An Extended Study with LLM-Generated Synthetic Data of ArSGam: A Framework for Cognitive-Affective Personalization and Gamification in Serious Game

Claudia Gómez
Universidad Francisco de Paula Santander, GIDISOFT,
Cúcuta, Colombia
claudiaygomez@ufps.edu.co

and

Paola Vallejo
Universidad EAFIT, GICOMP,
Medellín, Colombia
pvallej3@eafit.edu.co

and

Jose Aguilar
Universidad de Los Andes, CEMISID,
Mérida, Venezuela
IMDEA Network Institute,
Madrid, Spain
aguilar@ula.ve; jose.aguilar@imdea.org

Abstract

This article presents an in-depth analysis of ArSGam, a modular architecture designed to integrate gamification strategies into serious games for cognitive-affective training. The original work introduced four interconnected modules that guide the process from user profiling to the integration of game mechanics. This extended version strengthens the scientific contribution by conducting additional experiments with ArSGam and employing Large Language Models (LLMs) as tools for generating synthetic data to expand testing scenarios. Furthermore, new case studies and validation analyses were performed to examine the architecture's operational consistency and applicability in high-risk training contexts. A comparison with related architectures is also provided to contextualize the distinctive features of ArSGam. The results reaffirm the relevance of incorporating cognitive-affective profiling into gamification design and highlight the flexibility of the architecture for supporting personalized serious game experiences. Consequently, this study consolidates ArSGam as a robust and versatile framework for integrating cognitive-affective personalization and gamification within serious games.

Keywords: Software Architecture, Gamification, Serious Games, Cognitive-Affective Model.

1 Introduction

In training environments, classical methods such as slideshows and manuals are ineffective for learning processes [1], which affects the motivation and interest of the learners. Serious games (SG) and Gamification are strategies. SG and gamification are strategies that allow learning to be transformed into an attractive and enjoyable experience. SGs go beyond entertainment and focus on achieving purely educational goals, allowing them to play an important role in learning processes [2], [3]. In turn, gamification is understood as the use of game design elements in non-game contexts [4], and has been used in software development to improve its usability [5]. Gamification allows increasing student interest and performance, especially in educational games and mobile environments [6]. Ortiz-Rojas et al. [7] highlight that gamification, through elements such as badges, can strengthen motivation in learning processes by satisfying basic psychological needs, including autonomy, competence, and belonging.

The use of gamification has recently gained prominence in various fields because it enables psychological and behavioral improvements [8]. In particular, its use in conjunction with social learning systems facilitates participatory learning processes. This work analyzes a modular architecture for integrating gamification into SGs developed in previous work [9], in order to facilitate user profiling based on cognitive-affective aspects and learning styles. The architecture goal is to enable a personalized design of game mechanics, dynamics, and elements aligned with users' needs, focused on enhancing their learning experience, and particularly, their cognitive-affective skills. Specifically, this article extends the study of the modular architecture that integrates gamification into SGs by analyzing other use cases of cognitive-affective training in diverse learning contexts, using LLMs to generate data from these new contexts. Data generation using LLMs is based on the COGAF model [9], which establishes the cognitive, affective, and motivational characteristics to consider in the design of SGs. The results of these new case studies confirm its adaptability and positive impact on learning outcomes, based on the ISO/IEC 25010 standard [10], which establishes quality criteria at the usability and maintainability levels, which are of particular interest to us in this case.

Furthermore, an extensive literature review is conducted to highlight the contributions of ArSGam. This literature review emphasizes that while numerous studies on gamification have explored the use of badges, points, and leaderboards, few have addressed the integration of cognitive-affective profiles into modular architectures. This gap underscores the relevance of ArSGam, which provides a structured approach that aligns user characteristics with pedagogical goals. This literature review reinforces its scientific standing and clarifies its unique role in advancing gamification within SGs. [11], [12], [13]. Emerging trends such as adaptive learning systems, AI-driven personalization, and big data analytics further underscore the importance of modular architectures that integrate gamification in a scalable and context-aware manner. By placing ArSGam within this broader landscape, this paper not only validates its applicability but also highlights its potential for future research on adaptive gamification and cross-domain applications [14], [15] , [16].

The structure of this paper is as follows. Section 2 reviews the related work. Section 3 outlines the theoretical framework that underpins the proposed approach. Section 4 details the architecture, including its overall structure and individual modules. Section 5 describes a case study using LLMs for the generation of synthetic data in that case, beginning with the contextual background, then the synthetic data generation, and concluding with the architecture's instantiation. Section 6 presents the validation process, including exploratory reliability analyses, descriptive and statistical comparisons between real and synthetic profiles, and an assessment of the operational consistency of the proposed architecture. Section 7 presents and discusses the results. Finally, Section 8 summarizes the main conclusions and outlines directions for future work.

2 Related Work

Yousefi and Mirkhezri [17] defined a conceptual framework for integrating gamification and system learning (SG) into a holistic framework that considers game design principles, mechanics, and dynamics to enrich educational processes. The framework shows how gamification and SG can be combined to improve motivation, engagement, and learning effectiveness, with a structured approach for implementation in different educational contexts. Drakoulis et al. [18] implemented an on-demand public transport service in Trikala, Greece, with a gamification layer to motivate users to participate and behave appropriately within the system. They developed a ubiquitous computing architecture with several types of SG oriented towards these purposes, which has a main component that is the processing of game logic that manages the participation of users in the SGs in order to improve their mobility habits, which are evaluated by the application.

With respect to gamification, Elaish et al. [6] presented a taxonomy to classify gamification elements into two main categories based on content type and learning technique, aiming to simplify and accelerate the educational game design process: tangible and intangible. The authors identified a key challenge: categorization often focuses on elements that can be represented by physical objects or the user interface, without considering the more abstract complexity and dynamics of game design. Minas et al. [19] integrated gamification into learning environments by incorporating various game elements into the RAISE virtual environment, a learning environment designed to foster environmental awareness. The authors combined interactive scenarios, challenges, immediate feedback, and interaction with non-player characters (NPCs) to promote student engagement through experiential learning, thus demonstrating the potential of combining multiple elements to improve learning outcomes. Antunes et al. [20] defined an integrated gamification scheme for the design of therapeutic games for the rehabilitation of children with special needs. The authors specified a conceptual framework and a computational platform for the real-time personalization of therapeutic games, with levels and action sequences adapted to the specific needs of each patient. According to their results, their approach allows for greater involvement and motivation in children during therapy. Putra et al. [21] proposed a gamification architecture for children with attention deficit hyperactivity disorder (ADHD), incorporating logic games such as brain, logic, think, sport, move, and body. Their main contribution lies in a structured creation framework that integrates these components into games designed to engage, motivate, and enhance the cognitive and physical skills of children with ADHD, thereby supporting effective and enjoyable learning.

Recent studies have expanded the scope of gamification frameworks by emphasizing adaptive and AI-driven personalization. For instance, Chiotaki et al. [11] conducted a systematic review highlighting the importance of adaptive game-based learning systems in education, while the authors in [12] proposed a framework for user-based adaptive gamified systems. Similarly, Abbes et al. [13] explored integrating generative AI into gamification for personalized learning, identifying challenges and opportunities for future research. These contributions reinforce the relevance of modular and adaptive architectures like ArSGam, which align cognitive-affective profiling with gamification design. Other recent works have focused on scalability and cross-domain applications. Naseer et al. [14] presented a framework combining game mechanics with AI personalization to support adaptive learning systems, while Loizou & Andreou [15] introduced a digital twin-based architecture for SG with real-time self-adaptation. Mitsea et al. [16] provided a systematic review of SG in the era of AI and immersive technologies, emphasizing their transformative potential for meta-skills training. These studies illustrate the growing demand for architectures that integrate gamification with advanced personalization and adaptability, situating ArSGam within a broader landscape of evolving frameworks.

3 Theoretical Framework

This section introduces the core theories behind the proposed architecture: the COGAF model, which links cognitive and emotional aspects of learning, and the VARK model, used to personalize learning based on individual preferences.

3.1 Cognitive – Affective Model (COGAF)

COGAF is a model that allows the design of SG for user training using cognitive and affective dimensions, each with its own notation, components, and behavior. The cognitive dimension allows for the specification of cognitive functions, complementary activities, abilities, and mechanics that will be considered during the design of the SGs. Furthermore, this dimension establishes the psychological tasks integrated into cognitive tests for the evaluation of each cognitive function [22]. The affective dimension allows for the specification of emotional states and behaviours [23], [24], [25].

Beyond its theoretical specification, COGAF aligns with recent advances in adaptive gamification frameworks that emphasize integrating cognitive and affective dimensions for personalization [11], [13]. Unlike models focused solely on motivational or learning-style aspects, COGAF explicitly incorporates emotional states and behavioral characteristics, offering a more holistic approach to user profiling. Its validation through VR-based SG [9], illustrates its applicability in immersive environments, while this research highlights its potential extension toward AI-driven personalization and large-scale adaptive learning systems [14], [16]. Within ArSGam, COGAF serves as a core theoretical foundation, operationalized through modular components that enable coherent and personalized gamification design.

3.2 VARK Learning Style Model

The VARK model was developed to improve teacher training and help students become better learners [26]. It consists of four sensory modalities: visual, auditory, reading/writing, and kinesthetic, which identify how each student prefers to receive and process information [27], [28], [29]. For example, kinesthetic learners learn best through experience and practice, while visual learners prefer graphical information. The VARK model has been applied to analyze how different learning preferences influence student engagement, aiming to adapt teaching methods accordingly [29]. It has also been used to adapt pedagogical methodologies and educational technologies. Therefore, we selected VARK for our research to explore the relationship between learning styles, student engagement, and motivation.

Particularly, recent studies have revisited the relevance of learning-style models such as VARK in combination with adaptive technologies. For instance, Loizou & Andreou [15] demonstrated how SG can integrate real-time self-adaptation mechanisms to align with learner preferences, while Tene et al. [30] highlighted the role of SG in STEM education, showing that tailoring content to VARK modalities improves engagement and knowledge retention. These findings suggest that VARK remains a useful framework when embedded into modular architectures, ensuring that personalization extends beyond cognitive-affective profiling to encompass diverse learning preferences. In this sense, ArSGam leverages VARK not only to classify learners but also to guide the selection of gamification strategies that resonate with individual styles, reinforcing adaptability and inclusiveness in SG.

4 ArSGam Architecture

4.1 General Architecture Structure

ArSGam's architecture allows for the structured integration of gamification mechanisms into SGs, focusing on cognitive and emotional training [31]. It consists of modules that interact to analyze the user's context, recommend actions, select appropriate gamification strategies, and coherently integrate these components into the SG. ArSGam is adaptable to various pedagogical goals and user profiles. The architecture is user-centered, aiming to create gamified experiences that are motivating and pleasant, aligned with the cognitive-affective training objectives. ArSGam allows covering everything from establishing the learning context and developing the user profile to designing and integrating gamification elements into SG. Figure 1 illustrates its architecture: Module 1 – Contextual Analysis (orange), Module 2 – Recommendation (green), Module 3 – Gamification Scheme Design (blue), and Module 4 – Gamification Integration (purple). The diagram details the flow of information and processes within and between these modules.

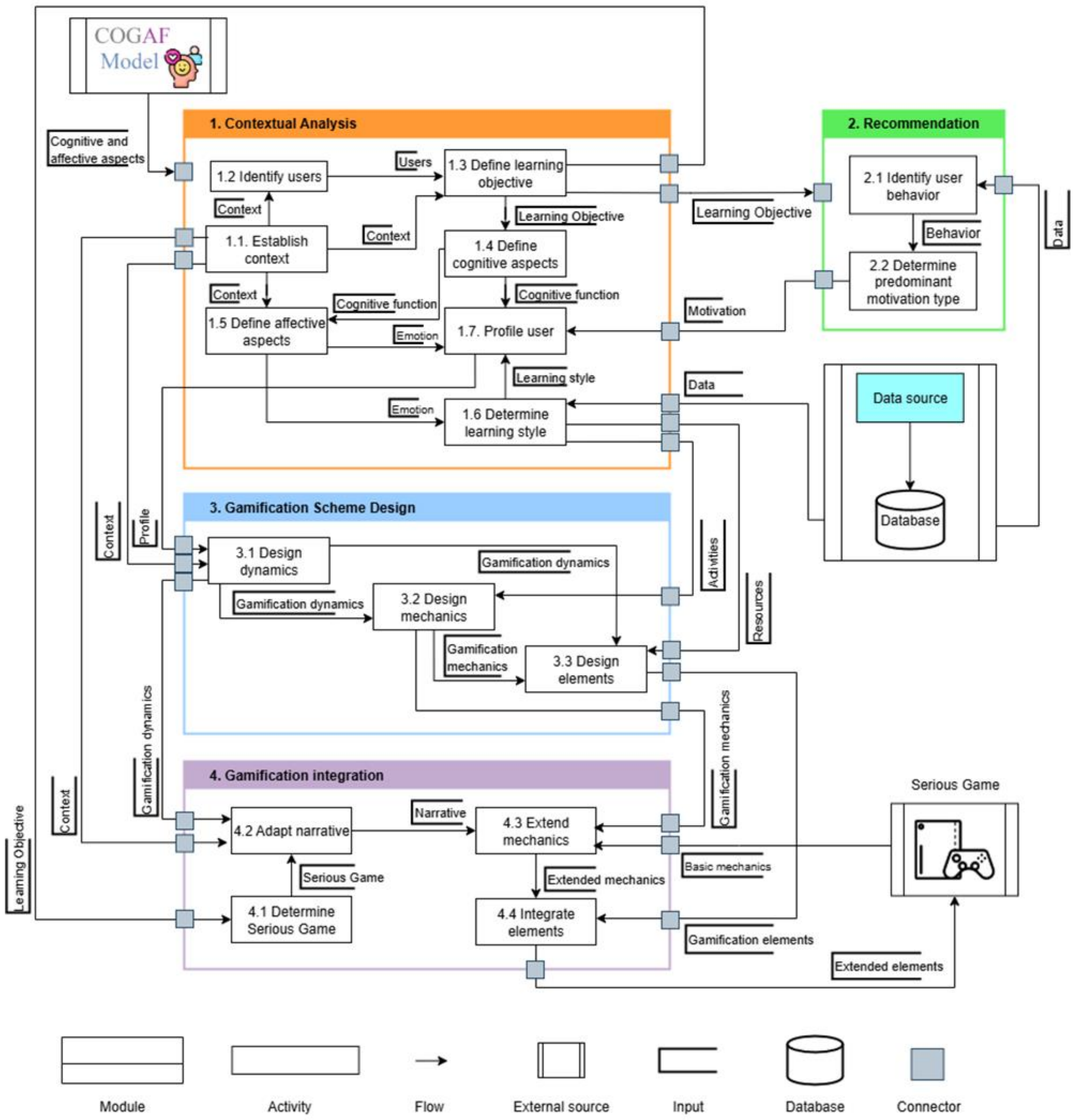

Figure 1: General structure of the ArSGam architecture

The COGAF model is used to define components related to the cognitive and emotional dimensions of the user, enabling the architecture to implement gamification dynamics and mechanics aligned with each user's needs. On the other hand, the VARK model provides an approach focused on learning styles, allowing the architecture to adapt activities and stimuli to visual, auditory, reading/writing, or kinesthetic preferences. The use of the COGAF and VARK models in the architecture ensures that learning games are coherent from an emotional and pedagogical perspective, and customizable, which reinforces participation, motivation, and effectiveness in cognitive-affective training processes. Each activity in Figure 1 corresponds to specific actions outlined in the macro-algorithm presented in Table 1. They provide an overview of the ArSGam architecture, guiding you through each module from contextual analysis to final gamification integration.

Table 1: Macroalgorithm of the ArSGam architecture

Step	Action	Module
1	Analyze the context, users, learning objective, and cognitive-affective aspects	Contextual Analysis
2	Define a data input source capable of handling multiple data types from various modalities (such as LMS, social networks, and games), providing a comprehensive understanding of the user and their context	Recommendation
3	Analyze user behavior based on incoming data to identify relevant patterns and trends	Recommendation
4	Determine the user's predominant motivation by analyzing their behavior and collected data using a framework that enables motivation classification for personalized experiences	Recommendation
5	Determine the user's learning style using data from sources such as LMS, social networks, or other platforms (based on step 2)	Contextual Analysis
6	Profile user (based on step 4)	Contextual Analysis
7	Design gamification dynamics	Gamification Scheme Design

Step	Action	Module
8	Design gamification mechanics and elements (based on step 5)	Gamification Scheme Design
9	Select the SG with basic mechanics from the learning objective (based on step 1)	Gamification Integration
10	Integrate the SG with the gamification schema (based on steps 7 and 8)	Gamification Integration

4.2 Architecture Modules Description

Each of the architecture modules is described below.

4.2.1 Module 1 - Contextual Analysis

This module analyses the learning context and user to collect essential cognitive and affective information for customizing the gamification system. It helps define the user characteristics, learning objectives, context, and cognitive-emotional aspects to tailor the gamification design. Its workflow is illustrated in Figure 2. The process begins with defining the context (1.1), which feeds into user identification (1.2). Based on the identified users, the learning objective is defined (1.3), followed by the characterization of cognitive aspects (1.4). The arrow labeled COGAF indicates that this module uses the COGAF model to specify cognitive and emotional dimensions. Affective aspects are then identified using information from steps 1.1 and 1.4 (1.5). Subsequently, the learning style is determined (1.6), drawing on insights from steps 1.4, 1.5, and relevant learning resources and activities. Finally, the user profiling stage (1.7) integrates all collected data, including the predominant motivation type (2.2) provided by Module 2. The main inputs to this module include cognitive and affective aspects, motivation, and data —such as interactions, grades, resources used (e.g., readings, quizzes, and videos), preferences, and likes—gathered from external sources like learning management systems (LMS) or social networks. Its outputs are a defined context, user profiling, learning objectives, and associated resources and activities.

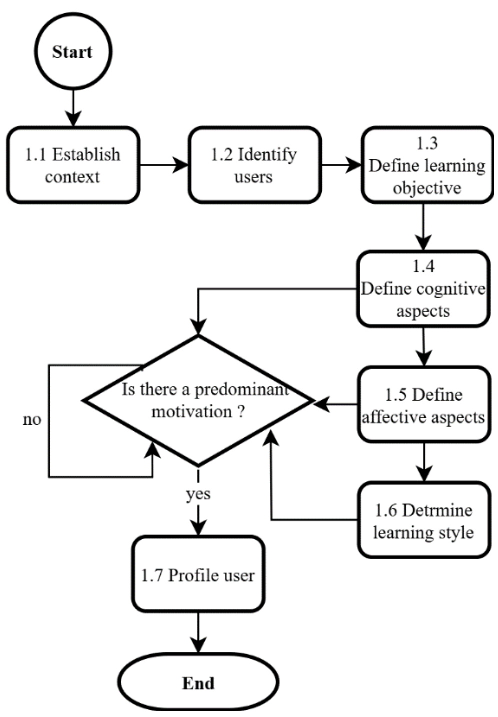

Figure 2: Contextual Analysis Flowchart

4.2.2 Module 2 - Recommendation

This module studies user data to identify the motivational factors that guide the gamification strategy. Its main utility is determining user preferences, behaviors, and motivations to recommend appropriate gamification elements and learning content. It takes as inputs the learning objectives provided by Module 1 and user data obtained from external sources such as LMS, social networks, video games, or other platforms. Its workflow is illustrated in Figure 3. The process starts with input validation and then analyzes user behavior about learning activities (2.1). Based on this analysis, the module determines the user's predominant type of motivation (2.2) using a framework that supports motivation classification. The output of this module is the predominant user motivation, which is integrated into the user profiling process in Module 1.

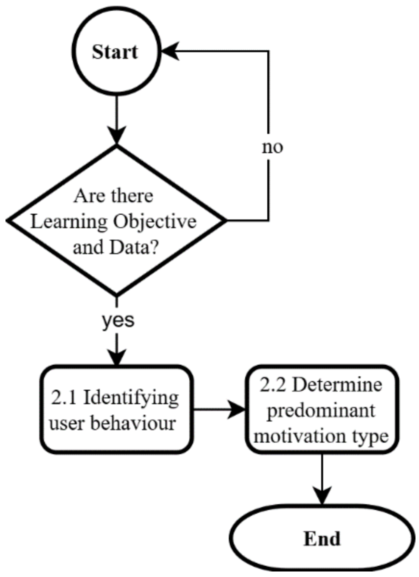

Figure 3: Recommendation Flowchart

4.2.3 Module 3 – Gamification Scheme Design

This module creates gamification elements tailored to the defined learning objectives and the user profile. Its utility lies in designing a structured set of mechanics, dynamics, and elements of gamification that align with user needs and pedagogical goals. It receives as inputs the context and user profile information, and details about activities and resources obtained from Module 1. This process is illustrated in Figure 4. The design process starts with the development of gamification dynamics (3.1), continues with the mechanics design (3.2), and concludes with the design of specific gamification elements (3.3). These elements (gamification dynamics, gamification mechanics, and gamification elements) constitute the results of the module and the inputs for the next stage.

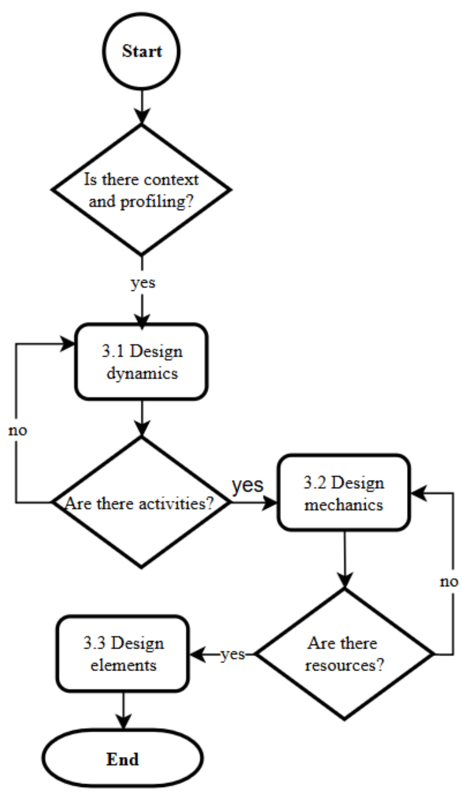

Figure 4: Gamification Scheme Design Flowchart

Although Figure 4 illustrates the sequential workflow of Module 3, the decision process that transforms a cognitive-affective profile into specific gamification recommendations may remain difficult to interpret. Therefore, Table 2 presents an illustrative traceability matrix that exemplifies how the architecture can be operationalized by mapping cognitive functions, emotions, and learning styles to corresponding gamification dynamics, mechanics, and elements. This matrix should not be interpreted as a set of rules, but rather as a representative example of the deterministic mappings that enable the execution of Module 3 and demonstrate the explainability of the recommendation process.

Table 2: Illustrative Traceability Matrix for the Operationalization of Module 3 in ArSGam

Cognitive Function	**Emotion**	**Learning Style**	**Gamification Dynamics**	**Mechanics**	**Elements**
Perception	Surprise, Disgust [32]	Visual	Exploration and discovery	Image identification, environmental scanning	Maps, visual cues, indicators
Attention	Surprise [32]	Visual	Progressive challenge	Time-based observation tasks	Timers, signals, checkpoints
Memory	Happiness [33]	Reading/Writing	Knowledge reinforcement	Recall and matching activities	Badges, notes, progress bars
Praxia	Fear [34]	Kinesthetic	Motor challenge and exploration	Object manipulation and navigation	Interactive objects, movement contminrols
Language	Happiness [33]	Aural	Social interaction	Dialogue and verbal tasks	Narration, audio cues
Inhibitory Control	Anger [35]	Visual	Self-regulation under pressure	Decision making and response inhibition	Timers, warning panels
Working Memory	Fear [36]	Reading/Writing	Sequential problem solving	Information retention tasks	Checklists, instructions
Cognitive Flexibility	Fear [37], Sadnees [38]	Kinesthetic	Adaptive exploration	Scenario switching and multiple pathways	Maps, branching paths

Table 2 illustrates an example of the traceability mappings used to operationalize Module 3. To improve the transparency and reproducibility of the recommendation process, these mappings can be represented through IF–THEN production rules that explicitly translate combinations of cognitive functions, emotions, and learning styles into gamification recommendations. The following rules are illustrative examples and do not constitute an exhaustive knowledge base of ArSGam; rather, they demonstrate the deterministic reasoning employed by the architecture to generate pedagogically grounded recommendations.

Rule 1:
IF Cognitive Function = Perception
AND Emotion = Surprise OR Disgust
AND Learning Style = Visual
THEN
Dynamics = Exploration and Discovery
Mechanics = Image Identification and Environmental Scanning
Elements = Maps, Visual Cues, and Risk Indicators

Rule 2:
IF Cognitive Function = Attention
AND Emotion = Surprise
AND Learning Style = Visual
THEN
Dynamics = Progressive Challenge
Mechanics = Time-Based Observation Tasks
Elements = Timers, Signals, and Checkpoints

Rule 3:
IF Cognitive Function = Memory
AND Emotion = Happiness
AND Learning Style = Reading/Writing
THEN
Dynamics = Knowledge Reinforcement
Mechanics = Recall and Matching Activities
Elements = Badges, Notes, and Progress Bars

Rule 4:
IF Cognitive Function = Praxia
AND Emotion = Fear
AND Learning Style = Kinesthetic
THEN
Dynamics = Motor Challenge and Exploration
Mechanics = Object Manipulation and Navigation
Elements = Interactive Objects and Movement Controls

Rule 5:
IF Cognitive Function = Language
AND Emotion = Happiness
AND Learning Style = Aural
THEN
Dynamics = Social Interaction
Mechanics = Dialogue and Verbal Tasks
Elements = Narration and Audio Cues

Rule 6:
IF Cognitive Function = Inhibitory Control
AND Emotion = Anger
AND Learning Style = Visual
THEN
Dynamics = Self-Regulation under Pressure
Mechanics = Decision Making and Response Inhibition
Elements = Timers and Warning Panels

Rule 7:
IF Cognitive Function = Working Memory
AND Emotion = Fear
AND Learning Style = Reading/Writing
THEN
Dynamics = Sequential Problem Solving
Mechanics = Information Retention Tasks
Elements = Checklists and Written Instructions

Rule 8:
IF Cognitive Function = Cognitive Flexibility
AND Emotion = Fear OR Sadness
AND Learning Style = Kinesthetic
THEN
Dynamics = Adaptive Exploration
Mechanics = Scenario Switching and Multiple Pathways
Elements = Maps and Branching Paths

The previous rules demonstrate that the recommendation process implemented in Module 3 follows explicit and explainable guidelines based on cognitive-affective and pedagogical theory. Consequently, the generation of gamification strategies is reproducible and is not based on subjective design decisions.

4.2.4 Module 4 – Gamification Integration

This module manages the incorporation of the gamification scheme into the SG, ensuring a personalized learning experience. Its utility lies in embedding the dynamics, mechanics, and elements designed in the previous module into an SG using the context and learning objectives as a foundation. The module takes as inputs the context, learning objectives, the complete gamification scheme (including gamification dynamics, gamification mechanics, and gamification elements), and the basic mechanics of the SG. This process is illustrated in Figure 5. The integration process starts with the selection of the SG (4.1), followed by the adaptation of the game narrative to align with the gamified structure (4.2), the extension of the game's existing mechanics (4.3), and finally, the integration of the specific gamification elements (4.4). The output is an adapted SG enriched with gamification features, offering the learner a personalized experience.

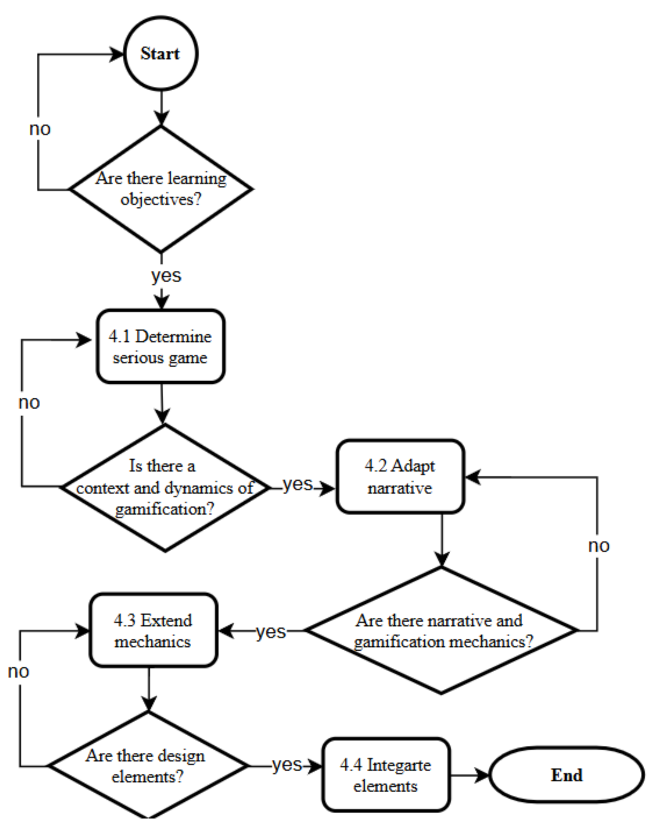

Figure 5: Gamification Integration Flowchart

5 Case Study

This section presents a new case study to evaluate the versatility of the ArSGam architecture [31]. The first case study, published in [31], addressed a mining context focused on spatial memory training to facilitate emergency evacuation and reduce fear in underground environments, using a 2D SG based on the COGAF model and a virtual reality mining explosion simulation as data sources to infer user profiles and support the gamification process. In contrast, the proposed case study differs from the case presented in [31] because it focuses on the design of gamification mechanisms for an SG that facilitates the training of *selective attention and perception*, considered fundamental cognitive functions. The training context is defined using lessons learned from the Colombian National Mining Agency (ANM), ensuring alignment with real safety challenges in mining environments.

5.1 Case Study Context

This case study focuses on an underground mining company that must train its workers in risk assessment before each shift. The lack of such training has been repeatedly linked to mining accidents, as reported by the ANM. The company aims to enhance workers' selective attention and ability to recognize warning signs, improving early detection of unsafe conditions and supporting preventive decisions in underground settings. In this new case study, the COGAF model specifies both emotional and cognitive dimensions. Unlike the original case, which analyzed university students using the SG "Labyrinth" and evaluated a virtual reality mining simulation (see [10][31]), this study uses a synthetic dataset generated by an LLM simulating mining workers' interaction with an online course on risk and hazard identification before starting their shifts. To characterize user motivation, the Octalysis gamification framework is again adopted [39], as it enables a detailed analysis of the fundamental motivational factors that need to be reinforced and sustained during training. In this case, motivational factors are inferred from users' observable behaviors, emotional states, and interaction patterns, as represented in the synthetic dataset. As in the original study, Octalysis organizes these factors into intrinsic and extrinsic motivation dimensions, enabling a consistent comparison between the two case studies based on how each motivational dimension manifests and is reinforced during training.

5.2 Protocol

This case study aims to validate the process defined in the ArSGam architecture [31] by designing an SG based on the COGAF model to achieve the learning objective and to verify its integration with that architecture. In addition, we propose defining a gamification scheme for SG and evaluating the reasoning process's ability to infer the user's learning style based on interactions in virtual learning environments. Therefore, an exploration was carried out to retrieve data from the Mining Learning Center (CAM), which offers more than 190 training courses. The CAM does not offer an open data repository on user interaction with these courses [40]. We also investigated learning platforms in the Colombian mining and energy sector, such as the Colombian Institute of Mining, Agricultural, and Environmental Sciences (ICCMAA), which offers informal education programs, advanced courses, and refresher courses; however, they do not have open data repositories for research purposes [41]. Similarly, the Integrame platform serves as a database for the Colombian mining and energy sector to consult, process, and download relevant information on mining, hydrocarbons, and energy, but it does not include data on mining training [42]. In addition, the ANM's open data section was reviewed, but no datasets associated with training courses on mining risk and hazard prevention were found [43].

Due to the lack of open datasets on user interaction in mining safety training courses, it was decided to generate a synthetic dataset using LLMs to validate the functionality and versatility of ArSGam. LLMs have become powerful knowledge aggregation tools, with applications in clinical decision support and education [44] - [45], among others, as described in [46]. Specifically, GPT-5 was used as the LLM to generate a synthetic dataset simulating user interactions in an online training course on "mining risks and hazards". For this article, the generated dataset allowed us to validate ArSGam's architecture and its profiling and gamification process, as presented in the [45]following section.

5.3 Synthetic Dataset Generation

The synthetic dataset was generated using a structured generation procedure designed to validate the profiling and gamification processes of a gamified SG architecture. Thus, the synthetic dataset was generated to verify the operational consistency of the extended ArSGam modules under controlled conditions. This procedure was guided by a predefined generation schema that specifies the application context, learning objectives, cognitive focus, and gamification principles relevant to high-risk training environments. The dataset represents interactions of workers from the mining sector participating in an online training course focused on identifying mining risks and hazards prior to the start of the working day. The training scenario emphasizes cognitive functions related to selective attention and perception, which are critical for early risk detection in hazardous environments. Dataset generation was supported by the ArSGam gamified SG architecture. Next, we discuss the process followed for the generation of synthetic data using LLMs. Below, we describe the different aspects used to define the prompt.

5.3.1 Input Data Specification

The input data used for dataset generation includes the application sector (mining), the learning environment (online training course), the training topic (mining risks and hazards before the working day), and the targeted cognitive emphasis (selective attention and perception). The dataset scheme defines five core variables: User, Emotion, Behavior during the course, Core Drive (Octalysis), and Motivation.

5.3.2 Reference Values and Variables

Users are represented by anonymized identifiers ranging from U01 to U30. Emotional states are defined according to six basic emotions: anger, disgust, fear, happiness, sadness, and surprise. Behavioral data describe typical actions and responses exhibited by miners during an online safety training course. Gamification drivers are modeled using the Octalysis framework, including Meaning, Achievement, Empowerment, Ownership, Social Influence, Scarcity, Curiosity, and Avoidance [47]. We emphasize that the Octalysis framework was adopted in our first case study [31] due to its well-established contributions to gamification in virtual learning environments [48]. Furthermore, the type of motivation (intrinsic or extrinsic) is inferred from the Octalysis framework.

5.3.3 Conceptual Constraints

The dataset generation process was guided by conceptual constraints to ensure semantic coherence. Behavioral patterns were assigned consistently with the emotional state and training context. Core Drives from the Octalysis framework were inferred based on observed behavior patterns, and motivation type (intrinsic or extrinsic) was assigned in alignment with the corresponding Core Drive.

For more details on the technical aspects of the prompt used to generate this synthetic dataset, please refer to the following repository: https://github.com/claudiagomezll/Prompt-for-Synthetic-Mining-Risk-Interaction-Data. This allows for its reuse and reproduction in research related to this article. Table 3 shows 10 of the 30 records generated by the LLM with the variables specified in section 5.3.2.

Table 3: Example of the synthetic dataset generated of a mining risk course

User	Emotion	Behavior during the course	Core Drive Octalysis	Motivation
U01	fear	Replays hazard-identification clips	Avoidance	Extrinsic
U02	surprise	Explores unexpected-scenario simulations	Curiosity	Intrinsic
U03	happiness	Completes micro-quizzes early	Social Influence	Extrinsic
U04	sadness	Writes reflective notes linking case studies to past incidents on site	Meaning	Intrinsic
U05	anger	Expresses frustration with the interface but persists	Achievement	Extrinsic
U06	disgust	Requests content warnings for graphic imagery	Avoidance	Extrinsic
U07	fear	Double-checks each answer using the checklist tool	Achievement	Extrinsic
U08	surprise	Clicks through interactive mine-map layers	Empowerment	Intrinsic
U09	happiness	Voluntarily shares a personal 'spotting blind spots' tip framed as protecting the crew.	Meaning	Intrinsic
U10	sadness	Learns slowly after emotive accident narratives.	Ownership	Intrinsic

5.4 Architectural Specification

This section presents the simulation of the ArSGam architecture, verifying the execution of the activities defined in each module and the corresponding input and output flows. The synthetic dataset generated and presented in Section 5.3, composed of the variable's user, emotion, behavior during the course, Octalysis core drive, and motivation, provides the necessary inputs for Modules 1 and 2, as illustrated in Figure 6. These inputs allow the flow of activities throughout the architecture to be validated, including the context, learning objective, and data source. Figure 6 also outlines the sequence of activities for each module and highlights the specific tasks performed in each.

The results of the *Module 1* activities are presented below.

1.a. Context: The mine needs to train its miners in risk assessment before they start their workday
1.b. Users: Miners
1.c. Learning objective: Improve early risk detection and preventive decision-making in underground mining
1.d. Cognitive aspects: Selective attention (Executive Function) and perception (Basic Cognitive Function)
1.e. Affective aspects: Fear
1.f. Learning style: Aural and visual
1.g. User profile: The learning style is aural and visual, the experienced emotion is fear, the targeted cognitive function is selective attention and perception, and the predominant motivation is extrinsic.

The synthetic data allows us to validate the reasoning process for the creation of profiles and gamification proposed in the ArSGam architecture. The results of the activities in *Module 2* are described below:

2.a. User behavior: Active participation in learning activities to improve immediate reaction to changes in the environment and perceive risks in underground mining.
2.b. Predominant motivation type: Intrinsic

To support the activities flow in *Module 3*, gamification components (dynamics, mechanics, and elements) are extracted from selected SG to achieve the context learning objective. The results of the *Module 3* activities are described below.

3.a. Dynamics

Following the ArSGam architecture, we present the gamification dynamics proposed for an SG that reinforces selective attention and perception of risks and warning signs in underground mining.

Progression dynamics:

- Users advance through increasing levels of complexity following risk assessment, sequentially unlocking mine areas as warning signs are correctly identified.
- The game must have progressive timers that reduce the time available for decision-making.

Emotion dynamics:

- Users must control fear through unexpected visual and audible alerts.
- The user must be exposed to danger signals, such as flashing lights and alarm sounds, and receive an immediate response that transforms fear into focused attention.

Cognitive Function Dynamics:

- Users must identify relevant signals among distractors (Selective attention).

- Users must visually and auditorily discriminate warning signals using colors, sounds, or symbols (Perception).

  Motivational Dynamics (Extrinsic):
- Users receive external rewards, such as points for correctly identifying risks, badges for sustained error-free performance, and individual rankings.
- Users incur minor penalties for ignoring warnings, such as time or point deductions.

3.b. Mechanics:
- Decide under time pressure (Time challenge)
- Confirm and alert decisions (Feedback)

3.c. Elements:
- Points for each risk sign correctly identified (Motivational resource)
- Visual and auditory alerts (Learning resource): key stimuli for an aural–visual style
- Progress indicator (Tracking resource): shows daily training progress.

Defining the context and the basic dynamics of SG is essential to support the activity flow in *Module 4*. The details of these activities are presented below.

4.a Determination of the SG:
Based on the learning objective and context, and considering the COGAF model, the study presented in [31] serves as a reference to determine the type of SG that will reinforce selective attention and perception, yielding the ADVENTURE (Write + Match) pattern as the metabrick. Complementarily, to strengthen the cognitive function of perception, the specification defined by COGAF for this function is used, selecting the LABYRINTH pattern (Move + Avoid), whose mechanics aim to improve the user's ability to detect dangers in the environment [10]. Therefore, the proposed SG is conceived as an integration of the maze and adventure types, aligned with the target cognitive functions.

4.b. Narrative: The player assumes the role of a mining supervisor who must conduct a risk inspection before the start of the workday. Before allowing their team to enter, they must tour the mine, identify critical warning signs, and make decisions under time pressure. During the inspection, warning sounds, flashing lights, and visual signals appear unexpectedly, creating a controlled, stressful environment that triggers fear as a functional emotion. Each correct decision strengthens the team's safety, while mistakes delay the operation and reduce the score. The ultimate goal is to ensure a safe start to the day, reinforcing selective attention and the perception of critical signals through progressive training motivated by external rewards.

4.c. Mechanics:
- Basic mechanics: Moving around the map while avoiding risks and writing and matching risk signs
- Mechanics with gamification: Decide under time pressure (Time challenge) and confirm and alert decisions (Feedback)

4.d. Integrate Elements:
- Timer to create urgency and focus attention
- Map of the mine to reinforce spatial awareness
- Points for each correctly identified hazard (Gamification element)
- Visual and audible alerts (Gamification element)
- Progress indicator (Gamification element)

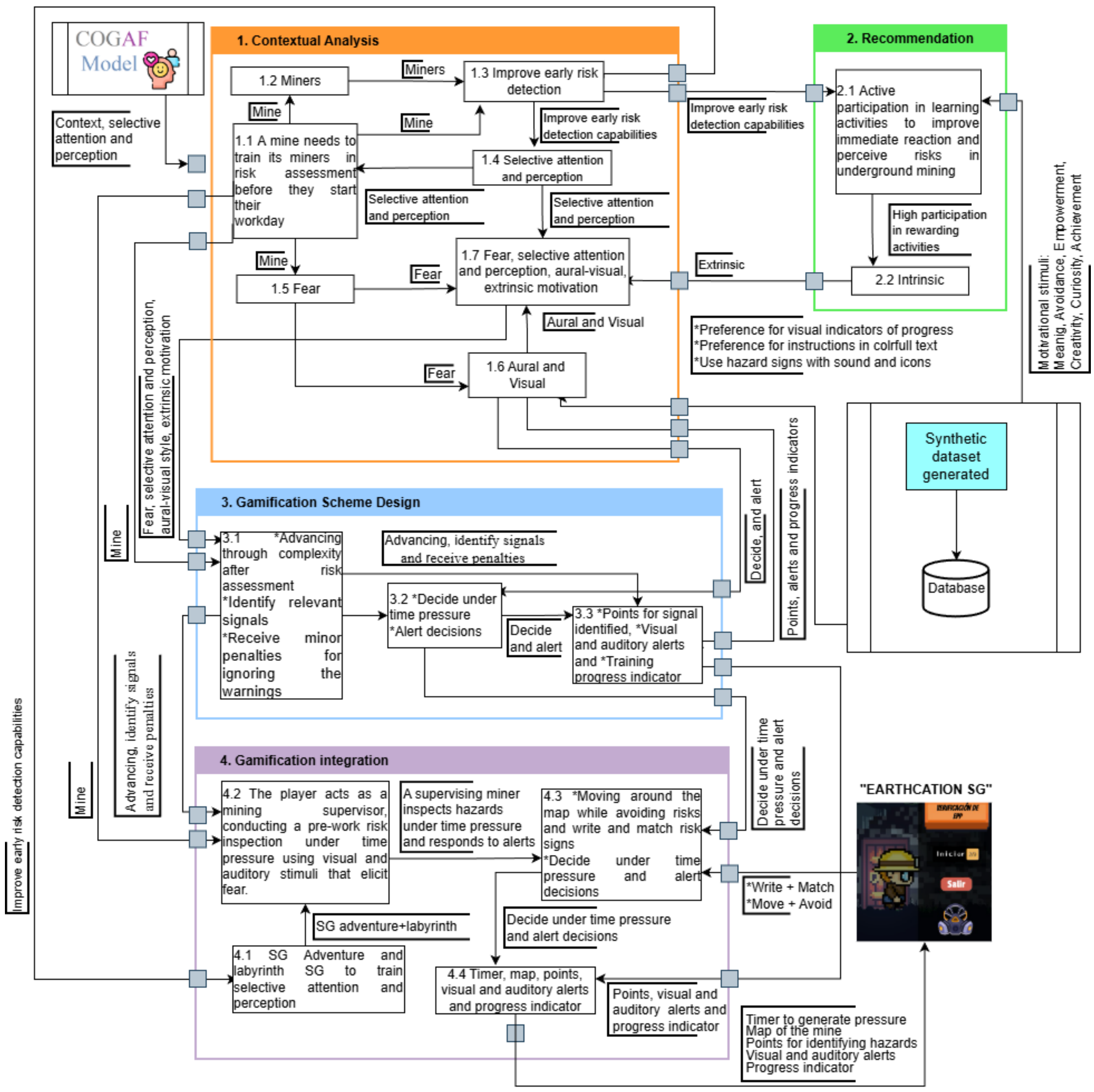

Figure 6: Architecture instantiation using an SG for selective attention and perception training

6 Validation Process

The ArSGam architecture was previously validated through expert judgment using the ISO/IEC 25010 quality model [10], [31]. Since this study employs the same architecture without structural modifications, no additional architectural validation was conducted. Therefore, the present study focuses on validating the synthetic data generated to support the recommendation process. The objective of the synthetic dataset was not to validate a classifier's predictive capability or to demonstrate automatic inference of cognitive-affective profiles. Instead, its purpose was to verify the operational consistency of the extended ArSGam modules under controlled conditions. In this process, GPT-5 was employed to generate input profiles following predefined constraints, whereas the interpretation and operationalization of these profiles remained under the responsibility of the architecture modules. Consequently, the present validation does not assess the accuracy of predictive inference; rather, it assesses the quality of the synthetically generated profiles and then analyzes whether the expanded modules can consistently process and operationalize cognitive-affective profiles under controlled conditions. This validation was conducted through an exploratory reliability analysis and a comparison of distributions against real student profiles.

6.1 Exploratory analysis

For this analysis, a dataset composed of 15 real student profiles reported in [45] and 30 synthetic profiles presented in Table 2 was considered. Due to the categorical nature of the variables Emotion, Core Drive, and Motivation, a numerical encoding process was performed by assigning a unique value to each category to enable the calculation of Cronbach's Alpha coefficient [49] according to Eq. (1).

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum Var(i)}{Var(Total)}\right) \quad (1)$$

Where (k) corresponds to the number of analyzed variables, Var(i) represents the variance of each variable, and Var(Total) denotes the variance of the total score. In this study, (k=3), corresponding to the variables Emotion, Core Drive, and Motivation. For the Emotion variable, the following numerical identifiers were assigned: Fear = 1, Sadness = 2, Disgust = 3, Anger = 4, Surprise = 5, and Happiness = 6. For Core Drive, the encoding Achievement = 1, Empowerment = 2, Meaning = 3, Ownership = 4, Social Influence = 5, Curiosity = 6, Scarcity = 7, and Avoidance = 8 was adopted. Finally, the Motivation variable was encoded as Extrinsic = 1 and Intrinsic = 2.

This transformation was performed solely for exploratory reliability purposes and does not imply the existence of ordinal relationships among categories. Since the analyzed variables are categorical rather than items on a traditional psychometric scale, Cronbach's Alpha was not used to infer the dataset's psychometric properties. Instead, it was used as an exploratory indicator to examine whether the synthetic dataset preserved variability patterns comparable to those observed in the original data. The results showed Cronbach's Alpha values of 0.369 for the real dataset and 0.344 for the synthetic dataset, yielding a difference of 0.025 between the two. Although both values indicate low internal consistency, this behavior is expected given the categorical nature of the analyzed variables. Nevertheless, the proximity between the coefficients suggests that the synthetic dataset preserves variability patterns similar to those observed in the original profiles. Consequently, these findings provide preliminary evidence that the synthetic generation process maintains structural characteristics comparable to those in the reference profiles and may serve as a useful complementary source of information for controlled testing and validation within the proposed architecture.

6.2 Descriptive Comparison

In this section, a descriptive comparison of the frequency distributions of the categorical variables Emotion, Core Drive, and Motivation is carried out between the real and synthetic datasets to assess the degree of similarity between the two. For each variable, the absolute and relative frequencies of its categories were calculated in both datasets. This analysis enabled the determination of whether the synthetic profiles retained proportions similar to those observed in the real profiles. The relative frequency was obtained by dividing the absolute frequency of each category by the total number of records in the corresponding dataset. Subsequently, the results were expressed as percentages to facilitate comparison between the 15 real profiles and the 30 synthetic profiles.

This analysis enabled assessment of whether the proportions of the synthetically generated categories retain patterns similar to those observed in the reference real data. Accordingly, Tables 4–6 summarize the descriptive comparison between the real and synthetic datasets by presenting the frequency distributions of Motivation, Emotion, and Core Drive Octalysis, respectively. The results for each variable are discussed individually below.

Table 4: Motivation Distribution in Real and Synthetic Data

Category	realData Frequency	Real Percentage	Synthetic Frequency	Synthetic Percentage
Extrinsic	12	80	14	46,6
Extrinsic, Intrinsic	1	6,67	0	0
Intrinsic	2	13,33	16	53,33

Table 5: Emotion Distribution in Real and Synthetic Data

Category	realData Frequency	Real Percentage	Synthetic Frequency	Synthetic Percentage
Anger	1	6,67	5	16,67
Disgust	1	6,67	5	16,67
Fear	3	20	5	16,67
Happiness	5	33,33	5	16,67
Sadness	1	6,67	5	16,67
Surprise	4	26,67	5	16,67

Table 6: Core Drive Distribution in Real and Synthetic Data

Category	Real Frequency	Real %	Synthetic Frequency	Synthetic %
Achievement	0	0	4	13,33
Avoidance	4	26,67	4	13,33
Avoidance, Impatience	1	6,67	0	0
Creativity	1	6,67	0	0
Creativity, Social	1	6,67	0	0
Curiosity	0	0	3	10

Empowerment	5	33,33	4	13,33
Impatience	2	13,33	0	0
Meaning	0	0	4	13,33
Ownership	0	0	5	16,67
Scarcity	0	0	3	10
Social Influence	1	6,67	3	10

Table 4 shows that Extrinsic motivation predominates in the real dataset, accounting for 80% of the profiles, whereas this category represents 46.6% of the synthetic profiles. Conversely, Intrinsic motivation increases from 13.33% in the real data to 53.33% in the synthetic dataset. Although the distributions are not identical, the synthetic dataset preserves the presence of both motivational categories while exhibiting a more balanced distribution between intrinsic and extrinsic motivation. These findings suggest that the generation process is capable of representing different motivational profiles while reducing the concentration observed in the original sample.

Table 5 shows that the real dataset presents a heterogeneous distribution of emotions, with Happiness (33.33%), Surprise (26.67%), and Fear (20%) being the most frequent categories. In contrast, the synthetic dataset distributes the six considered emotions uniformly, assigning approximately 16.67% to each category. Despite these differences in proportions, all emotional categories identified in the real profiles are preserved in the synthetic dataset. This behavior suggests that the generation process retains the coexistence of the emotional states observed in the original data while promoting a more balanced representation of the categories and reducing potential biases associated with the small size of the real sample. These findings are consistent with the subsequent Chi-square analysis, which revealed no statistically significant differences between the emotional distributions of the real and synthetic profiles.

Regarding the Core Drives (see Table 6), the real dataset is primarily concentrated in Empowerment (33.33%) and Avoidance (26.67%), indicating that these motivational drivers predominate in the reference profiles. In contrast, the synthetic dataset exhibits a broader distribution across the Octalysis drivers, particularly Ownership (16.67%), Achievement (13.33%), Avoidance (13.33%), Empowerment (13.33%), Curiosity (10%), Scarcity (10%), and Social Influence (10%). Although some categories identified in the real profiles are not represented in the synthetic dataset, the generation process preserves the presence of the most representative drivers, namely Empowerment and Avoidance, while introducing additional variability in the distribution of motivational drivers. Consequently, the synthetic profiles capture a wider range of gamification motivators, although these distributional differences should be interpreted cautiously and further examined through statistical comparison.

In summary, the frequency distribution analysis indicates that the synthetic profiles preserve several structural characteristics observed in the real data, particularly the coexistence of different emotional states, motivational types, and Core Drives. Although differences were observed in the proportions of specific categories, the synthetic dataset maintains the diversity of the original profiles and avoids excessive concentration in any single category. Therefore, these findings provide preliminary evidence that the synthetic generation process can produce profiles structurally comparable to the reference data and may serve as a complementary source of information for controlled testing and validation processes in cognitive-affective personalization scenarios in SGs.

6.3 Statistical Comparison

Although the descriptive analysis provides an initial understanding of the similarities between the real and synthetic datasets, an additional statistical analysis was conducted to determine whether the observed differences in category frequencies were statistically significant. Since Emotion, Motivation, and Core Drive Octalysis correspond to categorical variables, tests of normality such as Kolmogorov-Smirnov were not considered appropriate. Therefore, Chi-square tests of homogeneity were performed to compare the distributions of the real and synthetic profiles, as this test is recommended for evaluating differences in frequency distributions between categorical populations [50].

For each categorical variable (Emotion, Motivation, and Core Drive Octalysis), the following hypotheses were formulated:

- Null hypothesis (H0): The frequency distribution of the variable is the same in the real and synthetic datasets.
- Alternative hypothesis (H1): The frequency distribution of the variable differs between the real and synthetic datasets.

The null hypothesis was rejected when the p-value was lower than the significance level of $\alpha = 0.05$.

Table 7: Results of the Chi-Square Tests of Homogeneity for Real and Synthetic Profiles

Variable	**χ^2**	**p-value**	**Interpretation**
Emotion	4.0625	0.5405	No significant differences
Motivation	7.9231	0.1901	No significant differences
Core Drive Octalysis	22.6250	0.0200	Exploratory due to assumption violations

As shown in Table 7, no statistically significant differences were found for the Emotion variable ($\chi^2 = 4.0625$, $p = 0.5405$), suggesting that the synthetic profiles preserved the distribution of emotional states observed in the real data. That is similar for the case of Motivation ($\chi^2 = 7.9231$, $p = 0.1901$). Specifically, the synthetic dataset exhibited a more balanced representation of motivational profiles than the original sample. In contrast, statistically significant differences were identified for Core Drive Octalysis ($\chi^2 = 22.6250$, $p = 0.0200$), indicating that the synthetic generation process produced alternative distributions for this dimension. However, due to the small sample size and the large number of expected frequencies below five for the Core Drive Octalysis variable, violating one of the assumptions of the Chi-square test, these latter findings should be interpreted with caution and considered as exploratory and complementary evidence to the descriptive analysis, rather than conclusive evidence of population representativeness.

7. Results

This section presents the main findings obtained. It includes a discussion of the results, an analysis of the evaluation process, and a comparison with related works. The ArSGam architecture was previously validated through expert judgment using the ISO/IEC 25010 quality model [10], [31]. As this study employs the same architecture without modifications, no further architectural validation was conducted.

7.1 Results Discussion

The modular architecture (see Figure 1) provides a structured and coherent workflow for integrating gamification from module 1 to module 4. Its successful application in two different case studies, without the need for architectural modifications, demonstrates both its flexibility and robustness. In the first case study [22], the architecture was applied in a mining context, focusing on *spatial memory and emotional regulation* during emergency evacuation situations. In contrast, the new case study emphasizes different cognitive aspects, specifically *selective attention and perception*, during pre-shift mining hazard inspections in an online training environment. A comparative analysis of the COGAF specification tables in both case studies shows that, despite differences in learning objectives, cognitive functions, emotional triggers, and game mechanics, the architecture consistently supports the characterization of COGAF components. While the original case emphasized *navigation, fear regulation, and spatial awareness*, the new case focuses on *signal discrimination, hazard detection, and time-constrained decision-making*. This contrast highlights that the architecture can systematically adapt COGAF elements, such as the cognitive functions, emotional s tates, mechanics, and interaction patterns, to the training context without altering their internal structure.

The results of the new case study confirm that the architecture enables the design of SGs that are meaningful on both cognitive and emotional levels. By profiling users based on observed behaviors, inferred motivations, and dominant gamification drivers, the architecture avoids the arbitrary use of rewards or points. Instead, each component of the game is directly linked to a pedagogical, cognitive, or emotional objective. This human-centered approach ensures consistency across modules, aligning user needs with game mechanics and emotional outcomes. Overall, the consistent application of the architecture in two heterogeneous mining scenarios validates its flexibility and generalizability. The results indicate that the architecture is not limited to a single cognitive function but can be confidently adapted to other contexts, such as health, education, politics, or safety-critical industries, while maintaining methodological rigor and design coherence.

7.2 Objectivity Verification

As mentioned earlier, an evaluation based on the ISO/IEC 25010 standard was conducted, which had been previously validated by experts on the original ArSGam architecture [31]. Since this study used the same architectural proposal, expert validation was not repeated. The results of the first case study showed a high level of quality, with a total score of 47 out of 50 in functionality, usability, maintainability, compatibility, and portability, with maintainability standing out with a perfect score.

7.3 Comparison with Related Works

Below, we present a comparison with previous work, considering the following criteria: In what **domain** it has been applied, what its **contribution** is, whether it proposes a **formal architecture**, and what is the **challenge** it aims to address. This Table extends the comparative analysis done in paper [31].

Table 8: Comparison of Related Work

Work	Domain	Contribution	Formal architecture	Challenge
[19]	Environmental awareness in students	Integration of gamification elements in SG environment (RAISE)	Not	Does not distinguish structured scheme; mixes elements in an empirical way
[20]	Rehabilitation of children with special needs	Computational platform + conceptual framework for therapeutic games	Parcial (conceptual + plataforma)	Complex design of attractive and effective games for multiple clinical profiles
[6]	Mobile learning	Gamification elements taxonomy	Not	Difficulty in representing abstract dynamics in game design
[17]	General Education	Conceptual framework for integrating SG and gamification	Not	Balancing game mechanics with educational objectives
[21]	Attention to children with ADHD	Gamification architecture with integrated game types and methods	Yes	ADHD friendly and adaptable design, maintain motivation in the long term
[21]	Urban mobility	Ubiquitous architecture with SG and gamification layer	Yes	Design attractive and effective games for different user profiles in real contexts.
[11]	Adaptive game-based learning	Systematic review of adaptive game-based learning	Not	Need for personalization frameworks integrating cognitive-affective profiling

Work	Domain	Contribution	Formal architecture	Challenge
[13]	Personalized learning	Literature review on generative AI + gamification	Not	Challenge of integrating AI personalization into gamification strategies
[14]	Adaptive learning systems	Framework combining game mechanics with AI personalization	Yes	Scalability and adaptability across diverse learning contexts
[15]	SG with digital twins	Architecture for real-time self-adaptation	Yes	Ensuring effective adaptation to learner preferences in immersive environments
[16]	AI and immersive technologies	Systematic review of SG in AI/Metaverse contexts	Not	Need for meta-skills training and integration of advanced technologies
ArSGam	**SG for cognitive-affective training**	**Integration of gamification scheme**	**Yes**	**Profiling users to establish a customized gamification scheme**

In general, these studies contribute to integrating gamification elements and developing specific platforms, conceptual schemes, and architectures oriented to particular domains. However, most of them lack a formalized, adaptable architecture that allows SG customization across different contexts, user profiles, and pedagogical objectives, an aspect our proposal seeks to address in a structured manner. As shown in Table 8, there is a clear need for an architecture that integrates a gamification scheme into SG aimed at cognitive-affective training. While the reviewed works contribute significantly to educational, therapeutic, or awareness-raising contexts, few address the structured personalization of the gaming experience based on user profiling.

Our proposal addresses this challenge by incorporating user profiling to personalize the gamification scheme. This approach dynamically adapts game mechanics, dynamics, and elements to the player's cognitive and emotional states, fostering a more effective and engaging learning experience. Furthermore, by situating ArSGam alongside recent contributions such as adaptive frameworks [11], AI-driven personalization [14], and self-adaptive architectures [15], the distinctiveness of our proposal becomes evident. Unlike purely conceptual or domain-specific approaches, ArSGam operationalizes cognitive-affective profiling within a modular architecture, bridging theoretical models like COGAF with practical implementation. This positions ArSGam as a robust and versatile framework that not only addresses personalization but also anticipates integration with emerging technologies such as AI and immersive environments.

7 Conclusions

This article extends the analysis of ArSGam as a modular architecture that allows the integration of gamification schemes into SGs for cognitive-affective training. The proposal achieves its objective by considering a novel case study generated using synthetic data produced with LLMs and by conducting an in-depth analysis of related previous work. The results and findings of the article lead to the conclusion that ArSGam provides a formalized and adaptable structure applicable in different contexts. ArSGam enables the design and integration of gamification into SGs based on user profiles, allowing for the personalization of experiences based on cognitive, affective, emotional, and motivational profiles.

Beyond these contributions, this extended version situates ArSGam within the broader research landscape, comparing it with recent frameworks that emphasize adaptive gamification, AI-driven personalization, and immersive technologies. This comparison highlights ArSGam's distinctive role in operationalizing cognitive-affective profiling through a modular architecture that bridges theoretical models such as COGAF with practical implementation. By doing so, ArSGam not only addresses personalization challenges but also anticipates integration with emerging trends in SG, reinforcing its relevance as a robust and versatile framework.

As a limitation, although the architecture is flexible with respect to learning styles, its current implementation is based solely on the VARK model, so it would be necessary to incorporate and validate additional models. In addition, the architecture has been validated in a specific context (mining sector), suggesting the need to evaluate it in other domains to confirm its general applicability.

As future work, ArSGam will be extended to incorporate multiple user profiling models, including psychometric instruments that allow the collection of specific assessment items and enable more rigorous analyses of internal consistency and distributional similarity through measures such as Cronbach's Alpha and, where appropriate, non-parametric tests. Additionally, adaptive learning technologies and AI-driven personalization mechanisms will be integrated to dynamically adjust game mechanics and content based on continuous data analysis. Another research direction involves extending the architecture to support immersive environments such as Virtual Reality and Augmented Reality, where cognitive-affective profiles can be combined with multimodal interaction to promote deeper engagement. Furthermore, large-scale validations will be conducted across domains, including STEM education, healthcare, and professional training, to further assess the framework's generalizability. The interoperability of ArSGam with digital twin architectures will also be explored to enhance adaptability across heterogeneous platforms and user populations. Future studies will also include blind tests with educational psychology experts to examine whether synthetic and real profiles can be distinguished, thereby providing complementary qualitative evidence regarding the fidelity of the generated profiles. Finally, mechanisms for simulating psychometric noise, incomplete information, and conflicting cognitive-affective profiles will be investigated to evaluate the architecture's robustness and reliability under more realistic, heterogeneous conditions.

References

[1] Z. Liang, K. Zhou, and K. Gao, "Development of Virtual Reality Serious Game for Underground Rock-Related Hazards Safety Training," *IEEE Access*, vol. 7, 2019, doi: 10.1109/ACCESS.2019.2934990.

[2] F. R. López, M. Arias-Oliva, J. Pelegrín-Borondo, and L. M. Marín-Vinuesa, "Serious games in management education: An acceptance analysis," *International Journal of Management Education*, vol. 19, no. 3, 2021, doi: 10.1016/j.ijme.2021.100517.
[3] M. Carrion, M. Santorum, M. Perez, and J. Aguilar, "A participatory methodology for the design of serious games in the educational environment," in *2017 Congreso Internacional de Innovacion y Tendencias en Ingenieria, CONIITI 2017 - Conference Proceedings*, 2018. doi: 10.1109/CONIITI.2017.8273363.
[4] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: Defining 'gamification,'" in *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, MindTrek 2011*, 2011. doi: 10.1145/2181037.2181040.
[5] G. Baptista and T. Oliveira, "Gamification and serious games: A literature meta-analysis and integrative model," 2019. doi: 10.1016/j.chb.2018.11.030.
[6] M. M. Elaish, M. H. Hussein, L. Shuib, W. F. Wan Ahmad, and K. Becker, "A Proposed Gamification Elements of Educational Games," in *Proceedings - International Conference on Computer and Information Sciences: Sustaining Tomorrow with Digital Innovation, ICCOINS 2021*, 2021. doi: 10.1109/ICCOINS49721.2021.9497179.
[7] M. Ortiz-Rojas, K. Chiluiza, and M. Valcke, "Gamification in computer programming: Effects on learning, engagement, self-efficacy and intrinsic motivation," in *Proceedings of the 11th European Conference on Games Based Learning, ECGBL 2017*, 2017.
[8] A. Matallaoui, P. Herzig, and R. Zarnekow, "Model-driven serious game development integration of the gamification modeling language GaML with unity," in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2015. doi: 10.1109/HICSS.2015.84.
[9] C. Gómez, P. Vallejo, and J. Aguilar, "Specification of users' cognitive functions and emotions to promote their training through Serious games," *Entertain. Comput.*, vol. 53, p. 100921, Mar. 2025, doi: https://doi.org/10.1016/j.entcom.2025.100921.
[10] ISO/IEC 25000, "ISO 25000 PORTAL." Accessed: May 10, 2025. [Online]. Available: https://www.iso25000.com/index.php/normas-iso-25000/iso-25010
[11] D. Chiotaki, V. Poulopoulos, and K. Karpouzis, "Adaptive game-based learning in education: a systematic review," 2023. doi: 10.3389/fcomp.2023.1062350.
[12] W. Elgammal, N. Sherief, and W. Abdelmoez, "A framework for developing user-based adaptive gamified systems," *Advances in Computing and Engineering*, vol. 3, no. 2, 2023, doi: 10.21622/ace.2023.03.2.024.
[13] F. Abbes, S. Bennani, and A. Maalel, "Generative AI and Gamification for Personalized Learning: Literature Review and Future Challenges," 2024. doi: 10.1007/s42979-024-03491-z.
[14] F. Naseer, M. N. Khan, A. Addas, Q. Awais, and N. Ayub, "Game Mechanics and Artificial Intelligence Personalization: A Framework for Adaptive Learning Systems," *Educ. Sci. (Basel).*, vol. 15, no. 3, 2025, doi: 10.3390/educsci15030301.
[15] S. Loizou and A. S. Andreou, "A Framework for Standardizing the Development of Serious Games with Real-Time Self-Adaptation Capabilities Using Digital Twins," *Technologies (Basel).*, vol. 13, no. 8, 2025, doi: 10.3390/technologies13080369.
[16] E. Mitsea, A. Drigas, and C. Skianis, "A Systematic Review of Serious Games in the Era of Artificial Intelligence, Immersive Technologies, the Metaverse, and Neurotechnologies: Transformation Through Meta-Skills Training," 2025. doi: 10.3390/electronics14040649.
[17] B. H. Yousefi and H. Mirkhezri, "Toward A Game-based Learning Platform: A Comparative Conceptual Framework for Serious Games," in *Proceedings of the 2019 International Serious Games Symposium, ISGS 2019*, 2019. doi: 10.1109/ISGS49501.2019.9046979.
[18] R. Drakoulis *et al.*, "A Gamified Flexible Transportation Service for On-Demand Public Transport," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 3, 2018, doi: 10.1109/TITS.2018.2791643.
[19] D. Minas, M. Xenos, and D. Papatheodorou, "A Serious Game for Environmental Education Utilizing 3D Virtual Worlds," in *2024 IEEE Gaming, Entertainment, and Media Conference, GEM 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/GEM61861.2024.10585526.
[20] A. Antunes, R. N. Madeira, I. Tóth, and C. Bernardo, "Modeling serious games design towards engaging children with special needs in therapy," in *CHI PLAY 2021 - Extended Abstracts of the 2021 Annual Symposium on Computer-Human Interaction in Play*, Association for Computing Machinery, Inc, 2021, pp. 192–197. doi: 10.1145/3450337.3483493.
[21] A. S. Putra, H. L. H. S. Warnars, B. S. Abbas, A. Trisetyarso, W. Suparta, and C. H. Kang, "Gamification in the e-Learning Process for children with Attention Deficit Hyperactivity Disorder (ADHD)," in *1st 2018 Indonesian Association for Pattern Recognition International Conference, INAPR 2018 - Proceedings*, 2018. doi: 10.1109/INAPR.2018.8627047.
[22] C. Gómez, P. Vallejo, and J. Aguilar, "A Generic Metamodel for Cognitive-Affective Training of Users Using Serious Games," in *2023 XLIX Latin American Computer Conference (CLEI)*, IEEE, Oct. 2023, doi: 10.1109/CLEI60451.2023.10346132.
[23] J. Cordero, J. Aguilar, K. Aguilar, D. Chávez, and E. Puerto, "Recognition of the driving style in vehicle drivers," *Sensors (Switzerland)*, vol. 20, no. 9, 2020, doi: 10.3390/s20092597.
[24] J. Monsalve-Pulido, J. Aguilar, E. Montoya, and C. Salazar, "Autonomous recommender system architecture for virtual learning environments," *Applied Computing and Informatics*, vol. 20, no. 1–2, 2024, doi: 10.1016/j.aci.2020.03.001.
[25] J. Aguilar and G. Garcia, "An Adaptive Intelligent Management System of Advertising for Social Networks: A Case Study of Facebook," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 1, pp. 20-32, 2018, doi: 10.1109/TCSS.2017.2759188
[26] A. Kanojiya, "The Learning Style Preferences of the Undergraduate Computer Applications and IT Students," *International Journal of Scientific Research in Engineering and Management (IJSREM)*, vol. 07, no. 03, 2023.

[27] J. Aguilar, F. Díaz, J. Altamiranda, J. Cordero, D. Chavez, and J. Gutierrez, "Metropolis: Emergence in a Serious Game to Enhance the Participation in Smart City Urban Planning," *Journal of the Knowledge Economy*, vol. 12, no. 4, 2021, doi: 10.1007/s13132-020-00679-5.
[28] L. N. M. Velandia and L. A. Plazas-Gómez, "Agentes inteligentes y Modelo VARK, proponen estrategias de aprendizaje según la manera en que asimila un individuo," *Revista Educación en Ingeniería*, vol. 13, no. 26, 2018.
[29] C. K. Chang, T. Y. Chuang, and W. L. Kuo, "Relationships between engagement and learning style for using VPL on game design," in *Proceedings 2012 4th IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning, DIGITEL 2012*, 2012. doi: 10.1109/DIGITEL.2012.43.
[30] T. Tene, D. F. Vique López, P. E. Valverde Aguirre, N. I. Cabezas Oviedo, C. Vacacela Gomez, and S. Bellucci, "A systematic review of serious games as tools for STEM education," 2025. doi: 10.3389/feduc.2025.1432982.
[31] C. Gómez, P. Vallejo, and J. Aguilar, "ArSGam: A Modular Architecture for Integrating Gamification into Serious Games," in *2025 LI Latin American Computer Conference (CLEI)*, 2025.
[32] L. Itti and P. Baldi, "Bayesian surprise attracts human attention," *Vision Res.*, vol. 49, no. 10, 2009, doi: 10.1016/j.visres.2008.09.007.
[33] C. M. Tyng, H. U. Amin, M. N. M. Saad, and A. S. Malik, "The influences of emotion on learning and memory," 2017. doi: 10.3389/fpsyg.2017.01454.
[34] T. 22; Steimer, "The biology of fear- and anxiety-related behaviors.," *Dialogues Clin. Neurosci.*, vol. 4, no. 3, 2002.
[35] Y. Liu *et al.*, "The trait anger affects conflict inhibition: A Go/Nogo ERP study," *Front. Hum. Neurosci.*, vol. 8, no. JAN, 2015, doi: 10.3389/fnhum.2014.01076.
[36] S. L. Beilock and T. H. Carr, "When high-powered people fail: Working memory and 'Choking under pressure' in math," *Psychol. Sci.*, vol. 16, no. 2, 2005, doi: 10.1111/j.0956-7976.2005.00789.x.
[37] L. Pessoa, "On the relationship between emotion and cognition," 2008. doi: 10.1038/nrn2317.
[38] A. Diamond, "Executive Functions," *Annu. Rev. Psychol.*, vol. 64, p. 135, 2013, doi: 10.1146/annurev-psych-113011-143750.
[39] F. Marisa, S. S. Syed Ahmad, Z. I. Mohd Yusoh, A. L. Maukar, R. D. Marcus, and A. A. Widodo, "Evaluation of Student Core Drives on e-Learning during the Covid-19 with Octalysis Gamification Framework," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 11, 2020, doi: 10.14569/IJACSA.2020.0111114.
[40] MINENERGÍA, "Ministerio de Minas y Energía de Colombia." Accessed: Jan. 22, 2026. [Online]. Available: https://www.minenergia.gov.co/es/sala-de-prensa/noticias-index/plataforma-digital-ha-formado-a-más-de-1000-mineros-en-colombia/
[41] ICCMAA, "Instituto Colombiano de Ciencias Mineras, Agrarias y Ambientales." Accessed: Jan. 22, 2026. [Online]. Available: https://www.iccmaa.com/
[42] Intégrame, "Infraestructura de datos del sector Minero Energético colombiano." Accessed: Jan. 22, 2026. [Online]. Available: https://www.integrame.gov.co/
[43] ANM, "Datos Abiertos Agencia Nacional Minera Colombiana." Accessed: Jan. 22, 2026. [Online]. Available: https://www.anm.gov.co/seleccion-de-datos-abiertos
[44] A. Pandita, A. Keniston, and N. Madhuripan, "Synthetic data trained open-source language models are feasible alternatives to proprietary models for radiology reporting," *NPJ Digit. Med.*, vol. 8, no. 1, 2025, doi: 10.1038/s41746-025-01658-3.
[45] J. Huang *et al.*, "A critical assessment of using ChatGPT for extracting structured data from clinical notes," *NPJ Digit. Med.*, vol. 7, no. 1, 2024, doi: 10.1038/s41746-024-01079-8.
[46] E. Ko and P. Kang, "Evaluating Coding Proficiency of Large Language Models: An Investigation Through Machine Learning Problems," *IEEE Access*, vol. 13, 2025, doi: 10.1109/ACCESS.2025.3553870.
[47] F. Marisa *et al.*, "Customer motivation analysis on retail business with octalysis gamification framework," *J. Theor. Appl. Inf. Technol.*, vol. 99, no. 13, 2021.
[48] R. Rodrigues Major and M. Mira da Silva, "Evaluating a gamified MOOC," *Cogent Education*, vol. 12, no. 1, 2025, doi: 10.1080/2331186X.2025.2479400.
[49] M. Tavakol and R. Dennick, "Making sense of Cronbach's alpha," 2011. doi: 10.5116/ijme.4dfb.8dfd.
[50] A. Agresti, *Book - An Introduction to Categorical Data Analysis*. 2019.