capabilities for the development of innovative and immersive user experiences [6]. Adaptive virtual environments that can analyze user behavior to enhance interaction, intelligent agents that perceive, think, act, and interact with users and their surroundings, social agents equipped with voice recognition, speech synthesis, language understanding, real time text generation to facilitate interactions between agents and humans are only some indicative examples [3]. Especially with the recent significant advancements in the field of generative AI, this area has become a focal point of research. Technologies like OpenAI's ChatGPT and NVIDIA's ChatRTX demonstrate the capabilities and growing potential of generative AI to contribute in several aspects of the Metaverse realisation [7]. Generative AI can support the development of avatars and Non-Player Characters (NPCs), content creation, virtual world generation, automatic digital twin creation, and personalization, being an important asset in the Metaverse application development toolkit [8]. As generative AI continues to evolve, the future outlook within the Metaverse domain is expected to focus on enhancing technical research to improve customization and content quality, broadening the use of generated content across multiple domains, and the integration of generative AI with other technologies, to amplify its application impact and usage [7].

B. Cyber-Physical-Social Systems

Particular complex computing system architectures that leverage advancements in hardware and software, capable of blending the real and digital worlds, are facilitated by Cyber-Physical Systems (CPS) and Cyber-Physical-Social Systems (CPSS). CPS are engineered systems constructed from the seamless integration of computational algorithms and physical components through interconnected frameworks of infrastructure, smart computational devices, physical environments, and human interactions [9], [10]. CPS utilize sensing, actuating, and communication capabilities to merge physical operations with digital information systems [10], [11]. However, due to the fact that CPS are primarily focused on interfacing the physical world with computational processes and the cyberspace, they lack mechanisms for incorporating human interaction and influence within the system [12], a key element especially when attempting to realise Metaverse applications. The need for introducing human and social elements into complex computing systems has led to the development of CPSS, introducing the human dimension into CPS. CPSS are drawing significant academic and industrial interest due to their ability to integrate human interactions and social inputs into the system framework [13], and have been extensively applied across various sectors including energy, power grids, smart vehicles, management, transportation, anti-terrorism measures, manufacturing, smart cities, social computing, and other architectures (see the comprehensive review by Zhang et al. [14] for more). The convergence of emerging technologies within CPS and CPSS frameworks lays the foundation for the development of a new type of Cyber-Physical-Social Eco-Society of Systems (CPSeS) [3], [15], connecting our lives and blending realities, bringing us a step closer to the Metaverse.

However, designing such systems presents significant challenges due to their inherent complex computing nature as a result of heterogeneity of networks, and complexities of hardware and software. Furthermore, key challenges exist in the way of integrating the social dimension, humans and their influence into the systems, with a notable lack of effective methodologies that can address these issues [16]. Considering these challenges, this paper proposes a system architecture for deploying AI-driven interactive systems within the context of CPSS, aiming to contribute to the current body of knowledge in the development of immersive Metaverse environments (see comprehensive analysis of technological singularity, virtual ecosystems, and evolving research agenda review on the Metaverse by Lee et.al [17]).

III. Proposed System Architecture

"Fig. 1" illustrates the proposed system architecture designed within a CPSS framework, leveraging advanced visualisation and AI technologies to support immersive user interaction and intelligent system decision making.

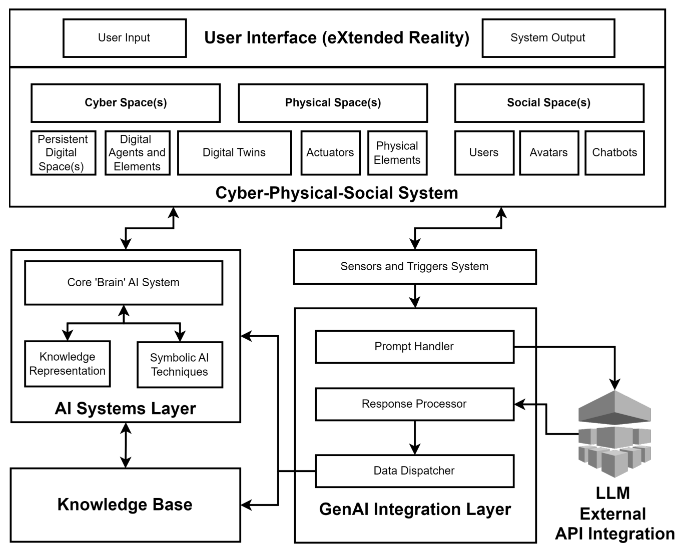

Fig. 1. Proposed System Architecture.

A. UI and Cyber-Physical-Social System Layers

The ***User Interface Layer*** enable users to interact with the system and with each other, and is responsible for presenting visual information to users and obtaining their requests and system interaction through XR Modes. It captures and processes user inputs into actionable data establishing a bi-directional flow of information with the system. The ***CPSS Integration Layer*** connects the real with digital world(s) and social spaces, facilitating the concept of the Metaverse. The cyber-space consists of persistent virtual spaces, equipped with virtual agents, digital twins, virtual objects, information, and other digital elements for rich and immersive user interactions. The physical space features real life elements and actuators that interact with and respond to cyber elements mediated by digital twins and interconnected actuators, seamlessly connecting the cyber with the physical spaces, and enabling real-time data exchange and synchronization between realities. The social component of the system introduces the human aspect, which is fundamental to the interactive and immersive experience of any Metaverse application. It encompasses users, their avatars, virtual agents, chatbots, and other interactive social entities, developing a dynamic social network within the system offering a Metaverse place for social interactions.

B. Artificial Intelligent Systems Layer

To drive the CPSS system behaviour both in the digital and the real world, the proposed architecture suggests a dedicated ***AI Systems Layer***. This layer encompasses a collection of AI components responsible to manage the system/user/entities interactions and decision-making, adapting the CPSS according to their actions based on real-time data and pre-defined set of rules. The successful integration of this layer is guided by the need for efficiency in real-time processing to coordinate the components of the CPSS to ensure that the system executes all processes efficiently and the system output provides access to usable XR user experiences. A recommended approach is the use of Symbolic AI techniques commonly used in the video games development and research domain. These techniques are typically less computationally intensive compared to other AI techniques such as Machine Learning (ML), Large Language Models (LLMs) etc. which require extensive data processing. Symbolic AI operates on logic and rules that can be executed quickly and deterministically, enabling to maintain high performance and a smooth user experience in real-time virtual environments, and is the most commonly used method within the context of real time virtual environments development.

Core to the AI systems layer, a central 'brain' is proposed, acting as the orchestrator for all AI-driven processes. The AI systems layer should be keeping track of facts and information on the environment, user interactions, preferences, system settings, and the actions of physical and virtual entities within the environment. An example of such system is the 'Director AI' system, a common game AI technique [18], capable to coordinate various AI components at the highest level to provide dynamic distribution of tasks, evaluation and execution of system actions and plans to achieve pre-defined goals. A central 'brain' AI approach can be used for adjusting the environment's behaviour based on user interactions and system status in real-time to accommodate changing conditions and user needs. This component will help to manage the overall user experience by dynamically adjusting the environment's pace, challenge, even audio, and respond to events and triggers based on the actions and performance, while also helping with resources management and optimisation.

Directly connected to the core AI brain should be a *Knowledge Representation* system that serves as a central repository for all dynamic and static information relevant to system and user operations, aggregating data from various sources and making it accessible to different subsystems within the AI architecture for information sharing across system processes. Blackboard AI system is another applicable example which is a common game AI technique, serving as a shared knowledge representation base. In such system, multiple independent processes and AI components have the capability to read from and write to a central blackboard, where instead of directly communicating with each other between processes, they interact implicitly by identifying and reacting to content on blackboard relevant to their specific operations [19]. The Knowledge Representation system in this architecture should provide the ability of coordinating AI systems and processes without the need for direct interaction to support collaborative problem-solving and decision-making among multiple AI agents and system functionalities. Considering the need for processing efficiency and based on the system's interactive requirements, several other Game AI techniques can be used to handle complex decision-making processes and managing rule-based operations within the system, such as Expert Systems, Finite State Machines, Decision and Probability trees, Influence Maps, Goal-Oriented Action Planning, and other.

Further to the AI systems layer, the implementation of a ***Knowledge Base Layer*** serving as the central repository of all contextual data necessary for the environment and AI systems to support decision making, disseminate contextual knowledge, and provide personalized interactions is recommended. The dedicated Knowledge Base system would be responsible for dynamically supporting contextual knowledge relevant to the application domain, processing and updating data inputs from various sources within the system, and from newly generated knowledge as described by the system below.

C. Generative AI Integration Layer

A key element on the proposed architecture is the implementation of a ***Generative AI (GenAI) Integration layer*** to support and complement AI-driven actions and further development of the system's contextual knowledge. Due to the computational requirements associated with running advanced generative AI models like LLMs, the proposed architecture suggests the use of an API integration to connect the system to external high-performance computing service for real time textual generation to be used for system content.

The functionalities of the GenAI Integration Layer are requested by a *'Sensors and Triggers System'* that initiates interactions with external generative AI tools (such as ChatGPT or ChatRTX) by generating specific prompts based on system and user data. Prompts will be processed by the GenAI to dynamically generate responses tailored to the current context, through a dedicated *Prompt Handler* component responsible for formulating and submitting prompts based on sensed data. An indicative prompt example in the context of visiting a particular area within a virtual museum using ChatGPT service is: *"User [userID] has entered [areaID]. They have spent [sessionDuration] in the [environment] and previously visited [previousInteractions]. Given their interests in [userPreferences.interests], provide detailed information about the [element of interest]. Formulate the reply in the following structure: {context}:, {fact}:, {nextAction}:.*

An example of expected ChatGPT response following the predefined format would be:

{context}: The [element] in [areaID] represents a significant aspect of [cultural/scientific context], exemplifying [characteristics]. This piece is particularly noted for its [attributes].

{fact}: An intriguing fact about the [element of interest] is that it [provide an interesting fact about the item or its background that aligns with the user's interests].

{nextAction}: To further explore the [element of interest], consider [suggest a follow-up action]. This will enhance your understanding and appreciation of [related topic/feature].

The received AI-generated content will be handled by the *Response Processor* component, responsible to break the structured information into objects/data structures. Each segment of the AI response will be parsed into data elements and handed over to the *Data Dispatcher* component responsible for feeding them into the AI Systems and to the Knowledge Base layers. For example, a *userBlackboardSystem* can be updated with key user information (such as: *currentPosition, areasVisited, listOfInteractedItems*, etc.) that will directly influence AI-driven actions and decision-making processes. Furthermore, generated content can modify how the Knowledge Base functions, for example, updating content in the virtual world or altering the information shared by virtual agents.

IV. Conclusions, Limitations and Future Work

This paper outlines a complex computing system architecture for AI-driven intelligent reality systems that foster the concept of the Metaverse within a CPSS framework. The proposed architecture converges multiple technologies and integrates advanced AI to support dynamic user, system, and environment interactions, setting the foundations for future advancements in intelligent interactive environments, offering a comprehensive platform for blending the digital with the real world effectively. The proposed infrastructure can be used as a guide for deploying Metaverse systems and applications, contributing to existing research efforts of seamlessly blending the real with digital worlds for improved user experience and intelligent system interactions.

However, this architecture is not without its limitations. Key system challenges include scalability, interoperability, data privacy, quality, and ethical considerations, all of which should be taken into deep consideration. A significant limitation of the system is its reliance on generative AI for producing information. These models can occasionally generate misleading, incorrect, or even harmful information. Additionally, due to their probabilistic nature, LLMs may not provide consistent or reliable responses, potentially offering different answers to the same query and contradicting previous information. Moreover, incorporating commercial LLM services like ChatGPT into the system can lead to substantial costs.

Future work aims to implement and evaluate the proposed architecture, focusing on refining and demonstrating the capabilities of such a system within a CPSS framework for Metaverse applications. Current emphasis is on developing a functional system prototype to test the integration of various technological components, evaluate, and validate the technical feasibility and operational robustness of the proposed architecture. Comprehensive evaluations to measure the system's performance will be conducted together with user experience evaluation studies focusing on usability, acceptance and overall user experience to guide further refinements.

Acknowledgment

In memory of my dear colleague, mentor, and friend Dr Lyuba Alboul.

References

[1] N. Stephenson, *Snow crash*. Bantam Books, 1992.

[2] Y. K. Dwivedi, L. Hughes, A. M. Baabdullah, S. Ribeiro-Navarrete, M. Giannakis, *et al.*, "Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy," *International Journal of Information Management*, vol. 66, p. 102542, 2022.

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

[4] M. Handa, E. G. Aul, and S. Bajaj, "Immersive technology–uses, challenges and opportunities," *International Journal of Computing & Business Research*, vol. 6, no. 2, pp. 1–11, 2012.

[5] S. Alizadehsalehi, A. Hadavi, and J. C. Huang, "From BIM to eXtended reality in AEC industry," *Automation in Construction*, vol. 116, p. 103254, 2020.

[6] T. J. L. Garcia, R. M. Rodriguez-Aguilar, J. A. Alvarez-Cedillo, and T. Alvarez-Sanchez, "Development of software architecture for a 3d virtual environment with the incorporation of a reactive intelligent agent," *Journal of Theoretical and Applied Information Technology*, vol. 97, no. 17, pp. 4589–4599, 2019.

[7] Z. Lv, "Generative artificial intelligence in the metaverse era," *Cognitive Robotics*, vol. 3, pp. 208–217, 2023.

[8] H. X. Qin and P. Hui, "Empowering the metaverse with generative AI: Survey and future directions," in *2023 IEEE 43rd International Conference on Distributed Computing Systems Workshops (ICDCSW)*, pp. 85–90, IEEE, 2023.

[9] 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.

[10] 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.

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

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

[13] G. Xiong, F. Zhu, X. Liu, X. Dong, W. Huang, S. Chen, and K. Zhao, "Cyber-physical-social system in intelligent transportation," *IEEE/CAA Journal of Automatica Sinica*, vol. 2, no. 3, pp. 320–333, 2015.

[14] 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, no. 3, pp. 829–840, 2018.

[15] 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.

[16] 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.

[17] L.-H. Lee, T. Braud, P. Zhou, L. Wang, D. Xu, Z. Lin, A. Kumar, C. Bermejo, and P. Hui, "All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda," *arXiv preprint arXiv:2110.05352*, 2021.

[18] M. Booth, "The AI systems of Left 4 Dead," in *Artificial Intelligence and Interactive Digital Entertainment Conference*, 2009.

[19] A. Liapis, G. N. Yannakakis, M. J. Nelson, M. Preuss, and R. Bidarra, "Orchestrating game generation," *IEEE Transactions on Games*, vol. 11, no. 1, pp. 48–68, 2019.