A Learning Analytics Research Agenda: Addressing Measurement and Evaluation Challenges in GenAI-Enhanced Smart Learning Environments

Pedro Manuel Moreno-Marcos[1,*], Pedro J. Muñoz-Merino[1] and Carlos Delgado Kloos[1]

[1]*Department of Telematic Engineering, Universidad Carlos III de Madrid, 28911 Leganés (Madrid), Spain, ROR: 03ths8210*

Abstract

Generative Artificial Intelligence (GenAI) is widely used in education and while there are reports about their potential use to enhance Smart Learning Environments (SLEs), there is more need of research around this area. Particularly, this paper identifies three main challenges where researchers should put the focus with GenAI: (1) extract information from non-structured data, (2) monitor students' interactions with GenAI, and (3) evaluate the impact of materials/activities generated by AI. In addition, some preliminary works are presented where researchers have worked around these ideas to detect errors in language and generate analytics, detect urgent messages in a forum, analyze students' activity with GenAI when solving questions, a chatbot to support learning, a tool to generate questions, and a GenAI-based tool to learn cybersecurity. These works show the potential of GenAI in SLEs, and invite researchers to further discuss and research in this area.

Keywords

Generative Artificial Intelligence, Smart Learning Environments, Learning Analytics, Students' interactions, Learning tools,

1. Introduction

The use of Generative Artificial Inteligence (GenAI) is very common among students, with some reports claiming that nearly 90% of the students use it [1]. This increasing use of GenAI can be beneficial and different stakeholders can improve the efficiency of their processes based on that [2]. For example, teachers can use it to generate materials or support class orchestrations; students to solve doubts when learning or receive useful feedback. Moreover, GenAI may open new opportunities to innovate and enhance existing learning environments.

Particularly, Smart Learning Environments (SLEs) are characterized by the use of technology to transform education, and they have three main core functions: sense, analyze and react [3]. In the past years, there have been extensive research on how learners behave in different SLEs [4], including hybrid scenarios (e.g., [5]). However, these SLEs did not include GenAI. Therefore, it is necessary an analysis on how students' behave when they interact with GenAI or with activities designed with GenAI. In addition, given that new learning tasks might be defined with the use of GenAI, there is lack of research on how GenAI can enhance the materials/activities.

In this direction, this position paper aims to provide some ideas on the main research needs around the use of GenAI in SLEs, and particularly in the sense and analyze core functions of SLEs related to learning analytics (LA) research about students' behaviours while engaging with GenAI.

LA4GenAISLE'26: Learning Analytics for Human-Centered Generative AI-Enhanced Smart Learning Environments Workshop, April 28, 2026, Bergen, Norway

*Corresponding author.

pemoreno@it.uc3m.es (P. M. Moreno-Marcos); pedmume@it.uc3m.es (P. J. Muñoz-Merino); cdk@it.uc3m.es (C. Delgado Kloos)

https://www.it.uc3m.es/pemoreno/ (P. M. Moreno-Marcos); http://www.it.uc3m.es/~pedmume/ (P. J. Muñoz-Merino); http://www.it.uc3m.es/cdk/ (C. Delgado Kloos)

0000-0003-0835-1414 (P. M. Moreno-Marcos); 0000-0002-2552-4674 (P. J. Muñoz-Merino); 0000-0003-4093-3705 (C. Delgado Kloos)

published 2026-09-26

2. Challenges of LA in GenAI-enhanced SLEs

The introduction of GenAI in SLEs has produced significant changes in the way courses are taught and how students learn. As suggested in [6], GenAI can serve to mitigate some of the existing challenges in SLEs, although they also face some limitations. In connection with LA, it also affects students' behaviors and opens the opportunity for new research in LA. Particularly, based on an analysis of current literature and the authors' experience in developing SLEs, this work identifies three main challenges:

- Extract information from non-structured data. The analysis of SLEs can be used to generate students' models (analyze core function) that could be potentially used to provide feedback (react core function) [7]. However, these models are developed using mainly quantitative data and the non-structure data that has been used has been very limited (e.g., length of documents, number of editions, etc. [8]). Thus, the use of GenAI opens the opportunity to collect new variables (sense core function). For example, the content of texts produced by the students could be further analyzed to get more insights about the writing style, register, perceived knowledge, etc. that cannot be easily obtained from the data. While traditional Natural Language Processing (NLP) methods have often focused on more limited or surface-level indicators, GenAI allows for a deeper analysis of these complex text structures. This provides a clearer understanding of students' work that goes beyond what was previously possible with traditional techniques. Moreover, this could also be applied to programming education, and programming code could also be analyzed to obtain more detailed information about it.
- Monitor students' interactions with GenAI. One typical research line in LA is the analysis of students' behaviours when interacting with learning platforms, and this has often led to research on predictive models [4]. These analyses are based on analysis of how students interact with the platform, their interactions with videos, exercises, class attendance, etc. [9]. However, the introduction of GenAI in SLEs poses new questions about how students interact with GenAI (e.g., frequency, reasons to use GenAI, moments when GenAI is used more frequently, etc.). Given that GenAI contexts are new, a lot of reserach can be done to characterize new behaviors with GenAI in SLEs, and to enhance or modify current predicitve models to reflect the new possibilities that GenAI is opening for research.
- Evaluate the impact of materials/activities generated by AI. GenAI can also be used to support the content generation in the different phases: design of learning objectives, learning materials, support class orchestration, design evaluations, and design certifications [10]. While the potential is widely acknowledged and there are some preliminary studies around this (e.g., [11]), there is a lot of room for further research in this area, and many new applications should be evaluated to better understand how tools generated with AI should be in the different aforementioned phases to maximize learning in SLEs. In this regard, existing evaluation frameworks often fail to capture the specific nuances of GenAI-based interventions. Identifying new indicators that go beyond general performance or satisfaction is essential to better understand the actual influence of these tools.

3. Research lines of LA in GenAI-enhanced SLEs

This section presents some research lines in GenAI-enhanced SLEs that aim to address the aforementioned challenges. They represent preliminary findings or on-going projects or work in these areas. While these cases address specific challenges independently, they are envisioned as complementary components of a holistic LA framework. For instance, text-based feature extraction (Section 3.1) can feed the behavioral models discussed in Section 3.2, which in turn inform the evaluation of the tools described in Section 3.3.

3.1. Extraction of text-based features using GenAI

In relation to the first challenge, GenAI offers new possibilities to analyze text to extract new features. In this line, PALABRIA project [12] aims to support learners in their language use by detecting errors that students often have in their written language. In a first analysis, they focused on the "impersonal you" phenomenon in Spanish [13], which occurs when directly addressing the person in written language (as opposed to spoken language in which this is natural and appropriate). For example, "You must read this paper before discussing new ideas" should be rewritten as "It is mandatory to read this paper before discussing new ideas" to avoid using the second person. In order to detect this error, "Mistral-7B-Instruct-v0.1" model and results of the evaluation of how this model was detected and corrected were measured with GLEU metric (Generalized Language Understanding Evaluation). Their preliminary analyses showed a GLEU of 0.85 when detecting this error. The implication of this research means that it is possible to detect more complext language errors using GenAI, and this allows generating new features about students' learning based on their text. This preliminary analysis is based on one specific phenomenon, but more complex language structure or grammatical errors could be incorporated to have a linguistic profile of the student. While the project is focused on native speakers who may have errors when writing reports, this may also enable LA analyses to language learners.

Another case study in this line is the analysis of forum data in Massive Open Online Courses (MOOCs). Although there have been several works to analyze forum data based on the quantity and quality of their messages (e.g., [14]), the analyzes are mainly based on Natural Language Processing techniques (NLP), which focus on the frequency of certain words. However, they have limitations in the way text is analyzed because the context is not often considered. In this direction, GenAI opens new opportunities to gather new insights on forum activity in the MOOC. For example, Méndez-Peñalosa Portillo [15] focused on the detection of urgent messages in a MOOC. He applied a heuristic method and manual analysis to classify 7,633 messages in urgent and non-urgent. Then, he splitted the dataset in 80% for training+validation and 20% for test, and the 80% in 80% for training and 20% for validation. For the training, he used the heuristic value with metadata and the result of applying TF-IDF (Term Frequency – Inverse Document Frequency) with Regression (RG) and Random Forest (RF). In order to evaluate the potential of Large Language Models (LLMs), he also used both algorithms with the embeddings produced by the "Sentence-Transformers all-MiniLM-L6-v2" model instead of the TF-IDF. His results show that embeddings improved the detection of urgent messages (recall improves from 0.93 to 0.94 in RG and from 0.92 to 0.94 in RF), but the precision dropped (from 0.78 to 0.75, and 0.83 to 0.70, respectively) in both cases as there were more false positives. This highlights a context-precision trade-off for LLM embeddings in learning analytics, and it supposes that the LLM approach might not always be the best, although further research is needed to better leverage the potential.

3.2. Analysis of students' behaviours when using GenAI

The analysis of new behaviors when using GenAI can be divided in two possibilities: (1) engagement with learning tools that use GenAI internally, or (2) use of GenAI to solve learning activities that could be done without using GenAI.

In the first case, it is possible to also refer to the blended project and the tool to detect language errors. When students engage with these tool, which internally uses GenAI, it is possible to collect new indicators, e.g., frequency of errors in one text submission, frequency of errors considering several text submissions, number of errors the students fix after using the tool, index of improvement of errors in consecutive submissions, etc. While this indicators are not specific about the use of GenAI, these indicators can be obtained because of the use of GenAI, so GenAI enables the analysis of new behaviors to enhance LA models.

The second approach is to analyze the interactions that students have with GenAI when solving tasks that are could also be solved without GenAI. In this line, Moreno-Marcos [16] presented a poster of a teaching innovation experience where students had to solve questions about C programming and they could use hints generated by AI or prompts. Results showed that 70% used hints to solve the

exercises and 80% used additional prompts. Moreover, 70% reported that they feel they have enough capabilities to make prompt, 65% indicated that the hints provided by the AI were useful/very useful and 60% admitted that hints helped their reflection.

3.3. Development of tools with GenAI to support learning in SLEs

The last research line is about the development of new tools that involve GenAI to enhance learning in SLEs. In this line, it is possible to evaluate the impact of new tools to measure their impact. Coming back to the PALABRIA project, an example of tool to detect language errors with AI, generate indicators about how students interact with the tool to use LA, and provide feedback to learners also reflects the three main core functions of SLE: (1) sense, with the generation of indicators; analyze, with the analysis of those indicators; and (3) react, with the provision of feedback.

Another example is the chatbot presented in [17]. This chatbot was used in a course about Java programming and it used the RAG (Retrieval-Augmented Generation) technique to provide course-specific answers by using the videos, quizzes programming exercises and other materials. They collected 1,059 messages sent by the students, and their results showed that the main use of the chatbot was to clarify concepts and explain code snippets. In this case, this tools allows enhancing the learning experience, while also promoting LA with the analysis of messages.

Other works have focused on question generation based on GenAI. For example, authors in [18] developed a tool to generate multiple-choice and short questions using Gemini. In their analysis, they found that the quality of multiple-choice questions was higher in terms of appropriateness for the course, exams and correctness of questions. In addition, it was found that while questions could be related to the course, their suitability for exams could be improvable. They also found that the inclusion of the syllabus in the prompts barely improved the appropiateness of the questions and answers but significantly improved the suitability for exams. This also proves that teachers can benefit from GenAI with the support of question generation.

This topic of question generation with GenAI was also addressed by Pérez López [19]. He developed CyberX, a mobile application to support learning of the cybersecurity basics. He developed several functionalities. The first one includes simulations, where users could practice with real scenarios to detect phising, scams, fraudulent websites and insecure passwords. For that functionality ChatGPT is used to generate new scenarios so that users can practice. The second functionality includes questionnaires, with multiple-choice questions and the use of ChatGPT to increase the pool of questions. The third functionality is related to the react function of SLEs, where some pieces of advice are shown depending on the user activity. This is implemented by combining neural networks to detect the areas where the user needs to improve and ChatGPT to generate feedback. The fourth functionality includes a chatbot so learners can ask questions, the fifth one includes recent news about the topics and the last one provides some charts and statistics about the performance of the user (LA). Therefore, the tool combines LA and GenAI to enhance the learning experience of cybersecurity.

4. Conclusions

This paper has illustrated the potential of GenAI to enhance SLEs, and it has provided several challenges and research lines in which researchers could work further. Particularly, it has been identified that more research is needed to extract information from non-structured data, monitor students' interactions with GenAI and evaluate the impact of tools which involve GenAI. In this direction, some preliminary works have been presented in which text is analyzed to detect language errors and obtain LA indicators, detect urgent messages in MOOCs, analyze how students solve programming questions with GenAI, interactions with a chatbot about programming, a tool to generate questions and a tool to enhance learning about cybersecurity.

The preliminary cases presented in this work are not isolated efforts but represent stages of an integrated SLE cycle. For instance, the analysis of language errors and LA indicators (as implemented in the PALABRIA tool) demonstrates how extracting insights from unstructured text (Challenge 1) is the

prerequisite for meaningful behavioral monitoring (Challenge 2). An integrated scenario would involve using these GenAI-driven linguistic analytics to provide personalized scaffolding, while simultaneously measuring the students' cognitive load and the accuracy of their progress (Challenge 3). Future research should focus on closing this loop to provide a holistic support system for learners.

However, several difficulties remain when addressing these challenges, such as the high cost of manually validating GenAI outputs to ensure pedagogical reliability and the lack of standardized metrics to measure the quality of student-AI interactions. As concrete next steps, the research community should prioritize the creation of shared evaluation frameworks and the development of real-time analytics that can be easily integrated into existing SLEs. Addressing these technical and methodological hurdles is essential to move from experimental pilots to scalable, evidence-based educational tools.

Ultimately, the works mentioned in this paper show the potential of GenAI, and they invite researchers to work around these areas to improve SLEs, with a crucial focus on defining new LA metrics and models designed to effectively measure the impact of GenAI-generated outputs and student-GenAI interactions.

Acknowledgments

This work was supported by FEDER / Ministerio de Ciencia, Innovación y Universidades - Agencia Estatal de Investigación through the grant PID2023-146692OB-C31 (GENIE Learn project) funded by MICIU/AEI/10.13039/501100011033 and by ERDF/UE. This was also funded by Comunidad de Madrid through the grant agreement for the promotion of research and technology transfer at Universidad Carlos III de Madrid (PALABRIA-CM-UC3M). Moreover, it received support from the UNESCO Chair of "Scalable Digital Education for All" at UC3M and by the grant RED2022-134284-T funded by MICIU/AEI/10.13039/501100011033.

Declaration on Generative AI

During the preparation of this work, the author(s) used Gemini Flash 2.5in order to: Grammar and spelling check. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] Fundación CYD, Uso y percepción de la IA en el entorno universitario, Technical Report, Fundación CYD, 2025. URL: https://www.fundacioncyd.org/wp-content/uploads/2025/05/PUBLICACION-Inteligencia-Artificial-y-universidad-8MAI.pdf, disponible en línea.

[2] C. Delgado Kloos, C. Alario-Hoyos, I. Estévez-Ayres, P. Callejo-Pinardo, M. A. Hombrados-Herrera, P. J. Muñoz-Merino, P. M. Moreno-Marcos, M. Muñoz-Organero, M. B. Ibáñez, How can generative ai support education?, in: 2024 IEEE Global Engineering Education Conference (EDUCON), IEEE, 2024, pp. 1–7. doi:10.1109/EDUCON60312.2024.10578716.

[3] B. Tabuenca, S. Serrano-Iglesias, A. C. Martín, C. Villa-Torrano, Y. Dimitriadis, J. I. Asensio-Pérez, C. Alario-Hoyos, E. Gómez-Sánchez, M. L. Bote-Lorenzo, A. Martinez-Mones, et al., Affordances and core functions of smart learning environments: A systematic literature review, IEEE Transactions on Learning Technologies 14 (2021) 129–145. doi:10.1109/TLT.2021.3067946.

[4] N. Sghir, A. Adadi, M. Lahmer, Recent advances in predictive learning analytics: A decade systematic review (2012–2022), Education and information technologies 28 (2023) 8299–8333. doi:10.1007/s10639-022-11536-0.

[5] A. C. Martín, C. Alario-Hoyos, P. M. Moreno-Marcos, C. Delgado Kloos, Student and teacher impact on the use of telepresence classrooms, Education and Information Technologies 29 (2024) 21355–21378. doi:10.1007/s10639-024-12618-x.

[6] C. Delgado Kloos, J. I. Asensio Pérez, D. Hernández Leo, P. M. Moreno-Marcos, M. L. Bote Lorenzo, P. Santos, C. Alario-Hoyos, Y. Dimitriadis, B. Tabuenca, GENIE learn: human-centered generative AI-enhanced smart learning environments, in: 17th International Conference on Computer Supported Education - Volume 2: CSEDU, Scitepress, 2025, pp. 15–26. doi:10.5220/0013076000003932.

[7] S. Serrano-Iglesias, E. Gómez-Sánchez, M. L. Bote-Lorenzo, J. I. Asensio-Pérez, A. Ruiz-Calleja, G. Vega-Gorgojo, Demonstration of scarlett: a smart learning environment to support learners across formal and informal contexts, in: European Conference on Technology Enhanced Learning, Springer, 2021, pp. 404–408. doi:10.1007/978-3-030-86436-1_44.

[8] W. Suraworachet, C. Villa-Torrano, Q. Zhou, J. I. Asensio-Pérez, Y. Dimitriadis, M. Cukurova, Examining the relationship between reflective writing behaviour and self-regulated learning competence: A time-series analysis, in: European Conference on Technology Enhanced Learning, Springer, 2021, pp. 163–177. doi:10.1007/978-3-030-86436-1_13.

[9] P. M. Moreno-Marcos, C. G. Antolín, C. Alario-Hoyos, P. J. Muñoz-Merino, C. Delgado Kloos, Integration of multiple sources to anticipate student performance using learning analytics, in: 2025 IEEE International Conference on Advanced Learning Technologies (ICALT), IEEE, 2025, pp. 117–119. doi:10.1109/ICALT64023.2025.00039.

[10] C. Delgado Kloos, C. Alario-Hoyos, R. Kemcha, P. M. Moreno-Marcos, I. Estevez-Ayres, P. Callejo-Pinardo, P. J. Munoz-Merino, M.-B. Ibanez, M. Munoz-Organero, How challenges become opportunities: Micro-credentials and artificial intelligence, in: 2025 IEEE Global Engineering Education Conference (EDUCON), IEEE, 2025, pp. 1–10. doi:10.1109/EDUCON62633.2025.11016509.

[11] A. Mikroyannidis, N. Sharma, A. Ekuban, J. Domingue, Using generative ai and chatgpt for improving the production of distance learning materials, in: 2024 IEEE International Conference on Advanced Learning Technologies (ICALT), IEEE, 2024, pp. 188–192. doi:10.1109/ICALT61570.2024.00061.

[12] P. M. Moreno-Marcos, M. Serrano-Marín, P. J. Muñoz-Merino, N. Centeno Alejandre, R. Fernández Castillejos, M. V. Pavón Lucero, C. Delgado Kloos, R. Conde Rubio, A. Suárez Hernández, PALABRIA: Language Learning Enhancement through Linguistic Rule-Based Analytics and Artificial Intelligence, in: XVII Jornadas de Ingeniería Telemática, Sociedad Científica de Ingeniería Telemática, 2025, pp. 1–8.

[13] N. Centeno Alejandre, P. M. Moreno-Marcos, M. Serrano-Marín, Mejora de la escritura en español mediante inteligencia artificial y reglas explícitas en contextos académicos, in: VII Congreso de la Sociedad de Humanidades Digitales Hispánicas, Universitat de València, 2025, pp. 1–8. doi:10.7203/PUV-OA-9788491338307.

[14] R. L. Pérez-Nicolás, C. Alario-Hoyos, I. Estévez-Ayres, P. M. Moreno-Marcos, P. J. Muñoz-Merino, C. Delgado Kloos, Evaluation of an algorithm for automatic grading of forum messages in mooc discussion forums, Sustainability 13 (2021) 9364. doi:10.3390/su13169364.

[15] Á. Méndez-Peñalosa Portillo, Sistema de Clasificación Automática de Mensajes en Foros de MOOCs, Bachelor's thesis, Universidad Carlos III de Madrid, Madrid, España, 2025.

[16] P. M. Moreno Marcos, Empoderando al alumnado en el buen uso de la inteligencia artificial generativa y el valor de los datos abiertos con soporte de analítica de aprendizaje, 2025. URL: https://www.uc3m.es/uc3mdigital/media/uc3mdigital/doc/archivo/doc_posteres-2025---14/014_moreno_marcos_pedro-manuel_memoria_innov2425_anexoii.pdf, póster/Memoria del Proyecto de Innovación Docente 2025, UC3M.

[17] C. Alario-Hoyos, R. Kemcha, C. Delgado Kloos, P. Callejo, I. Estévez-Ayres, D. Santín-Cristóbal, F. Cruz-Argudo, J. L. López-Sánchez, Tailoring your code companion: Leveraging llms and rag to develop a chatbot to support students in a programming course, in: 2024 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE), IEEE, 2024, pp. 1–8. doi:10.1109/TALE62452.2024.10834365.

[18] P. M. Moreno-Marcos, J. Gil Santiuste, P. J. Muñoz-Merino, C. Alario-Hoyos, C. Delgado Kloos, Analysis and evaluation of gemini for question generation in two engineering courses, in: Proceedings of the 2025 IEEE Frontiers in Education Conference, Nashville, Tennessee, United States,

November 2025 (FIE '25), IEEE, Nashville, Tennessee, United States, 2025. doi:10.1109/FIE63693.2025.11328589.

[19] A. A. Pérez López, Aplicación web para la concienciación de la población sobre la ciberseguridad, Master's thesis, Universidad Carlos III de Madrid, Madrid, España, 2025.