MILSDeM: Guiding immersive learning system development and taxonomy evaluation

Khaleel Asyraaf Mat Sanusi[1,2] · Daniel Majonica[1,2] · Deniz Iren[2] ·
Nardie Fanchamps[2] · Roland Klemke[1,2]

Abstract

Developing immersive learning systems is challenging due to their multidisciplinary nature, involving game design, pedagogical modelling, computer science, and the application domain. The diversity of technologies, practices, and interventions makes it hard to explore solutions systematically. A new methodology called Multimodal Immersive Learning Systems Design Methodology (MILSDeM) is introduced to address these challenges. It includes a unified taxonomy, key performance indicators, and an iterative development process to foster innovation and creativity while enabling reusability and organisational learning. This article further reports on applying design-based research to design and develop MILSDeM. It also discusses the application of MILSDeM through its implementation in a real-life project conducted by the research team, which included four initiatives and eight prototypes. Moreover, the article introduces a unified taxonomy and reports on the qualitative analysis conducted to assess its components by experts from different domains.

✉ Khaleel Asyraaf Mat Sanusi
   ks@colognegamelab.de

   Daniel Majonica
   dm@colognegamelab.de

   Deniz Iren
   deniz.iren@ou.nl

   Nardie Fanchamps
   nardie.fanchamps@ou.nl

   Roland Klemke
   roland.klemke@ou.nl

[1] Cologne Game Lab, TH Köln, 51063 Köln, Germany

[2] Open University of the Netherlands, 6419 AT Heerlen, The Netherlands

⚫ Springer

1 Introduction

Immersion refers to the state of being fully engaged and absorbed in a task, losing track of time and external reality (Jennett et al., 2008). It has been widely used as a learning approach, known as immersive learning (IL), where learners feel fully engaged in the learning activity. IL leverages digital technologies, such as Augmented Reality (AR), Virtual Reality (VR), and Mixed Reality (MR), to simulate real-life learning scenarios and facilitate skill acquisition (Ibáñez and Delgado-Kloos, 2018; Menin et al., 2018; Motejlek and Alpay, 2021). Research suggests that immersion can increase engagement and enhance the learning experience, further improving the learning outcome (Georgiou and Kyza, 2018). Such systems have been applied to train individuals in various domains (Limbu et al., 2019).

IL Systems (ILS) come in various forms, depending on the technology, learning objectives, and purpose (Ibáñez et al., 2011; Menin et al., 2018). ILS allow the integration of sensor technologies, supporting different communication channels called modalities, which provide interactive learning for learners (Limbu et al., 2018). When concepts are presented in multiple modalities (i.e., multimodality), the communication conveys more detailed and comprehensive information (Di Mitri et al., 2018). Immersive technologies can reproduce multimodal sensory information, including visual, auditory, and haptic inputs (Martin et al., 2022), and can recognize user inputs from different modalities (Cohen et al., 1999). This approach creates a sense of heightened presence in the learning environment (Martin et al., 2022). Additionally, ILS incorporate game elements into learning tasks to increase engagement, fun, and excitement (Plass et al., 2015).

Developing ILS relies on a multidisciplinary approach, involving software components, game elements, and pedagogical models (Ibrahim and Jaafar, 2009). ILS development relies on diverse skills and expertise from various fields (Luo et al., 2021). Due to the novelty of IL technologies, and the availability of a plethora of game design practices and learning interventions, it is challenging to make appropriate design decisions and achieve desired outcomes. To increase the likelihood of project success, organisations commonly use rapid prototyping to explore multiple design options.

The multimodality and multidisciplinarity of IL, coupled with the vast selection of design possibilities, make the development of ILS particularly challenging. Despite the existence of various software development methodologies (Sommerville, 2011), there is a lack of systematic methods that offer creative freedom while providing guidance and constraints to achieve project goals and allow for optimal exploration of solution possibilities. Thus, we lay out our research questions as follows.

RQ1 What are essential components of ILS that make it specifically challenging to design and develop them?

RQ2 How can we define a process for ILS development that balances creative freedom with systematic supervision?

RQ3 How can we systematically guide ILS developers to comprehensively explore design options while ensuring the creation of reusable (learning) components from prototyping outcomes?

RQ4 Which components emerge as prominent, and what additional components can be integrated to further enhance the taxonomy, according to the experts?

The contribution of this paper is a novel methodology (MILSDeM) that addresses the challenges of ILS development and offers systematic guidance to the ILS community. MILSDeM builds upon the best practices of management and software development methodologies such as agile management and rapid prototyping. MILSDeM is designed to guide the development of ILS systematically and overcome its unique challenges. It comprises a unified taxonomy of ILS, an iterative development process, and key performance indicators (KPIs) to measure production outcomes, lessons learned, and exploration comprehensiveness. MILSDeM prescribes a development approach that involves initiating multiple rapid prototyping efforts in parallel, guiding the execution of prototyping endeavours, and consolidating the learning experiences towards the ultimate project goal. Furthermore, we report on findings of how experts from multiple domains evaluate the unified taxonomy of ILS and their respective components by conducting a qualitative analysis.

This paper is organised into several sections. The literature review and related work on the specific challenges of ILS development and existing management and software development methodologies are presented in Section 2. The research method is described in Section 3. Details of the proposed MILSDeM are presented in Section 4. An example of the instantiation of MILSDeM is provided in Section 5. The experts' evaluation of the proposed taxonomy is reported in Section 6. A discussion on how MILSDeM addresses the research questions, contributions, limitations, future work, and the conclusion is presented in Section 7.

2 Literature review and related work

Multidisciplinarity of ILS: ILS involve software development, game design, and pedagogical modelling, which can be used in different application domains (Fig. 1). This multidisciplinary nature makes ILS development challenging, and the usage of these components is influenced by the application domain.

Software development practices have a predominantly computer science, engineering, and programming focus (Devedžić et al., 2010). Game design encompasses many aspects such as mechanics, aesthetics, story, and technology (Ahmad, 2019). Pedagogical modelling covers a wide variety of practices, such as setting learning objectives and designing interventions based on learning theories (Bloom, 1956). Each pairwise combination of these three components is possible and has many occurrences in the literature and practice. These three components can be combined in many ways, and each combination has been observed in the literature and practice. Beemer et al. (2019) used gamification to promote the physical activity of students during classroom breaks without the use of software technology. Intelligent tutor studies combine software and learning, but may not include game design elements (Diziol et al., 2010). Any video game not designed for learning purposes can be considered as a combination of software development and game design (Bethke, 2003). ILS is a unique combination of software development, game design, and pedagogical modelling (Motejlek and
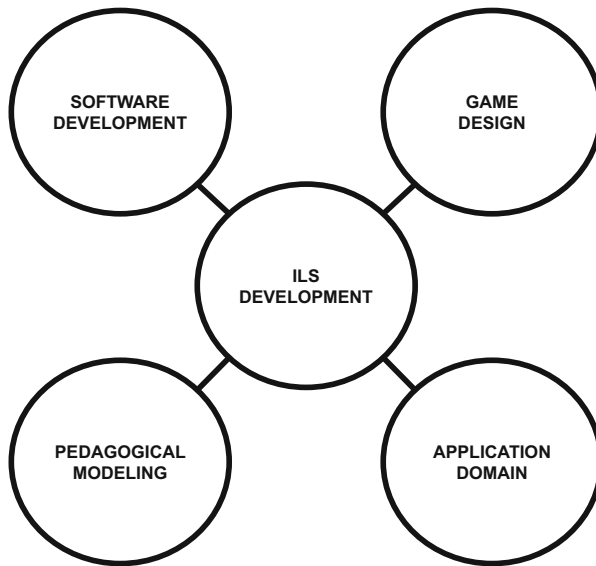
Fig. 1 Main components of ILS development

Alpay, 2021). Immersive systems like AR/VR require advanced software development techniques as well as environmental design and abstract representation of reality. Pedagogical modelling is also necessary to combine ILS components based on learning theories and practices to serve educational purposes. Thus, developing ILS requires multidisciplinary efforts (Kim et al., 2019).

Rubio-Tamayo et al. (2017) conducted a review on IL and identified several key factors that are associated with it, including storytelling, representation of reality, game-play, human-computer interaction, interaction design, and user experience. Depending on the application domain, the interaction may also involve additional factors such as multimodal feedback and sensor data (Mat Sanusi et al., 2021). These aspects correspond to the main components of ILS development (as shown in Fig. 1).

Software development methodologies: Software development methodologies have evolved over time. Early approaches, such as the waterfall method, used rigid sequential steps but were criticized for not offering enough feedback to improve ongoing projects (Royce, 1987). The spiral model was proposed as an improvement, allowing practitioners to manage risks and define goals and constraints with each spiral in the process (Boehm, 1988). Later, agile methodologies, such as Scrum, gained popularity by promoting adaptability, simplicity, and self-reflection (Beck et al., 2001; Schwaber and Beedle, 2002). Scrum uses time-boxed iterations called sprints and includes sprint planning, review, and retrospectives to demonstrate progress, receive feedback, and identify areas for improvement based on recent lessons learned.

In the design of MILSDeM, we build on the fundamental principles and good practices offered by many software development methodologies. Specifically, we adopt an iterative approach to foster continuous learning, early goal-setting, and rapid prototyping to facilitate creative freedom and exploration comprehensiveness.

Open Innovation: Open Innovation is a strategy that combines knowledge from both internal and external sources to enhance an organisation's innovation capabilities (Chesbrough, 2006). It fosters organisational learning by identifying opportunities to tap into external knowledge (Chesbrough, 2006). Examples of open innovation approaches include hackathons, application development contests, and crowdsourcing (Almirall et al., 2014). Studies show that competitions and hackathons are effective pedagogical activities that help participants learn and produce prototypes (Lara and Lockwood, 2016; Suominen et al., 2018; Taylor and Clarke, 2018). Hackathons are also widely used in corporations to foster organisational learning and improve productivity (Brede Moe et al., 2021). Crowdsourcing, a form of open innovation, involves a large number of potentially anonymous individuals contributing to the solution of a problem (Iren and Bilgen, 2014). However, the anonymity of contributors raises concerns about the quality of work and hidden costs. Open competitions, such as Innocentive, are a special type of crowdsourcing that allows organisations to publicly define problems and offer rewards to anyone who sufficiently solves them (Brabham, 2008).

By design, MILSDeM allows the identification and utilisation of opportunity-driven open innovation practices such as hackathons, competitions, and crowdsourcing. The prescribed process advocates the development of reusable product components as well as the dissemination of cross-initiative learning.

IL methodologies: Numerous studies have been conducted to establish a framework for developing and evaluating IL. De Freitas and Oliver (2006) propose a four-dimensional evaluation framework for game- and simulation-based learning that encompasses pedagogic considerations, learner specifications, mode of representation, and context. The framework emphasizes the importance of considering all four dimensions cohesively. In a subsequent study, the authors extended the framework to evaluate IL experiences in a virtual world (de Freitas et al., 2009). In contrast, MILSDeM incorporates evaluation steps in every iteration.

The exploratory learning model is an iterative process that comprises five steps: *experience, exploration, reflection, forming concepts,* and *testing* (de Freitas et al., 2009). It emphasizes group meta-reflection and knowledge sharing in IL environments, aiming to transform the traditional one-way information flow in classrooms into a bi-directional interactive one. Luo et al. (2021) propose a method for IL that helps practitioners identify learning opportunities in existing video games through a systematic approach. By mapping the desired learning behaviour with game components, this study ensures that learning goals are achieved.

The existing IL frameworks generally focus on a particular aspect of IL for specific purposes. MILSDeM addresses the entire development cycle, thus it significantly differs from the existing IL frameworks, thus bridging a potentially impactful gap in the literature.

Taxonomy of ILS: The examples of ILS implementation in the literature show great variance in terms of the underlying interaction technology, game design approaches, pedagogical modelling, and specific purpose. Studies exist that put in an effort to curate the various ILS components and form taxonomies. ILS developers might benefit from a unified taxonomy of ILS when defining the scope of their projects and selecting design trajectories to explore.

Several taxonomies have been proposed in the literature to describe the dimensions and factors related to IL environments. De Freitas and Oliver (2006) define four dimensions that include context, learner, mode of representation, and pedagogical considerations. Rubio-Tamayo et al. (2017) identify narrative, interactivity, representation, gameplay, and technological mechanics as the factors related to the design of immersive environments. In his paper, Ahmad (2019) lists mechanics, aesthetics, story, and technology as the basic elements of educational games. Another taxonomy focuses on specific aspects of IL, such as the dimensions of task and modality in AR interaction techniques (Hertel et al., 2021). Motejlek and Alpay (2021) propose a taxonomy of IL applications that includes production and delivery technology, user and system interaction, gamification, and education purposes. In their review, Menin et al. (2018) classify the literature on IL using the dimensions; display device, level of immersion, interaction, feedback, serious purpose, and target participants. The recent paper of Park and Kim (2022) list the components of a Metaverse world that consists of hardware, multimodal content representation, multimodal and embodied user interaction, and educational purpose. Finally, Bloom proposes Bloom's taxonomy, which covers learning objectives in three domains: cognitive, affective, and psychomotor (Bloom, 1956).

The study of taxonomies and categorizations in the literature, which is summarized and grouped in Table 1, paves the way to curating a unified taxonomy of ILS.

Table 1 ILS factors from the literature; gathered and categorised

Reference	Technology	Game elements	Pedagogy	Modality
Bloom (1956)			cognitive, affective, psychomotor	
De Freitas and Oliver (2006)	context		pedagogic considerations	mode of representations
Rubio-Tamayo et al. (2017)	interactivity, technological mechanics	narrative, gameplay		representation
Ahmad (2019)	mechanics, technology	aesthetics, story		
Hertel et al. (2021)			task (partially)	mode of representations
Motejlek and Alpay (2021)	production and delivery technology, user and system interaction	gamification	educational purpose	interaction mode
Menin et al. (2018)	display device, interaction (partially)	level of immersion	feedback, serious purpose	
Park and Kim (2022)	hardware		educational purpose, multimodal content representation	multimodal and embodied user interaction

3 Research method

In this study, we adopted the Design-based Research (DBR) method to address our research questions (Easterday et al., 2014). The process combines design and scientific approaches to create practical products and develop meaningful theories that address educational challenges. More specifically, we followed the steps: *focus, understand, define, conceive, build, and test* (Fig. 2). We iterated over DBR steps, and with each iteration we clarified the research problem more and refined our solution (i.e., MILS-DeM) further to better address the defined problem.

In the **focus** phase, we identified the stakeholders relevant to our study and set the direction of the project. Subsequently, in the **understand** phase, we studied the literature comprehensively to examine the related work that we can build upon and to identify gaps (see Section 2). In the **define** phase, we set the research goals and formulated the research questions. In the **conceive and build** phases, we sketched MILSDeM and refined its components with each iteration in Section 4. These include the MILSDeM concepts, unified taxonomy of ILS, key performance indicators, and development process.

Finally, in the **test** phase, we instantiated MILSDeM in a project that consists of four initiatives and eight prototypes for a duration of nine months (see Section 5). The reflections were then reported. Following that, we evaluated the taxonomy with experts from various domains to expand it (see Section 6). This paper presents a comprehensive description of our study and introduces MILSDeM as the proposed solution.

4 Multimodal immersive learning system development methodology (MILSDeM)

In this section, we introduce MILSDeM which aims at guiding ILS development in a way that facilitates creative freedom while ensuring measurable progress. Specifically, we define the concepts of MILSDeM, introduce a conceptual taxonomy of ILS to serve as a guideline for comprehensive exploration, describe the underlying development process, and propose KPIs for evaluation.
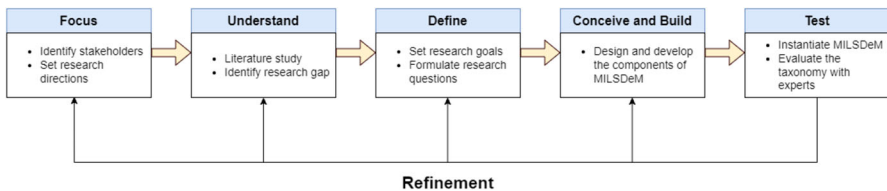

Fig. 2 Research model

4.1 MILSDeM concepts

MILSDeM approaches the ILS development at three levels; *projects, initiatives,* and *prototypes* (Fig. 3) which are defined in the following paragraphs.

A *project* is an endeavour in which resources are allocated to create a product or a service (Westland, 2007). A project has a defined beginning and an end and has resource constraints (e.g., time, funds, human resources) and operational limitations (e.g., laws, regulations, organisational procedures, and culture). MILSDeM can be employed to ensure the optimal utilization of the project in terms of developed components, lessons learned, and explored design options. According to our terminology, a project consists of one or more initiatives.

Initiatives are efforts that arise from opportunities to partially or fully achieve project goals (Schwalbe, 2009). They can take different forms, such as hackathons or competitions, and are always aligned with project goals. Initiatives involve one or more prototypes and teams working together.

A *prototype* is a set of tasks that aim to achieve goals aligned with the project and initiative goals (Houde and Hill, 1997). It is the smallest unit of management in MILSDeM. Prototypes within an initiative share a common theme but may have different goals and constraints. Teams working on prototypes have creative freedom; outcomes are reusable components and lessons learned. Prototyping contributes to exploring design options, guided by project learning goals.

4.2 Unified taxonomy of ILS

In this subsection, a unified ILS taxonomy is introduced, derived from various literature sources (see Section 2) and presented in Table 1. This taxonomy serves as a comprehensive framework for categorizing and summarizing the different components of ILS. It helps define the scope of an ILS project, set project or prototype goals, and guide exploration throughout the initiatives.

MILSDeM recommends the use of this unified ILS taxonomy, which encompasses technology, pedagogy, and modality perspectives, to establish goals and constraints for prototypes and define exploration objectives. The taxonomy allows for exploring various design options. While not exhaustive, developers can customize the taxonomy by adding or removing categories and entries to align with their project requirements. By comparing explored and unexplored dimensions at the end of each initiative, developers can establish exploration goals for subsequent initiatives.
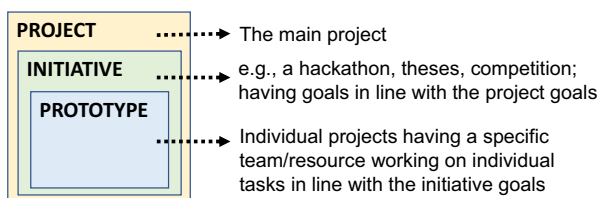

Fig. 3 Hierarchical representation of development levels

Our proposed unified taxonomy ILS consists of dimensions (black-coloured boxes), categories (grey-coloured boxes), and entries (white-coloured boxes) (Fig. 4). The dimensions comprise technology, pedagogy, and modality. We classify the technology dimension into two subcategories; experience technologies enable users to experience immersion and interaction technologies comprise the capabilities to recognize different modalities of user inputs (Hertel et al., 2021; Motejlek and Alpay, 2021). The pedagogy dimension is divided into two categories; types of learning (Bloom, 1956) and intervention techniques that describe when certain informative actions are taken by ILS. Such techniques include before (instructions), during (real-time, immediate, delayed feedback), and after (post-session feedback). The immediacy of "Feedback: During" can be provided in three methods: real-time, immediate, and delayed. Real-time feedback offers continuous information that aligns with ongoing task execution, providing insights and allowing for instant revisions (Geisen and Klatt, 2022). Immediate feedback furnishes learners with information about their performance after a task is completed, allowing them to evaluate their actions and rectify errors promptly (Patil et al., 2015). Apart from receiving instant information, feedback during sessions can also be delivered as delayed feedback. In this form, cues or guidance are not provided during task execution but are instead given after a predefined number of task repetitions or with some delay afterwards (Metcalfe et al., 2009). Finally, the modality dimension is categorised into different sensory channels and interaction modes that can be used by ILS (Menin et al., 2018).

The exploration goals and the progress can be displayed using data visualisation techniques that are appropriate to represent a hierarchical, part-of-whole relationship. To this end, we use a sunburst diagram/chart because the radials and slices can represent the exploration goals in their hierarchical structure. In the chart, we designated the radials as follows: the inner circle represents the dimensions of ILS, the middle circle signifies categories within these dimensions, and the outer circle represents individual
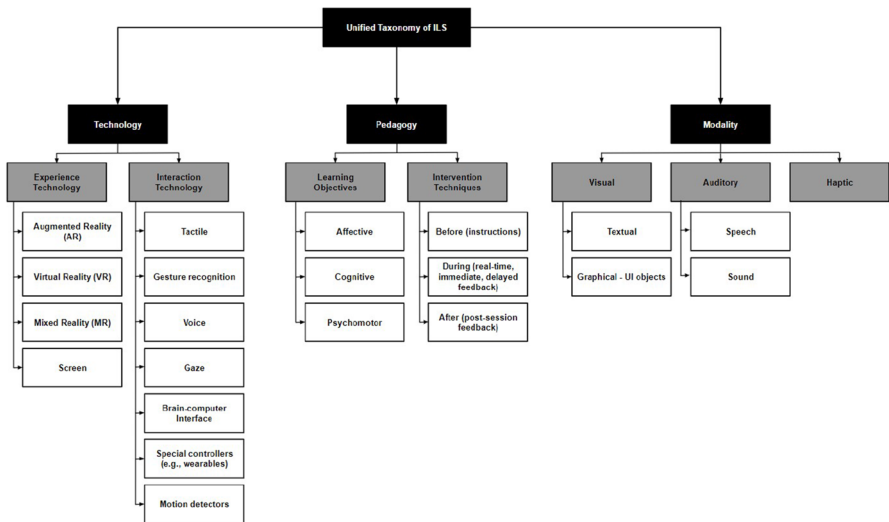

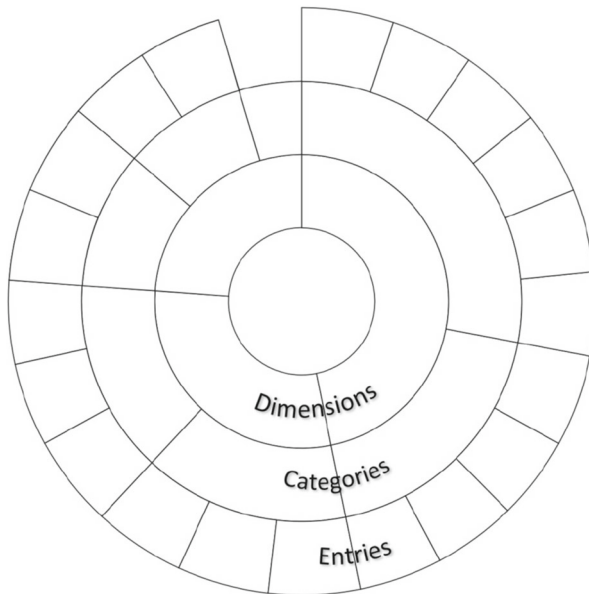**Fig. 4** Overview of the unified ILS taxonomy

Fig. 5 The sunburst chart illustrates the exploration goals, comprising the dimensions of ILS (inner circle), categories (middle circle), and entries (outer circle)

entries (see Fig. 5). In addition to the radials, we applied three distinct colours (blue, orange, green) to the three dimensions (technology, pedagogy, modality), representing their corresponding categories and entries, as shown in Fig. 6.

Figure 7 (left) depicts an example of a comprehensive exploration chart that includes all elements of the unified ILS taxonomy. However, most ILS projects rarely aim at exploring such wide coverage. Figure 7 (right) shows an example with a limited scope.

4.3 Key performance indicators (KPIs)

MILSDeM prescribes several KPIs to monitor and control the project's progress (Kerzner, 2017). Projects may need and benefit from many KPIs other than the ones
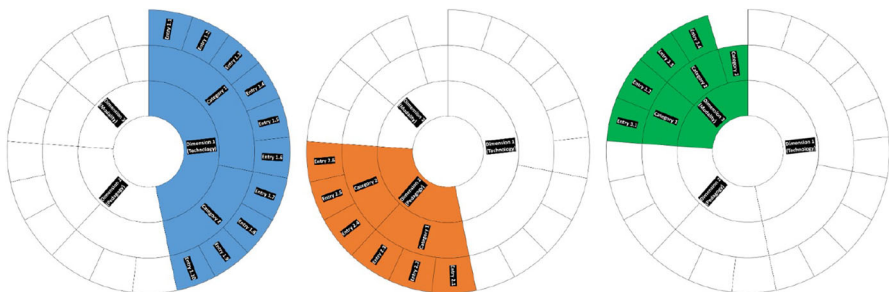

Fig. 6 The dimensions of ILS - technology (blue), pedagogy (orange), and modality (green) - alongside their corresponding categories and entries

Fig. 7 *Example chart 1 (left):* The components aligned with the unified ILS taxonomy. The lack of coloured background indicates that none of the areas have been explored yet. *Example chart 2 (right)*: Areas that have been explored (coloured) and areas that still need to be explored (no colour)

mentioned here. Thus, project teams must decide which KPIs are relevant to their specific requirements. The KPIs that are prescribed by MILSDeM aim at facilitating reusable component development, comprehensive exploration, and organisational learning (Chouseinoglou et al., 2013). Due to the changes in the nature of work in which traditional measurement methods are insufficient to enable continuous organisational learning, current measurement should be dynamic, collective, localized, strategy-based, and future-oriented (Pöyhönen and Hong, 2006).

Project-specific KPIs: At the project's outset, project-specific KPIs are established and ideally monitored consistently throughout the project (Kerzner, 2017). MILSDeM recommends defining, monitoring, and controlling project-specific KPIs, but does not prescribe a specific approach for these practices.

Process KPIs: MILSDeM recommends tracking three important process KPIs and how they are utilised in the context of our paper:

1. *Reusable components* (Gill, 2003):
   - System components that can be reused in the next initiative's prototypes. The more reusable components used, the better.
   - Reducing the time needed to develop new features for the next initiative's prototypes instead of reinventing the wheel. Teams can leverage existing components, accelerating the development process.
   - Minimizing duplication of resources, thereby saving costs associated with development.

2. *Lessons learned* (Wiewiora and Murphy, 2015; Wang-Trexler et al., 2021):
   - A common metric in project management and software development. Teams gain insights and knowledge during prototyping that can benefit future prototypes.

- A foundation for continuous development, enabling teams to refine processes, strategies, and approaches for the next iteration of their prototypes/projects.
- Sharing valuable tacit knowledge across the team will foster a learning culture, preventing the repetition of mistakes and encouraging the adoption of successful practices.
- Enabling teams to anticipate challenges and implement preventive measures to mitigate risks early in new prototype development.

3. *Exploration comprehensiveness* (Kerzner, 2017):

- A pivotal factor in deciding whether to finalize the project or continue the exploration by beginning a new initiative.
- Guiding teams in making informed decisions about resource allocation and project continuation based on the achieved exploration milestones.
- It encourages teams to explore new ideas and technologies, fostering a culture of adaptability and innovation within the team or organisation.

4.4 Development process

MILSDeM includes a development process that defines the activities of a project (Fig. 8). These steps consist of goal setting, expanding, execution, evaluation, supervision, collapsing, and closing. Table 5 shows an overview of the MILSDeM steps (see Appendix A).

Goal Setting: In this step, the project goals, scope, rules, and procedures are defined, including exploration and learning objectives, as well as content-specific objectives such as requirement specifications. The unified taxonomy of ILS is recommended to define the scope of exploration, and developers may tailor it to explore relevant design options. MILSDeM does not prescribe how to manage content-specific objectives but guides developers to define exploration possibilities comprehensively. The organisation also identifies opportunities to create initiatives in this step.

Expanding: In this step, an initiative is defined and divided into prototypes, and the theme of the initiative is established. The project goals are refined and translated into the goals of each initiative and prototype. Necessary resources are allocated, and prototype development commences.
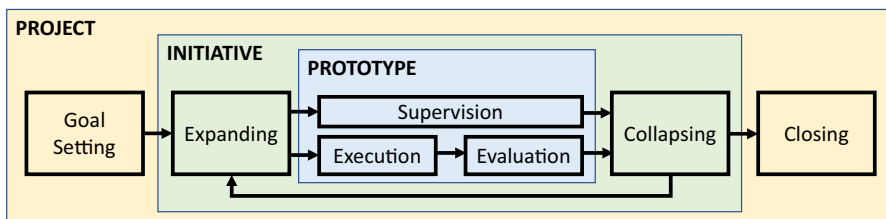

Fig. 8 The iterative development process

Execution: During this step, prototype development teams work on their assigned tasks with creative freedom. They are encouraged to develop prototype components that are reusable and share their lessons learned throughout the development process.

Evaluation: In the evaluation step, prototypes are assessed using project-specific and process-related KPIs. These include project goals, reusable component development, lessons learned, and exploration comprehensiveness. The overall initiative is also evaluated based on the outcomes of the prototypes created within it.

Supervision: During an initiative, teams developing prototypes are supervised to ensure adherence to the plan. This includes monitoring and controlling development, managing risks, sharing insights, and ensuring prototypes are created within operational and resource limitations. Corrective actions are implemented, if needed, to maintain progress.

Collapsing: In this step, the prototypes' outcomes are collected, and the initiative is wrapped up. Reusable components and lessons learned are shared with all teams and management. If the initiative has met the project goals, the project is closed. If further exploration is warranted, the process begins again with a new initiative.

Closing: At the end, when all project goals are achieved, the project is formally closed.

5 Instantiation of MILSDeM

In this section, we will show how MILSDeM was used in a real-life project setting to develop ILS components for multiple initiatives. The project had four initiatives that resulted in eight prototypes. We will describe the process using the Input/Activities/Outputs/KPI structure and explain the prototype development activities, lessons learned, reusable components, and exploration comprehensiveness for each phase of the project. Figure 9 shows the flow of the project from the first to the last initiative.
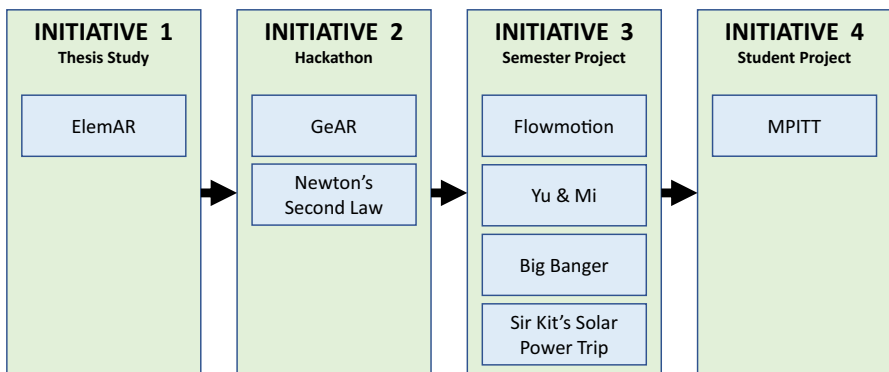
Fig. 9 The initiatives with their respective prototypes

5.1 Goal setting

Input: The instantiation of MILSDeM is driven by two parallel research projects (see Acknowledgments). These projects collaborated to create ILS components for diverse applications. The project was conducted within a higher education and research institution, using the unified taxonomy of ILS as a guiding framework to establish exploration objectives.

Activities: The MILSDeM instantiation was planned jointly, defining the types of initiatives to be used based on the joint effort of the two research projects. The initiatives included the involvement of students of different levels due to the nature of the research projects and their embedding into higher education and research.

Output: A rough project plan specifying the planned initiatives, their dependencies, timing, and development objectives.

KPIs: Initiative project goal match.

5.2 Initiative I

Input: The initial goals and the project plan defined in the goal-setting step required defining a theoretical, technical, and practical concept with an initial prototype. The initial prototype has aimed to set the ground in terms of technical feasibility.

Activities: To gather detailed information on theories, technologies, and design decisions, the project conducted a thesis study. The study focused on the use of AR for chemistry education, considering technical, didactic, and design aspects.

Output: The thesis study presented a theoretical concept, a technical concept, a prototype, and an introductory presentation as a foundation for further development. The prototype, ElemAR (Fig. 14a), explored the technical, didactic, and design-oriented aspects of using AR technology for chemistry education.

Reusable components: The prototype yielded several reusable components, including an AR marker generator for image and 3D shape tracking, scannable cards, a teaching guide worksheet for learners, and interactive UI elements. These components are detailed in Table 6.

KPIs: The reusable components mostly complied with the project goals.

Lessons learned: The utilization of AR technology and scannable physical cards allowed for the visualisation of molecular combinations and their chemical formulas, providing valuable information. This prototype demonstrated the potential of this technology, which can be further explored and applied in various domains.

Exploration comprehensiveness: In Initiative I (Fig. 10), we explored various design choices across the dimensions of the ILS taxonomy. The prototype utilised AR and tactile technologies to enhance cognitive skill training. Instructions were provided before the learning task, and feedback was given during and after, incorporating visual cues in textual and graphical formats. Although the outcomes were promising, a second initiative was established to expand the exploration.
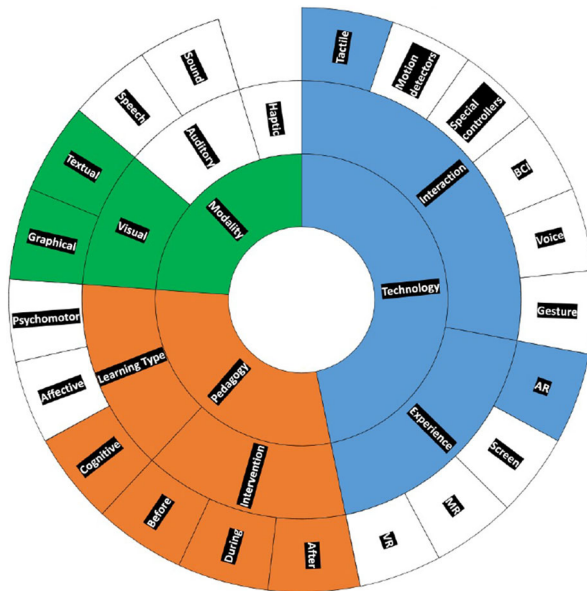
Fig. 10 Exploration comprehensiveness chart for Initiative I

5.3 Initiative II

Input: The initial phase of this initiative benefited from reusable components and lessons learned from Initiative I. The goals and project plans were refined according to the MILSDeM process to guide the development activities.

Activities: During the expanding step, the initiative's objectives and theme were established. To foster creativity and diversity in the development process, an interdisciplinary hackathon was organised. Teams worked on their prototypes during the hackathon to achieve the development objectives.

Output: The hackathon resulted in two prototypes, GeAR (Fig. 14b) and Newton's Second Law (Fig. 14c), which contribute theoretically and technically and serve as a foundation for future development. GeAR focused on engineering, while Newton's Second Law explored physics, both utilizing AR technology to address technical, didactical, and design aspects.

Reusable components: Table 6 presents the development of new reusable components in Newton's Second Law, including a hint system and sound effects for audio feedback. However, no new reusable components were developed in GeAR.

KPIs: These components mostly complied with the project goals.

Lessons learned: Two prototypes were developed for engineering and physics, utilizing the same components from the previous initiative. The GeAR prototype employed AR technology and scannable cards to demonstrate cogwheels and gear mechanisms visually.

Exploration comprehensiveness: In Initiative II, a new design trajectory for feedback modalities, specifically auditory, was explored (Fig. 11). The hackathon yielded
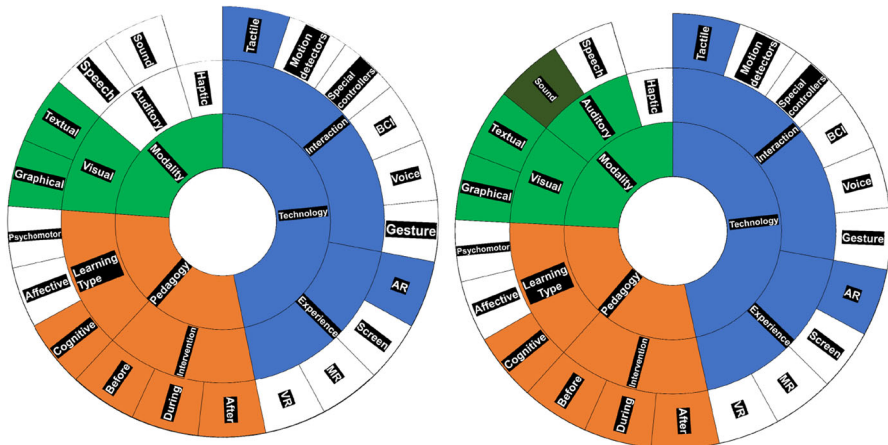
Fig. 11 The comparison charts for Initiatives I (left) and II (right)

promising outcomes, but there were still unexplored areas within the initial exploration goals. As a result, Initiative III was introduced.

5.4 Initiative III

Input: The initial project goals stated that more design trajectories could be explored.

Activities: The teams have worked on their prototypes throughout a semester project towards developing educational game prototypes. The teams consisted of individuals with diverse backgrounds (e.g., artists, designers, programmers.

Output: Four prototypes have been developed for multiple application domains, namely The Big Banger (physics) (Fig. 14d), Sir Kit's Solar Power Trip (electricity) (Fig. 14e), Yu & Mi (human-robot interaction) (Fig. 14f), and Flowmotion (yoga) (Fig. 14g).

Reusable components: Several reusable components were developed (see Table 6) such as the modules for localization and mapping (The Big Banger and Yu & Mi), body tracking (Flowmotion), mini-game (Yu & Mi), puzzle system (The Big Banger and Sir Kit's Solar Power Trip), and score points (The Big Banger and Flowmotion).

KPIs: The developed components mostly complied with the project goals.

Lessons learned: Several new AR prototypes were developed, each focusing on different aspects of learning. Big Banger explored physics with a visual representation of the solar system. Sir Kit's Solar Power Trip used a card game to teach about solar power. Yu & Mi and Flowmotion incorporated various feedback mechanisms. The prototypes showcased the potential for reusing and adapting components across different domains and learning goals.

Exploration comprehensiveness: The four prototypes explored new design trajectories, including screen displays, gesture recognition, and motion detectors in the technology dimension. The psychomotor domain was explored in the type of learning dimension. New modalities such as speech, graphical, and textual were also explored.
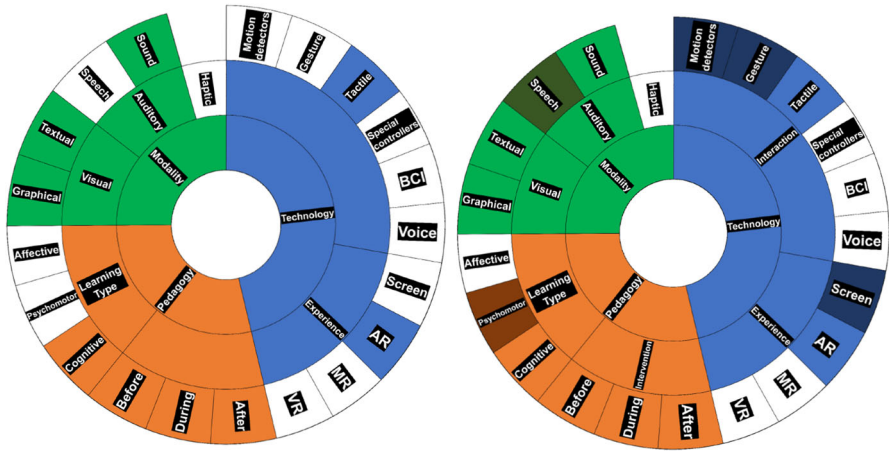
Fig. 12 The comparison charts for Initiatives II (left) and III (right)

Figure 12 visualises the newly explored areas in Initiative III. Despite significant progress, further exploration was needed, leading to the establishment of the expanding phase.

5.5 Initiative IV

Input: Initiative IV has extended the development of the prototypes in the previous initiatives. Recurring participants from the previous initiative took part. Similarly, the project goals were refined and used as objectives of this initiative to explore more design trajectories.

Activities: The developers were tasked to create a prototype that serves as a training toolbox that can be applied in multiple psychomotor domains. Components from the previous prototypes were combined and improved in the execution phase.
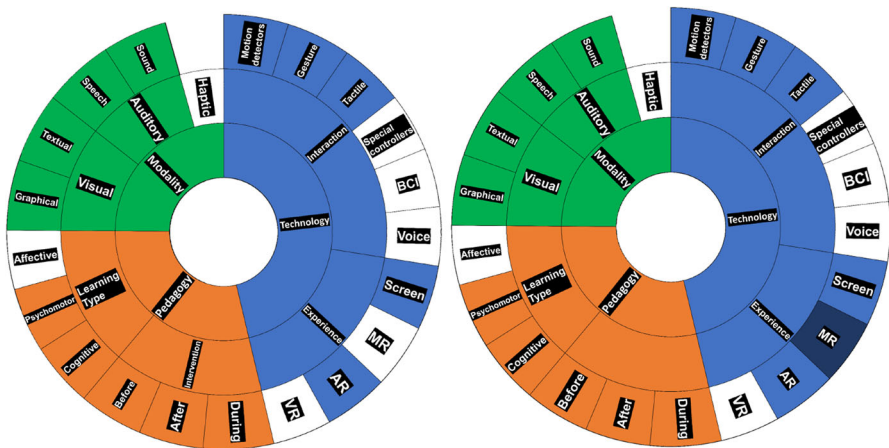

Fig. 13 The comparison charts for Initiatives III (left) and IV (right)

Output: A prototype named MPITT (Multimodal Psychomotor Immersive Training Toolbox) (Fig. 14h) was developed.

Reusable components: Visual instruction and feedback components from the previous initiative were significantly improved. colour vision deficiency UI was developed as a reusable component (see Table 6).

KPIs: The components complied with the project goals.

Lessons learned: We observed that the improved version of the prototype could potentially be applied in multiple psychomotor domains.

Exploration comprehensiveness: In Initiative IV, as a design trajectory, the MR technology was explored. Figure 13 compares the newly explored areas using solid background colours in Initiatives III and IV.

5.6 Closing

Input: In simpler terms, the decision to end the project was based on the fact that the project had achieved its exploration objectives, as demonstrated by the charts and Table 7 summarizing the project's progress.

Activities: This project consisted of four initiatives that resulted in many reusable components and valuable lessons learned. Most of the exploration goals were achieved, and the project objectives were completed, leading to the finalization of the project.

Output: The finalization entailed the documentation of achieved results and the release of the developed components.

KPIs: Project goals achieved.

5.7 Reflections

The MILSDeM framework was instrumental in managing and guiding the development of the project. The use of different levels allowed for efficient exploration of multiple design trajectories, while the unified ILS taxonomy provided a clear and systematic approach to setting exploration goals and guiding the prototyping teams. The iterative development process facilitated learning from previous iterations, and the KPIs recommended by MILSDeM helped to maintain a focus on performance and
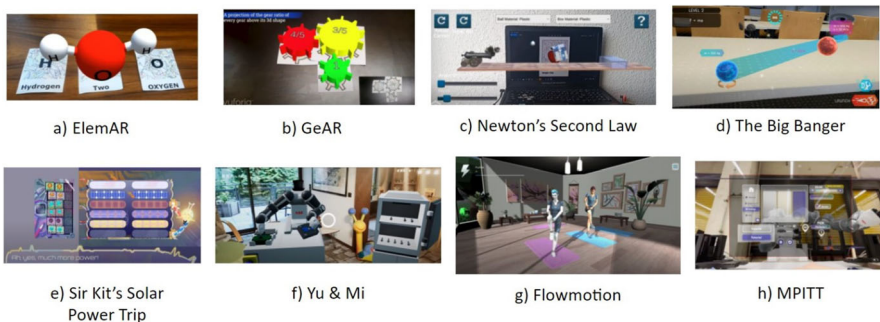

Fig. 14 Screenshots of the ILS prototypes

purposeful exploration of design options. Figure 14 shows the eight ILS prototypes that were developed across four initiatives. Table 6 shows the summary information about each prototype and the reusable components they produced, and Table 7 displays the design options explored in different prototypes based on the unified taxonomy of ILS (see Appendix A).

6 Evaluation of the unified taxonomy of ILS

To evaluate the unified taxonomy of ILS, we conducted multiple focus group studies. Participants were presented with the eight developed ILS prototypes in the form of presentation slides, and their feedback was collected to explore potential extensions of the taxonomy.

6.1 Design

The presentation slides were prepared with the following contents: a definition of ILSs, the introduction of our proposed taxonomy, and video demonstrations of ILS prototypes that were developed throughout the instantiation of MILSDeM. We selected five ILS prototypes with the most components from the taxonomy; 1) Flowmotion, 2) Yu & Mi, 3) The Big Banger, 4) Sir Kit's Solar Power Trip, and 5) MPITT. In doing so, we provide concrete examples to the participants.

The presentation was given in four different sessions; An online training program for teachers and students with different backgrounds (Session 1), an on-site meeting amongst the learning scientists (Session 2), an online meeting amongst the game development experts (Session 3), and an online meeting amongst the AI, social science, and management experts (Session 4). The presentation materials were shared with the participants afterwards.

6.2 Instrument

An online survey was conducted to gather data on participants' perspectives regarding the components of future ILS development. The survey consisted of two sections: demographic information and ideal components for ILS development. Participants were asked to select the necessary ILS components in the taxonomy and provide suggestions for additional components and improvements. The survey remained open for one week to allow participants ample time to complete it. Data analysis involved using the Atlas.ti tool to code and analyze transcripts, create network diagrams, and visualise the data.

6.3 Participants

A total of 42 people participated in the survey. However, five responses had to be excluded due to the lack of detail, or insufficiently following instructions, leaving 37 responses to be analyzed. There were 39% students, 34% teachers, 2% with two roles

(teacher and researcher), and 2% with three roles (teacher, researcher, and manager). The backgrounds consist of; computer science (CS) (43%), educational technology (ET) (16%), game design (GD) (9%), physics (PH) (7%), science and technology (ST) (7%), psychology (PS) (5%), management (MGMT) (5%), art (5%), mathematics (MT) (2%), and language (LG) (2%). Nearly half of the participants (45.2%) had never tried ILS before our study, while the rest either have (35.7%) or are unsure (19%).

In Session 1, 30 participants (CS, PH, MT, GD, CS, LG, A, and ST) answered the survey. Following the session, several participants raised a common issue: the survey design was overwhelming and confusing, particularly the diagram of the unified ILS taxonomy and the images of the prototypes. Hence, we excluded the images but kept the unified taxonomy of ILS as it is an essential component of the survey and to ensure the consistency of the collected data for the next sessions.

In Session 2, five participants from the technology-enhanced learning background answered the survey. The distribution of the survey was altered due to the on-site session. Hence, a QR code was created, linking to the survey. Sessions 3 and 4 were conducted remotely. Five participants (A, CS, and GD) from Session 3 and three participants (PS and MGMT) from Session 4 answered the survey.

6.4 Results of the evaluation

Table 2a & b report the components that were selected by the experts, addressing the technology dimension of the taxonomy. "AR" recorded the highest percentage (34.09%) for the experience technology while "Voice" had the highest percentage (15.91%) for the interaction technology.

Table 2 The dimension of Technology

Experience	Percentage	Interaction	Percentage
(a) Experience Technology.		(b) Interaction Technology.	
AR	34.09%	Voice	15.91%
VR	11.36%	Gesture recognition	9.09%
MR	9.09%	Motion detectors	6.82%
Screen	2.27%	Gaze	4.55%
		BCI, tactile, wearables	2.27%

Table 3 The dimension of Pedagogy

Learning Objective	Percentage	Intervention	Percentage
(a) Learning Objectives.		(b) Intervention Techniques.	
Cognitive	58.62%	Feedback: During	10.34%
Psychomotor	13.79%	Feedback: After	6.90%
Affective	6.90%	Instructions	3.45%

Table 4 The dimension of Modality

Auditory	Percentage	Visual	Percentag
(a) Auditory Modality.		(b) Visual Modality.	
Speech	37.50%	Graphical	31.25%
Sound	25.00%	Textual	6.25%

The selected components from the pedagogy dimension are reported in Table 3a and b. In the learning objectives, the "Cognitive" aspect recorded the highest percentage value (58.62%). For the intervention technique, the highest count recorded was "Feedback: During" with a percentage of 10.34%.

For the modality dimension, the entries are reported in Tables 4a and b. We observed that the "Graphical" (31.25%) and "Speech" (37.50%) were the most preferred visual and auditory modalities, respectively.

6.5 Additional components for the unified taxonomy of ILS

In this section, we report the findings on the additional ILS components (dark green and light green boxes) suggested by the experts and which of the existing components can be improved for refining our taxonomy. Figure 15 shows the overview of the extended version of our proposed taxonomy.

Participants suggested adding audio technology to the experience technology and several sensors such as emotion, temperature, eye-tracking, and odor to the interaction technology (Fig. 15). AI technology was also suggested for personalized learning. Only olfactory was suggested as an ILS component for the modality dimension.
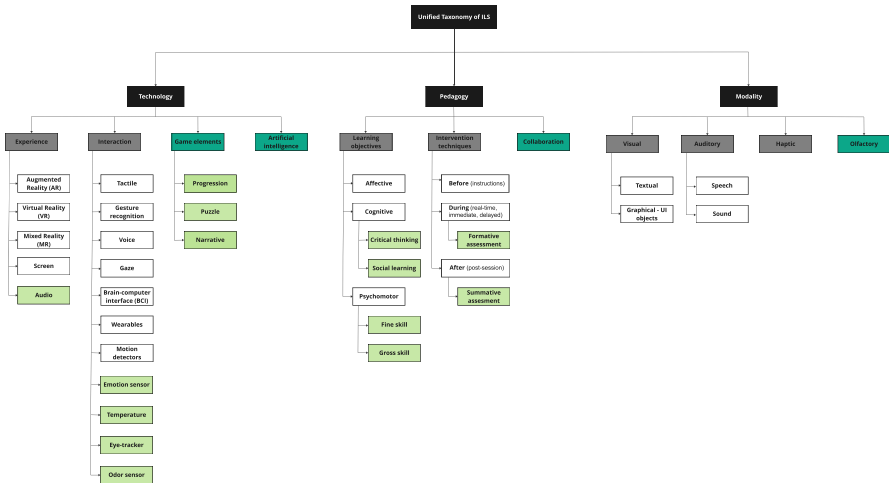

Fig. 15 Overview of the unified taxonomy of ILS; Our proposed subcategories (grey) and entries (white); Experts' suggested subcategories (dark green) and entries (light green)

For the pedagogy dimension, participants suggested learning domains should be divided into smaller sub-categories. Sub-types of skills focusing on specific and generic human muscles for the psychomotor domain were recommended. For the cognitive domain, critical thinking and social learning were suggested as sub-types. Formative evaluation and tactile feedback, such as real-life tangible objects, were proposed to refine the feedback. Participants also suggested a new sub-category for collaboration among learners to recreate the traditional classroom experience.

7 Discussion and conclusion

This paper introduces a methodology called MILSDeM, which is designed to guide developers in creating ILS by overcoming challenges such as multidisciplinarity, balancing creative freedom and systematic development, and abundant design choices. MILSDeM incorporates a unified taxonomy, an iterative development process, and specialized KPIs. The methodology was designed using the DBR method and evaluated in a scenario where 37 individuals developed eight prototypes over a nine-month period. The study identified many IL components for various application domains. The paper concludes with reflections on the study and a discussion of how the research questions were addressed.

RQ1 What are essential components of ILS that make it specifically challenging to design and develop them?

Based on a literature review and our experiences, we identified several components that make ILS development uniquely challenging. Firstly, ILS development is inherently multidisciplinary as it comprises components from multiple disciplines; software development, game design, and pedagogical modelling. Each of these disciplines represents a distinct know-how and skill-set, and different requirements in their work approaches, such as creativity and systematic practices. Secondly, the set of design choices that can be used in ILS development is extensive. One could use many different interaction technologies that operate in multiple modalities and embody various pedagogical interventions to achieve the goals of a project. Our proposed unified taxonomy provides an overview of various components of ILSs, laying out potential design decisions and explorations when defining the scope of an ILS development project. Consequently, we advocate that ILS developers employ this taxonomy as an initial reference and tailor it by incorporating, omitting, or refining components to suit the specific requirements of their projects.

RQ2 How can we define a process for ILS development that balances creative freedom with systematic supervision?

ILS development is a multidisciplinary process that demands a balance between creativity and systematic supervision. MILSDeM addresses this challenge by organizing a real-life project that consists of four initiatives, resulting in eight educational prototypes in two different domains (see Table 6). This approach allows teams to freely explore creative ideas while receiving systematic guidance at the initiative level (see

Table 5). In our instantiation phase, diverse development teams benefited from the MILSDeM approach, leveraging their skills and decision flexibility to unlock their full potential.

RQ3 How can we systematically guide ILS developers to comprehensively explore design options while ensuring the creation of reusable (learning) components from prototyping outcomes?

The path to achieving ILS project goals is complex, with numerous design choices and limited guidance in the literature. Investing in one choice may overlook better alternatives, leading to sub-optimal outcomes. Multiple prototyping endeavours can address this challenge, but without systematic guidance, inefficiencies may arise. MILSDeM tackles this by offering a unified taxonomy of ILS, providing a comprehensive overview of dimensions for guided exploration and improved learning outcomes. Visual charts track exploration progress and set goals for future initiatives.

ILS development offers organisational learning through its multidisciplinary nature and diverse design possibilities. MILSDeM facilitates this learning with an iterative development process that encourages multiple prototyping teams to explore different areas. To promote reusability and knowledge creation, MILSDeM utilises three KPIs: lessons learned, reusable components, and exploration comprehensiveness. Tracking these KPIs at each stage enhances progress, learning, and reusability within the project. For instance, we observed that some components from prototypes of Initiative II were reused for the prototypes within Initiative III, involving a shift in the learning focus from cognitive to psychomotor domains (see Table 7).

RQ4 Which components emerge as prominent, and what additional components can be integrated to further enhance the taxonomy, according to the experts?

During the evaluation of the taxonomy by experts, several observations were made. AR technology was found to be the dominant component in the experience technology category, likely due to its widespread use in cognitive skills development, facilitated by the prevalence of smartphones and tablets with built-in AR capabilities (Ibáñez and Delgado-Kloos, 2018). Conversely, AR technology sees limited utilization within the psychomotor domain. Therefore, it's valuable to explore whether similar pedagogical benefits can be attained by employing AR in this domain. In the interaction technology category, voice input emerged as a preferred component, possibly reflecting the popularity of voice-based AI assistants like Amazon Echo (Alexa) and Apple Siri. Such technology could aid learners with disabilities by providing an alternative method for inputting information, enabling them to engage with educational content more effectively (Bain et al., 2002). It should be noted that some participants may not be familiar with these technologies, so no counts were recorded in such cases. In the pedagogy dimension, the cognitive domain was the most prominent component, which could be attributed to the participants' backgrounds. Real-time feedback was the preferred intervention technique, known for its effectiveness in immediate error correction (Geisen and Klatt, 2022).

Participants expressed the significance of audio technology in delivering instructions and feedback. They also suggested integrating AI technology to track learner

performance and provide personalized feedback. The inclusion of game elements as a sub-category was proposed to enhance the enjoyment and engagement of ILS. In the pedagogy dimension, participants emphasized the importance of the psychomotor and cognitive domains targeting specific skills that can be learned within ILS. Refinement of feedback was deemed essential, particularly by educators who accounted for over 30% of the participants. Collaborative learning practices in ILS to foster team-building skills were also highlighted by the participants, enabling learners from diverse locations to converge in a shared virtual learning environment, facilitating mutual learning, knowledge exchange, and collective progress toward common educational objectives.

In summary, MILSDeM makes important theoretical and practical contributions. Theoretically, the methodology addresses the unique challenges of ILS development, filling a gap in the literature. The unified taxonomy offers a systematic approach to exploring design choices and the iterative development process balances creativity with systematic guidance. MILSDeM's KPIs provide novel metrics for setting goals related to reusability, learning, and comprehensive exploration.

The MILSDeM approach allows for the development of ILS by multidisciplinary teams, by providing a structured process. This encourages the utilisation of diverse skills within the team, leading to better project outcomes. The use of MILSDeM can also result in the development of reusable components, which can save time and resources in future projects. Overall, MILSDeM offers a practical framework for ILS development that addresses the challenges posed by the unique components of ILS, grounded on Bloom's taxonomy (i.e., cognitive, psychomotor, affective), and enables organisations to achieve better project outcomes.

7.1 Limitations

This study has several limitations. Firstly, the iterative approach of MILSDeM requires significant time and effort, which limits the number of projects that can be evaluated within a given timeframe. The study was able to evaluate the methodology using one project over a period of nine months, with different initiatives being carried out. Secondly, the COVID-19 pandemic forced the team to work remotely. While remote work offered flexibility, the absence of in-person interactions hindered spontaneous discussions and knowledge-sharing among team members. The natural exchange of ideas that typically occurs in a shared workspace was limited, potentially impeding the depth of learning and innovative thinking. Further, the shift to remote work altered the dynamics of mentorship offered by the supervision team, which is integral for casual conversations, guidance, and on-the-job training. Next, while deriving learning objectives from Bloom's taxonomy, which encompasses the three domains of psychomotor, cognitive, and affective, the process encountered a limitation. The taxonomy, albeit robust, presents a constraint due to its focus on these specific domains, potentially omitting other crucial dimensions of learning. Finally, the study was conducted within an academic research group and may not be directly applicable to commercial organisations, which may have different considerations that require the tailoring of MILSDeM to fit their specific needs.

7.2 Future work

Our future endeavours are directed toward expanding the application of MILS-DeM across diverse ILS projects, spanning various organisational structures and non-academic settings. The intention is to systematically gauge its efficacy as a methodology beyond its current scope. In doing so, our future objectives will be centered around the continuous enhancement of MILSDeM, aiming to refine its taxonomy through iterative improvements. This process involves an evaluation of its components, structure, and applicability to optimize its effectiveness and relevance in diverse project settings. Moreover, we aim to observe the instantiation of MILSeM in a shared physical environment where teams have close communication with each other, as team communication might yield different results in terms of organisational learning and exploration. Furthermore, we aim to adapt and integrate various learning taxonomies that offer a broader spectrum of domains beyond Bloom's taxonomy, enriching our approach by encompassing a more comprehensive range of learning dimensions and frameworks.

Appendix A Tables

Table 5 Overview of the MILSDeM steps

Step	Inputs	Activities	Outputs	KPIs
Goal Setting				
	• Idea	• Define project goals and constraints	• Project goals and constraints	• Initiative-project goal match
	• Unified Taxonomy of ILS	• Define the exploration objectives	• Exploration objectives	
		• Identify initiative opportunities	• Initiative opportunities	
Expanding				
	• Project goals and constraints	• Set up the initiative	• Initiative	• Prototype-initiative goal match
	• Exploration objectives	• Define the initiative theme	• Initiative/prototype goals	
		• Refine project goals and clarify initiative/prototype goals	• Prototyping teams	
		• Assign resources	• Resource assignments	
		• Initiate prototyping endeavours		
Execution				
	• Initiative/ prototype goals	• Prototyping teams carry out the prototyping activities	• Developed prototype components	• Number of reusable components

Table 5 continued

Step	Inputs	Activities	Outputs	KPIs
	• Resource assignments	• Explore, exploit, and learn	• Lessons learned	• Lessons learned
Evaluation				
	• Developed prototype components	• Compare the execution outcomes against the project goals	• Reported KPIs	• Number of reusable components
	• Lessons learned	• Check compliance with the project constraints	• Evaluation report/decisions	• Lessons learned
	• Project goals and constraints		• Exploration comprehensive-ness	
Supervision				
	• Prototyping teams	• Monitor and control the execution	• Corrective actions	• Mitigated risks
	• Lessons learned	• Take corrective action if necessary	• Inter-team communication	• Number of corrective actions
	• Risks identified	• Manage shared risks		
		• Ensure communication among prototype teams		
Collapsing				
	• Exploration objectives	• Gather prototype outcomes	• Prototype outcomes	• Number of reusable components
	• Developed prototype components	• Share reusable components and the lessons learned with all teams	• Reusable components	
	• Lessons learned	• Evaluate the comprehensiveness of exploration	• Lessons learned	
		• Decide to finalize the project or reiterate by defining a new initiative	• Exploration status	• Lessons learned
			• Decision on the next steps	• Exploration comprehensive-ness
Closing				
	• Decision to finalize the project	• Formally close the project • Project report	• Finalized project	• Project goals achieved

Table 6 The prototypes with their respective reusable components

Prototype (Initiative)	Topics	Team size (Number of individuals)	Project duration (Week)	Supervision team (Number of supervisors)	Reusable components	Reused components from the previous initiative/s
ElemAR (Initiative I)	Chemistry	1	24	1	• Image and 3D shape tracking (AR Marker generator) • Scannable cards • Teaching guideline (worksheet) template • Interactive UI elements	-
GeAR (Initiative II)	Engineering	3	1	6	-	• Image and 3D shape tracking (AR Marker generator) • Scannable cards • Teaching guideline (worksheet) template • Interactive UI elements
Newton's Second Law (Initiative II)	Physics	4	1	6	• Audio feedback • Hints	• Image and 3D shape tracking (AR Marker generator) • Scannable cards • Teaching guideline (worksheet) template • Interactive UI elements

Table 6 continued

Prototype (Initiative)	Topics	Team size (Number of individuals)	Project duration (Week)	Supervision team (Number of supervisors)	Reusable components	Reused components from the previous initiative/s
The Big Banger (Initiative III)	Solar systems and physics	6	8	5	• Localization and mapping • Puzzle • Score points • Visual instructions • Visual feedback	• Interactive UI elements • Audio feedback
Sir Kit's Solar Power Trip (Initiative III)	Electric circuit	6	8	5	• Puzzle • Visual instructions • Visual feedback • Audio instructions	• Image and 3D shape tracking (AR Marker generator) • Scannable cards • Interactive UI elements • Hints • Audio feedback
Yu & Mi (Initiative III)	Human-robot interaction	5	8	5	• Localization and mapping • Mini-game • Visual instructions • Visual feedback	• Interactive UI elements • Audio feedback

Table 6 continued

Prototype (Initiative)	Topics	Team size (Number of individuals)	Project duration (Week)	Supervision team (Number of supervisors)	Reusable components	Reused components from the previous initiative/s
Flowmotion (Initiative III)	Yoga	6	8	5	• Body tracking • Score points • Visual instructions • Visual feedback • Audio instructions	• Interactive UI elements • Audio feedback
MPITT (Initiative IV)	Yoga	6	8	4	• Improved and added new visual instructions • Improved and added new visual feedback • Visual instructions • colour vision deficiency UI	• Localization and mapping • Interactive UI elements • Visual instructions • Visual feedback

Table 7 Unified ILS Taxonomy dimensions explored in the initiatives

Prototype (Initiative)	Technology	Types of Learning	Instructions - Modalities	Instructions - Frame of interaction	Feedback - Modalities	Feedback - Frame of interaction
ElemAR (Thesis)	• AR* • Tactile*	• Cognitive*	• Visual (textual)*	• Before*	• Visual (graphical - UI objects: a combination of molecules)*	• During* • After*
(Initiative Cumulative I)	2	1	1	1	1	1
GeAR (Hackathon)	• AR • Tactile	• Cognitive	• Visual (textual)	• Before	• Visual (graphical - UI objects: reaction of the gear systems)	• During
Newton's Second Law (Hackathon)	• AR • Tactile	• Cognitive	• Visual (textual)	• Before	• Visual (graphical - UI objects: collision between objects) • Audio (sound)*	• During • After
(Initiative Cumulative II)	2	1	1	1	2	1
The Big Banger (Semester Project)	• AR • Tactile	• Cognitive	• Visual (textual)	• Before	• Visual (graphical - UI objects: collision of the planets) • Visual (textual)* • Audio (sound)	• During • After
Sir Kit's Solar Power Trip (Semester Project)	• AR • Tactile	• Cognitive	• Visual (textual) • Audio (speech)*	• Before	• Visual (graphical - UI objects: electric flow) • Audio (sound)	• During • After

Education and Information Technologies

Table 7 continued

Prototype (Initiative)	Technology	Types of Learning	Instructions - Modalities	Instructions - Frame of interaction	Feedback - Modalities	Feedback - Frame of interaction
Yu & Mi (Semester Project)	• AR • Tactile	• Cognitive • Psychomotor	• Visual (textual)	• Before	• Visual (graphical - UI objects: sequence of robot movements) • Audio (sound)	• During
Flowmotion (Semester Project)	• Screen* • Gesture* • Motion detectors*	• Cognitive • Psychomotor	• Visual (graphical - UI objects: avatar)* • Audio (speech)	• Before	• Visual (graphical - UI objects: correct/wrong indicator)* • Audio (sound) • Audio (speech)*	• During • After
(Initiative III) Cumulative	5	2	3	1	4	2
MPITT (Student Project)	• MR* • Gesture	• Cognitive • Psychomotor	• Visual (textual)	• Before • During*	• Visual (textual)	• During • After
Total	6	2	3	2	4	2

Data avaliability statement The datasets generated during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethical approval The Ethical Research Board (cETO) of the Open University of the Netherlands has assessed the proposed research and concluded that this research is in line with the rules and regulations and the ethical codes for research in Human Subjects (reference: U2023/02316).

Conflict of interest None.

References

Ahmad, M. (2019). Categorizing game design elements into educational game design fundamentals. *Game Design and Intelligent Interaction*, 1–17.

Almirall, E., Lee, M., & Majchrzak, A. (2014). Open innovation requires integrated competition-community ecosystems: Lessons learned from civic open innovation. *Business horizons, 57*(3), 391–400.

Bain, K., Basson, S. H., & Wald, M. (2002). Speech recognition in university classrooms: Liberated learning project. In *Proceedings of the fifth international ACM conference on Assistive technologies* (pp. 192–196)

Beck, K., Beedle, M., Van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R.C., Mellor, S., Schwaber, K., Sutherland, J., Thomas, D. (2001). Manifesto for Agile software development. Agile Alliance. Available at https://agilemanifesto.org/

Beemer, L. R., Ajibewa, T. A., DellaVecchia, G., & Hasson, R. E. (2019). A pilot intervention using gamification to enhance student participation in classroom activity breaks. *International journal of environmental research and public health, 16*(21), 4082.

Bethke, E. (2003). *Game development and production*. Wordware Publishing, Inc.

Bloom, B. S. (ed.), Engelhart, M. D., Fürst, E. J., Hill, W. H., Krathwohl, D. R. (1956). *Taxonomy of Educational Objectives: Handbook I: Cognitive Domain*. New York: David McKay

Boehm, B. W. (1988). A spiral model of software development and enhancement. *Computer, 21*(5), 61–72.

Brabham, D. C. (2008). Crowdsourcing as a model for problem solving: An introduction and cases. *Convergence, 14*(1), 75–90.

Brede Moe, N., Ulfsnes, R., Stray, V., & Smite, D. (2021). Improving productivity through corporate hackathons: A multiple case study of two large-scale agile organizations. *arXiv e-prints*, arXiv:2112.05528

Chesbrough, H. (2006). Open innovation: A new paradigm for understanding industrial innovation. *Open innovation: Researching a new paradigm*, 400, 0–19

Chouseinoglou, O., Iren, D., Karagöz, N. A., & Bilgen, S. (2013). Aiolos: A model for assessing organizational learning in software development organizations. *Information and Software Technology, 55*(11), 1904–1924.

Cohen, P., McGee, D., Oviatt, S., Wu, L., Clow, J., King, R., Julier, S., & Rosenblum, L. (1999). Multimodal interaction for 2d and 3d environments [virtual reality]. *IEEE Computer Graphics and Applications, 19*(4), 10–13.

De Freitas, S., & Oliver, M. (2006). How can exploratory learning with games and simulations within the curriculum be most effectively evaluated? *Computers & education, 46*(3), 249–264.

de Freitas, S., Rebolledo-Mendez, G., Liarokapis, F., Magoulas, G., & Poulovassilis, A. (2009). Developing an evaluation methodology for immersive learning experiences in a virtual world. In *2009 Conference in games and virtual worlds for serious applications* (pp. 43–50). IEEE

Devedžić, V., et al. (2010). Teaching agile software development: A case study. *IEEE transactions on Education, 54*(2), 273–278.

Di Mitri, D., Schneider, J., Specht, M., & Drachsler, H. (2018). From signals to knowledge: A conceptual model for multimodal learning analytics. *Journal of Computer Assisted Learning, 34*(4), 338–349.

Diziol, D., Walker, E., Rummel, N., & Koedinger, K. R. (2010). Using intelligent tutor technology to implement adaptive support for student collaboration. *Educational Psychology Review, 22*(1), 89–102.

Easterday, M. W., Lewis, D. R., & Gerber, E. M. (2014). *Design-based research process: Problems, phases, and applications*. Boulder: International Society of the Learning Sciences.

Geisen, M., & Klatt, S. (2022). Real-time feedback using extended reality: A current overview and further integration into sports. *International Journal of Sports Science & Coaching, 17*(5), 1178–1194.

Georgiou, Y., & Kyza, E. A. (2018). Relations between student motivation, immersion and learning outcomes in location-based augmented reality settings. *Computers in Human Behavior, 89*, 173–181.

Gill, N. S. (2003). Reusability issues in component-based development. *ACM SIGSOFT Software Engineering Notes, 28*(4), 4.

Hertel, J., Karaosmanoglu, S., Schmidt, S., Bräker, J., Semmann, M., & Steinicke, F. (2021). A taxonomy of interaction techniques for immersive augmented reality based on an iterative literature review. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 431–440). IEEE.

Houde, S., & Hill, C. (1997). What do prototypes prototype? In *Handbook of human-computer interaction* (pp. 367–381). Elsevier

Ibáñez, B. C., Marne, B., & Labat, J.-M. (2011). Conceptual and technical frameworks for serious games. In *Proceedings of the 5th European conference on games based learning* (pp. 81–87).

Ibáñez, M.-B., & Delgado-Kloos, C. (2018). Augmented reality for stem learning: A systematic review. *Computers & Education, 123*, 109–123.

Ibrahim, R., & Jaafar, A. (2009). Educational games (eg) design framework: Combination of game design, pedagogy and content modeling. In *2009 international conference on electrical engineering and informatics* (vol. 1, pp. 293–298). IEEE

Iren, D., & Bilgen, S. (2014). Cost of quality in crowdsourcing. *Human Computation*, 1(2)

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International journal of human-computer studies, 66*(9), 641–661.

Kerzner, H. (2017). *Project management metrics, KPIs, and dashboards: A guide to measuring and monitoring project performance*. John Wiley & Sons.

Kim, L., Markovina, S., Van Nest, S. J., Eisaman, S., Santanam, L., Sullivan, J. M., Dominello, M., Joiner, M. C., & Burmeister, J. (2019). Three discipline collaborative radiation therapy (3dcrt) special debate: Equipment development is stifling innovation in radiation oncology. *Journal of applied clinical medical physics, 20*(9), 6.

Lara, M., & Lockwood, K. (2016). Hackathons as community-based learning: A case study. *TechTrends, 60*(5), 486–495.

Limbu, B., Vovk, A., Jarodzka, H., Klemke, R., Wild, F., & Specht, M. (2019). Wekit. one: A sensor-based augmented reality system for experience capture and re-enactment. In *European Conference on Technology Enhanced Learning* (pp. 158–171). Springer

Limbu, B. H., Jarodzka, H., Klemke, R., & Specht, M. (2018). Using sensors and augmented reality to train apprentices using recorded expert performance: A systematic literature review. *Educational Research Review, 25*, 1–22.

Luo, V., Klinkert, L.J., Foster, P., Tseng, C.-Y., Adams, E., Ketterlin-Geller, L., Larson, E.C., & Clark, C. (2021). A multidisciplinary approach to designing immersive gameplay elements for learning standard-based educational content. In *Extended Abstracts of the 2021 Annual Symposium on Computer-Human Interaction in Play* (pp. 67–73)

Martin, D., Malpica, S., Gutierrez, D., Masia, B., & Serrano, A. (2022). Multimodality in VR: A survey. *ACM Computing Surveys (CSUR), 54*(10s), 1–36

Mat Sanusi, K. A., Iren, Y., & Klemke, R. (2021). Immersive multimodal environments for psychomotor skills training. In *Sixteenth European Conference on Technology Enhanced Learning* (16)

Menin, A., Torchelsen, R., & Nedel, L. (2018). An analysis of vr technology used in immersive simulations with a serious game perspective. *IEEE computer graphics and applications, 38*(2), 57–73.

Metcalfe, J., Kornell, N., & Finn, B. (2009). Delayed versus immediate feedback in children's and adults' vocabulary learning. *Memory & cognition, 37*(8), 1077–1087.

Motejlek, J., & Alpay, E. (2021). Taxonomy of virtual and augmented reality applications in education. *IEEE transactions on learning technologies, 14*(3), 415–429.

Park, S.-M., & Kim, Y.-G. (2022). A metaverse: Taxonomy, components, applications, and open challenges. *IEEE access, 10*, 4209–4251.

Patil, S., Hoyle, R., Schlegel, R., Kapadia, A., & Lee, A. J. (2015). Interrupt now or inform later? comparing immediate and delayed privacy feedback. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 1415–1418)

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). *Foundations of game-based learning. Educational psychologist, 50*(4), 258–283.

Pöyhönen, A., & Hong, J. (2006). Measuring for learning–outlining the future of organizational metrics. In *Submitted to OLKC 2006 Conference at the University of Warwick*. https://pdfs.semanticscholar.org/5ad8/02fd0026eb959f38e5af233ba2abe220ab71.pdf. Accessed on 29/11/2023

Royce, W. W. (1987). Managing the development of large software systems: Concepts and techniques. In *Proceedings of the 9th international conference on Software Engineering* (pp. 328–338)

Rubio-Tamayo, J. L., Gertrudix Barrio, M., & García García, F. (2017). Immersive environments and virtual reality: Systematic review and advances in communication, interaction and simulation. *Multimodal Technologies and Interaction, 1*(4), 21.

Schwaber, K., & Beedle, M. (2002). *Agile software development with Scrum* (vol. 1). Prentice Hall Upper Saddle River

Schwalbe, K. (2009). *Introduction to project management*. Course Technology Cengage Learning Boston

Sommerville, I. (2011). Software engineering 9th edition. *ISBN-10*, 137035152, 18

Suominen, A. H., Jussila, J., Lundell, T., Mikkola, M., & Aramo-Immonen, H. (2018). Educational hackathon: Innovation contest for innovation pedagogy. In *LUT Scientific and Expertise Publications, Reports* (78). Lappeenranta University of Technology; ISPIM

Taylor, N., & Clarke, L. (2018). Everybody's hacking: Participation and the mainstreaming of hackathons. In *CHI 2018* (pp. 1–2). Association for Computing Machinery

Wang-Trexler, N., Yeh, M. K., Diehl, W. C., Heiser, R. E., Gregg, A., Tran, L., & Zhu, C. (2021). Learning from doing: Lessons learned from designing and developing an educational software within a heterogeneous group. *International Journal of Web-Based Learning and Teaching Technologies (IJWLTT), 16*(4), 33–46.

Westland, J. (2007). *The project management life cycle: A complete step-by-step methodology for initiating planning executing and closing the project*. Kogan Page Publishers

Wiewiora, A., & Murphy, G. (2015). Unpacking 'lessons learned': Investigating failures and considering alternative solutions. *Knowledge Management Research & Practice, 13*(1), 17–30.