GENIE Learn: Human-Centered Generative AI-Enhanced Smart Learning Environments

Carlos Delgado Kloos[1][a], Juan I. Asensio-Pérez[2][b], Davinia Hernández-Leo[3][c], Pedro Manuel Moreno-Marcos[1][d], Miguel L. Bote-Lorenzo[2][e], Patricia Santos[3][f], Carlos Alario-Hoyos[1][g], Yannis Dimitriadis[2][h] and Bernardo Tabuenca[4][i]

[1]Universidad Carlos III de Madrid, Av. Universidad 30, 28911 Leganés (Madrid), Spain
[2]Universidad de Valladolid, Paseo de Belén 15, 47011 Valladolid, Spain
[3]Universitat Pompeu Fabra, Roc Boronat 138, 08018 Barcelona, Spain
[4]Universidad Politécnica de Madrid, Calle Alan Turing sn, 28031 Madrid, Spain

Keywords: Smart Learning Environments, Hybrid Learning, Generative Artificial Intelligence, Human Centeredness.

Abstract: This paper presents the basis of the GENIE Learn project, a coordinated three-year research project funded by the Spanish Research Agency. The main goal of GENIE Learn is to improve Smart Learning Environments (SLEs) for Hybrid Learning (HL) support by integrating Generative Artificial Intelligence (GenAI) tools in a way that is aligned with the preferences and values of human stakeholders. This article focuses on analyzing the problems of this research context, as well as the affordances that GenAI can bring to solve these problems, but considering also the risks and challenges associated with the use of GenAI in education. The paper also details the objectives, methodology, and work plan, and expected contributions of the project in this context.

1 INTRODUCTION

Recent advances in the interdisciplinary field of Technology-Enhanced Learning (TEL) have made possible novel models of teaching and learning based on a mixture or fusion of traditional approaches along different dimensions or dichotomies: learning in physical/digital spaces, informal/formal learning, face-to-face/online learning, individual/collaborative active learning, etc. (Hilli et al., 2019). These novel TEL models, which showcased their relevance during the COVID pandemic, are studied under the theoretical umbrella of the so-called **Hybrid Learning (HL)** (Cohen et al., 2020, Gil et al., 2022), which blurs the boundaries of those dichotomies.

Smart Learning Environments (SLEs) can be useful as technological support for HL (Delgado Kloos et al., 2018 and 2022). SLEs are conceived to provide personalized support to learners, considering learning needs and context. SLEs "collect data from the learning context (**sense**), decode, process the data collected (**analyze**), and coherently suggest actions to ease learning constraints toward improved learning performance (**react**)" (Tabuenca et al., 2021). SLEs expand the work from context-aware, ubiquitous learning, and adaptive learning systems, actively supporting students according to their learning situation, across physical and virtual learning spaces (Gross, 2016), based on data-driven interventions (Hernández-Leo, et al., 2023a).

[a] https://orcid.org/0000-0003-4093-3705
[b] https://orcid.org/0000-0002-1114-2819
[c] https://orcid.org/0000-0003-0548-7455
[d] https://orcid.org/0000-0003-0835-1414
[e] https://orcid.org/0000-0002-8825-0412
[f] https://orcid.org/0000-0002-7337-2388
[g] https://orcid.org/0000-0002-3082-0814
[h] https://orcid.org/0000-0001-7275-2242
[i] https://orcid.org/0000-0002-1093-4187

Delgado Kloos, C., Asensio-Pérez, J. I., Hernández-Leo, D., Moreno-Marcos, P. M., Bote-Lorenzo, M. L., Santos, P., Alario-Hoyos, C., Dimitriadis, Y. and Tabuenca, B.
GENIE Learn: Human-Centered Generative AI-Enhanced Smart Learning Environments.
DOI: 10.5220/0013076000003932
In *Proceedings of the 17th International Conference on Computer Supported Education (CSEDU 2025) - Volume 2*, pages 15-26
ISBN: 978-989-758-746-7; ISSN: 2184-5026

SLEs rely on advances in both **Learning Design (LD)** (Wasson & Kirschner, 2020) and **Learning Analytics (LA)** (Long & Siemens, 2011). Thanks to LD tools, teachers can make their pedagogical intentions explicit, and even represent them in computer-interpretable formats, thus enabling SLEs to use them as contextual inputs. LA constitutes a key component of SLEs collecting data from both physical and virtual spaces. SLEs are aimed at modelling students in context to provide adequate personalized support, in many cases making use of Artificial Intelligence (AI) algorithms (Buckingham et al., 2019a). The synergetic relationship between LD and LA is key for supporting data-driven interventions in SLEs: interventions based on LA indicators aligned with LDs are typically more meaningful; and the effectiveness of LDs can be better understood in the light of LA indicators. This relationship between LD and LA is also reflected in the so-called Design Analytics (metrics of design decisions and related aspects characterizing LDs) and Community Analytics (metrics and patterns of LD activity) (Hernández-Leo, et al., 2019). Finally, academic analytics, when supporting educational decision making at institutional level, can be seen as a variation of LA also in interplay with complementary data layers (Misiejuk et al., 2023). Thus, academic analytics can also benefit from the core functions of SLEs, while differing in the requirements of the stakeholders (educational managers, beyond teachers) and the scale of the data potentially relevant for its analysis (Hernández-Leo et al., 2019, Ortiz Beltrán et al., 2023). Figure 1 summarizes the research context, which includes technological, pedagogical and human contexts.

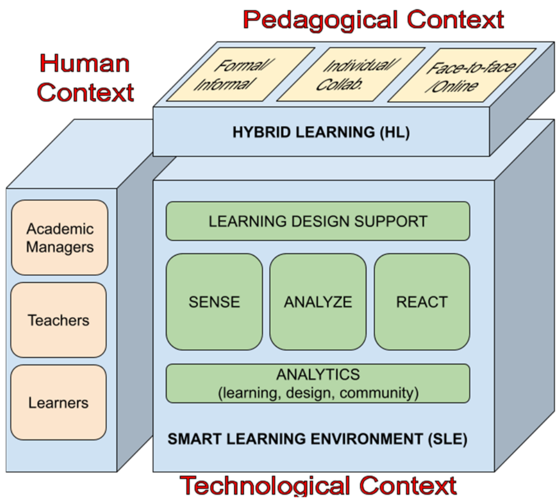

Figure 1: Research context (human, pedagogical, technological).

This paper builds on this research context and presents the theoretical foundations and research plan of the GENIE Learn project, a coordinated research project funded by the Spanish Research Agency and with the participation of three Spanish Universities (Universidad Carlos III de Madrid, Universidad de Valladolid, and Universitat Pompeu Fabra). GENIE Learn is aimed at improving SLEs for HL support by integrating Generative Artificial Intelligence (GenAI) tools in line with the preferences and values of the key stakeholders (teachers, learners, and academic managers). The remaining of this paper is as follows: Section 2 analyzes potential problems present in the research context as well as the affordances and challenges of GenAI tools to address these problems. Section 3 presents the initial hypothesis and objectives of the GENIE Learn project. Section 4 discusses the methodology and work plan. Finally, Section 5 draws the conclusions of this article.

2 PROBLEMS, AFFORDANCES, AND CHALLENGES

2.1 Problems in the Research Context

There are currently several significant problems (P) to be addressed when trying to improve the support to HL provided by state-of-the-art SLEs.

P1. Improving the support for LD in connection with LA. Educators can use multiple methods and tools to support the process of LD in a variety of HL contexts (Wasson & Kirschner, 2020). However, the complexity of the LD process has been a recognized challenge, hindering the widespread adoption of LD methods and tools, despite the field's significant relevance (Dagnino, et al., 2018). This complexity arises from the limitations of the tools in meeting the theoretical ambitions of the field. The goal is to enable multiple stakeholders to get involved in the creative inquiry process of data-driven co-designing for learning, generating design artifacts that evolve through time and feed the successive phases (Michos & Hernández-Leo, 2020). A new generation of LD tools for SLE is necessary to achieve its goals. These LD tools should facilitate the collaboration of various stakeholders towards pedagogically-sound LDs and enable a thorough exploration of the educational context's needs and consideration of LA from previously implemented designs. These tools should also support data-driven inspiration drawing on analytics from related design artifacts (Hernández-Leo, et al., 2019), predominantly available in unstructured formats and created by teacher communities (Gutiérrez-Páez et al., 2023). In summary, enhanced support is needed for needs analysis, brainstorming, inspiration, and creative ideation, informed pedagogical decision-making, scaffolding and integration of stakeholder conversations, and improved interpretation and actionability of advanced (LD community) analytics.

P2. Supporting a wider diversity of learners (universal design for learning) and ethical design aspirations. The Universal Design for Learning (UDL) framework is structured around three main principles: a) Multiple means of representation; b) Multiple means of engagement; c) Multiple means of expression. However, the seamless integration of the principles into educational practices brings challenges, demanding a reconsideration of teaching methodologies and materials. The implementation of UDL by educators demands specific knowledge as well as substantial time and effort (Rose and Meyer, 2002; Evmenova, 2018). There is a lack of LD tools (Gargiulo and Metcalf, 2010) that support the adaptation of pedagogical and educational material design considering UDL. Such tools would enable stakeholders to more efficiently and effectively embrace UDL principles.

P3. Reducing the burden on teachers associated with the real-time orchestration demands of HL environments. An important challenge in SLEs refers to the "orchestration load" (Amarasinghe et al., 2022). This concept reflects the teachers' attentive processing during the real-time management of complex learning scenarios. Designing tools to lower this burden should be a top priority. The results from a recent international workshop (Amarasinghe et al., 2023a) highlight the need to provide orchestration tools that consider different needs depending on the delivery mode, pedagogical method (individual/collaborative), teacher characteristics, and content knowledge being considered in the learning activity (Hakami et al., 2022, Raes et al., 2020). This has implications in the type of data collected from the stakeholders, as well as in its analysis and presentation to support teachers in e.g., orchestration dashboards. The main data source is quantitative data related to student engagement (LA presented to teachers) and teachers' actions in the dashboard (for researchers to study orchestration load) (Amarasinghe, et al., 2021). A problem is that understanding the orchestration load involves the use of additional data collection sources and more advanced approaches for data analysis and presentation in SLEs. This includes real-time analysis of students individual and collective self-explanations in their answers to learning activities (e.g., to lower the burden when intervening if a problem is identified, or during debriefing sessions), as well as in the collection and triangulated analysis of teachers' data coming from sensors in SLE that goes from physiological sensors to reflection diaries and video recordings.

P4. Large-scale analysis of SLE data to support decision making in academic management. LA (from learners) and design analytics (from teachers) collected in SLE have the potential to support academic managers in institutional decision making (e.g., evolution of the educational model, identification of needs for teacher training) (Hernández-Leo, et al., 2019). Yet, this potential has not been exploited. Only limited, but relevant approaches considering student ratings regarding their teaching satisfaction, e.g., depending on the HL modalities (Ortiz Beltrán et al., 2023), and classifications of course designs approaches have been developed so far (Toetenel & Rienties, 2018; Misiejuk et al., 2023). The problem is that the envisaged potential requires an analysis at large scale involving multiple courses, data types and formats (including e.g., course and program descriptions, student feedback on course design) and an integrated analysis of the different relevant data sources.

P5. Limited types of automatically generated personalized context-aware learning tasks and feedback interventions. SLEs can generate automatic, data-driven interventions in HL scenarios with minimum or no teacher involvement, thus opening the possibility of personalized learning at different scales (e.g., when many students participate in the learning scenario). For example, Ruiz-Calleja et al. (2021) propose the use of Linked Open Data in the Web for automatically generating large numbers of contextualized learning tasks in the domain of cultural heritage. However, that approach relies on the use of a small and rigid set of teacher-generated task templates that only allow the generation of some very specific types of learning tasks (e.g., asking about the architectural style of a church, or suggesting taking a photo of a generic architectural element) that cannot always be automatically assessed. The generation of more complex learning tasks such as the explanation of the characteristics of a building in relation to its historical context cannot be generated with the current technological solutions. Similarly, the provision of personalized feedback in large-scaled educational settings has also been explored (e.g., Topali et al., 2024) resulting in the development of technological support tools (Ortega-Arranz et al., 2022). However, current technological solutions for personalized feedback in SLEs only support a limited number of types of feedback interventions (e.g., predefined email messages encouraging the students to carry out certain tasks) based also on a limited set of LA indicators (e.g., the score in an online quiz). More elaborated types of feedback interventions (e.g., providing an automatic explanation about the reasons for a low score in a text-based learning task) are also not possible for current SLEs.

P6. Students' models automatically generated at large scales are mostly based on quantitative/structured data. SLEs can generate individual students' models used as inputs for automatic personalized interventions (e.g., Serrano-Iglesias et al., 2021). However, state-of-the-art SLEs base those students' models mostly on LA indicators that make use of structured data coming from Virtual Learning Environment, physical sensors, etc. LA indicators derived from other, non-structured learning data such as text-based learning outcomes have also been proposed but typically based on quantitative features such as temporal evolution of the length of the documents, the number of editions, etc. (e.g., Suraworachet et al., 2021). The building of students' models based on unstructured learning data (e.g., the actual contents of a document produced by a student) is a desirable feature not found in current SLEs.

P7. Improving personalized and more effective learning in scenarios that use conversational interfaces and (self-selected) artificial intelligence (AI) tools. SLEs have been incorporating the use of conversational interfaces to promote active and authentic learning experiences (e.g., in social media education, Theophilou, et al., 2023a). Personalization in these contexts is often achieved through LA, adapting activity flow to individual needs, and considering students use their own self-selected tools (e.g., Ognibene et al., 2023). However, despite the advantages these scenarios offer for enhanced and efficient learning, several challenges remain. These include: a) the restricted free interaction capabilities of available conversational-oriented SLEs and the identified effects in the limited writing quality of students' submissions to conversational interfaces (Theophilou et al., 2023b); b) the necessity of scaffolding functions to bolster student self-regulation and collaboration quality, considering socio-emotional aspects (Hadwin, 2017); c) the effects of individual attitudes (openness) influence attitudes in the use of supporting tools, especially AI tools (Sánchez-Reina et al., 2023); d) the significant concerns related to academic integrity in education related to the use of AI tools (Kasneci, 2023).

2.2 Affordances of GenAI

GenAI is the branch of AI aimed at creating realistic content such as text, images, audio, or video, based on a given input or prompt (Jovanovic & Campbell, 2022). There are many types of GenAI: Generative Adversarial Networks, Generative Diffusion Models, Generative Pre-trained Transformers (GPTs), etc. Although modern GenAI is based on decades-old models and techniques (e.g., neural networks, back propagation, etc.), the availability of improved computational capabilities, the increase in the size of models, and huge training datasets have made possible that some recent incarnations of GenAI systems, such as OpenAI's GPT-4, show a performance that is "strikingly close to human-level performance" (Bubeck at al., 2023). Previous efforts on AI in Education (AIED) and LA did not anticipate the public availability of more powerful GenAI tools (e.g., OpenAI's ChatGPT or DALL-E, Google Gemini, etc.) capable of carrying out a wide range of tasks with "zero-shot" or "few-shot" training (i.e., without the need to provide a lot of additional training data for fine-tuning the model) and by simply using textual prompts as inputs (Brown et al., 2020). These new capabilities have enabled students and teachers to explore with little effort creative ways of integrating GenAI tools in their learning and teaching tasks (Kasneci et al., 2023). Despite conflicting perceptions (both euphoric and worrisome, or even apocalyptic) of this rushed introduction of GenAI in education (Rudolph et al., 2023), the TEL community is making important efforts to understand its impact, opportunities and challenges. Specifically, in the context of the research problems previously identified in HL support using SLEs, new GenAI tools may bring significant opportunities and affordances (A):

A1 in relation to P1. GenAI may help SLEs provide better support to communities of teachers during the process of creating LDs for HL. For instance, Demetriadis & Dimitriadis (2023) illustrate how to use GPT-3 for creating a conversational agent that reuses design knowledge extracted from existing design conversations. Hernández-Leo (2023b) proposes speculative functions in which GenAI integrated with analytics layers (Hernánez-Leo et al., 2019) may support the LD life cycle. Also, some preliminary results were published on the enrichment of LD tools with Design Analytics (Albó et al., 2022) automatically generated by GenAI. For instance, Pishtari et al. (2024) study the impact of LLM-generated feedback on the quality of LDs produced by teachers.

A2 in relation to P2. UDL principles emphasize the importance of accommodating diverse learning styles and needs in educational environments. GenAI emerges as a potentially highly helpful tool for educators, facilitating the application of UDL principles in educational settings in two ways: a) by reducing the time and effort demanded from teachers through the provision of specialized knowledge and the automation of tasks (Lim et al., 2023); and b) by enhancing the overall quality of teaching and learning experiences. GenAI can play several roles in the support of UDL including the recommendation of effective communication strategies for students with special needs (Garg and Sharma, 2020), or the suggestion of diverse, pedagogically sound content tailored to the students' needs (Mizumoto, 2023).

A3 in relation to P3. GenAI conversational approaches offer the potential to support teachers while implementing learning scenarios (Sharples, 2023), prompting for different analysis and human-readable explanations (Susnjak, 2023) of student progress for guiding interventions and feedback on-the-fly. GenAI may also help to advance the analytics used in orchestration dashboards so textual descriptions and visualizations about the students' actions and unstructured answers are presented to support teacher-led debriefing in HL (Hernández-Leo, 2023b). Multimodal analytics based on sensor data (EEG, heart rate, eye-tracking) and (fine-tuned) GPTs can also be used to analyze classroom orchestration data (Crespi et al., 2022; Amarasinghe et al., 2023b; Tabuenca et al., 2024)

A4 in relation to P4. GenAI may help the large-scale analysis and integration of relevant data sources that can potentially support holistic education decision making to academic managers. The automatic collection of relevant data (sense), including qualitative description (e.g., course descriptions), can take advantage of approaches used in web analytics (e.g., Calvera-Isabal et al., 2023), which facilitates the creation of rich datasets for an integrated analysis of relevant data sources. GenAI can play pivotal roles in the analysis of unstructured data and in advancing interactive and explanatory analytics, which needs to be approached ethically (Susnjak, 2023; Yan et al, 2023). Examples of the potential includes from text similarity analysis across course designs (e.g., extracting information about the pedagogical model to be triangulated with students' performance and satisfaction) to the generation of text explaining insights to the stakeholders about the alignment of data indicators with institutional priorities (e.g., development of desired common competences across study programs).

A5 in relation to P5. Some early research results (see, e.g., Kalo et al., 2020) suggest that LLMs might improve the accuracy and flexibility of SLEs that make use of Linked Open Data in the Web. GenAI may also help SLEs improve the way personalized and context-aware reactions are delivered to human stakeholders (students, teachers, etc.) in HL settings. For example, Dai et al. (2023) gathered promising empirical results about the effectiveness of LLM to automatically generate feedback for learners. In this way, LLMs could potentially be used to compare a student's response with the information available on the Web of Data and generate a reaction that explains what aspects of the response were incorrect.

A6 in relation to P6. GenAI may help SLEs collect (sense) and analyze learning data sources based on natural language, thus widening the range of HL situations that might be supported by SLEs. For example, Amarasinghe et al. (2023b) fine-tuned GPT-3 to automatically code text data from a learning setting and provided evidence of its performance with respect to alternative approaches. This type of GenAI applications might be used by SLEs to automatically assess which concepts are not adequately covered by students in, e.g., a written essay, and thus trigger personalized feedback interventions and recommendations of additional learning tasks to reinforce those concepts (Pereira et al., 2023).

A7 in relation to P7. Advances in language learning models with zero-shot learning capabilities suggest a new possibility for developing educational chatbots for personalized learning using a prompt-based approach. Preliminary tests were already conducted in a case study on effective educational chatbots with ChatGPT prompts (Koyuturk et al., 2023). The results are encouraging, although more research is needed as challenges are posed by the limited history maintained for the conversation and the highly structured form of responses by ChatGPT, as well as their variability. It would be possible to use automatized prompting engineering methodologies (Pryzant et al., 2023) to advance this line of research. On the other hand, providing features embedded in learning platforms that offers adaptive (depending on students' needs) teachable moments (Ognibene et al., 2023; Hernández-Leo, 2022) related AI literacy (e.g. learning to prompt) is expected to improve attitudes and quality of interactions with AI-driven systems (Theophilou et al., 2023c). Finally, the definition of new constructs for LA considering the new requirements imposed using self-selected AI tools in learning processes would enable the development of systems facilitating solutions to address concerns related to academic integrity in education.

2.3 Challenges of GenAI in Education

Despite the promising advantages of using GenAI, the SLE-based support to HL cannot be oblivious to the significant risks and challenges (C) that GenAI pose to education (see, e.g., European Commission, 2023; UNESCO, 2023) and that need to be addressed by the research initiatives, educational institutions and policy makers. Some of these challenges include:

C1. Lack of alignment of GenAI with human values for learning. Although the ethical implications of AI in education have been for years a strong concern for the research community (e.g., Akgun & Greenhow, 2021) and for policy makers (e.g., HLEG-AI, 2019), many of the recently released GenAI tools may be contributing to the worsening of the situation. UNESCO (2023) has already identified ethical risks of GenAI in education and research: a) concentration of GenAI usage in technologically advanced countries, b) lack of democratic control of GenAI companies, c) data protection and copyright issues, d) use of unexplainable models, e) biases in model training and generated content, f) reduction of diversity, and g) manipulation of content. The lack of transparency, privacy and the diminishing of equality have also been identified by Yan et al. (2023) in recent research about GenAI in education. However, the same authors also agree on the need for "adopting a human-centered approach throughout the developmental process" of GenAI tools so as "to protect human agency and genuinely benefit students, teachers and researchers" (UNESCO, 2023), thus trying to overcome the posed ethical challenges. Interestingly, the GenAI research community itself is also paying attention to the challenge of making GenAI tools (with special emphasis on LLMs) "aligned" with human preferences and values, i.e., making them more helpful, honest, and harmless (HHH framework) (Askell et al., 2021). To address the alignment problem, recent research suggests going beyond the mere scaling up of GenAI models and incorporating model fine-tuning based on human feedback (Ouyang et al., 2022). In the case of SLEs for the support of HL, incorporating GenAI solutions may contribute, especially in the "react" function, to less explainable and more biased automatically generated interventions, which can increase the barrier for adoption of this technology among human stakeholders (Serrano-Iglesias et al., 2023).

C2. Challenges to current forms of Human-AI collaboration. The low effort required to use GenAI tools may have a negative effect on the creativity and critical thinking skills of the human actors in education, eventually causing a heavy reliance on GenAI (Kasneci et al., 2023). Beyond the ongoing debate on whether GenAI tools should be either banned or fostered in education (see, e.g., Rudolph et al., 2023), there seems to be a consensus on the need for teachers to develop new skills to incorporate GenAI tools into their practice (Baidoo-Anu & Ansah, 2023). Kasneci et al. (2023) suggest several ways to address this goal: develop new education theory, provide adequate guidance and teacher training, create resources and guidelines for educators and institutions, nurture communities of educators to share and reuse knowledge in applying these new AI tools, among others. All these changes in education should be accompanied by new models of the so-called Human-AI collaboration (also known as hybrid AI-human approaches): finding the proper balance when sharing tasks among humans and AI at different moments of the teaching-learning processes. Recent proposals (see, e.g., Molenaar, 2022, and Järvelä et al., 2023) explore AI-human approaches in TEL settings, although their proposals do not consider the specific case of GenAI. In the case of SLEs for the support of HL finding a right balance between automation and human autonomy and decision-making (also known as "agency") is particularly challenging. This balance has been explored through educators' agency (e.g., Alonso-Prieto, 2023) mainly from the perspective of "orchestration". Also, learners' agency has been explored (see, e.g., Villa-Torrano et al., 2023) from the perspective of self, socially shared and co-regulation of learning (Hadwin et al., 2017) in which learners take metacognitive control of their individual and/or collective cognitive, behavioral, motivational and emotional processes. Although the use of previous generations of AI for detecting and supporting regulation of learning has been widely explored in the LA field (see, e.g., Järvelä, et al., 2023), capabilities of GenAI in terms of natural language processing (e.g., for detecting regulatory episodes in group conversations) and text generation (e.g., for automatically generating personalized feedback regarding learning regulation issues) are still under explored (Gamieldien, 2023). In any case, the impact of novel GenAI solutions in the agency of human stakeholders involved in HL settings supported by SLEs needs to be re-assessed (Hernández-Leo, 2022): will GenAI tools increase teachers' orchestration load (in the already challenging environment of HL)?; how does GenAI-enhanced support for LD affect efficiency but also perspective taking and pedagogical creativity of educators?; how may GenAI tools affect learners' metacognitive processes now that certain learning tasks can be easily solved by those tools?.

C3. There is scarce research evidence about the benefits of GenAI for education in authentic educational practice. For instance, Yan et al. (2023) did not identify any research work showcasing the use of LLMs in "successful operations" (TRL-6 level in the Technology Readiness Level scale, ISO, 2013). Moreover, evidence is needed as it is imperative for ethical GenAI solutions supporting education to be aware of their limitations (Sharples, 2023). Those limitations need to be continuously considered in the responsible conception of AI for education so they are transparently presented in the designed functions for users for their mindful use (Hernández-Leo, 2022). Future research on GenAI and education also needs to pay careful attention to the way empirical results are reported (Yan et al., 2023), trying to provide as many details as possible (e.g., employed prompts, models, source code, data for fine-tuning, etc.) with the ultimate goal of fostering replicability as much as possible in a context in which most training datasets and algorithms are not disclosed. This scarcity of empirical evidence about the impact of GenAI also affects the support of HL with SLEs. However, HL and the associated technological support in the form of SLEs provide a relevant and wide pedagogical and technological context for researching on the impact and risks of GenAI in TEL settings.

3 INITIAL HYPOTHESIS AND OBJETIVES

The GENIE Learn project focuses on the technological support for HL under the initial hypothesis that *the integration of GenAI tools into SLEs may help to overcome important limitations in the current state-of-the-art*. Therefore, the main goal of this project can be formulated as: "**to improve SLEs for HL support by integrating GenAI tools in a way that is aligned with the preferences and values of human stakeholders**". Alignment is here understood using the HHH framework (helpful, honest, and harmless) (Askell et al., 2021) as a starting point and addressed following human-centered principles (Buckingham Shum et al., 2019b; HLEG-AI, 2019) and hybrid AI-human approaches (Järvelä et al., 2023; Siemens et al., 2022). More specifically, GENIE Learn applies Value Sensitive Design (VSD) principles and techniques to consider human values when designing such GenAI-enhanced SLEs. Considering the value alignment perspective, human-centered principles, AI-human collaboration approaches and VSD principles as transversal requirements for the project, the main objectives (O) of this project can be formulated as follows:

- O1: **To define a research framework**, consisting of: 1) a systematic analysis on the use of GenAI tools for supporting SLE functions, as well as on novel Human-AI collaboration models in education; 2) a pedagogical model for HL that considers the affordances and impact of GenAI in the different stakeholders; 3) the definition of a set of HL scenarios in the context of SLEs, co-designed with collaborating teachers and educational institutions, that illustrate the affordances of GenAI tools and their eventual lack of alignment; 4) the definition of human-AI collaboration models applicable to the project scenarios that take into account educational goals as well as the agency of teachers (focused on orchestration) and learners (focused on regulated learning); and 5) the definition of a set of research instruments and a methodology for reporting GenAI-related research results.
- O2: **To design and develop GenAI-enhanced solutions for teachers and academic managers improving learning design and academic decision making in SLEs for HL support** by addressing the problems: 1) improving the support for LD in connection with LA; 2) supporting a wider diversity of learners (UDL) and ethical design aspirations; 3) reducing the burden on teachers associated with real-time orchestration demands of HL environments; and, 4) large-scale analysis of SLE data to support decision making in academic management.
- O3: **To design and develop GenAI-enhanced solutions to improve support for learners in SLEs for HL support** by addressing the problems: 1) limited types of automatically generated personalized context-aware learning tasks and feedback intervention; 2) students' models automatically generated at large scales are mostly based on quantitative/structured data and may include several types of biases (e.g., gender bias); and, 3) improving personalized and more effective learning in scenarios that use conversational interfaces and AI tools.
- O4: **To define a technology framework as an integrated infrastructure**, consisting of: 1) a selection of GenAI platforms and tools (according to O1); 2) the architecture and a development of the integrated infrastructure of an advanced SLE that makes use of the affordances of GenAI tools in ways that are aligned with the stakeholder values and Human-AI collaboration models (O1), and that provides the infrastructure for the outputs

from the tasks derived from O2 and O3; and; 3) technical guidelines for the integration of existing educational datasets with the proposed GenAI-enhanced architecture, so as to improve current approaches to data-driven interventions in SLEs.

- O5: **To design, implement and evaluate pilot experiences** in real settings, of the outcomes of the project following a human-centered approach. The pilots demonstrate the potential of the project contributions considering several educational topics and levels e.g., primary/secondary education, higher education, and lifelong learning. The HL scenarios (O1) are used as a starting point for the design of the pilot experiences, and the infrastructure (O4) is the basis for their technological implementation.

4 METHODOLOGY AND WORK PLAN

4.1 Methodology

The main goal of the GENIE Learn project involves enhancing SLEs for HL by incorporating GenAI in a manner consistent with the values of educational stakeholders. The project aims to develop solutions that address multiple challenges in particular educational settings (see objectives O1 and O5), while simultaneously expanding the understanding of effective technology design (O2, O3, and O4). Given these aims, the Design Science Research Methodology (DSRM) is a suitable methodological framework for guiding the project. DSRM (Peffers et al., 2007) is a widely used iterative methodology in information systems research. DSRM consists of six phases: 1) problem identification and motivation, 2) definition of the objectives for a solution, 3) design and development of the artifact, 4) demonstration of the artifact in a relevant context, 5) evaluation of the artifact and its outcomes, and 6) communication of the research results.

DSRM is also appropriate for the objectives of the GENIE Learn project due to its humanistic dimension (i.e., the issue of aligning GenAI improvements with human stakeholder values), since DSRM is designed to “create things that serve human purposes” with a research-oriented perspective (Peffers et al., 2007). The human dimension of the present project also calls for the use of human-centered research and educational technology design methods (see, e.g., Buckingham Shum et al., 2019), and a human-centered AI perspective (HLEG-AI, 2019). Particularly, the project follows a hybrid human-AI approach (Järvelä et al., 2023; Siemens et al., 2023) in which the AI elements are not meant to fully automate the activities of educational stakeholders, but rather complement their abilities in a way that preserves human agency.

The GENIE Learn project also relies on value-sensitive design (VSD) to align human and AI. VSD (Friedman et al., 2017) is a theoretically grounded approach to the design of technology that explicitly inquires and models human values, their trade-offs, and tensions. In terms of human stakeholder values and their impact on the design of the project’s conceptual and technological proposals, the project uses Askell’s (2021) HHH framework (helpful, honest, and harmless) as the starting point, which is expanded by the state-of-the-art activities needed to define the project’s research framework (O1), and is further aligned with the more specific human stakeholder values to be elicited during early engagement activities with stakeholders from the educational settings and domains of application.

Another important issue in the project’s methodology is the complexity of both the concerned technologies (SLEs, and their GenAI enhancements) and their contexts of application (HL). The understanding of this complexity demands a mixed methods strategy (Johnson et al., 2007) in the data collection and analysis, going beyond the use of purely quantitative/qualitative approaches to obtain a more holistic picture, e.g., of the pilot experiences of use of the technology.

4.2 Work Plan

The work plan for GENIE Learn is organized in 6 work packages (WPs) following the DSRM (see Figure 2 with more details below); each of the first 5 WPs pursues one of the 5 objectives (O1-O5) while the last WP deals with the coordination of the project (Figure 3).

- *WP1: Research Framework.* This WP includes tasks related to the state of the art on GenAI tools for supporting SLE functions and on Human-GenAI collaboration models, the scenarios that are used to illustrate and evaluate the outcomes of the project, the pedagogical model that considers how GenAI might affect teaching-learning processes that need to be aligned with the values of the different human stakeholders, and the definition of research methods and instruments for HL supported by GenAI-enhanced SLEs.
- *WP2: GenAI-enhanced support for management and design of learning*. This WP includes tasks

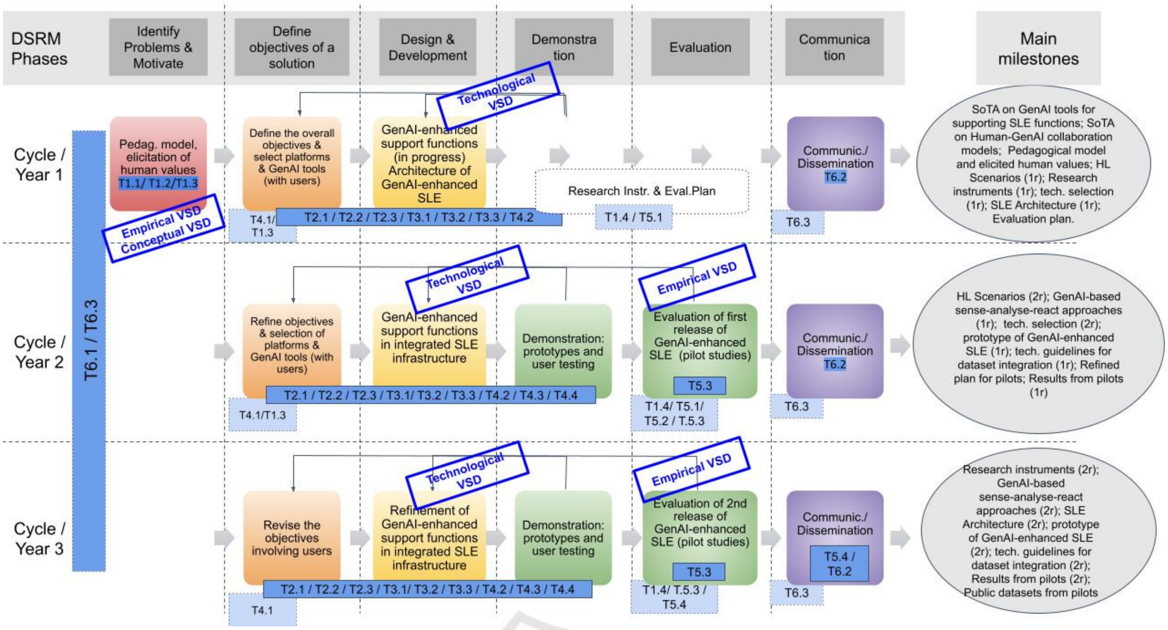

Figure 2: Evolution of the project following the DSRM phases in three cycles. Blue labels represented project tasks.

related to the proposal, development, and testing of approaches aimed at sensing, analyzing, and reacting for GenAI-enhanced support for teachers and academic managers.

- *WP3: GenAI-enhanced support for learning*. This WP includes tasks related to the proposal, development, and testing of approaches aimed at supporting the sensing, analyzing, and reacting core functions of SLEs in HL that are relevant to support learners considering the affordances of GenAI.
- *WP4: Technological Framework: integrated infrastructure*. This WP includes tasks related to the selection of GenAI platforms and tools, and platforms, tools, and devices in HL where GenAI is integrated, the design of an architecture of a GenAI-enhanced SLE for HL, the integration of the proposed technologies in a prototype of a GenAI-enhanced SLE for HL, and the proposal of technical guidelines for the integration of existing educational datasets in the proposed architecture.
- *WP5: Pilot experiences*. This WP includes tasks related to the design of the evaluation plan, the co-design, implementation, and evaluation of pilot experiences, and the sharing of datasets.
- *WP6: Coordination, dissemination, and data management*. This WP is transversal to the GENIE Learn project and includes specific tasks for coordination, dissemination, and data management.

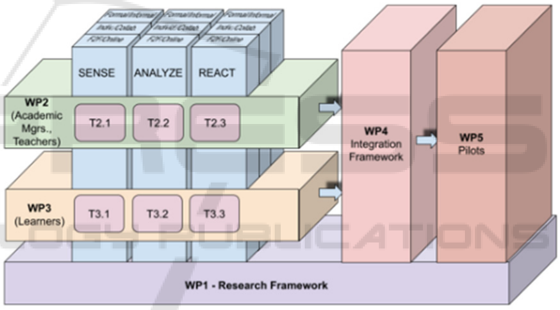

Figure 3: Structure of WPs of the project.

Figure 2 shows the adaptation of the DSRM methodology for the specific case of the GENIE Learn project considering the WPs and tasks to be addressed. The project is structured into three main iterations, one for each year. Each iteration follows the three DSRM phases, outlining the tasks relevant to each phase. Two phases are particularly crucial: the initial phase, where the objectives are set at the start of the project and for each cycle; and the evaluation phase, especially important in the second and third cycles when proposals are assessed.

5 CONCLUSIONS

The GENIE Learn project aims to enhance SLEs for HL by incorporating GenAI tools in a manner that aligns with the values and preferences of human stakeholders. The current problems presented by the research context have been analyzed, considering the human, technological and pedagogical contexts, the affordances of GenAI, but also the risks and challenges of this technology. The expected contributions of GENIE Learn include: 1) a Research Framework with the pedagogical model, human values, and scenarios for HL supported by GenAI-enhanced SLEs; 2) GenAI-enhanced approaches for management and design of learning; 3) GenAI-enhanced approaches for learning; 4) a Technological framework as an integrated infrastructure; and 5) Pilot experiences co-designed with educational stakeholders. The complexity of the research context can only be effectively addressed through projects of this nature, supported by an interdisciplinary approach.

ACKNOWLEDGEMENTS

This work was supported by grants PID2023-146692OB-C31, PID2023-146692OB-C32 and PID2023-146692OB-C33 funded by MICIU/AEI/10. 13039/501100011033 and by ERDF/EU, project GENIELearn. DHL (Serra Húnter) acknowledges support by ICREA Academia.

REFERENCES

Akgun, S. & Greenhow, C. (2021). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, 1-10.

Albó, L., et al. (2022). Knowledge-based design analytics for authoring courses with smart learning content. *International Journal of Artificial Intelligence in Education, 32*(1), 4-27.

Alonso-Prieto, V., et al. (2023). Studying the Impact of Orchestrating Intelligent Technologies in Hybrid Learning Spaces for Teacher Agency. In *European Conference on Technology Enhanced Learning (EC-TEL)* (pp. 636-641).

Amarasinghe, I., et al. (2021). Deconstructing orchestration load: comparing teacher support through mirroring and guiding. *International Journal of Computer-Supported Collaborative Learning, 16*(3), 307-338.

Amarasinghe, I., et al. (2022). Teacher orchestration load: What is it and how can we lower the burden? *Digital Promise and the International Society of the Learning Sciences* (pp. 1–13).

Amarasinghe, I., et al. (2023a). Orchestrating Hybrid Learning Scenarios: Challenges and Opportunities. In *ISLS Annual Meeting 2023* (p. 72).

Amarasinghe, I., et al. (2023b). Generative Pre-trained Transformers for Coding Text Data? An Analysis with Classroom Orchestration Data. In *European Conference on Technology Enhanced Learning (EC-TEL)* (pp. 32-43).

Askell, A., et al. (2021). A general language assistant as a laboratory for alignment. *arXiv preprint arXiv:2112.00861*.

Baidoo-Anu, D. & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI, 7*(1), 52-62.

Brown, T., et al. (2020). Language models are few-shot learners. *Advances in neural information processing systems, 33*, 1877-1901.

Bubeck, S., et al. (2023). Sparks of artificial general intelligence: Early experiments with gpt-4. *arXiv preprint arXiv:2303.12712.*

Buckingham Shum, S., et al. (2019a). Learning analytics and AI: Politics, pedagogy and practices. *British journal of educational technology, 50*(6), 2785-2793.

Buckingham Shum, S., et al. (2019b). Human-centred learning analytics. *Journal of Learning Analytics, 6*(2), 1-9.

Calvera-Isabal, M., et al. (2023). How to automate the extraction and analysis of information for educational purposes. *Comunicar, 31*(74), 23-35.

Cohen, A., et al. (2020). Hybrid learning spaces—Design, data, didactics. *British journal of educational technology, 51*(4), 1039-1044.

Crespi, F., et al. (2022). Estimating orchestration load in CSCL situations using EDA. In *2022 International Conference on Advanced Learning Technologies (ICALT),* (pp. 128-132).

Dagnino, F. M., et al. (2018). Exploring teachers' needs and the existing barriers to the adoption of Learning Design methods and tools: A literature survey. *British journal of educational technology, 49*(6), 998-1013.

Dai, W., et al. (2023). Can large language models provide feedback to students? A case study on ChatGPT. In *2023 International Conference on Advanced Learning Technologies (ICALT 2023)* (pp. 323-325).

Delgado Kloos, C., et al. (2018). SmartLET: Learning analytics to enhance the design and orchestration in scalable, IoT-enriched, and ubiquitous SLEs. In *Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM* 2018) (pp. 648-653).

Delgado Kloos, C., et al. (2022). H2O Learn-Hybrid and Human-Oriented Learning: Trustworthy and Human-Centered Learning Analytics for Hybrid Education. In *2022 IEEE global engineering education conference (EDUCON)* (pp. 94-101).

Demetriadis, S. & Dimitriadis, Y. (2023). Conversational Agents and Language Models that Learn from Human Dialogues to Support Design Thinking. In *International Conference on Intelligent Tutoring Systems (ITS)* (pp. 691-700).

European Commission (2023), A European approach to artificial intelligence. https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence

Evmenova, A. (2018). Preparing teachers to use universal design for learning to support diverse learners. *Journal of Online Learning Research, 4*(2), 147-171.

Friedman, B., et al. (2017). A survey of value sensitive design methods. *Foundations and Trends in Human–Computer Interaction, 11*(2), 63-125.

Gamieldien, Y. (2023). *Innovating the Study of Self-Regulated Learning: An Exploration through NLP, Generative AI, and LLMs* (Doctoral dissertation, Virginia Tech).

Gargiulo, R. M. & Metcalf, D. (2010). *Teaching in today's inclusive classrooms: A universal design for learning approach*. Cengage Learning.

Gil, E., et al. (Eds.). (2022). *Hybrid learning spaces*. Springer.

Gros, B. (2016). The design of smart educational environments. *Smart Learning Environments, 3*, 1-11.

Garg, S., & Sharma, S. (2020). Impact of artificial intelligence in special need education to promote inclusive pedagogy. *International Journal of Information and Education Technology, 10*(7), 523-527.

Gutiérrez-Páez, N. F., et al. (2023). A study of motivations, behavior, and contributions quality in online communities of teachers: A data analytics approach. *Computers & Education, 201*, 104829.

Hadwin, A., et al. (2017). Self-regulation, co-regulation, and shared regulation in collaborative learning environments. In *Handbook of self-regulation of learning and performance* (pp. 83-106).

Hakami, L., et al. (2022). Exploring Teacher's Orchestration Actions in Online and In-Class Computer-Supported Collaborative Learning. In *European Conference on Technology Enhanced Learning (EC-TEL)* (pp. 521-527).

Hernández-Leo, D., et al. (2019). Analytics for learning design: A layered framework and tools. *British journal of educational technology, 50*(1), 139-152.

Hernández-Leo, D. (2022). Directions for the responsible design and use of AI by children and their communities: Examples in the field of Education, In *Artificial Intelligence and the Rights of the Child: Towards an Integrated Agenda for Research and Policy*, EUR 31048 EN, pp. 73-74.

Hernández-Leo, D., et al. (2023a). Editorial: Technologies for Data-Driven Interventions in Smart Learning Environments. *IEEE Transactions on Learning Technologies, 16*(3), 378-381.

Hernández-Leo, D. (2023b). ChatGPT and Generative AI in Higher Education: user-centered perspectives and implications for learning analytics, In *LASI Spain*, Madrid.

Hilli, C., et al. (2019). Designing hybrid learning spaces in higher education. *Dansk Universitetspædagogisk Tidsskrift, 15*(27), 66-82

HLEG-AI (High-Level Expert Group on Artificial Intelligence) (2019), Ethics Guidelines for Trustworthy AI. https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines/1

ISO 2013. International Standardization Organization (2013), 16290:2013 Space systems — Definition of the Technology Readiness Levels (TRLs) and their criteria of assessment. https://www.iso.org/standard/56064.html

Järvelä, S., et al. (2023). Advancing SRL Research with Artificial Intelligence. *Computers in Human Behavior*, *147*, 107847.

Johnson, R. B., et al. (2007). Toward a definition of mixed methods research. *Journal of mixed methods research*, *1*(2), 112-133.

Jovanovic, M. & Campbell, M. (2022). Generative artificial intelligence: Trends and prospects. *Computer*, *55*(10), 107-112.

Kalo, J. C., et al. (2020). KnowlyBERT-Hybrid query answering over language models and knowledge graphs. In *19th International Semantic Web Conference (ISWC)* (pp. 294-310).

Kasneci, E., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, 103, 102274.

Koyuturk, C., et al. (2023). Developing Effective Educational Chatbots with ChatGPT: Insights from Preliminary Tests in a Case Study on Social Media Literacy, In *31st International Conference on Computers in Education (ICCE)* (pp. 1-58).

Lim, W. M., et al. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *International journal of management education, 21*(2), 100790.

Long, P. & Siemens, G. (2011), What is Learning Analytics? In *1st International Conference Learning Analytics and Knowledge (LAK)*. ACM.

Michos, K. & Hernández-Leo, D. (2020). CIDA: A collective inquiry framework to study and support teachers as designers in technological environments. *Computers & Education, 143*, 103679.

Misiejuk, K., et al. (2023). Changes in online course designs: Before, during, and after the pandemic. In *Frontiers in Education (FIE)* (Vol. 7, p. 996006).

Mizumoto, A. (2023). Data-driven Learning Meets Generative AI: Introducing the Framework of Metacognitive Resource Use. Applied Corpus *Linguistics, 3*(3), 100074.

Molenaar, I. (2022). The concept of hybrid human-AI regulation: Exemplifying how to support young learners' self-regulated learning. *Computers and Education: Artificial Intelligence, 3*, 100070.

Ognibene, D., et al. (2023). Challenging social media threats using collective well-being-aware recommendation algorithms and an educational virtual companion. *Frontiers in Artificial Intelligence, 5*, 654930.

Ortega-Arranz, A., et al. (2022). e-FeeD4Mi: Automating Tailored LA-Informed Feedback in Virtual Learning Environments. In *European Conference on Technology Enhanced Learning (EC-TEL)* (pp. 477-484).

Ortiz Beltrán, A., et al. (2023). Surviving and thriving: How changes in teaching modalities influenced student

satisfaction before, during and after COVID-19. *Australasian Journal of Educational Technology*, *39*(6), 72–88.
Ouyang, L., et al. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems, 35*, 27730-27744.
Peffers, K., et al. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of management information systems, 24*(3), 45–77.
Pereira, F. D., et al. (2023). Towards human-AI collaboration: a recommender system to support CS1 instructors to select problems for assignments and exams. *IEEE Transactions on Learning Technologies, 16*(3), 457:472.
Pishtari, G., et al. (2024). Mirror on the wall, what is missing in my pedagogical goals? The impact of an AI-driven feedback system on the quality of teacher learning designs. In *14th Learning Analytics and Knowledge Conference (LAK)* (pp. 145-156).
Pryzant, R., et al. (2023). Automatic prompt optimization with "gradient descent" and beam search. *arXiv preprint arXiv:2305.03495*.
Raes, A. (2022). Exploring student and teacher experiences in hybrid learning environments: Does presence matter? *Postdigital Science and Education, 4*(1), 138–159.
Rose, D. H. & Meyer, A. (2002). *Teaching every student in the digital age: Universal design for learning*. Association for Supervision and Curriculum Development.
Rudolph, J., et al. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?. *Journal of applied learning and teaching*, *6*(1), 342-363.
Ruiz-Calleja, A., et al. (2021). Supporting contextualized learning with linked open data. *Journal of Web Semantics, 70*, 100657.
Sánchez-Reina, R., et al. (2023). "Shall we rely on bots?" Students' adherence to the integration of ChatGPT in the classroom. In *International Conference on Higher Education Learning Methodologies and Technologies Online* (HELMeT) (pp. 128-130).
Serrano-Iglesias, S., et al. (2021). Demonstration of SCARLETT: A Smart Learning Environment to Support Learners Across Formal and Informal Contexts. In EC-TEL 2021 (pp. 404-408).
Serrano-Iglesias, S., et al. (2023). Exploring Barriers for the Acceptance of Smart Learning Environment Automated Support. In *European Conference on Technology Enhanced Learning (EC-TEL)* (pp. 636-641).
Siemens, G., et al. (2022). Human and artificial cognition. *Computers and Education: Artificial Intelligence, 3*, 100107.
Sharples, M. (2023). Towards social generative AI for education: theory, practices and ethics, Learning: *Research and Practice, 9*(2), 159-167.
Susnjak, T. (2023). Beyond Predictive Learning Analytics Modelling and onto Explainable Artificial Intelligence with Prescriptive Analytics and ChatGPT, *International Journal of Artificial Intelligence in Education*, *34*(2), 452-482.
Suraworachet, W., et al. (2021). Examining the relationship between reflective writing behaviour and self-regulated learning competence: A time-series analysis. In *European Conference on Technology Enhanced Learning (EC-TEL)* (pp. 163-177).
Tabuenca, B., et al. (2021). Affordances and core functions of smart learning environments: A systematic literature review. *IEEE Transactions on Learning Technologies*, *14*(2), 129-145.
Tabuenca, B., et al. (2024). Greening smart learning environments with Artificial Intelligence of Things. *Internet of Things, 25*, 101051.
Theophilou, E., et al. (2023a). Gender-based learning and behavioural differences in an educational social media platform. *Journal of Computer Assisted Learning*. Early view.
Theophilou, E., et al. (2023b). Students' generated text quality in learning environment: effects of pre-collaboration, individual, and chat-interface submissions. In *International Conference on Collaboration Technologies and Social Computing (CollabTech 2023)* (pp. 101-114).
Theophilou, E., (2023c). Learning to Prompt in the Classroom to Understand AI Limits: A pilot study, In *International Conference of the Italian Association for Artificial Intelligence*, (pp. 481-496).
Toetenel, L., & Rienties, B. (2016). Analysing 157 learning designs using learning analytic approaches to evaluate the impact of pedagogical decision making. *British journal of educational technology, 47*(5), 981-992.
Topali, P., et al. (2024). Unlock the Feedback Potential: Scaling Effective Teacher-Led Interventions in Massive Educational Contexts. In *Innovating Assessment and Feedback Design in Teacher Education* (pp. 1-19). Routledge.
UNESCO (2023). Guidance for generative AI in education and research. https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research
Villa-Torrano, C. et al., (2023). Towards Visible Socially-Shared Regulation of Learning: Exploring the Role of Learning Design. In *16th International Conference on Computer-Supported Collaborative Learning (CSCL 2023)* (pp. 289-292).
Wasson, B. & Kirschner, P. A. (2020). Learning design: Eu approaches. *TechTrends, 64*, 815-827.
Yan, L., et al. (2023). Practical and ethical challenges of large language models in education: A systematic literature review. *British Journal of Educational Technology*, *55*(1), 90-112.