Social Augmentation in Cross-Platform XR: Enabling Collaborative Learning

Letícia Lucas[1], Carla Costa[1] and Pedro Mendes Jorge[1,2]

[1]Instituto Superior de Engenharia de Lisboa, Instituto Politécnico de Lisboa, Rua Conselheiro Emídio Navarro 1, 1959-007 Lisboa, Portugal

[2]NOVA-LINCS ISEL, Rua Conselheiro Emídio Navarro 1, 1959-007 Lisboa, Portugal

Abstract

While Extended Reality (XR) has the potential to transform learning by turning abstract concepts into something tangible, it often introduces new barriers to equitable access, limiting the potential for inclusive social interaction. Significant challenges arise when students using diverse hardware - which may range from affordable mobile Augmented Reality (AR) to high-end Virtual Reality (VR) headsets - attempt to collaborate within the same environment, leading to fragmented social experiences.

This paper presents a modular, cross-platform architecture designed to address these challenges, enabling real-time collaboration across heterogeneous XR devices and preserving socially meaninful interaction. The proposed system employs a dual-server architecture, combining Unity-based real-time multiplayer synchronization with a Python server responsible for custom object detection. This approach supports both the technical requirements of cross-platform interoperability and the domain-specific needs of educational applications.

The system's viability is demonstrated through a collaborative solar energy learning module, in which users equipped with Magic Leap 2 and Meta Quest 3 devices (examples of AR and VR technologies, respectively) interact with shared virtual objects anchored to physical, real-world markers. This setup allows students to observe synchronized environmental effects on energy production, independent of their hardware constraints, allowing for shared referential communication and joint attention.

Preliminary evaluations indicate low synchronization latency and accurate marker tracking, providing the fluid interaction necessary for maintaining social cues and engagement.

Overall, this work contributes with a scalable solution for XR interoperability, aiming to democratize socially augmented experiences and ensure that educational collaboration remains inclusive, regardless of the user's hardware constraints.

Keywords

Collaborative Learning, Cross-Platform Development, Device Interoperability, Educational Experiences, Extended Reality, Mixed Reality, Social Augmentation, System Design

1. Introduction

The exponential evolution of Extended Reality (XR) technologies (encompassing Virtual, Augmented, and Mixed Reality paradigms) mandates the redefinition of human-machine interaction paradigms. These technologies enable the creation of immersive experiences previously considered unattainable by blending digital content with the physical world or creating entirely virtual environments [1].

The educational sector offers a clear example of this transformative potential: XR allows students to visualize invisible phenomena and conduct complex experiments that would otherwise be too dangerous, costly, or altogether impossible in traditional settings. Recent studies have shown that these approaches enhance learners' motivation, deepen conceptual understanding, support long-term retention, and foster critical thinking and problem-solving skills [2, 3, 4]. However, XR's educational benefits extend beyond realistic visualization, as they have the potential to include a crucial social dimension: the ability to foster collaborative exploration and shared discovery. Learning is enriched by social processes, and immersive environments can amplify peer-to-peer engagement and collective problem-solving [5].

1st ACM CHI 2026 Workshop on Shaping Future Human Connection: Social Augmentation through XR Technologies, April 13–17, 2026, Barcelona, Spain

leticialucas@deetc.isel.ipl.pt (L. Lucas); carla.costa@isel.pt (C. Costa); pedro.mendes.jorge@isel.pt (P. M. Jorge)

0009-0007-6668-3165 (L. Lucas); 0000-0003-4222-4734 (C. Costa); 0000-0003-1597-410X (P. M. Jorge)

1.1. The Challenge of Device Fragmentation and the Equity Gap

Despite its potential, XR remains far from the everyday reality in classroom settings, with one of the primary obstacles being the technological fragmentation [6]. The current landscape suffers from a fundamental interoperability problem: platforms operate within "walled gardens" - isolated hardware ecosystems that lack cross-compatibility.

This fragmentation produces a social hierarchy of access, and creates barriers in the educational field. Institutions cannot be expected to guarantee a completely homogeneous catalog of devices, yet a student with a Meta Quest (VR) is unable to meaningfully collaborate with a peer using, for instance, a Magic Leap (AR) or a smartphone, due to its different technical characteristics. This incompatibility produces three core forms of educational inequity:

- **Economic Barriers:** High costs for uniform hardware deployments are prohibitive for most institutions, granting wealthier schools exclusive access to state-of-the-art immersive tools [7].
- **Capability Disparities:** Incompatibility prevents collaboration even among students with capable devices, as different hardware generations cannot share the same virtual space.
- **Pedagogical Limitations:** Developers are forced to design for the "lowest common technical denominator," often abandoning collaborative goals due to hardware constraints [8].

The technology that promised to democratize education now risks reinforcing existing inequalities. If XR is to fulfill its transformative potential, it must ensure **social inclusion**, where the student's headset choice does not determine the quality of the collaborative experience.

1.2. Research Objectives and Contributions

The present work addresses these barriers by designing and implementing a modular, cross-platform architecture capable of **enabling equitable, real-time collaboration across heterogeneous XR devices**. Primary contributions include:

- **Cross-Platform Collaborative Infrastructure:** A dual-server architecture separating experience management (Unity-based) from specialized computer vision tasks (Python-based).
- **Flexible Marker-Based Alignment:** Integration of YOLO [9] object detection, allowing the use of pedagogically relevant physical objects (e.g. solar panel replicas) as spatial anchors for the immersive experiences.
- **Extensible Educational Framework:** A modular system supporting parallel development of learning modules and device-agnostic interactions.
- **Demonstrated Pedagogical Application:** A collaborative solar energy module proving system viability for social learning.

This work rests on the principle that infrastructure and educational application deserve equal emphasis. Technical interoperability alone is insufficient, as the architecture must facilitate pedagogically meaningful experiences that transform how students engage with complex concepts through social, collaborative exploration. By bridging the gap between incompatible hardware, this project provides an open-source pathway toward a more inclusive and socially connected immersive classroom.

The following sections focus on providing a comprehensive overview of the platform's conceptual foundation, the supporting infrastructure, and a brief use case illustrating system operation. For further implementation details and extended discussions, please refer to previous related publications and conference proceedings [10, 11].

2. Related Work

2.1. Social and Educational XR and the Interoperability Gap

Research in social XR has established that collaborative immersive experiences significantly enhance learning engagement and knowledge retention. While platforms like VRChat or Meta Horizon Worlds support shared spaces [12, 13], they typically assume hardware uniformity or sacrifice device-specific capabilities for basic compatibility. Recent studies have explored enriching social presence through expressive avatars [14] or gaze visualization [15], yet these rarely address the foundational challenge of enabling collaboration across fundamentally different device types (e.g. bridging optical AR with fully immersive VR). The present work aims to complement this research by providing infrastructure where such social augmentations can function consistently across different platforms.

2.2. Fragmentation in STEM

In STEM education, tools like Nanome [16] or zSpace [17] demonstrate high value - these examples allow students to engage in virtual collaborative molecular visualization or participate in Mixed Reality laboratories for engineering education. However, like many other similar applications, these also remain confined to specific hardware ecosystems.

2.3. Challenges in Cross-Platform Development

Efforts to address XR fragmentation have primarily focused on abstraction layers. Current standards like OpenXR [18] or WebXR [19] have significantly simplified device tracking and input handling but do not inherently solve higher-level challenges such as spatial synchronization or session management across heterogeneous devices. Similarly, networking frameworks, such as Photon [20] and Unity Netcode [21], support multiplayer XR experiences. These, however, still require substantial custom logic to handle device-specific visualizations and interaction paradigms.

2.4. Spatial Alignment and Pedagogical Meaning

Spatial alignment is crucial for shared referential communication. While SLAM-based approaches are valuable [22], they demand high computational resources and stable environmental conditions - constraints often violated in dynamic classroom settings. Furthermore, while traditional marker-based systems are reliable alternatives, the usage of abstract patterns lacks pedagogical relevance and can increase cognitive load by introducing visually distracting elements unrelated to the subject matter.

By integrating customizable YOLO-based object detection, the proposed architecture allows the use of any domain-specific physical artifacts as spatial anchors, turning learning materials themselves into facilitators of shared digital-physical interaction.

2.5. Research Gaps

In summary, current literature reveals a lack of systems that:

1. **Support true real-time collaboration** between AR and VR users without degrading social cues.
2. **Utilize pedagogically meaningful objects that support learning objectives** for spatial alignment, rather than relying on marker-based systems that use abstract fiducial patterns.
3. **Provide an accessible infrastructure for educators**, without requiring programming backgrounds.

The developed architecture addresses these gaps by enabling device-agnostic collaboration through a modular, marker-based approach.

3. Understanding the Collaborative Experience Paradigm

To better comprehend the technical requirements of the architecture, it is necessary to define how users interact with the platform and collaborate within shared immersive environments.

3.1. Individual Exploration and Collaborative Sessions

Upon launching the application, users are presented with a menu showcasing the educational modules, which may span diverse educational domains (Fig. 1). While the platform supports individual exploration, its primary design emphasis is on collaborative learning through "teams" - which can be seen as digital study groups.

Figure 1: Main menu interface - at present time, only the "Solar" module is made available.

By joining a team, participants - regardless of their physical location - enter a shared synchronous virtual environment. Within this space, participants manipulate content and observe each other's contributions in real-time, mirroring the collaborative dynamic of a traditional classroom while augmenting it with immersive tools. This setup promotes teamwork and co-construction of knowledge.

3.2. Bridging the Virtual and Physical Worlds through Marker-Based Anchoring

To leverage the complementary strengths of both AR and VR technologies, educational modules may incorporate the use of physical markers as spatial anchors. These markers establish consistent reference points to connect physical and virtual spaces.

Consider a scenario where a group gathers around a physical microprocessor. Through their devices, they can perceive and interact with virtual components - such as circuit pathways, data flow visualizations or voltage indicators - overlaid on or anchored to the hardware . The system must handle this visualization in a way that accommodates distinct interaction modalities while maintain social equity and spatial consistency:

- **AR Users:** Camera-equipped devices (optical or video see-through) enable participants to **directly perceive the physical marker** in their environment. Virtual content is overlaid atop the physical object, enabling tactile, "hands-on" manipulation where users can physically reposition markers and observe corresponding updates to anchored virtual elements.
- **VR Users:** Fully immersed participants cannot perceive physical markers directly. Instead, the system provides a **synchronized virtual twin** - that is, a digital replica of the physical marker positioned in the shared coordinate space. Virtual components maintain identical spatial relationships to this twin, ensuring VR users manipulate the same content at the same relative positions as their AR collaborators.

3.3. Maintaining a Shared Referential Space

The platform's fundamental objective is to create a convincing "illusion of a shared space" [23]. Effective collaborative learning requires joint attention [24]: when a student in one geographic location manipulates a virtual component anchored to their physical marker, a remote peer in a different location must observe the same manipulation applied to the spatially corresponding element in their own view.

By anchoring experiences to pedagogically meaningful physical objects, the platform transforms everyday educational materials into facilitators of hybrid physical-digital interaction. All participants share a unified spatial understanding, enabling fluid social dialogue and coordination within the immersive learning environment.

4. Proposed Architecture

A modular client-server architecture is proposed, designed to offload intensive tasks from the headsets, ensuring that low-power devices remain responsive and lightweight - a critical factor for maintaining co-presence and immersion. Fig. 2 illustrates the implemented solution.

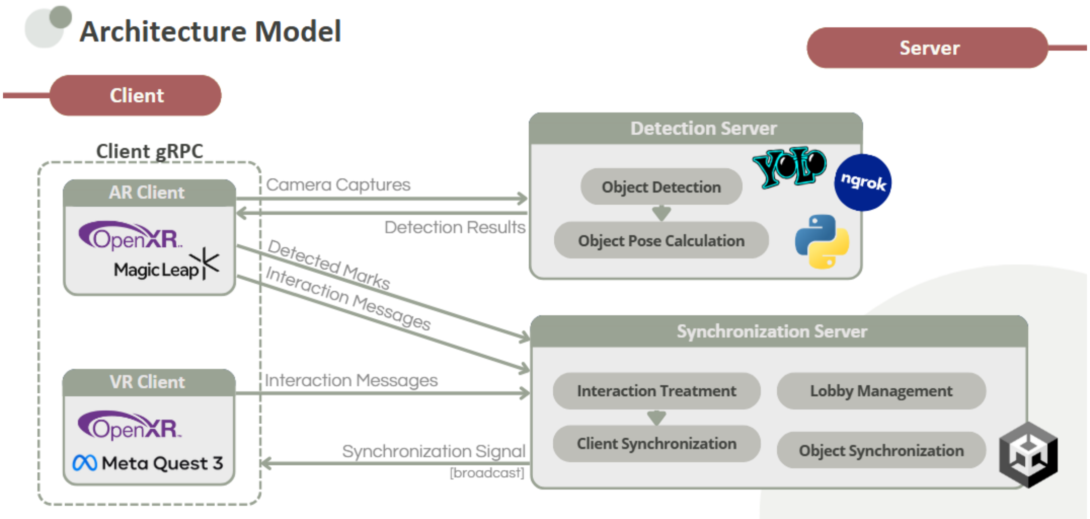

Figure 2: Diagram representing the implemented architecture (detailed further in [10, 11]).

The architecture is guided by the following core principles:

- **Separation of Infrastructure and Content:** Core networking and device management (static concerns) should be cleanly separated from educational content (dynamic content), enabling independent development and updates.
- **Extensibility and Accessibility:** New devices and learning modules should integrate through well-defined interfaces requiring minimal technical expertise.

4.1. System Components

4.1.1. Client Layer: Device Management and Interaction

The client layer handles user interaction and visualization - from high-end AR glasses and VR headsets to “simple” mobile tablets. As each of these devices offer vastly different characteristics, OpenXR is used to avoid the “silo effect” of device-specific programming.

Put simply, OpenXR acts as an abstraction layer, providing a unified interface to access the devices’ various subsystems (such as the camera or motion-tracking). While some device-specific optimization is still required, this abstraction ensures that new hardware can be integrated with minimal friction, directly supporting our goal of a democratized educational ecosystem where students are not excluded based on their available hardware.

4.1.2. Python Server: Custom Object Detection

A dedicated Python server handles real-time computer vision tasks, specifically the identification and spatial localization of the physical markers. It is characterized by:

- **YOLO-based Detection:** AR clients stream frames via asynchronous WebSockets. The server utilizes custom-trained YOLO models identify domain-specific physical objects and return 2D coordinates.
- **Asynchronous Processing:** Non-blocking sockets handle concurrent requests from multiple users without latency spikes, maintaining fluid shared experiences.
- **Network Accessibility:** To ensure the system works across different physical locations and institutional networks, the server is exposed via Ngrok [25]. This platform provides secure public endpoints that bypass complex firewall and NAT restrictions common in institutional networks.

4.1.3. Unity Server: Social Coordination

While the Python server manages the physical-to-digital link, the Unity-based server coordinates the collaborative experience, acting as the central authority of the session. It handles:

- **Real-time Synchronization:** Whenever a student interacts with a virtual object, the change must be transmitted across the network. This is attained using Remote Procedure Calls (RPCs) and Unity's multiplayer services. With these tools, the server replicates state changes across all concerned clients near-instantaneously.
- **Authoritative State:** Server-side state management ensures all participants perceive coherent, synchronized environments regardless of device or location, establishing the foundation for joint attention and social interaction.

4.2. Communication and Spatial Alignment Workflow

The synchronization flow follows a three-step cycle to ensure spatial consistency:

1. **Registration:** An AR client sends a frame to the Python server. YOLO identifies the object, and the client reconstructs its 3D pose using local depth data.
2. **Global Update:** The client transmits the 3D pose to the Unity Server via a *ServerRPC*.
3. **Social Synchronization:** The Unity Server broadcasts the update via a *ClientRPC*, with all users within the same team receiving a consistent synchronous visualization of the shared content.

This authoritative model prevents race conditions and ensures that joint attention remains accurate across all device types.

4.3. Architectural Advantages

This dual-server approach offers three strategic benefits:

- **Optimization:** By offloading vision and networking logic, the performance of standalone headsets is optimized, reducing thermal throttling and motion sickness [26].
- **Specialization:** The best tool for each task is leveraged: Python's extensive computer vision libraries and Unity's robust 3D environment management.
- **Scalability and Fault Tolerance:** Each component can be scaled independently.

5. System Validation and Performance Analysis

To validate the proposed architecture's technical feasibility, along with its pedagogical potential, a collaborative solar energy learning module was implemented, to be deployed and tested with heterogeneous XR devices. This subject addresses a fundamental challenge in STEM education: making abstract physical relationships tangible and explorable through social interaction, while simultaneously engaging learners with knowledge regarding sustainability transitions, aligned with the objectives framed by the European Green Deal [27].

5.1. The Solar Energy Learning Module

Photovoltaic energy production represents a critical sustainability topic, yet traditional instruction often reduces its complex physical phenomena to static equations - an approach that may fail to foster intuition regarding *how* individual variables truly affect energy output. To address this, the implemented module was developed with the intent of allowing students to isolate variables, experiment with configurations, and observe causal relationships in real-times: essential skills to build an intuitive understanding.

Although real-world photovoltaic modeling typically involves more complex equations, Equation 1 provides a deliberately simplified representation of the fundamental relationships for foundational learning purposes:

$$P = I_{DNI} \cdot \cos(\theta) \cdot A \cdot \eta \tag{1}$$

where: I_{DNI} := direct normal irradiance, θ := angle of incidence, A := panel area, η := conversion efficiency

Building upon this, the module facilitates collaborative inquiry by allowing users to manipulate three core parameters:

- **Manipulate Sun Position:** Students interact with a virtual Sun to vary the angle of incidence (θ), observing immediate shifts in efficiency through the user interface.
- **Atmospheric Conditions:** Adjusting cloud density allows students to observe the impact of direct normal irradiance (I_{DNI}) on energy production.
- **Internal Structural Exploration:** Users can "dive into" the solar panel layers to understand their functional importance.

This design embodies constructivist learning principles, where knowledge emerges through experimentation and social dialogue rather than passive reception.

5.2. Object Detection Preparation

This experience creates a "bridge" between a physical marker (a miniature solar panel replica) and the digital content (the Sun and clouds). In order to enable this, a YOLO segmentation model was trained on 37 annotated images of solar panel replicas (80/20 train-validation split), achieving **99.5% mean Average Precision (mAP)** under varied classroom lighting conditions. The lightweight model (5.89 MB) enables real-time detection with **45–55ms inference latency**, ensuring responsive marker tracking during collaborative manipulation.

5.3. Evaluation Setup and Metrics

Preliminary technical validation was conducted using Magic Leap 2 (AR) and Meta Quest 3 (VR) to simulate heterogeneous classroom deployments. Fig. 3 illustrate both visualization paradigms.

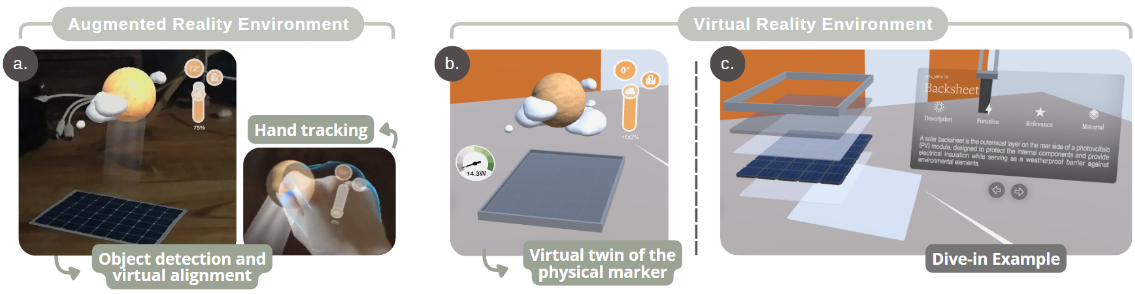

Figure 3: Comparison between (a) augmented and (b) virtual reality environments; (c) virtual dive-in for both devices.

To note, object manipulation is dependent on the device the user has available. For example, and as pictured above, on Magic Leap and Quest devices, users can use gestures or controllers; on a tablet, users can use the touchscreen; on a desktop, users can use the mouse. The OpenXR layer abstracts this hardware dependency.

This testing setup confirmed the following:

- **Synchronization Latency:** State propagation (e.g. one user repositioning the Sun, others observing the change) occurred consistently in **under 50ms**, through varying proximities and network conditions - this is well below the ~100ms threshold for imperceptible lag in real - time collaboration [28]. This sustains the illusion of shared presence essential for fluid social interaction.
- **Marker Tracking Robustness:** Detection pipeline (frame capture → server inference → 3D alignment → object positioning) completed in **under 100ms total**, enabling continuous tracking without perceptible delays. Virtual objects maintained spatial alignment through physical marker repositioning, adapting dynamically to pose updates.
- **Resource Optimization:** Offloading YOLO inference to the Python server enabled client devices to maintain stable **72 FPS** rendering throughout extended sessions, preventing thermal throttling and preserving user comfort - critical for longer collaborative learning activities.
- **Cross-Device Consistency:** Both AR (hand tracking) and VR (controller input) interaction paradigms successfully manipulated shared virtual objects. The pose of the digital elements remained synchronized across heterogeneous clients, with AR users perceiving physical markers overlaid with virtual content while VR users interacted with spatially consistent virtual twins.
- **Network Resilience:** The system operated successfully across separate WiFi networks and mobile hotspots, validating Ngrok tunneling effectiveness for bypassing institutional firewall restrictions without requiring specialized network configuration.

To better see the module in action, a short video showing a VR walkthrough of the application is made available - VR Demonstration Video[1].

In short, these preliminary results demonstrate that the architecture successfully addresses the fragmentation barrier: students using fundamentally different devices are able to collaborate on equal footing within the same immersive space. Importantly, neither experience is degraded to accommodate the other: device-specific strengths (physical marker manipulation for AR, immersive environments for VR) are preserved while maintaining synchronized shared state.

6. Conclusions and Future Work

This work addresses a critical barrier to equitable educational XR: hardware fragmentation that prevents collaboration between students using different devices. The proposed modular, cross-platform architecture demonstrates that social augmentation does not have to be constrained by device ownership or institutional budgets.

To review, the system establishes the following key foundations:

- **Democratized Access:** The dual-server architecture enables real-time collaboration across heterogeneous devices without degrading capabilities.
- **Pedagogically Grounded Anchoring:** YOLO-based detection transforms learning materials into spatial anchors, eliminating abstract markers while enhancing the connection between the physical and digital world.
- **Maintained Social Presence:** Sub-50ms synchronization and stable tracking preserve joint attention and shared communication regardless of participants' geographic proximity, both essential for collaborative learning.

[1]Video link: `https://youtu.be/UCXooBNJHHA`

- **Extensible Framework:** Separation of infrastructure from content raises the potential for educator-driven module development, even without deep technical expertise.

Furthermore, the proposed architecture's core components - encompassing the dual-server model, OpenXR-based device abstraction, and real-time state synchronization - are inherently domain-agnostic and transferable to other collaborative XR contexts beyond education, such as remote industrial maintenance or training simulations, for example.

The Python-based object detection server may be the component most sensitive to a domain shift, as the underlying YOLO models would require retraining on domain-specific objects (e.g. machinery parts or anatomical models). However, even in such cases, the architectural role of spatial anchoring - and how it is handled - would remain the same. The modular separation of infrastructure from content was designed precisely to accommodate such adaptations with minimal friction.

Finally, it is worth noting that high-stakes contexts - such as surgical training or safety-critical industrial applications - would warrant stricter latency guarantees and fault-tolerance mechanisms in the synchronisation server than those required in classroom settings.

6.1. Future Directions

While a strong foundation has already been established, future work may focus on incorporating further enhancements:

- **High-Load Scalability:** Future efforts should evaluate system performance with an increased number of concurrent users. This will allow for a precise characterization of bandwidth requirements and synchronization stability under intensive, real-world network loads.
- **Expanded Ecosystem:** Similarly, a wider spectrum of hardware should continue to be tested and evaluated, including the potential development of a browser-based port to ensure cross-platform availability without specialized hardware.
- **Educational Evaluation:** A core priority involves conducting controlled classroom studies, assessing learning outcomes, conceptual understanding, and collaborative skill development across different levels.
- **Personalization and Gamification:** User-level features such as progress tracking, customizable profiles, avatars, and achievement systems could deepen engagement and motivation, creating a more rewarding and personalized learning experience.
- **Adaptive Learning:** Leveraging user data to enable real-time module adaptation through predictive models could provide tailored exercises, identify struggling areas, and offer instructor feedback to support personalized learning paths.
- **Modular Expansion:** The platform's infrastructure should support continuous addition of new topics, challenges, and pedagogical approaches, ensuring ongoing relevance and adaptability to evolving educational needs.

Acknowledgments

The authors gratefully acknowledge support from the ICSE Science Factory project (2023–2026), funded by the European Union's Horizon Europe Research and Innovation Programme under the Call HORIZON-WIDERA-2022-ERA-01 (Grant Agreement No. 101093387). DOI: 10.3030/101093387.

Declaration on Generative AI

During the preparation of this work, Overleaf's built-in AI Assist was used **exclusively** for spell checking and equation formatting. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] P. Milgram, H. Takemura, A. Utsumi, F. Kishino, Augmented reality: A class of displays on the reality-virtuality continuum, Telemanipulator and telepresence technologies 2351 (1994) 282–292.

[2] M. Akçayır, G. Akçayır, Advantages and challenges associated with augmented reality for education: A systematic review of the literature, Educational Research Review 20 (2017) 1–11.

[3] A. Banjar, X. Xu, M. Z. Iqbal, A. Campbell, A systematic review of the experimental studies on the effectiveness of mixed reality in higher education between 2017 and 2021, Computers & Education: X Reality 3 (2023) 100034.

[4] M. C. Johnson-Glenberg, C. Megowan-Romanowicz, D. A. Birchfield, C. Savio-Ramos, Embodied learning with multi-sensorial augmented reality: A large field study in physics education, Frontiers in Virtual Reality 4 (2023) 1047833.

[5] L. Yang, S. Ni, Y. Wang, A. Yu, J. Lee, P. Hui, Interoperability of the metaverse: A digital ecosystem perspective review, arXiv preprint arXiv:2403.05205 (2024).

[6] T. Bezmalinovic, Virtual reality is facing an old threat again: fragmentation, https://mixed-news.com/en/vr-industry-fragmentation-threat/, 2024. Accessed: 2026-03-15.

[7] J. Radianti, T. A. Majchrzak, J. Fromm, I. Wohlgenannt, A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda, Computers & Education 147 (2020) 103778.

[8] N. Sala, Virtual reality, augmented reality, and mixed reality in education: A brief overview, in: Virtual and Augmented Reality in Education, Art, and Museums, IGI Global, 2021, pp. 48–69.

[9] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 779–788.

[10] L. Lucas, C. Costa, P. M. Jorge, Cross-device platform for collaborative and immersive experiences in mixed reality, 2025. URL: DOI:10.1109/YEF-ECE66503.2025.11117510.

[11] L. Lucas, C. Costa, P. M. Jorge, Cross-device platform for collaborative and immersive experiences in mixed reality, 2025. URL: https://repositorio.ipl.pt/entities/publication/dd88a27f-1541-4a1e-984e-8438c8582f0f.

[12] K. Kim, G. Bruder, D. Maloney, G. F. Welch, Social presence in virtual reality (vr) remote collaboration: An exploratory study on user experience, Presence: Teleoperators and Virtual Environments 27 (2018) 93–120.

[13] S. Mystakidis, Metaverse, Encyclopedia 2 (2022) 486–497.

[14] G. Freeman, D. Acena, D. Maloney, J. Cherba, Collaborative virtual environment to facilitate avatar-avatar interaction in shared physical workspaces, in: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp. 1–8.

[15] T. Piumsomboon, D. Altimira, H. Kim, A. Clark, G. Lee, M. Billinghurst, Exploring natural eye-gaze-based interaction for immersive virtual reality, IEEE Symposium on 3D User Interfaces (3DUI) (2017) 36–39.

[16] Nanome Inc., Nanome vr software for molecular design, https://nanome.ai/, 2020. Accessed: 2026.

[17] zSpace Inc., zspace ar/vr learning platform, https://zspace.com/, 2019. Accessed: 2026.

[18] Khronos Group, The OpenXR specification, https://www.khronos.org/openxr/, 2019. Accessed: 2026.

[19] W3C, WebXR device api, https://www.w3.org/TR/webxr/, 2021. Accessed: 2026.

[20] Exit Games GmbH, Photon unity networking, https://www.photonengine.com/, 2020. Accessed: 2026.

[21] Unity Technologies, Unity netcode for gameobjects, https://docs-multiplayer.unity3d.com/netcode/, 2025. Accessed: 2026.

[22] M. McGill, J. Gugenheimer, E. Freeman, A quest for co-located mixed reality: Aligning and assessing slam tracking for same-space multi-user experiences, in: Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology, Association for Computing Machinery, New York, NY, USA, 2020. URL: https://doi.org/10.1145/3385956.3418968, accessed: 2026.

[23] M. Slater, Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments, Philosophical Transactions of the Royal Society B: Biological Sciences 364 (2009) 3549–3557.
[24] M. Tomasello, Joint attention as social cognition, Joint attention: Its origins and role in development (1995) 103–130.
[25] ngrok Inc., ngrok Documentation: Secure Tunnels to Localhost, ngrok Inc., ???? URL: https://ngrok.com/docs, accessed: 2026.
[26] L. Rebenitsch, C. Owen, Review on cybersickness in applications and visual displays, Virtual Reality 20 (2016) 101–125.
[27] European Commission, The european green deal, 2019. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2019%3A640%3AFIN.
[28] R. Albert, A. Patney, D. Luebke, J. Kim, Latency requirements for foveated rendering in virtual reality, ACM Transactions on Applied Perception, Vol. 14, No. 4, Article 25, 2017. URL: https://doi.org/10.1145/3127589.