

Integrating Game Learning Analytics and Local LLMs for Automated Interpretation of Player Behaviour in Games

Guilherme Silva, Íris Campanella, Júlia Pessoa, Karla Vittori, André L. Brandão

CMCC – Universidade Federal do ABC (UFABC)
Santo André – SP – Brasil

{guilherme.hugo, iris.campanella, julia.pessoa}@aluno.ufabc.edu.br,
{karla.vittori, denise.goya, andre.brandao}@ufabc.edu.br

Abstract. Introduction: Digital games are valuable educational tools that generate interaction data for assessing student learning. However, analysing large volumes of gameplay logs remains challenging for educators due to their highly technical nature. **Objective:** This paper proposes an architecture that integrates Game Learning Analytics (GLA) with Large Language Models (LLMs) to automatically interpret player interactions in a serious game developed with Unity. **Methodology:** Player actions are captured using the GLA framework, generating structured logs that are analysed by a locally deployed LLM to produce natural-language reports. **Preliminary Results:** Initial tests show that the system can interpret player behaviour, identifying gameplay patterns, engagement characteristics, and potential conceptual difficulties without manual inspection of raw logs.

Keywords Serious Games, Game Learning Analytics, Large Language Models.

1. Introduction

Serious games have emerged as promising tools for education, combining engagement with intentional learning objectives [Connolly et al. 2012, Zeng et al. 2020, Yu et al. 2021]. However, simply presenting content in a game format is insufficient to ensure learning outcomes; the quality of feedback and the ability to monitor student progress play critical roles in fostering achievement.

In this context, Game Learning Analytics (GLA) has gained attention as a framework for collecting and interpreting data generated by players during gameplay [Calvo-Morata et al. 2019]. By capturing fine-grained interaction traces, GLA enables educators and researchers to identify patterns of engagement, diagnose learning difficulties, and inform pedagogical decisions. In practice, however, the interpretation of these data remains challenging. Traditional approaches rely on Visual Analysis [Silva et al. 2022], which often requires specialised expertise and does not scale well to the large volumes of logs produced by modern digital games.

Recent advances in Large Language Models (LLMs) offer a promising alternative for transforming structured interaction data into interpretable insights. By generating coherent natural-language explanations, LLMs can reduce the cognitive burden associated with manual log inspection and make analytics outputs accessible to non-technical stakeholders [Bastos et al. 2025].

The main contribution of this study is an end-to-end architecture that integrates xAPI-based Game Learning Analytics, structured gameplay log storage, privacy-preserving local LLM execution, and automated generation of explainable pedagogical

reports for serious games developed with Unity [Unity Technologies 2025]. Gameplay data are captured using the xAPI standard via XASU and subsequently processed by the Llama 3.2 model running through Ollama to produce evidence-based pedagogical insights while preserving data privacy.

The remainder of this paper is structured as follows: Section 2 reviews the related work, Section 3 describes the proposed architecture, Section 4 presents the methodology and Section 5, the preliminary results. Finally, Section 6 presents the next steps.

2. Background

Serious games have been applied across diverse educational domains, including STEM (Science, Technology, Engineering, and Mathematics), language learning, and health education. Research shows that game-based learning can enhance motivation and engagement, although improvements in learning outcomes vary depending on the context [Connolly et al. 2012].

In this context, GLA extends Learning Analytics to game-based environments by analysing player interactions and in-game behaviour, enabling the extraction of insights that support both learning assessment and game design improvement [Calvo-Morata et al. 2019, Alonso-Fernández et al. 2019]. The xAPI (Experience API) standard is widely used to record these interactions in a structured and interoperable format, allowing data to be shared across platforms and tools [Heinemann et al. 2022, Serrano-Laguna et al. 2017].

Telemetry tools, such as trackers integrated into game engines, are used to capture gameplay data according to this standard. For instance, XASU (xAPI Analytics Supplier) is a Unity-compatible framework that instruments game objects to automatically emit xAPI statements upon relevant in-game events [Calvo-Morata et al. 2025b]. Such tools enable researchers to monitor player actions, store interaction data, and support subsequent analysis of learning processes.

Recent research has explored the use of LLMs as assistants for analysing interaction data from educational games [Bastos et al. 2025], demonstrating their ability to interpret player behaviour, identify patterns such as trial-and-error strategies, and generate pedagogical insights, potentially reducing reliance on complex dashboards. These models are particularly effective for detailed analyses of individual player trajectories and for producing qualitative interpretations of gameplay data, although they still face limitations when handling large datasets, analysing groups of players simultaneously, or generating clear visual representations [Bastos et al. 2025].

To illustrate the practical application of these approaches, a small number of serious games have begun integrating GLA with LLMs, such as *Robô Euroi*, which records detailed interaction data and has been used to analyse gameplay logs with LLMs to extract pedagogical insights, and *Wake: Tales from the Aqualab*, where large-scale datasets have been analysed using hybrid approaches combining LLMs with traditional machine learning models [Melo et al. 2018, Bastos et al. 2025, Liu et al. 2025].

In contrast, a larger body of work focuses on serious games that rely exclusively on GLA pipelines without LLM integration. *Articoding* [Calvo-Morata et al. 2025a], for instance, is an open-source game designed to teach introductory programming concepts

through puzzle-based levels using a visual block language. The game incorporates a telemetry system based on the xAPI standard to capture fine-grained interaction data, such as time spent per level, code blocks used, and number of retries. While effective for supporting learning assessment and game design improvement, this approach still depends on manual interpretation or predefined metrics, limiting its ability to produce richer and more adaptive insights.

Despite recent advances, the integration of GLA with LLMs remains limited. Most existing approaches either rely solely on traditional analytics pipelines or use LLMs only as auxiliary tools. Few studies propose end-to-end solutions capable of transforming raw gameplay data into comprehensive pedagogical reports. This gap motivates the architecture proposed in this paper, which aims to automate data interpretation while preserving transparency, privacy, and educational relevance.

3. Architecture

The proposed architecture defines a general framework for collecting, storing, and analysing player interaction data in serious games. It is composed of three main functional components (Figure 1): (i) GLA capture module responsible for collecting player interaction data, (ii) a cloud infrastructure for storing interaction logs and metadata, and (iii) an analysis module for interpreting the collected data.

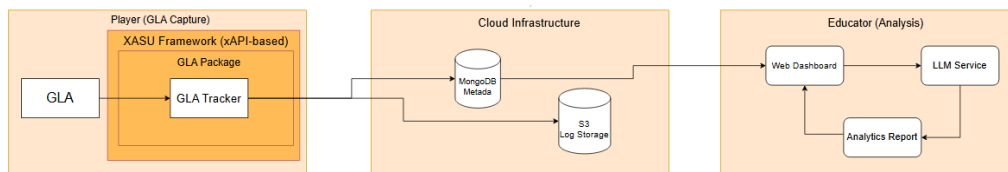


Figure 1. State diagram of the system architecture.

This architecture is technology-agnostic and can be instantiated using different tools, standards, and platforms depending on the application context.

4. Methodology

In this work, the methodology consists of instantiating the proposed architecture through a case study consisting of three operational phases, as shown in Figure 1: i) a serious game; ii) a data collection pipeline based on the xAPI standard using the XASU framework, and iii) an LLM-based analysis module implemented with Llama 3.2 via OllamaOS.

4.1. Serious Game: *Destino Cósmico*

Destino Cósmico is a book-game designed to foster interdisciplinary reasoning and scientific curiosity. The player creates or selects an organism and decides in which environments (planets, moons, and ecosystems) it will be tested, considering physicochemical, biological, and evolutionary factors (Fig 2). The game's outcomes depend on the coherence between the player's biological choices and the selected environments, reflecting their understanding of concepts such as adaptation, evolution, and organism-environment interaction across disciplines including biology, chemistry, physics, and astrobiology. The game was developed using the Unity Engine [Unity Technologies 2025].



Figure 2. (a) Selection of organism and (b) Selection of planet.

4.2. Phase 1 - Interaction Tracking with xAPI and XASU

Player interactions are captured during gameplay using the XASU framework which instruments game objects to automatically emit xAPI statements upon relevant in-game events (e.g., *player-selected-rigid_body*), enriched with timestamps and contextual metadata.

During the gameplay loop, generated statements are temporarily stored in memory by the XASU Tracker. When the application terminates, the XASU Manager serialises the in-memory queue into a local log file, which is then asynchronously uploaded to AWS S3 for persistent storage. In parallel, the corresponding log path and player metadata are registered in MongoDB. After successful upload, the local file is removed to minimise storage usage. This design ensures reliable data persistence without introducing significant runtime overhead, preventing loss of interaction data between sessions.

4.3. Phase 2 - LLM-Based Analysis with OllamaOS and Llama 3.2

Once the session log has been stored, a post-session analysis pipeline retrieves the file from AWS S3 using the path recorded in MongoDB. The log is then submitted to a locally deployed instance of the Llama 3.2 model, executed via OllamaOS [Ollama 2024], which generates a structured natural-language analysis of the player's behaviour.

Local deployment ensures that sensitive student data remain within the institution's infrastructure, avoiding transmission to external services and addressing privacy and regulatory concerns common in educational settings. The resulting analytical report can subsequently be accessed by educators through a web dashboard, supporting evidence-based interpretation of gameplay and learning processes.

5. Preliminary Results

The analysis module is driven by a structured system prompt (Modelfile), which operationalises the proposed architecture by defining the model's analytical role, output structure, and evidence-grounding strategy. The prompt embeds the educational context of *Destino Cósmico*, including its learning objectives and domain-specific knowledge, enabling gameplay events to be interpreted according to GLA principles and meaningful learning theory. It was iteratively refined based on xAPI semantics and successive evaluations of the generated reports to improve consistency and traceability.

Preliminary experiments were conducted using gameplay sessions representing different player interaction profiles. Throughout all experiments, the inference

configuration remained fixed (temperature = 0.2) to reduce output variability and improve consistency across generated reports.

The generated report follows a fixed five-section structure, summarised in Table 1. Each inference includes three supporting elements: *Motivation*, explaining the reasoning behind the conclusion; *Evidence/Data Used*, explicitly listing the supporting xAPI statements; and a confidence level (low, medium, or high), indicating the consistency of the available evidence. The reports are intended to support educators and domain specialists by providing interpretable feedback without requiring direct inspection of raw gameplay logs.

Table 1. Structure of the generated pedagogical report.

Section	Purpose	Supporting Evidence
Session Identification	Summarises player and session information.	Player metadata and session statistics.
Action Timeline	Presents gameplay events in chronological order with contextual interpretation.	Ordered xAPI statements.
Behavioural Analysis	Identifies gameplay patterns, response strategies, and consistency.	Event sequences and interaction frequency.
Pedagogical and Conceptual Analysis	Identifies learning evidence, conceptual gaps, and personalised feedback.	Gameplay decisions and domain-specific xAPI events.
Final Synthesis	Summarises overall performance and recommendations.	Aggregated evidence from previous sections.

This report structure was designed to reduce unsupported inferences by explicitly requiring evidence, justification, and confidence indicators for every conclusion, thereby improving transparency, facilitating human verification, and aligning the generated reports with principles of explainable AI [Ji et al. 2023].

6. Next Steps

Future work will involve a controlled empirical evaluation of the proposed architecture in authentic classroom settings using real student interaction data. The generated reports will be assessed by domain experts with respect to factual accuracy, pedagogical usefulness, and interpretability for non-technical educators, and compared with baseline approaches such as manual log inspection and traditional analytics dashboards. Further studies will also explore the applicability of the architecture to other types of serious games and interaction patterns.

Preliminary tests indicate that the depth and reliability of the generated analyses are constrained by the capacity of the locally deployed model, particularly when processing long or complex interaction logs.

References

- Alonso-Fernández, C., Calvo-Morata, A., Freire, M., Martínez-Ortiz, I., and Fernández-Manjón, B. (2019). Applications of data science to game learning analytics data: A systematic literature review. *Computers & Education*, 141:103612.
- Bastos, M., Honda, F., Lima, M., Pessoa, M., and Pires, F. (2025). How do llms analyze and interpret data from educational games? a study with gla experts. In *Anais do XXXVI Simpósio Brasileiro de Informática na Educação (SBIE 2025)*, SBIE 2025, page 1318–1331. Sociedade Brasileira de Computação - SBC.
- Calvo-Morata, A., Alonso-Fernandez, C., Freire-Moran, M., Martinez-Ortiz, I., and Fernandez-Manjon, B. (2019). Game learning analytics, facilitating the use of serious games in the class. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, 14(4):168–176.
- Calvo-Morata, A., Alonso-Fernández, C., Santilario-Berthilier, J., Martínez-Ortiz, I., and Fernández-Manjón, B. (2025a). Learning analytics to guide serious game development: A case study using articoding. *Computers*, 14(4):122.
- Calvo-Morata, A., Santilario-Berthilieres, J., Alonso-Fernández, C., Freire, M., Martínez-Ortiz, I., and Fernández-Manjón, B. (2025b). Full software support for game learning analytics. *Learning Analytics Summer Institute Spain*.
- Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., and Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, 59(2):661–686.
- Heinemann, B., Ehlenz, M., Görzen, S., and Schroeder, U. (2022). xapi made easy: A learning analytics infrastructure for interdisciplinary projects. *International Journal of Online and Biomedical Engineering (iJOE)*, 18(14):99–113.
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., and Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Comput. Surv.*, 55(12).
- Liu, X., Wei, Z., Baker, R. S., Metcalf, S. J., Zhang, J., Barany, A., Slater, S., Swanson, L., and Gagnon, D. J. (2025). Integrating large language models and machine learning to detect struggle in educational games. In *International Conference on Artificial Intelligence in Education*, pages 398–405. Springer.
- Melo, D., Júnior, R., Duarte, J., and Pires, F. (2018). Robô euroi: Estratégias matemáticas para desenvolver o pensamento computacional. In *Anais dos Workshops do Congresso Brasileiro de Informática na Educação*, page 242.
- Ollama (2024). Ollama’s documentation. <https://docs.ollama.com/>. Accessed: March 2026.
- Serrano-Laguna, Á., Martínez-Ortiz, I., Haag, J., Regan, D., Johnson, A., and Fernández-Manjón, B. (2017). Applying standards to systematize learning analytics in serious games. *Computer Standards & Interfaces*, 50:116–123.
- Silva, D., Pires, F., Melo, R., and Pessoa, M. (2022). Glboard: um sistema para auxiliar na captura e análise de dados em jogos educacionais. In *Anais Estendidos do XXI*

Simpósio Brasileiro de Jogos e Entretenimento Digital (SBGames Estendido 2022), SBGames Estendido 2022, page 959–968. Sociedade Brasileira de Computação.

Unity Technologies (2025). Unity real-time development platform. <https://unity.com/>. Accessed: March 16, 2026.

Yu, Z., Gao, M., and Wang, L. (2021). The effect of educational games on learning outcomes, student motivation, engagement and satisfaction. *Journal of Educational Computing Research*, 59(3):522–546.

Zeng, J., Parks, S., and Shang, J. (2020). To learn scientifically, effectively, and enjoyably: A review of educational games. *Human Behavior and Emerging Technologies*, 2:186–195.