SENEM: A Software Engineering-Enabled Educational Metaverse

Viviana Pentangelo, Dario Di Dario, Stefano Lambiase, Filomena Ferrucci, Carmine Gravino, Fabio Palomba

Software Engineering (SeSa) Lab - Department of Computer Science, University of Salerno (Italy)

Abstract

Context: The term metaverse refers to a persistent, virtual, three-dimensional environment where individuals may communicate, engage, and collaborate. One of the most multifaceted and challenging use cases of the metaverse is education, where educators and learners may require multiple technical, social, psychological, and interaction instruments to accomplish their learning objectives. While the characteristics of the metaverse might nicely fit the problem's needs, our research points out a noticeable lack of knowledge into (1) the specific requirements that an educational metaverse should actually fulfill to let educators and learners successfully interact toward their objectives and (2) how to design an appropriate educational metaverse for both educators and learners. **Objective:** In this paper, we aim to bridge this knowledge gap by proposing SENEM, a novel software engineering-enabled educational metaverse. We first elicit a set of functional requirements that an educational metaverse should fulfill. **Method:** In this respect, we conduct a literature survey to extract the currently available knowledge on the matter discussed by the research community, and afterward, we assess and complement such knowledge through semi-structured interviews with educators and learners. Upon completing the requirements elicitation stage, we then build our prototype implementation of SENEM, a metaverse that makes available to educators and learners the features identified in the previous stage. Finally, we evaluate the tool in terms of learnability, efficiency, and satisfaction through a Rapid Iterative Testing and Evaluation research approach, leading us to the iterative refinement of our prototype. **Results:** Through our survey strategy, we extracted nine requirements that guided the tool development that the study participants positively evaluated. **Conclusion:** Our study reveals that the target audience appreciates the elicited design strategy. Our work has the potential to form a solid contribution that other researchers can use as a basis for further improvements.

Keywords: Metaverse Engineering; Virtual Learning Environments; Human-Centered Studies; Software Engineering in Practice.

1. Introduction

Nowadays, the interest in the vast field of the metaverse has been steadily growing and is starting to find its place in various contexts and applications. A metaverse is a highly immer-

Email addresses: `vpentangelo@unisa.it` (Viviana Pentangelo), `ddidario@unisa.it` (Dario Di Dario), `slambiase@unisa.it` (Stefano Lambiase), `fferrucci@unisa.it` (Filomena Ferrucci), `gravino@unisa.it` (Carmine Gravino), `fpalomba@unisa.it` (Fabio Palomba)

sive three-dimensional digital world where users interact in real-time with the environment and others through their avatars, i.e., digital representations of the users [1]. The rapid technological advances, the increasing digitization of numerous daily activities, ranging from work to entertainment, and, notably, the impact of the COVID-19 pandemic that led to the need for digital solutions have all contributed to the significant surge in research interest in this field [2]. Indeed, recently, researchers started to design and develop metaverse for a plethora of tasks, ranging from entertainment to software development, passing through medical simulations and education contexts.

One of the most challenging use cases of the metaverse is education, where it may support educators and learners from many points of view. First, the remote nature of the metaverse could break down physical barriers between individuals, allowing geographically dispersed people to communicate and study together [3], letting individuals in less affluent living conditions to receive a good education. In addition, classrooms could benefit from students' diverse backgrounds by engaging in more interesting discussions [4]. Lastly, implementing artificial intelligence (AI) and visualization tools in the metaverse can further improve communication and collaboration, break down cultural and language barriers, and engage students with immersive instruments such as simulations and historical event reconstructions [4, 5].

Although there are various potentials for using the metaverse in educational settings, its limitations still impede its development and widespread adoption. The metaverse itself is a complex structure that needs to incorporate aspects across diverse fields [6], and this complexity becomes more and more evident when considering the educational context. Moreover, despite the possibility for the metaverse to offer immersive learning experiences and the large adoption of virtual technologies for education [7, 8], there is still a lack of investigations in the field, specifically an insufficiency of studies of optimal design for educational metaverse that aims to support educators and learners effectively. In this regard, one emerging primary issue is the lack of guidance on the specific functional requirements that an educational metaverse should possess. Indeed, when potential implementers of this technology do not have a clear set of criteria or capabilities to aim for, it can lead to inconsistent user experiences, limited effectiveness in educational outcomes, and ultimately, hesitation among educational institutions to adopt this innovative technology. These specific gaps in research outline our research problem, which points out **the lack of guidance on the specific functional requirements that an educational metaverse should enable**, which impacts the wider adoption of this technology in society.

This paper contributes to overcoming the limitations above through two main steps. First, we elicit and catalog a list of requirements following an engineering-focused approach. We conducted a literature survey followed by semi-structured interviews to validate our requirements list further. Second, we implement these requirements within an open-source metaverse platform named SENEM, i.e., Software Engineering-eNabled Educational Metaverse, tailored for remote educational and collaborative activities. We envisioned the tool as a platform that will be instrumental in defining further studies and tools. Lastly, we empirically validate SENEM to evaluate its learnability, efficiency, and user satisfaction.

Our findings are encouraging and represent a foundational starting point for future research, which can build both on top of the list of elicited requirements and the implemented platform. We believe that our work makes a series of solid contributions, specifically:

- a set of functional requirements *device-agnostic*—i.e., general requirements that are independent of the implementation device—that an educational metaverse should implement, elicited with a literature survey, followed by semi-structured interviews;

- the implementation of a new metaverse platform, coined SENEM, that addresses the requirements identified;
- a platform's evaluation through a Rapid Iterative Testing and Evaluation research approach to assess its usability and usefulness.

The paper is organized as follows. Section 2 presents a theoretical background and the analysis of the related work. Section 3 introduces the research questions of the study. Section 4 reports the research design of the literature survey process and the achieved results, i.e., the set of requirements. Furthermore, it describes the methodology for the interview study and the final set of refined requirements. Section 5 presents the prototype developed starting from the finding of the first research step. It describes the main characteristics and how each requirement has been mapped into a functionality of the tool. Section 6 reports the research design and the results of the final experiment operationalized to evaluate the usability of the prototype.

2. Background & Related Work

In this section, we briefly introduce the main concepts of our work. Subsequently, we conduct an analysis of related works found in the literature.

2.1. Background

The metaverse, a concept first introduced in Neal Stephenson's 1992 novel "Snow Crash" [9], has significantly evolved until it has become a feasible new reality. Deriving from the broader fields of Extended Reality (XR) and Virtual Reality (VR), with which it shares the action within virtual environments and a generally high level of immersivity, literature today is defining the metaverse as a three-dimensional virtual world, freely explorable, in which users interact with one another via their customizable digital representation, named *avatar* [1]. The metaverse further distinguishes itself for a series of key characteristics: persistence of the actions, high-level realism of the experience, interoperability among different virtual worlds, and a good scalable architecture [10]. It is synchronous and operates in real-time, supporting interactions among thousands or potentially millions of users simultaneously.

One of the fields where the metaverse is finding significant application is education. Indeed, its highly immersive and interactive nature opens up new educational possibilities, capable of breaking down physical and linguistic barriers. Especially after the COVID-19 pandemic [2], which has led to an increased general interest in developing digital solutions for teaching [11], several studies have laid the groundwork for designing and implementing such educational platforms [12–14]. Such an applicational context for the metaverse is also the field in which we frame our research. The metaverse for education is gaining momentum due to its great potential, but it must be more thoroughly investigated to provide new concrete solutions.

2.2. Related Work

In the following, we discuss the current state of the art concerning experiments conducted in an educational context that exploit metaverse platforms. Specifically, we will focus on two typologies of works: (1) works that have developed their learning environment on third-party platforms, specifically Second Life (SL), and (2) works that have presented prototypes of new metaverses for learning and collaboration purposes. While the systematic literature review conducted and reported in Section 4 has the goal of scanning existing articles to identify possible functional requirements to include in our prototype, this section aims at assessing the commonalities and differences between our education platforms and other, similar platforms previously proposed in the literature.

Second Life virtual environments. SL[1] is a virtual world online platform originally launched in 2003. It is a platform where users can create themselves and their virtual world, and it is possible to interact in real-time with other users, representing one of the metaverse-like platforms that have gained the most traction [15]. It is particularly relevant when analyzing the metaverse research field, as it has been the basis for many experiments in a research context due to its versatility. De Lucia et al. [16] presented a virtual campus for the Mathematics and Informatics Department at the University of Salerno created in SL, featuring collaborative zones, a student campus, lecture rooms, and recreational areas integrated with Moodle for session management and content. August et al. [17] developed the Virtual Engineering Sciences Learning Lab (VESLL) in SL to enhance STEM education with interactive tools and instant feedback, improving learning quality despite the platform's high hardware and connectivity requirements. Aydogan and Aras [18] used SL for a virtual laboratory in Programmable Logic Controllers, providing a 3D interface for learning tasks, but noted drawbacks such as synchronization issues without a fast connection.

All mentioned works utilize the SL platform to create customized learning environments, but they face limitations due to the platform's functionalities and lack of a requirements engineering process. Additionally, SL is also showing its age, with its outdated models reducing user engagement. Our work seeks to overcome these issues by developing a new metaverse-like platform, designed through a thorough requirements extraction process validated by literature and potential end-users, moving beyond the constraints of third-party platforms.

Metaverse prototypes for education. As mentioned above, we aimed to explore new solutions for interactive online learning environments not bound to third-party platforms to identify current trends and limitations of prototypes currently proposed in the literature. Duan et al. [19] introduced the CUHKSZ metaverse, a campus-based prototype, targeting campus students with smartphone sensor integrations for interaction and content sharing through lacking explicit academic spaces. Jovanović et al. [20] developed VoRtex, a Unity3D virtual learning environment focusing on avatar control and text communication but limiting other interactive features. Shen et al. [21] described V-Classroom, transforming physical spaces into 3D digital classrooms, focusing technically on representation without delving into broader metaverse implications. Sin et al. [22] prototyped K-Cube VR, an educational metaverse using head-mounted displays for exploring 3D knowledge graphs, which lacks personal avatars and real-time communication, focusing on educational impacts through interactive teaching methods. Finally, Fernandes et al. actively contributed to enhancing the current body on knowledge on the development of educational metaverses. In the first place, Fernandes and Werner [13] performed a scoping review targeting the main challenges and opportunities of a metaverse for software engineering education. Later on, Fernandes et al. [23] conducted a systematic literature review aiming at characterizing the state of the art of frameworks, models, and guidelines that assist learning through immersive platforms. Furthermore, Fernandes et al. [24] also proposed a Software Product Line-based approach to support the development of metaverse platforms.

[1] Second Life: https://secondlife.com/

As a result of the analysis of proposals found in the literature concerning the new prototypes, a crucial aspect that emerged is that the main emphasis is often placed on either the social aspect [19] or the educational aspect [20–22], with one aspect being prioritized over the other. In contrast, our work aims to draw from the insights gained from these studies and conduct a domain analysis with a literature survey to extract requirements that could effectively integrate the learning component with the collaborative and social elements among users. Our approach aims to achieve a balance and provide a holistic experience in the metaverse for educational and academic purposes. As such, our work complements the advances made so far, both in technological and scientific terms, e.g., by providing further insights that may complement the work by Fernandes et al. [13, 23, 24] on the development of metaverses.

3. Research Methodology

The goal of the study was to identify and validate a set of functional requirements for the development of a dedicated educational metaverse through a software engineering approach. The final purpose was to establish a foundational suite of features imperative for the well design and efficacy of educational metaverses. The study involves two perspectives: researchers and practitioners. On the one hand, researchers are interested in comprehensively understanding functional requirements in the educational field; they could also expand the user's needs using the metaverse and enhance its features within the educational context. On the other hand, practitioners are interested in developing educational tools that are founded on well-designed guidelines and principles in order to ensure a plethora of quality aspects within a highly interactive environment.

3.1. Research Questions and Method

Our paper is structured around the following main goal:

> ◎ **Research's Main Goal**
>
> *Define a comprehensive list of functional requirements for an educational metaverse, designing and developing the requirements within a novel and engaging prototype utilizing virtual reality technologies. Ultimately, it is essential to thoroughly assess the platform, focusing on usability parameters such as learnability, efficiency, and satisfaction.*

To reach the aforementioned goal, we identified two research questions that guided our research design.

> ⓘ **RQ$_1$**—What are the functional requirements that an educational metaverse should fulfill to let educators and learners reach their objectives?

The first research question aimed at identifying a set of starting requirements for an educational metaverse. We were interested in identifying functional requirements independent of the type of technology or device on which the metaverse can be implemented. We addressed the first research question (**RQ**$_1$) by means of two research methods. In the first place, we conducted a *literature survey* to extract the currently available scientific body of knowledge to extract an initial set of requirements using a traditional technique [25]. Afterward, we operationalized a *semi-structured interview* with seven researchers and Ph.D. students with mixed backgrounds and levels of experience in the field to validate the initial set of requirements and extend it. On the one hand, the *literature survey* approach allowed us to synthesize the current knowledge on the matter. On the other hand, the *semi-structured interviews* enabled a deeper and more robust analysis of the desirable characteristics of an educational metaverse.

After obtaining the set of requirements for the educational metaverse, we then designed and developed a prototype based on such requirements. We aimed to evaluate the tool from a usability perspective and obtain ulterior insights from a practical operationalization of our findings. For doing so, we formulated the second research question.

> ⓘ **RQ$_2$**—To what extent can an educational metaverse prototype empowered by software engineering satisfy the usability requirements of learners and educators concerning social and educational aspects?

To address the second research question (**RQ**$_2$), we adopt the *Rapid Iterative Testing and Evaluation* (RITE) research method [26, 27]. Through RITE, we evaluated and improved the tool's usability with 28 participants, resulting in a final platform—publicly available in our online appendix [28]—that can be used by other researchers to conduct further studies in the field. Participants were asked to complete a questionnaire to assess the prototype's usability, focusing on well-known metrics [29], i.e., *learnability*, *efficiency*, and *satisfaction* [30, 31]. The strategy was to incrementally enhance the usability of the prototype, achieving a stable version through a systematic process. This also involved refining the initial requirements in a continuous quality improvement cycle, ensuring the prototype's evolution towards significant quality improvement in functionality and user experience urging an integrative, multimodal approach to enrich usability evaluation through ongoing iterative methods [32].

Figure 1 outlines the three steps employed to answer the proposed research questions. In particular, the first step—highlighted in blue—focuses on eliciting and validating requirements to provide a software prototype for educators and learners to perform specific educational activities addressing **RQ**$_1$. In the second step, we utilize the requirements gathered in the first step to design and develop an educational metaverse, identifying actors and scenarios to be used. Thus, the third step—highlighted in yellow—employs the tool developed in the previous step to validate its usability with 28 participants, addressing the **RQ**$_2$. The aforementioned three steps have been described in Section 4, 5, and 6, thereby providing a comprehensive understanding of the research methodology.

3.2. Ethical Considerations

Since our study involved human participants, let us discuss some of the ethical considerations that guided all the phases of the study. At our university, studies involving human participants do not require approval from an Ethical Review Board yet. Nevertheless, the interview and survey design took into account numerous ethical and privacy concerns. Below, we outline the precautions we have taken to ensure full compliance with ethical considerations:

- All activities related to interviews and surveys were entirely anonymous. We recorded no identifying information of the involved participants.
- We made it clear to participants that they could withdraw their survey submission at any time and that no information entered up to that point would be tracked.

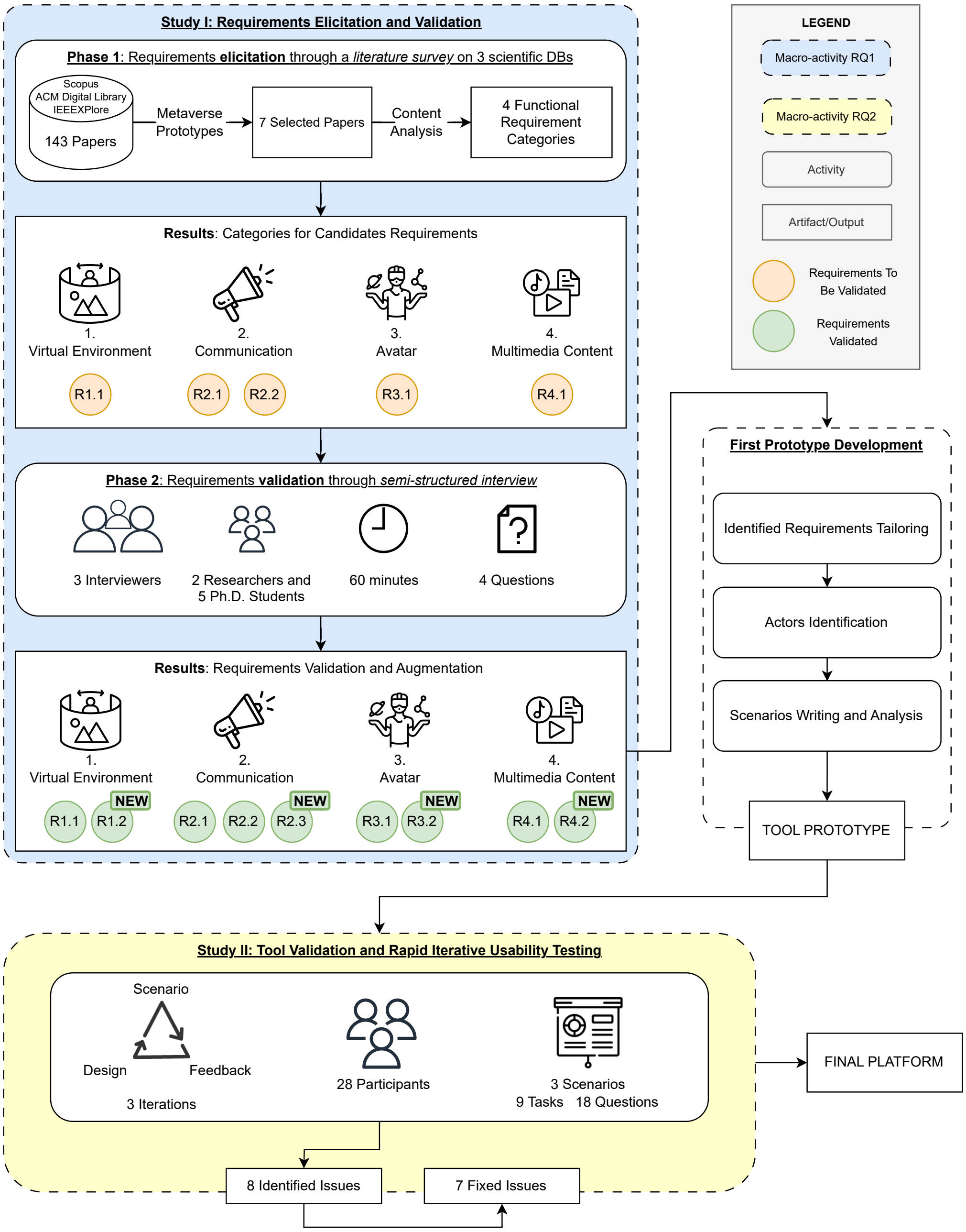

Figure 1: Overview of the research method.

- Before conducting an interview, we obtained explicit consent from each participant to record the session.
- Gender was the only sensitive personal data required in the tool's evaluation survey, which was collected for statistical purposes. However, it was an open-ended question, and participants were entirely free not to specify it.
- Access to the tool did not require any sensitive data. Participants were free to enter any name they wished and had complete freedom in customizing the appearance of their avatars.

These measures were all agreed upon and explicitly communicated by the authors of the paper. All of them were made clear to the participants before any surveys or interviews for our research were conducted.

4. Study I: Requirements Elicitation and Validation

In this section, we report the research method used to answer **RQ**$_1$, i.e., the requirements elicitation and validation for an educational metaverse. Specifically, we report the design and results for the literature survey and interview; since these methods are consecutive—i.e., the second depends on the results of the first—we decided to discuss them in the order of execution.

4.1. Literature Survey

As a first step to address **RQ**$_1$ we conducted a literature survey to collect the educational metaverse requirements that the research community has already identified in previous research. This step allowed us to have an initial set of requirements to extend and validate during the second step (i.e., the interviews). To conduct the literature survey, we followed the guidelines by Kitchenham and Charters [33], as detailed in the following sections.

4.1.1. Design

In the context of our literature survey, we performed three activities. First, we built a search string for identifying relevant literature; then, we collected the literature and checked if it was related to our objectives; in the end, we extracted knowledge to answer our RQ$_1$.

Starting from the first research question, we highlighted the main terms for our research. When defining the search string, we followed the well-established guidelines by Kitchenham and Charters [33] to ensure the completeness of the article collection phase. Indeed, we were interested in obtaining the requirements for a "*virtual environment*" to conduct "*educational activities*", leading to **metaverse** and **educational environment** as target domains. At this point, we derived a set of alternative spellings for broadening the scope of our research as follows:

Metaverse ("*metaverse*").

Educational Environment ("*distance learning*" OR "*online learning*" OR "*distance education*").

In constructing the query, the disjunction of synonyms was achieved by utilizing OR operators, whereas an AND operator established the conjunction of the target domain. This approach enables the query to be used effectively across different online data sources. The keywords of the second part of the strings focus mainly on the remote learning modality. This choice was motivated by two main reasons: (1) the surge in publications on the metaverse indeed started in the (post-)COVID-19 period, i.e., after 2020, when there was a growing need to devise digital solutions to continue remote education [2]; (2) literature advocated the role of the metaverse as a tool to support geographically distributed teaching since such distributed learning provides the opportunity for more people—facing various issues—to continue learning [34] and also shows outstanding commercial advantages [35].

> **General Search String**
>
> *((“metaverse”) AND (“distance learning” OR “online learning” OR “distance education”))*

We ran the search string against three well-known research engines, i.e., *Scopus*, *IEEEXplorer*, and *ACM Digital Library*, reporting 143 documents in total. Upon conducting a comprehensive review of the collected literature, we examined and chose papers pertaining to educational metaverse prototypes. This approach has allowed us to identify the most relevant and informative research papers, which have been subsequently selected for an in-depth analysis. The initial step in extracting data involves a *qualitative content analysis* [36], whereby researchers carefully examine the relevant data to extract meaning and concepts. This process entails two essential phases: (1) iterative development of a codebook and (2) application of the codes to the raw data. In the present study, the researchers who conducted the data extraction also developed the codebook, while the primary author applied the codes to the raw text. The data were sufficiently clear, and we did not require additional software tools during this phase. We report an example of how the first and second authors applied the data analysis procedure, passing from raw data to codes:

Raw Data: “*The VoRtex prototype contains a virtual environment classroom with 3D material that supports learning. The virtual environment classroom contains a table, chairs, projectors, a video and presentation panel, and a virtual agent that assists the avatars during lectures. [...] The users can communicate using the chat or voice systems.*”

Codes: *Virtual environment; using 3D learning materials; communication systems; avatar representation*.

All codes were extracted using an inductive approach from raw data. We analyzed the codes and their frequency to extract categories and functional requirements. However, it is important to emphasize that in cases of disagreement, we confidently engaged in discussions to resolve. Additionally, in an effort to minimize subjectivity in non-explicit requirements, the third author of the paper provided invaluable assistance in resolving any conflicts that may arise. As a final remark, the entire process was subject to continuous supervision by two senior researchers who diligently double-checked the activities performed to mitigate any inconsistencies that crept in.

4.1.2. Analysis of the Results

In our research, we conducted a thorough literature survey to identify the candidate requirements for developing a prototype metaverse tailored to educational purposes. In the following, we reported on the relevant literature obtained during our survey process. Then, we presented the identified requirements divided into categories extracted during our analysis.

Collected Literature Analysis. Table 1 summarizes the papers collected from the literature analysis.

Extracted Requirements. In the field of education, prototypes are built using different technologies and serve various purposes. In addition, research has emphasized the significance of active learning in virtual classrooms. This approach prioritizes learners' engagement in activities, allowing them to actively participate in these activities rather than relying on passive learning methods such as lectures or teacher-led instruction. This stimulates learners to collaborate, participate in group activities, and engage in discussions, increasing teamwork, decision-making, and communication skills [38]. These recommendations have inspired the design of our tool and, therefore, implement an active learning environment. On top of this active learning environment, we identified several unique characteristics to be considered in an educational metaverse. Their shared attributes help us to determine the needs categories that should be integrated into our prototype as candidate requirements. The following list describes the obtained categories and the relative requirements found:

Virtual Environment It is imperative to emphasize the importance of virtual environments as a fundamental element that distinguishes a metaverse from a traditional 2D platform, as pointed out by Mystakidis [41]. These environments facilitate user autonomy in exploring and navigating virtual settings. Creating such environments entails replicating familiar areas on a university campus, such as classrooms with desks and chairs. Users can interact with these elements to attend lectures or explore non-educational spaces.

- R1.1 *The system should allow users to explore all virtual environments. Exploration consists of interacting with objects, rotating the view, zooming in on components of interest, and walking.*

Communication Virtual learning environments should not only function as a means for learners to enhance their skills but also include features that facilitate efficient communication. Such communication should occur in both directions, among learners and among learners and instructors, and should be complemented by two separate modes: text chat and voice chat. Those methods will empower users to effectively exchange and deliberate on educational matters versatilely.

- R2.1 *The system should allow users to facilitate communication by allowing them to use a voice channel.*
- R2.2 *The system should allow users to facilitate communication by allowing them to use a text channel.*

Avatar Within the metaverse, avatars play a crucial role in crafting a digital identity and encouraging social connections in the virtual realm. They empower users to undertake, communicate, socialize, and participate in various activities tailored to their tastes, whether rooted in reality or imagination. As such, avatars are essential for personalization and interaction within the metaverse.

- R3.1 *The system should allow users to have a digital replication of themselves to attend a large number of activities.*

Table 1: Summary of the Collected Literature

Reference	Description
Duan et al. [19]	A metaverse prototype developed with the primary aim of allowing new learners to explore the campus and create shareable digital content. Hence, the focus is on building a **virtual environment** that closely replicates the physical environment, including classrooms, rooms, and offices, where users can move freely. Additionally, users can **communicate** with one another and establish social and collaborative relationships through voice or text chat.
Jovanovic et al. [20]	VoRtex, a Unity-based platform allowing multi-user interactions in a 3D **virtual environment**. With VoRtex, online classes can be conducted using **avatars** to interact with other users and access **multimedia contents**, such as teaching materials. The virtual classroom is designed to resemble a real-life classroom with desks, chairs, and projectors. Moreover, avatar's **communication** is allowed through the use of text chats that support speech-to-text, making it easier for users to complete writing tasks.
Shen et al. [37]	A prototype called V-Classroom that uses Internet of Things (IoT) devices such as sensors and high-performance cameras to convert physical university classrooms into 3D **digital environments**. V-Classroom does not use avatars to represent learners: they are automatically inserted into the virtual classroom with no possibility to explore the environment.
Sin et al. [22]	An educational metaverse based on a system of nodes. Each node represents an academic course and allows for **multimedia content** sharing. However, although the proposed prototype encourages exploration, the nodes only provide access to new environments without emphasizing user interactions.
Ng et al. [38]	A virtual platform to enrich learners' educational experiences. It allow users, even if not represented by an avatar, to explore and observe **virtual environments** with a 360-degree view of the entire developed environment. Moreover, exploration and observation were enhanced by connecting and **communicating** with other users to exchange information according to specific requirements.
Song et al. [39]	The Learningverse, a metaverse platform which is based on three key concepts: (1) the interaction in the platform through an **avatar**, which is customizable within the platform itself; (2) a 3D **virtual environment** in which users can **communicate** and cooperate to learn with many 3D objects and different **multimedia content**, such as images, videos, or text resources; (3) a peer-to-peer network to allow users to connect via Internet.
Schaf et al. [40]	The 3D AutoSysLab, a metaverse platform prototype, to assist students in hands-on electrical engineering lab sessions. The prototype presents a Mixed Reality system that integrates a 3D **virtual environment** and physical objects for the lab session. Users are represented by their **avatar**, and they must **communicate** and cooperate to complete the assigned task. The learning material is shown in the virtual world as **multimedia content**, such as images or slides.

Multimedia Content The effective management and representation of digital content plays a crucial role in the success of a metaverse. This allows users to easily create, share, and consume content. This functionality has added significance in an educational metaverse as it offers access to various educational resources, such as images, videos, presentations, audio, and interactive media. The availability of such content enhances the educational potential of the metaverse.

R4.1 *The system should allow users to projector and share educational multimedia resources.*

4.2. Interview

Our second investigation consisted of an interview study. Specifically, we conducted *semi-structured interviews* [42]; they combine specific questions (for studying the main topic to be covered by the research) and open-ended questions (to elicit unexpected types of information). The conduction of an interview holds significant importance for two fundamental reasons: (1) to validate the requirements that have already been established in the literature, thereby adding valuable contribution to the existing body of knowledge, and (2) to elicit various perspectives from educators based on their positive and negative experiences during the COVID-19 pandemic, but also to have their vision about the field.

4.2.1. Design

By following the guidelines by Hove and Anda [42], we crafted a semi-structured interview with straightforward and accessible questions for our participants. For the interview, we decided to utilize an online format via Zoom[2] as our communication tool. This choice was made due to its ability to record the meeting without the file expiring and its automatic transcription feature. However, as technical issues may arise, we had three interviewers present. Our primary interviewer was the first author, with the second and third authors manually transcribing concepts and offering support throughout the interview. It is worth noting that all participants present during the meeting were duly informed of the recording that was to take place and provided explicit consent.

The semi-structured interview consisted of three parts. It began with an introduction and ice-breaker where participants shared their research interests and demographic details to provide context on their career backgrounds. Then, the first author then outlined the study's goals and explained key metaverse concepts like interaction techniques and educational technologies, necessary because of the novelty of the topic. After confirming participant comfort, the main interview began with questions divided into four categories: virtual environment, communication, avatar, and multimedia content, as detailed in a previous section on candidate requirements.

All data collected were combined and handled in accordance with data regulations, which specified that information would be used just for research purposes (more details about the questions can be found in our online appendix [28]). Participants were informed of these policies to ensure transparency and clarity throughout the process. After completing all interviews, the first and second authors collected the text, comparing the automatic and manual transcriptions. Then, they conducted a quality assessment phase to evaluate the overall data quality for the content analysis phase.

[2] Zoom: `https://www.zoom.us`

4.2.2. Participants

Our study required participants with a certain level of proficiency in educational settings. We have defined specific criteria for each participant, including high skill and expertise in software engineering, human-computer interaction, and computer science education. To meet these requirements, we have chosen *convenience sampling*, a non-probabilistic approach that selects participants based on proximity, availability, and willingness to participate [43].

Our sample consisted of 2 highly experienced senior researchers and 5 Ph.D. students enthusiastic to participate in the study. The motivation behind the decision to recruit both faculty and students was that only by combining feedback from both was it possible to capture all the facets needed to achieve our goal. We carefully selected diverse participants, including 4 male and 3 female participants. Senior researchers have dedicated over 20 years to research and academia. Indeed, their areas of expertise range from software engineering, focusing on application design and usability engineering, to computer science education. The Ph.D. students' research topics are equally impressive and cover a broad range of software engineering domains, including software engineering for accessibility, software engineering education and training, and human and social aspects of software engineering.

4.2.3. Procedures

To initiate the process, we utilized a scheduling tool to arrange individual interviews with our 7 participants. We decided to conduct each interview individually to allow each participant to freely discuss their point of view. As a first step, we obtained permission from each interviewee to record the entire session. Once this permission has been explicitly obtained, we connected to the Zoom platform with the participant. Each participant's interview lasted for 60 minutes. The first part of the available time was used to break the ice with the interviewee and provide a useful informational overview of the topic. Then, one by one, questions were posed to the interviewee, who was free to discuss openly, guided by the interviewers. The call was recorded, but at the same time, the interviewers took note of the key points highlighted from the perspective of each of the interviewees. At the end of the discussion, the interviewee was allowed to add any final comments, after which the interview was terminated.

4.2.4. Analysis of the Results

All the semi-structured interviews confirmed that the candidate requirements obtained through the comprehensive literature survey were suitable for inclusion in an educational metaverse. However, the contribution of 7 respondents in each candidate requirements category was even more significant, as further reported in the following.

Virtual Environment When participants were asked to envision an educational virtual environment, it became clear that greater freedom to explore could ignite curiosity and facilitate learning, enabling learners to pursue their interests independently. Furthermore, participant P4 provided an insightful perspective: "*During the COVID-19 pandemic, I attended several meetings on Zoom, each with a different name and password. An educational metaverse should provide the same structure, with scalable rooms with passwords and the ability to facilitate multiple simultaneous sessions to foster interaction and social relationships between learners and educators.*". Considering this, we have added a scalable room system requirement to the criteria for virtual environments:

R1.2 *The system should allow users to create simultaneous sessions that are protected by a password.*

Communication All 7 participants confirmed during the interview that communication in a metaverse should incorporate textual and vocal techniques. However, an exciting reflection emerges from the interview with P1, who states: "*Utilizing both methods is crucial since the textual technique allows for retaining important information, whereas the vocal approach promotes social interaction and enhances communication efficiency. Moreover, the vocal approach creates a sense of presence akin to physically being with another individual. Additionally, we should include nonverbal interactions such as facial expressions and gestures to enhance the immersive experience and perception of presence.*". The authors have unanimously recognized this view as one of the most relevant, as it is intrinsically linked to the closest definition of a metaverse. However, it should be noted that the accurate representation of facial expressions in a metaverse is still a challenge. To overcome this limitation, the feature of "Emotes", which consist of predefined facial expressions or body movements activated at the user's discretion, was introduced:

R2.3 *The system should allow users to interact using non-verbal interaction, represented by "emotes".*

Avatar Regarding the creation of avatars, i.e., digital representations of users within the metaverse, all participants agreed on their relevance for facilitating interactions. However, when we asked what main aspects should characterize it, P3 pointed out,"*I think you should not classify features as more or less important. Everyone might have different opinions about it. From my point of view, I would like to have the ability to customize my avatar according to my preferences. In real life, I can decide how to comb my hair, what to wear according to my mood, and add any accessories. This flexibility could also include users with disabilities.*". This comment stressed the importance of allowing users to customize their avatar from two perspectives: (1) physical characteristics, such as skin, lip, and eye color; (2) personal characteristics, including hairstyle, clothing, and accessories:

R3.2 *The system should allow users to customize their avatars based on physical and personal characteristics.*

Multimedia Content Finally, participants were asked to share their preferences for multimedia content in an educational metaverse. Participants emphasized the importance of projecting and presenting slides or images for educational purposes. Moreover, the interview with P2 highlighted a key point: "*In my lectures, slides are essential to illustrate my talk and provide references to students about what I am explaining. However, using a whiteboard, I always try to change the use of slides with practical examples*". Based on this observation, it was decided to include the virtual whiteboard as an additional multimedia requirement:

R4.2 *The system should allow users to interact with a whiteboard, sharing all the content with other users.*

Summary of the Requirements Elicitation Study (RQ_1)

The requirements obtained during our literature survey (Section 4.1) were validated by the answers provided by the participants of our interview study (Section 4.2); moreover, the last study also corroborated the requirements with new ones. The overall set of elicited requirements is reported and described in Table 2.

Table 2: Summary of the Functional Requirements for Study I

Category	Functional Requirement	Source
Virtual Environment	R1.1: *The system should allow users to explore all virtual environments. Exploration consists of interacting with objects, rotating the view, zooming in on components of interest, and walking.*	Literature Survey
	R1.2: *The system should allow users to create simultaneous sessions that are protected by a password.*	Interview
Communication	R2.1: *The system should allow users to facilitate communication by allowing them to use a voice channel.*	Literature Survey
	R2.2: *The system should allow users to facilitate communication by allowing them to use a text channel.*	Literature Survey
	R2.3: *The system should allow users to interact using non-verbal interaction, represented by "emotes".*	Interview
Avatar	R3.1: *The system should allow users to have a digital replication of themselves to attend a large number of activities.*	Literature Survey
	R3.2: *The system should allow users to customize their avatars based on physical and personal characteristics.*	Interview
Multimedia Content	R4.1: *The system should allow users to projector and share educational resources.*	Literature Survey
	R4.2: *The system should allow users to interact with a whiteboard, sharing all the content with other users.*	Interview

5. SENEM: A Novel Software Engineering-Enabled Educational Metaverse

In this section, we describe the tool developed according to the requirements elicited in the first phase of the study (Section 4), while in the next Section 6 we present the evaluation of and improvement strategy for the tool. We decided to report in the following order because of the impossibility of understanding the evaluation experiment without having an overall idea of the tool composition and function.

The developed platform is publicly available on its GitHub repository.[3] The repository contains both the Unity3D project with the source code and the already built application for Windows and MacOS.

5.1. Requirements Implementation

In the following, we describe how the requirements identified in the first phase of the study (see Table 2) have been concretely implemented into the metaverse platform. Figure 2 shows an overview of the tool architecture and event flow—better discussed in Section 5.3—focusing on which component implements the identified requirements.

[3] SENEM Github Repository: `https://github.com/vipenti/SENEM_Metaverse`

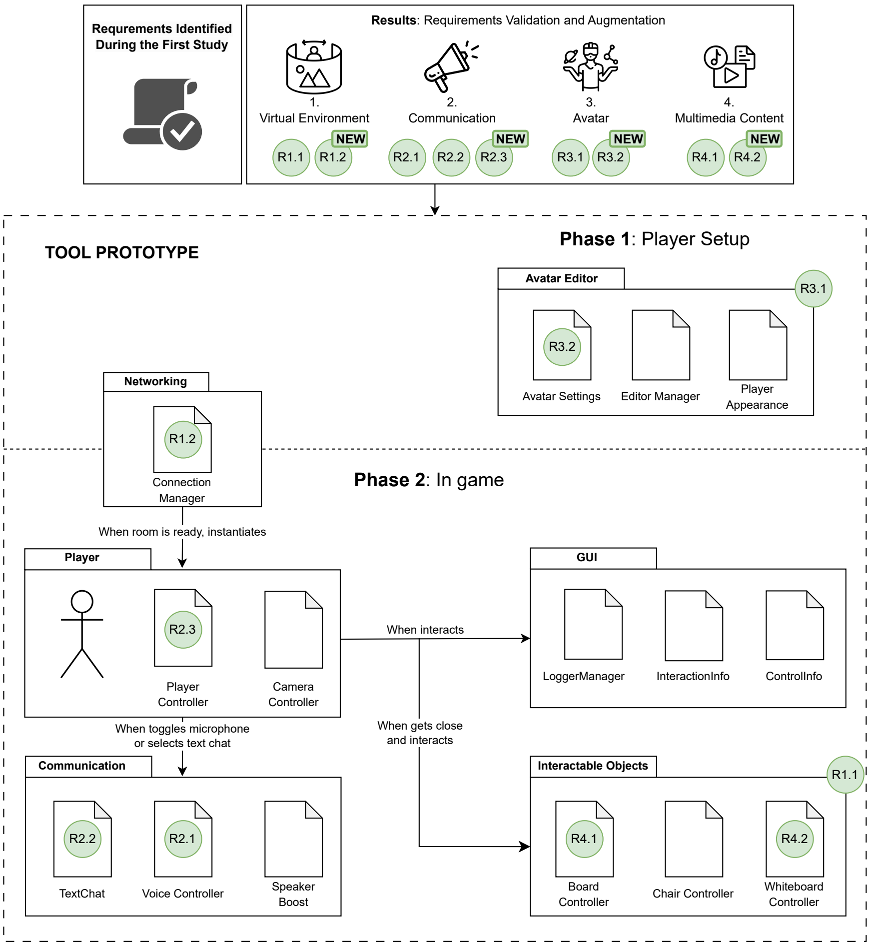

Figure 2: Overview of the prototype and requirements.

Requirements of Virtual Environment. Such a category refers to requirements to provide users with an environment to facilitate user autonomy in exploring and navigating virtual settings.

- *Scalable rooms system.* The platform offers real-time connection with other users as a critical feature. This connection has been implemented through a *room system*, allowing users to choose whom to connect with. The first user intending to connect will create the room, providing it with a password, and then share this information with the other interested people. They can use it to connect to the room and initiate real-time interaction with other users. When each user disconnects from the room and it becomes empty, it is automatically deleted.

- *Free three-dimensional exploration*. Interaction with the platform occurs freely, exploring the 3D environment through one's avatar. Users can walk, sit, and interact freely with various objects in the scene. They can rotate their view and adjust the zoom level of their perspective as well.

Requirements of Communication. Such a category refers to requirements to provide users with tools and methods for communicating with each other easily.

- *Voice communication*. Users can engage in real-time communication with each other through voice chat. Using their microphones and voices, they can make their avatars speak and hear others as in genuine verbal communication. The voice chat is also equipped with proximity and three-dimensionality features, allowing the tone of voice to vary based on the distance and position of the interlocutor.
- *Text communication*. Users can communicate through text chat, which is readily available and visible in the platform's user interface. Such a chat always displays the last ten sent messages on the screen. Additionally, users can review the entire chat history after the session, through a log file stored locally on their device—as notified upon accessing the virtual classroom.
- *Non-verbal interactions*. Users can use some gestures to interact with each other non-verbally. Avatars are equipped with communication animations, such as waving, clapping, or raising a hand. All users in the virtual classroom are notified when one of these actions is performed to decrease the likelihood of missing actions not directly seen by the avatar.

Requirements of Avatar. Such a category refers to the possibility of users creating an avatar.

- *Avatar's customization*. Users can customize the avatar's appearance through which they interact with the platform. The tool provides a dedicated editor to modify various avatar elements, including skin, hair, eyes, and more.

Requirements of Multimedia Content. Such a category refers to requirements to provide users with tools to easily create, share, and consume content.

- *Projection and presentation of multimedia content*. Users can broadcast multimedia content in the virtual classroom through the available projector. Educational materials like slides or images can be uploaded and displayed during the platform's runtime. Users can control the content shown on the projector and make it visible to everyone else.
- *Interaction with the whiteboard*. Users can write on the whiteboard inside the classroom using their keyboard. One user at a time can approach the whiteboard and start writing, and what they write will be visible to everyone, supporting interaction and communication.

5.2. Actors of the system

This section introduces the actors that can be found in the platform and what distinguishes them from each other. The functionalities accessible to each user vary slightly depending on their role in the virtual classroom. Such a division is not clear-cut within the environment, but it will be used to better describe potential interactions within the tool.

The *room owner* is the first user intending to connect, thus creating the room and setting its password. This user also has the complete chat and access log on their device. All other users connecting to the room are *attendees* and can interact with each other and use all the communication methods. If the room owner crashes or disconnects, the second user who joined takes over their role to prevent the entire room from being deleted. The new room owner has access to the room's log from when they assumed the role. When a user approaches the projector, they can take on the role of *presenter* and control the displayed content. There can be only one presenter at a time, and the role is passed when a new user approaches and interacts with the projector.

5.3. Architecture

In the following section, we describe the architecture behind the tool and the event-based system that governs user interaction with the platform.

The application was developed with the UNITY3D game engine, which shaped its architecture into three core components: (1) the game scene, (2) game objects, and (3) assets. The game scene forms a 3D virtual environment where all elements are integrated and interactions take place. It includes the virtual classroom scene, which houses most application logic, and the character editing scene. Game objects, such as avatars and GUI elements like buttons and text fields, consist of scripts, meshes, and textures that determine their functionality and appearance. Assets cover all resources used in the scene like sounds and animations.

The application operates on an *event-based* logic, where functions are triggered by specific events rather than a central management system. This involves code scripts attached to each component, activated by respective events. The code scripts can be logically organized into clusters that represent different functionalities of the platform. User interaction with the app is split into a setup phase, where users create avatars, and an in-game phase that starts when entering a room, with functions executing based on user actions. Throughout, networking features run continuously to manage user interactions and room dynamics.

5.4. Graphical User Interface (GUI)

In this section we illustrate the different screens the user encounters and the commands through which they can interact with the environment.

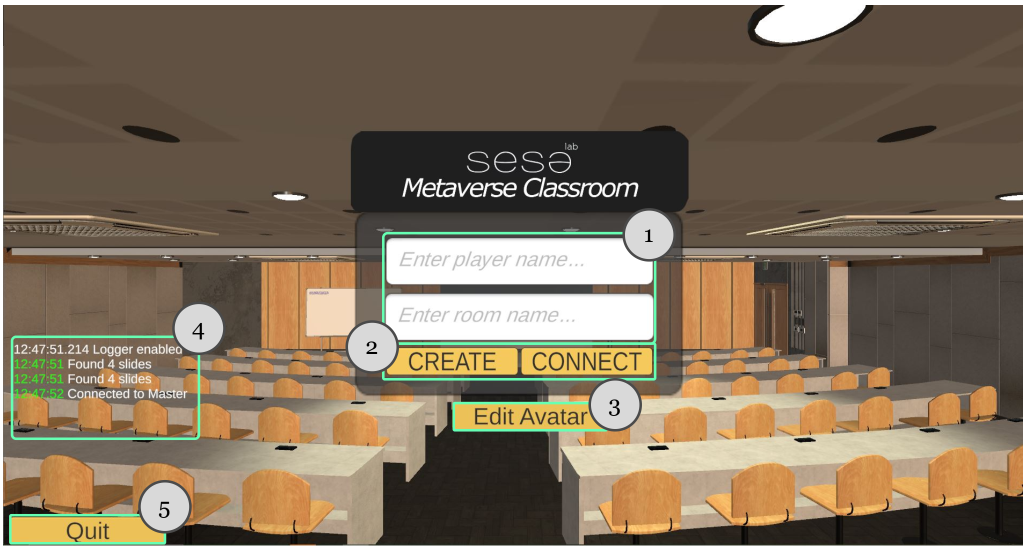

Figure 3: Starting screen of the tool.

Figure 3 displays the interface of the tool's startup screen. When users open the application, they encounter the main elements shown in the green squares. The enumerated components are explained below.

1. The form with login information: In the first text input area, the user enters the name that will be displayed to others in the classroom, and in the second, the password of the room they want to join.

2. *Create* or *Connect* to room buttons: If a user is the first to connect, after entering the data, they will press the *Create* button to create the room. Otherwise, they will press *Connect*. In both cases, if this operation is successful, the user will have access to the virtual classroom.

3. The *Edit Avatar* button: This will take the user to the avatar creation screen.

4. The logger: It displays information about the application's status. If the log shows "Connected to Master," the server connection was successful, and the user can create and access rooms.

5. The *Quit* button: It allows to exit the application.

Figure 4: Avatar editing screen.

When the *Edit Avatar* button is pressed, the user is redirected to the avatar editing screen. In such a screen (see Figure 4), the user can customize their avatar's appearance, which will then represent how other users see them within the virtual classroom. The enumerated components are explained below.

1. Editing window: The interface through which the user customizes their avatar's appearance. It is divided into sections, each controlling the color and shape of one of the customizable components: uniform, skin, lips, eyes, hair, beard, eyebrows, and accessories.

2. Avatar preview: The model that shows real-time changes to the appearance made through the editing window. It can be rotated and zoomed in to inspect it thoroughly.

3. *Load Game* button: It allows to save the changes and return to the application's initial screen.

4. Info label: It informs users of the commands to control the model during the editing session.

Figure 5: Virtual classroom's interface.

Upon entering a room, users find themselves in a 3D virtual environment, viewing the virtual classroom and other users in real-time from a first-person perspective, with the avatar's camera positioned at head level. The interface, shown in Figure 5, includes the following components:

1. The virtual environment and other users: Users see each other's avatars and names, observe real-time movements, and hear voice chats, with the intensity varying by distance.

2. Text chat: Displays the last ten messages along with the sender's name and timestamp. A text input field allows for sending new messages.

3. *Leave* button: It returns the user to the platform's initial screen.

4. Info label: Displays contextual information, such as whether the user is near the stage or a chair, and whether these are occupied or available

5. Voice chat info: Shows the detected input device and microphone's status—i.e., *muted* or *unmuted*.

6. Command info and *Options* button: Provides quick command info and access to settings for microphone sensitivity, resolution, and ambient sound adjustments.

Within the virtual environment, users can interact with various elements to access different features provided by the tool. Figure 6 displays the following elements that users can interact with, namely: (1) the whiteboard, where they can write as an editable text area shared by all users in the room; (2) desks and chairs where users can sit and observe, as well as interact with the presenter; (3) the stage and projector, where the presenter can display and control slides for other users to see, and view them themselves while speaking, thanks to the synchronized display in front of the projector.

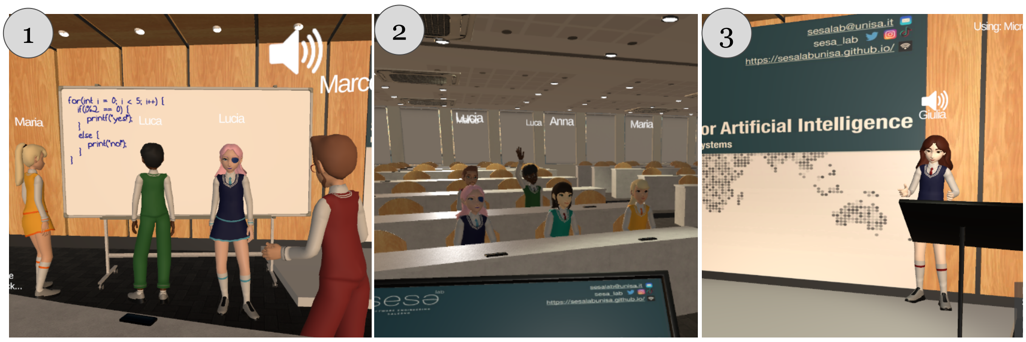

Figure 6: Users interacting with the virtual environment.

6. Study II: SENEM Validation with Rapid Iterative Testing and Evaluation (RITE)

In this section we describe the method used to address **RQ**$_2$, i.e., to evaluate the tool's usability and improving it until its final version. In particular, we first present an overview of the *Rapid Iterative Testing and Evaluation (RITE)* [26, 27] approach. Then, we describe the tailored version of RITE for our study.

6.1. The RITE Methodology

RITE was first introduced by Medlock et al. [26, 27] in video game development. Specifically, the approach was used to identify and fix issues—from a usability point of view—of the tutorial of Age of Empires II.[4] Concretely, RITE consists of evaluating a tool's usability through an iterative approach, where each iteration consists of (1) testing the tool with a set of potential users and (2) modifying the tool according to participants' feedback [26, 27]. The decision of what to fix or ulteriorly test is in charge of the research team, and the process should end when the product reaches enough quality.

6.2. The RITE Evaluation

We applied RITE to evaluate the usability of our tool and its functionalities. Specifically, we represented usability in terms of three well-known aspects [30, 31]:

- *Learnability*, i.e., the measure that represents how easy it is for users to learn how to use the tool.

[4] Age of Empires II (HD) official website (accessed on September 2023): `https://www.ageofempires.com`

- *Efficiency*, i.e., the trade-off between the resource used—expressed in terms of completion time, mental, and communication effort—and accuracy and completeness with which the users achieve their goals.
- *Satisfaction*, i.e., the user's satisfaction when using a tool in a specific context.

The choice of opting for these metrics rather than alternative ones, e.g., the System Usability Scale [44], was mainly driven by our willingness to use metrics that might have assessed SENEM under multiple perspectives, such as learnability, efficiency, and satisfaction. Indeed, the metrics by Nielsen [30] and by Constantine and Lockwood [31] allow to consider a broader spectrum of variables, hence providing insights that might have better reflected the overall experience of users interacting with SENEM. Furthermore, these metrics have been largely recognized in the field of usability assessment [29]. In the following, we present the experiment design for evaluating the above-mentioned aspects as well as the obtained results and implications.

6.2.1. Experiment and Survey Design

Our evaluation consisted of the execution of 9 tasks described—described below—using our platform by a set of participants.

T1 Creating an avatar using the built-in editor

T2 Performing the login in the application

T3 Using the projector and associated tools to share and present slides

T4 Sitting on chairs in the room

T5 Communicating performing an avatar's emote

T6 Communicating using the built-in voice system

T7 Communicating using the built-in textual chat

T8 Writing something on the whiteboard

T9 Writing something on the notes' tablet

The same tasks—or a subset—were executed at each iteration to identify and fix issues, and we selected the tasks to test all the platform functionalities. Such process continued until (1) no more critical issues were identified, i.e., issues preventing users from performing tasks, and (2) the participants agreed that the tool's usability was good. To evaluate the outcome of each iteration, we relied on surveys—administered to the participants after each test—and findings obtained during the iteration execution.

Experiment Design. According to the RITE approach guidelines, we evaluated the tool through a set of tasks encompassing real-world use scenarios. The first one consisted of a *Remote seminar*: a group of individuals was asked to attend a seminar in the metaverse. In particular, one of the participants was the speaker researcher, while the others were the attendees, composed by both students and other researchers. In the second one, a group of students was asked to deliver an *exam* project's presentation in the metaverse. This time students involved in the evaluation played the presenter role in turn and the teacher participated as an attendee. Last but not least, participants were asked to collaborate to perform a *group activity* in the metaverse. In particular, each group of students had to access the virtual classroom and find a space to work alone, sharing the available resources—i.e., the projector and the whiteboard. It is important to highlight that all the described scenarios were real tasks that had to be performed, i.e., in a completely real context rather than a simulation.

From a practical point of view, a set of participants were recruited for each iteration to participate in the scenario activity. Before the execution date, the authors contacted the participants and helped them set up the platform on their workstations. This allowed us to identify potential compatibility problems and rapidly fix them. Before the execution, all the participants were asked to start the platform and enter the session room. The participants were physically dispersed, i.e., they participated in the experiment remotely, while the authors distributed themselves in order to provide help if some problems—also technical—arose.

During each iteration, the authors of the paper participated in the experiment. Following RITE, we recorded all interesting data and issues identified during the experiment, as well as participants' feedback, and analyzed them afterward.

Survey Design. After each iteration, using Google Forms—we administered an anonymized survey to the participants in order to (1) evaluate their experience, (2) evaluate the tool's usability, and (3) collect follow-up feedback. Details about the administered questions are in Table 3.

The survey is structured into four sections. The first section gathers demographic data such as gender, role, and years of experience of the participants. For students, 'years of experience' refers to years of academic study. The survey is anonymous and does not require identifying information. At the end of this section, participants are asked if they acted as a presenter, using presentation tools, or as an attendee, observing others' presentations without using these specific features. Depending on their response, participants are directed to the second section if they were presenters, or to the third section if they were attendees. The second section, once completed, also leads to the third section. These sections feature closed-ended questions designed to measure the usability of the tool through metrics such as satisfaction, learnability, and efficiency, with responses recorded on a 5-point Likert scale, with 1 representing the lowest and 5 the highest. The fourth section includes an optional open-ended question to gather additional feedback, report bugs, and suggest improvements, made optional to avoid collecting uninformative responses.

Regarding the results analysis, we used descriptive statistics/plots for the closed-ended questions of the surveys, while content analysis [36] for extracting potential ideas from the open-ended feedback.

The selection of survey questions was a result of a collaborative effort. Such an activity was conducted during three meetings, each involving researchers with at least three years of experience in survey design and survey-based research. Each question underwent thorough discussion in these meetings until a consensus was reached. Subsequently, the initial survey design was validated through a pilot study involving three researchers who were neither involved in the tool's development nor the experiments. Adhering to the guidelines outlined by Flanigan et al. [45], a deliberate decision was made to maintain anonymity in the survey, ensuring that the influence of the authors on respondents' answers was minimized. At the end of this process, the survey was created using Google Forms[5], and the completion time was estimated to be ca 10 minutes.

[5]Google Forms website: https://www.google.com/forms/about/

Table 3: Follow-up survey questions.

ID	Question	Type	Metric
Q1	What gender do you identify with?	Open-ended question	Demographic
Q2	What is your academic role?	{Student, Researcher}	Demographic
Q3	How many years of experience do you have?	Numeric	Demographic
Q4	Were you a presenter or a listener?	{Presenter, Listener}	-
Q5	How satisfied are you with the overall usability of the tool?	Likert scale of five points from *Very dissatisfied* to *Very satisfied*	Satisfaction
Q6	How easily did you learn to use the tool for your presentation?	Likert scale of five points from *Very difficult* to *Very easy*	Learnability
Q7	How would you rate the presentation modality provided by the tool?	Likert scale of five points from *Very bad* to *Very good*	Efficiency
Q8	How easy is it to control your presentation and avatar during the exposition?	Likert scale of five points from *Very difficult* to *Very easy*	Learnability
Q9	How easy is to interact with your listeners and answer their questions?	Likert scale of five points from *Very difficult* to *Very easy*	Efficiency
Q10	How would you rate the quality of communication channels?	Likert scale of five points from *Very bad* to *Very good*	Satisfaction
Q11	How likely are you to use this tool compared to online communication platforms (e.g., Teams, Zoom, Meets...) to give your presentations?	Likert scale of five points from *Not at all likely* to *Very much likely*	Satisfaction
Q12	How engaged did you feel while experiencing the presentation?	Likert scale of five points from *Not engaged at all* to *Very engaged*	Satisfaction
Q13	How easily did you learn to control and position your avatar?	Likert scale of five points from *Very difficultly* to *Very easily*	Learnability
Q14	How satisfied were you with the interaction capabilities with the presenter and other participants during the presentation?	Likert scale of five points from *Not satisfied at all* to *Very satisfied*	Satisfaction
Q15	How well were you able to view and read the presentation?	Likert scale of five points from *Very difficult to see and read* to *Very easy to see and read*	Efficiency
Q16	How well were you able to listen to the presenter?	Likert scale of five points from *Very difficult to listen* to *Very easy to listen*	Efficiency
Q17	How likely are you to use this tool compared to online communication platforms (e.g., Teams, Zoom, Meets...) to attend to presentations and seminars?	Likert scale of five points from *Not at all likely* to *Very much likely*	Satisfaction
Q18	Please share any additional feedback or suggestions you have regarding the tool's features, usability, or any improvements you would like to see.	Open-ended question, Optional	Learnability, Efficiency, Satisfaction

6.2.2. Participants

Regarding the recruitment of participants, in the beginning, we adopted *convenience sampling*, i.e., a non-probability sampling method where the sample is taken from a group of people easy to contact [43]. We decided to adopt such a strategy because of the availability of participants from our university. However, we were aware that such a strategy is known to provide results that are not highly generalizable [43, 46]. For this reason, we identified a series of base criteria that the participants should meet (described in the following). Moreover, we plan to perform replications of the validation experiment in order to strengthen our results.

Following the RITE approach guidelines [26, 27], we defined some fundamental criteria that each participant had to meet. Specifically, we were looking for students and researchers with experience in human-computer interaction. Our choice was motivated by the aim of obtaining feedback supported by a theoretical knowledge foundation. These participants needed to be affiliated with the academic context for which the tool was intended, and they had to be individuals capable of genuinely representing users who would use the platform. Among these, we selected individuals with knowledge or experience about the metaverse to obtain opinions informed by other experiences. We did not define more stringent criteria to avoid bias in our results and tried to keep our set of participants as heterogeneous as possible.

In the end, 28 participants were recruited, comprising 5 researchers and 23 students of the master's and Ph.D. programs. As previously specified, each was already meant to participate in the activities later scheduled within the metaverse. Consequently, students for each activity were selected from the pool of participants available, aiming to align with the criteria mentioned before. 13 participants participated in the first activity (*remote seminar*), 5 participants in the second (*exam presentation*), and 10 participants in the third activity (*group activity*).

6.2.3. Procedure

We followed the same procedure for each of the three iterations performed with RITE. In the preparation phase, participants received the platform prototype to install on their devices at least 24 hours prior. This addressed technical issues beforehand, ensured everyone's participation, and gathered initial installation feedback. Organizers assisted in resolving any problems, and no issues preventing individual participation were encountered in any of the three iterations.

Participants and organizers were physically present at the university for the experiment setup. Participants were divided into small groups, ranging from 1 to 3 individuals, and they were physically separated to enable remote communication. Within each group, individuals were positioned far apart in the room to restrict communication to the platform. This ensured a controlled environment and organizer presence to guide participants. During the third iteration's group activity, participants in the same group collaborated within the platform despite being in different rooms.

The experiment began with an organizer accessing the platform first and creating a room password for all participants to connect. After participants created their avatars and accessed the platform, 10 minutes of free exploration allowed for connectivity checks and interface familiarization. Organizers then explained activity guidelines, e.g., project presentation methods or collaborative task instructions. Each activity lasted 45 minutes, during which platform behavior and technical issues were observed. Subsequent iterations addressed previously highlighted problems. The activity concluded after 45 minutes, followed by a 10-minute evaluation survey. Upon completing the survey, the session ended, and participants left the experiment location.

6.2.4. Analysis of the Results

The usability evaluation concluded after three iterations when the tool reached an acceptable state, and no issues were encountered that prevented the correct completion of the tasks and activities. As mentioned above, the activities carried out during the iterations were: (1) the *remote seminar*, with 13 participants, marking the first real test of the platform with multiple users connected in a single room; (2) the *exam presentation*, in which a group of 4 students presented their exam project to a professor using the platform, resulting in a total of 5 participants; (3) the *group activity*, during which three groups of three students each and a teacher connected to the platform to collaborate on a project, for a total of 10 participants.

Table 4 shows the issues encountered at each iteration and the implemented solutions. By the end of the third iteration, the platform was fully functional, allowing for more subjective feedback beyond technical or usability aspects. Data from all iterations were summarized in a graph showed in Figure 7, which depicts the changes in the mean—on the left—and in the standard deviation—on the right. In al phases, it can be observed an increase in average responses and a decrease in standard deviation, indicating more consistent positive feedback. This trend was also reflected in the declining number of issues reported, showed in Figure 8. showcasing the effectiveness of immediate solutions applied after each iteration under the RITE method.

In conclusion, the RITE method has been a valuable approach to refining the usability of the developed tool. While room for further enhancements remains, the platform's overall usability reached an acceptable level, prompting the conclusion of the evaluative iterations. The collected feedback will continue to inform future developments and refinements.

Table 4: Emerged issues and implemented solutions during the experiment.

ID	Name	Description	Iteration	Solution
I1	Synchronization bug	When a new user accessed the platform, they began controlling an avatar that wasn't theirs due to bad synchronization among network objects	1	We improved the synchronization logic between clients and how the platform handles new accesses
I2	Bad visibility	The visibility of the slides from the farthest seats from the projector was poor, and the names of the other avatars sitting in front covered parts of the presentation	1	We added The option to zoom the camera, allowing users to zoom in or out at any time. We also added the option to temporarily hide other users' names from above their avatars' to avoid visibility issues
I3	Animation lag	The animation synchronization was slow and laggy, causing many avatars to appear standing even though they were actually seated	1	We improved the animation trigger mechanism, preventing avatars from standing when they should be sitting
I4	Bad tablet interaction	The interaction with the tablet for taking notes was found to be not very usable and useful	1	We removed the note-taking functionality on the tablet available to each user
I5	Lack of interaction	Users would have liked to have other non-verbal interactions, such as applause that they could use at the end of a presentation	1	We added the ability to applaud and wave with the hand through one's avatar
I6	Low voice volume	Users found it difficult to hear each other using voice chat unless they were very close to each other	2	We increased the overall volume of the voice chat has been increased; additionally, we slightly reduced the spatial audio blending
I7	No chat history	The history of the text chat was lost, and there was no way to read previous messages beyond the 10 displayed on the screen	2	We implemented a text chat log feature, saving a history of all messages on the user's device, which can be accessed even after the session ends
I8	Limited choice in avatar customization	Users requested a greater variety of cosmetics during the avatar customization phase	3	*We have planned to address this feedback in future developments*

Summary of the RITE Evaluation (RQ_2)

The RITE evaluation successfully contributed to improving the platform's usability. At the end of the third iteration, all three metrics considered—i.e., efficiency, learnability, and satisfaction—showed significant improvement, and the number of issues encountered had decreased considerably from the beginning of the experiment.

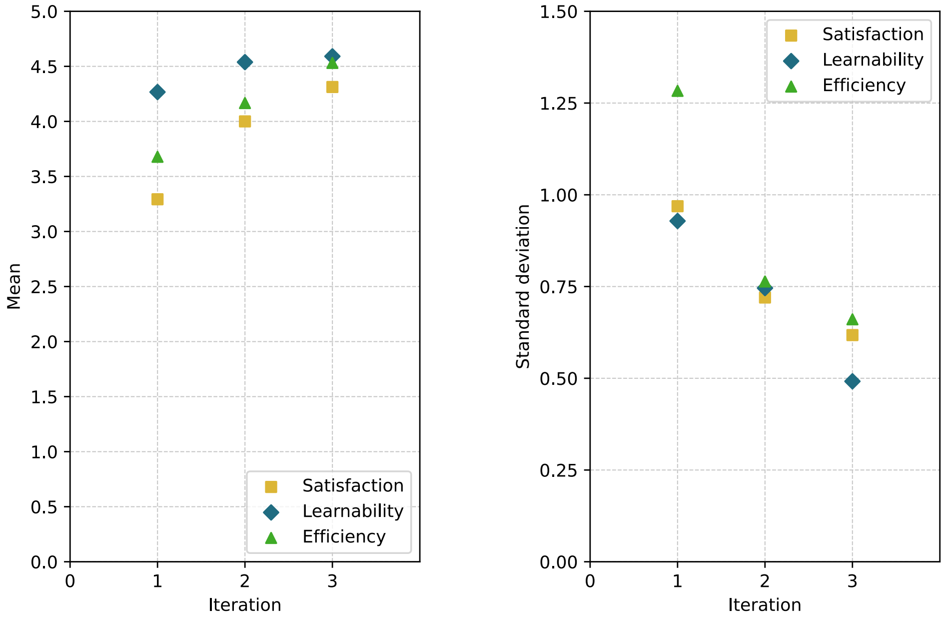

Figure 7: Mean and SD of survey's answers across iterations.

7. Limitations

The work has some limitations, which can be identified both in the requirements-gathering phase and due to the prototype nature of the developed tool and the constrained context in which it was developed and tested. In the following, we identify the main limitations and explain how they were addressed to improve the research.

Literature Selection and Analysis. We performed a *literature survey* to identify educational metaverse requirements from previous research. However, this approach shows the critical challenge of determining the complete set of primary studies. It is worth remarking that the research process reported in 4.1.1 has been applied following the guidelines provided by Kitchenham et al. [33], who defined a set of steps to collect the most significant amount of resources and to elaborate the results.

An additional limitation could arise from literature analysis, as the identified requirement categories may need to be more accurate. We tried to mitigate this risk by not confining everything to a single analysis session, but organizing multiple meetings and discussions. Multiple authors of this work focused on different aspects of the requirements to ensure each category received proper attention and to mitigate the risk of incorrect requirement extraction. In other words, applying these steps makes us confident of the process used. Our future research agenda includes a more thorough systematic search, including additional search engines and applying more rigorous selection criteria to the primary studies.

Semi-Structured Interview. Requirements validation and augmentation were obtained by

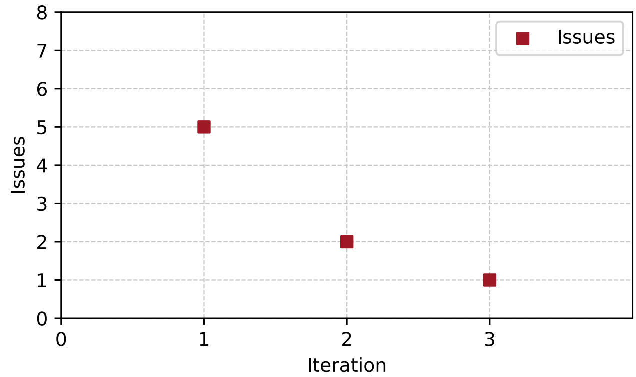

Figure 8: Number of found issues across iterations.

performing *semi-structured interviews* with 7 participants selected through a *convenience sampling*. Despite this, they have the characteristics described in 4.2.2—i.e., they have a certain level of proficiency in the educational setting, including high expertise in different domains close to software engineering, human-computer interaction, and computer science education—the results could be limited to the participants own perspective and opinion. Our future research agenda includes a focus on generalizing the obtained findings and assessing the transferability of the results to a broader population of both learners and educators.

Moreover, we faced a significant limitation concerning the design of the interview. To ensure the accuracy and comprehensiveness of the information gathered, we adhered to the recommendations provided by Hove and Anda [42] while formulating our questions. Furthermore, we conducted a pilot study involving four learners who identified certain biases and imperfections we addressed before performing the reported interviews.

Set of Requirements. Our set of requirements shows some limitations. Indeed, while our requirements aim to provide a broad set applicable to various platforms and devices, they do not specifically cover all the aspects of the metaverse, specifically the ones concerning the more immersive technologies—e.g., VR headsets and haptic feedback devices, widely considered in metaverse applications. The requirements resulting from our literature survey and interviews cover more generic and traditional aspects of metaverse platforms. While it ensures greater accessibility and applicability across standard technology platforms, this choice results in the underrepresentation of immersive technologies' unique capabilities and experiences, which could impact user engagement and realism in the metaverse. Furthermore, the focus of the extracted requirements is more concentrated on a remote education modalities, not currently allowing their effective application in face-to-face contexts. In our future work, we will seek to address these gaps, extending our framework to incorporate more requirements that allow the full potential of immersive technologies to be exploited and different teaching contexts.

Tool Development. Concerning the developed tool, it still exists as a small-scale prototype, which can be further developed to become a more comprehensive platform from both an educational and academic perspective, as well as a true metaverse. Furthermore, to keep the project completely open source, the server technologies used have limited capacity, preventing the loading and synchronization of the application once it is built. Currently, it is only possible to load and display multimedia content before building the executable. Furthermore, the tool currently lacks the feature for taking notes via a tablet. This feature was initially included but was later removed during the tool evaluation because it was deemed not very user-friendly in its implementation. As a result, participants in educational activities had to compromise their immersive experience and resort to external tools for note-taking. As a final concern, the limitation that emerged during the last iteration of the RITE —i.e., to increase the avatar customization options—will be addressed to improve and promote inclusivity within the platform.

8. Conclusions and Future work

Our effort to define a set of requirements for an educational metaverse led to the development of SENEM, our open-source application for supporting academics through a well-designed metaverse. After gathering an initial set of requirements from the literature, we validated it using semi-structured interviews; last, we developed an initial prototype of the application, evaluating and improving it using Rapid Iterative Usability Testing. Our final results convinced the participants to the evaluation and opened the way to new potential contributions—founded on our—to the field of remote and virtual teaching.

Our work can be largely expanded and further developed in future endeavors. First and foremost, the set of requirements can be further expanded and refined, taking into account new, more specific factors that allow SENEM to be used in broader contexts. A practical example of this is the possibility of expanding literature research to find ad hoc requirements for immersive technologies—which have the potential to increase the immersivity and engagement of the user experience significantly—that guide the integration of such technologies into the design of an educational metaverse. Moreover, the scope of the requirements currently focused more on remote teaching activities, can be broadened to consider face-to-face learning modalities to fully exploit the potential of the metaverse in this context as well.

As for further validation of SENEM and its usability assessment, we used well-known metrics that evaluated the tool in terms of learnability, efficiency, and satisfaction. Since the scope of the work was the evaluation of the tool itself resulting from the extracted functional requirements, considerations on the learning outcome of using the tool can be integrated in the future. Indeed, in addition to conducting further evaluations of SENEM using other well-established metrics to strengthen our results, such as the System Usability Scale (SUS) [44], it will be crucial to evaluate the platform from an educational standpoint. Future developments may involve participants in specific educational tasks to compare the use of SENEM with other teaching modalities—e.g., solely face-to-face or through 2D remote communication applications like Teams[6]—to measure and make important considerations about its learning outcome. Moreover, various immersive frameworks can be implemented and tested on SENEM to evaluate their learning outcomes and expand knowledge on the applicability and usefulness of such frameworks [23].

Lastly, the tool itself can be further developed to become a more comprehensive platform both from an educational and academic perspective. Among the initial limitations to address will be the addition of the ability to create and upload multimedia materials within the platform itself, which is currently absent, and the projector currently displays only pre-loaded content before running the platform. Furthermore, it will be essential to reconsider and refine the note-taking capability via the tablet, initially present but later removed during the tool evaluation. As a final

[6]Microsoft Teams: https://www.microsoft.com/microsoft-teams/group-chat-software

concern, the limitation that emerged during the last iteration of the RITE, i.e., to increase the avatar customization options, will be addressed to improve and promote inclusivity within the platform. Furthermore, the avatars and the general environment can be significantly improved to make them even more realistic and engaging. Some aspects and elements of the 3D environment can be brought to a higher level of detail and quality to make the experience with SENEM more appealing and credible. Last but not least, future developments will involve integrating SENEM with immersive technologies, such as HMD devices. This type of technology has yet to be integrated due to the more general focus of our requirements; however, it could enhance the immersivity, engagement, and sense of presence within the virtual environment. Such integration will enable SENEM to support new modes of educational and collaborative activities, leveraging VR and AR technologies for both remote and face-to-face modalities.

Acknowledgments

The authors would like to thank the Editor and anonymous Reviewers for their insightful comments on our paper. The authors would like to thank the participants who contributed to SENEM evaluation and requirements assessment.

References

[1] G. D. Ritterbusch, M. R. Teichmann, Defining the metaverse: A systematic literature review, IEEE Access (2023).
[2] S. Pokhrel, R. Chhetri, A literature review on impact of covid-19 pandemic on teaching and learning, Higher education for the future 8 (2021) 133–141.
[3] U. Bilotti, D. Di Dario, F. Palomba, C. Gravino, M. Sibilio, Machine learning for educational metaverse: How far are we?, in: 2023 IEEE International Conference on Consumer Electronics (ICCE), IEEE, 2023, pp. 01–02.
[4] J. BenedettDörr, L. BeatrysRuizAylon, A survey on the metaverse aspects and opportunities in education, in: 2023 International Conference on Intelligent Metaverse Technologies & Applications (iMETA), IEEE, 2023, pp. 1–8.
[5] R. Roy, M. D. Babakerkhell, S. Mukherjee, D. Pal, S. Funilkul, Development of a framework for metaverse in education: A systematic literature review approach, IEEE Access (2023).
[6] L.-H. Lee, T. Braud, P. Zhou, L. Wang, D. Xu, Z. Lin, A. Kumar, C. Bermejo, 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).
[7] J. O'Donoghue, G. Singh, L. Dorward, Virtual education in universities: a technological imperative, British Journal of Educational Technology 32 (2001) 511–523.
[8] M. Klement, Models of integration of virtualization in education: Virtualization technology and possibilities of its use in education, Computers & Education 105 (2017) 31–43.
[9] N. Stephenson, Snow Crash, Bantam Books, New York, NY, USA, 1992.
[10] D. et al., 3d virtual worlds and the metaverse: Current status and future possibilities, ACM Computing Surveys (CSUR) 45 (2013) 1–38.
[11] G. Basilaia, D. Kvavadze, Transition to online education in schools during a sars-cov-2 coronavirus (covid-19) pandemic in georgia., Pedagogical Research 5 (2020).
[12] Z. et al., The metaverse in education: Definition, framework, features, potential applications, challenges, and future research topics, Frontiers in Psychology 13 (2022) 6063.
[13] F. A. Fernandes, C. M. Werner, A scoping review of the metaverse for software engineering education: Overview, challenges, and opportunities, PRESENCE: Virtual and Augmented Reality 31 (2022) 107–146.
[14] G. A. M. López, H. R. Chaux, F. A. C. Alvarez, The university in the metaverse. proposal of application scenarios and roadmap model, in: 2022 Congreso de Tecnología, Aprendizaje y Enseñanza de la Electrónica (XV Technologies Applied to Electronics Teaching Conference), 2022, pp. 1–9. doi:10.1109/TAEE54169.2022.9840630.
[15] S.-M. Park, Y.-G. Kim, A metaverse: Taxonomy, components, applications, and open challenges, IEEE access 10 (2022) 4209–4251.
[16] A. De Lucia, R. Francese, I. Passero, G. Tortora, Development and evaluation of a virtual campus on second life: The case of seconddmi, Computers & Education 52 (2009) 220–233.
[17] S. E. August, M. L. Hammers, D. B. Murphy, A. Neyer, P. Gueye, R. Q. Thames, Virtual engineering sciences learning lab: Giving stem education a second life, IEEE Transactions on Learning Technologies 9 (2015) 18–30.

[18] H. Aydogan, F. Aras, Design, simulation and virtual implementation of a novel fundamental programmable logic controllers laboratory in a 3d virtual world, The International Journal of Electrical Engineering & Education 59 (2022) 266–281.
[19] H. Duan, J. Li, S. Fan, Z. Lin, X. Wu, W. Cai, Metaverse for social good: A university campus prototype, in: Proceedings of the 29th ACM international conference on multimedia, 2021, pp. 153–161.
[20] A. Jovanović, A. Milosavljević, Vortex metaverse platform for gamified collaborative learning, Electronics 11 (2022) 317.
[21] T. Shen, S.-S. Huang, D. Li, Z. Lu, F.-Y. Wang, H. Huang, Virtualclassroom: A lecturer-centered consumer-grade immersive teaching system in cyber–physical–social space, IEEE Transactions on Systems, Man, and Cybernetics: Systems (2022).
[22] Z. P. Sin, Y. Jia, A. C. Wu, I. D. Zhao, R. C. Li, P. H. Ng, X. Huang, G. Baciu, J. Cao, Q. Li, Towards an edu-metaverse of knowledge: Immersive exploration of university courses, IEEE Transactions on Learning Technologies (2023).
[23] F. A. Fernandes, C. S. C. Rodrigues, E. N. Teixeira, C. M. L. Werner, Immersive learning frameworks: A systematic literature review, IEEE Transactions on Learning Technologies 16 (2023) 736–747. doi:10.1109/TLT.2023.3242553.
[24] F. Fernandes, C. Werner, Software product line for metaverse: Preliminary results, in: 2022 IEEE Smartworld, Ubiquitous Intelligence & Computing, Scalable Computing & Communications, Digital Twin, Privacy Computing, Metaverse, Autonomous & Trusted Vehicles (SmartWorld/UIC/ScalCom/DigitalTwin/PriComp/Meta), IEEE, 2022, pp. 2413–2420.
[25] B. Nuseibeh, S. Easterbrook, Requirements engineering: A roadmap, in: Proceedings of the Conference on The Future of Software Engineering, ICSE '00, Association for Computing Machinery, New York, NY, USA, 2000, p. 35–46. URL: https://doi.org/10.1145/336512.336523. doi:10.1145/336512.336523.
[26] M. C. Medlock, D. Wixon, M. McGee, D. Welsh, The rapid iterative test and evaluation method: Better products in less time, in: Cost-justifying usability, Elsevier, 2005, pp. 489–517.
[27] M. C. Medlock, The rapid iterative test and evaluation method (rite), Games User Research (2018) 203–215.
[28] V. Pentangelo, D. Di Dario, S. Lambiase, F. Ferrucci, C. Gravino, F. Palomba, Senem: A novel software engineering-enabled educational metaverse, 2023. URL: https://drive.google.com/drive/folders/1w7KoOBapiC6ZYQQu8F7TRaJGk39NgYT5?usp=share_link.
[29] C. Jimenez, P. Lozada, P. Rosas, Usability heuristics: A systematic review, in: 2016 IEEE 11th Colombian Computing Conference (CCC), IEEE, 2016, pp. 1–8.
[30] J. Nielsen, The usability engineering life cycle, Computer 25 (1992) 12–22.
[31] L. L. Constantine, L. A. Lockwood, Software for use: a practical guide to the models and methods of usage-centered design, Pearson Education, 1999.
[32] J. Lu, M. Schmidt, M. Lee, R. Huang, Usability research in educational technology: A state-of-the-art systematic review, Educational technology research and development 70 (2022) 1951–1992.
[33] S. Keele, et al., Guidelines for performing systematic literature reviews in software engineering, 2007.
[34] S. Bali, M. Liu, Students' perceptions toward online learning and face-to-face learning courses, in: Journal of Physics: conference series, volume 1108, IOP Publishing, 2018, p. 012094.
[35] Q. Deng, A research on online education behavior and strategy in university, Frontiers in Psychology 13 (2022) 767925.
[36] S. Cavanagh, Content analysis: concepts, methods and applications., Nurse researcher 4 (1997) 5–16.
[37] T. Shen, S.-S. Huang, D. Li, Z. Lu, F.-Y. Wang, H. Huang, Virtualclassroom: A lecturer-centered consumer-grade immersive teaching system in cyber–physical–social space, IEEE Transactions on Systems, Man, and Cybernetics: Systems (2022).
[38] P. H. Ng, J. F. Kar, A. Wong, W. R. Leung, K. Lai, G. Ngai, Implement virtual reality tour in blended learning, in: International Conference on Blended Learning, Springer, 2022, pp. 74–84.
[39] Y. Song, J. Cao, K. Wu, P. L. H. Yu, J. C.-K. Lee, Developing 'learningverse'-a 3d metaverse platform to support teaching, social and cognitive presences, IEEE Transactions on Learning Technologies (2023).
[40] F. M. Schaf, S. Paladini, C. E. Pereira, 3d autosyslab prototype, in: Proceedings of the 2012 IEEE Global Engineering Education Conference (EDUCON), IEEE, 2012, pp. 1–9.
[41] S. Mystakidis, Metaverse, Encyclopedia 2 (2022) 486–497.
[42] S. E. Hove, B. Anda, Experiences from conducting semi-structured interviews in empirical software engineering research, in: 11th IEEE International Software Metrics Symposium (METRICS'05), IEEE, 2005, pp. 10–pp.
[43] J. F. Hair, A. H. Money, P. Samouel, M. Page, Research methods for business, Education+ Training 49 (2007) 336–337.
[44] J. Brooke, et al., Sus-a quick and dirty usability scale, Usability evaluation in industry 189 (1996) 4–7.
[45] T. S. Flanigan, E. McFarlane, S. Cook, Conducting survey research among physicians and other medical professionals: a review of current literature, in: Proceedings of the Survey Research Methods Section, American

Statistical Association, volume 1, 2008, pp. 4136–47.
[46] S. Baltes, P. Ralph, Sampling in software engineering research: A critical review and guidelines, Empirical Software Engineering 27 (2022) 94.