Where Does Value Take Shape? An AI–Human Serious Game Design Experiment on the Poème Électronique

Vittorio Murtas[1,*,†], Vincenzo Lombardo[1,†]

[1]Department of Computer Science, University of Turin, Via Pessinetto 12, 10149, Turin, Italy

Abstract

This article presents a design experiment investigating whether and how cultural heritage values may emerge in the development of a serious game created through generative AI. Centered on *Edgar à GoGo*—a game based on Edgar Varèse's *Poème Électronique*—the study explores whether meaningful aspects of the heritage item can take shape in the game's components, even in the absence of explicitly formalized values. The game concept was generated using ChatGPT within the Co.Lab design framework, followed by a prototyping phase in which the authors, acting as developers, posed implementation-driven clarification questions. The resulting prototype was evaluated by experts familiar with the original work, whose feedback highlighted both latent value alignments and critical omissions. While some values weakly emerged through the AI's design logic, others were absent, prompting reflective discussion. The study proposes a hybrid, iterative approach to value-sensitive design in cultural heritage games, positioning expert evaluation as a crucial layer for value articulation and refinement.

Keywords

generative AI, game design, serious games, cultural heritage

1. Introduction

Serious (or applied) games is the use of full-fledged games in a non-gaming context [1]. Their use has expanded across a wide range of domains—including health [2], heritage [3], social issues [4], and training [5]—thanks to their capacity to engage users through interactive and immersive experiences. In the cultural heritage field, they are increasingly used to promote tourism [6] and deepen public understanding of both tangible and intangible heritage [7]. At the heart of such games lies a key objective: to convey the *cultural heritage values* (CHValues) that communities and experts associate with heritage—whether historical, symbolic, aesthetic, or spiritual.

While the notion of value is rooted in economic and ethical theory, in the heritage context it plays a central—albeit contested—role. CHValues are used by institutions to justify conservation decisions through structured assessments that highlight dimensions such as historical, aesthetic, or spiritual relevance: the set of values attributed to heritage items [8]. For instance, the Arch of Constantine in Rome is valued not only for its historical importance but also for its symbolic and spiritual associations.

In the design of serious games for cultural heritage, the drive to maximize engagement and interactivity [9] can easily overshadow the deeper cultural, symbolic, or emotional meanings associated with heritage items. This creates a risk of what we might term *value dilution*, where heritage becomes a decorative backdrop. While other types of instructional content—such as historical facts or procedural knowledge—are routinely mapped onto game elements through structured design frameworks [10], values pose a different challenge. They are often implicit, context-dependent, and difficult to encode, yet they are crucial to understanding why heritage matters. Despite their importance, most design frameworks for serious games in cultural heritage—namely, guidelines and structures that assist designers in shaping the gameplay experience [11]—do not provide explicit mechanisms for incorporating cultural heritage values. Recent reviews have shown that these frameworks primarily emphasize knowledge

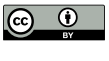
Proceedings of AI4HGI '25, the First Workshop on Artificial Intelligence for Human-Game Interaction at the 28th European Conference on Artificial Intelligence (ECAI '25), Bologna, October 25-30, 2025

*Corresponding author.

†These authors contributed equally.

vittorio.murtas@unito.it (V. Murtas); vincenzo.lombardo@unito.it (V. Lombardo)

0009-0006-8096-8207 (V. Murtas); 0000-0002-8166-9827 (V. Lombardo)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

transmission, player engagement, and technical development, while the systematic integration of values remains largely overlooked [12, 13, 10]. As a result, the inclusion of CHValues is typically left to the discretion of individual designers, which raises concerns about their inconsistent or superficial treatment [8]. This highlights a pressing need for a more deliberate and structured approach to meaningfully embed heritage values within the design process of serious games.

This paper presents an experiment to investigate further the values-mapping topic and, in particular, the emergence of values through the design of a serious game. We adopted the Co.Lab framework [11], a validated and generalizable model [14] applicable to a wide range of serious games, including those focused on heritage. Co.Lab takes a structuralist approach, dividing the design into discrete components—such as learning mechanics, game mechanics, rules, goals, and interface—each supported by clear definitions and practical guidelines. It is intended as a collaborative tool to guide multidisciplinary teams.

To ensure control, rapid iteration, and interpretive neutrality, we used ChatGPT as a surrogate game designer, guiding it to fill the Co.Lab framework. The use of AI allowed us to observe where and how values might emerge during concept generation in a reproducible and bias-controlled setting.

We applied this approach to a case study: *Edgar à Gogo*, a game based on Edgar Varèse's *Poème Électronique*, an electroacoustic work premiered at the 1958 Brussels World's Fair in the Philips Pavilion. The show presented in the pavilion included images and lighting effects (Le Corbusier), "organized sound" from 350 speakers on "sound paths" (Edgar Varèse), and Yannis Xenakis's hyperbolic paraboloid-shaped walls. As Varèse's only purely electronic composition, the piece was built through tape manipulation and spatial diffusion. Although it lacks a traditional score, its structure reflects many of the compositional principles found in his instrumental works, particularly in the use of contrast, gesture, and spatialized sound [15].

The game concept document was generated by ChatGPT based on its knowledge of the piece and later implemented by the authors. Our aim was to analyze whether and where cultural values could emerge in an AI-generated design without being explicitly supplied. The study followed a three-phase structure:

1. *Concept generation* via Co.Lab with ChatGPT;
2. *Technical prototyping* through clarification prompts;
3. *Expert evaluation* to identify emergent and missing values.

This study investigates whether a generative AI can autonomously identify and integrate cultural heritage values (CHValues) within the design of a serious game, or whether such values only emerge when experts reinterpret its outputs. Rather than aiming for best practices, we staged a deliberate absence: we instructed a generative AI (ChatGPT) to design a game using the Co.Lab framework, relying on general knowledge about the Poème Électronique. Our goal was to test whether cultural values might still emerge—and where—in the resulting design. By analyzing the prototype and gathering expert feedback, we examined both which values emerged and which were absent, treating both as indicators for guiding future design iterations.

2. Related Works

In the context of heritage, serious games offer diverse approaches to embedding values. In *Never Alone* (E-Line Media, 2014), developed in collaboration with the Iñupiat community (Alaska) and the Cook Inlet Tribal Council, the game conveys strong symbolic, spiritual, and social values through indigenous myths, language, and cosmology, fostering cultural transmission and identity recognition. Serious games like *Attentat 1942* and *Svoboda 1945* (Charles Games, 2017–2021) offer a more explicit integration of historical and communal values. Drawing on survivor interviews, archival materials, and interactive storytelling, they immerse players in ethically and emotionally complex histories of war and occupation. These games foreground the contested nature of memory and identity within society, articulating social and symbolic values that emerge through engagement with lived experiences and personal narratives

rather than abstract historical facts [16, 17]. The presented examples show the effective participation of communities (or their experiences) in the process of values mapping.

These examples illustrate that cultural heritage values can be communicated through a game's narrative, mechanics, visuals, or their integration. However, despite promising cases, such intentional and coherent mappings of values remain rare. Many serious games still prioritize factual accuracy or visual authenticity [9] while neglecting the deeper value systems that give heritage its cultural significance. Other games do not consider communities' values but only the experts' ones. Many other games use cultural heritage as an inspiration, leaving the values interpretation largely to the player. As Goud observes [17], key questions such as *"Why does this heritage matter?"* and *"Whose values are represented?"* are often left unaddressed. This points to a critical disconnect: representing cultural elements does not equate to engaging meaningfully with the values they entail. Without deliberate design strategies that surface, contextualize, and negotiate these values, their inclusion risks being superficial, incidental, or even distorted. There is, therefore, an urgent need for structured frameworks that support developers in systematically identifying, prioritizing, and embedding heritage values into game design—ensuring cultural transmission that is not only engaging but also meaningfully grounded.

In the field of serious games for cultural heritage, few design frameworks exist, and those available—such as those by Andreoli [13] and Antoniou [12]—tend to be high-level, offering minimal guidance, examples, or documentation. They often rely heavily on the designer's intuition and creative effort, leaving room for subjectivity and inconsistent communication of heritage values. In contrast, structured frameworks have emerged in the broader, more generic, serious games field. The Co.Lab framework proposes 21 structured design elements, beginning with problem definition and contextual constraints, followed by the game-based learning solution, and ending with assessment. Co.Lab has been validated, is well-cited, and is available in an online collaborative format. Its modular, systemic approach allows teams to adapt the framework to their needs, supported by a suggested workflow to guide less experienced designers. Crucially, its structure and interdependencies make it a strong candidate for future adaptation to explicitly and systematically integrate cultural heritage values. The output of Co.Lab is a concept document in which the content can be further explored for the development of a prototype.

While structured frameworks like Co.Lab offer a robust foundation for embedding values into serious game design, recent advances in generative AI—particularly Large Language Models (LLMs) such as ChatGPT—provide new tools for supporting specific stages of the design process. LLMs have been applied in procedural content generation, as seen in *MarioGPT* and *GAVEL*, and as design assistants aiding in ideation, prototyping, and mechanic refinement [18, 19, 20]. In these contexts, LLMs are not intended to replace human creativity but to assist it—particularly by reducing the influence of external biases, such as personal preferences toward certain themes or genres (e.g., a designer's attitude toward electronic music). Their ability to respond consistently to structured prompts allows for more controlled experimentation, especially when dealing with culturally sensitive or contested content [21, 22]. For this reason, LLMs offer a useful lens for investigating how cultural values surface in design pipelines, including which components of the Co.Lab framework are engaged when value inputs are made explicit versus when they are omitted, as explored in the experiment presented in this paper.

3. A Serious Game on the *Poème Électronique*

To investigate the mapping of cultural values in game design, our experiment focused specifically on a case study, the musical dimension of the *Poème Électronique* by Edgard Varèse. The piece was composed for the Philips Pavilion at the 1958 Brussels World's Fair, one of the first ever multimedia environments designed by Le Corbusier and Xenakis to integrate architecture, sound, and image. The pavilion itself functioned as a resonant instrument: over 350 loudspeakers were embedded in its curved surfaces, enabling Varèse's electronic sounds to move dynamically through space. Although full documentation of the spatial control is missing, recent reconstruction efforts—such as those by the VEP Project [15, 23]—have relied on archival recordings, sketches, and modern digital tools to approximate the intended spatial effects. These efforts suggest that Varèse approached spatialization as a means of

structuring the piece, using spatial distribution to highlight contrasts in sound material and to define distinct sonic regions within the performance environment. The piece itself consists of a meticulously organized sequence of "sound objects," divided into three mono tapes categorized by Varèse and later by Philips engineers into distinct types with different (often onomatopoeic) labels that reference the original recorded sound sources [15, 24]. A possible taxonomy of the 142 sound fragment names, their duration, and their "sound route" [15] was realized in the context of the VEP project (see an excerpt in Table 1).

Start	Name	Tape	End	Dur	Where
0	bell-1	1	5.4	5.4	grpN+LT
5.4	bell-2	1	11.733	6.333	grpN+LT
11.733	lowbell	1	22.733	11	LT
16.466	wblock-1	3	17.6	1.134	J
...	...	...	...	...	...

Table 1
Sound objects data with timing, source tape, and routing information.

The significance of the musical piece is an ideal testbed for prompting ChatGPT to generate design responses informed by the Co.Lab framework. In this initial phase, the model was provided with minimal factual information about the composition (the fact that it can be segmented into single sounds) and project-specific knowledge from the VEP project (the table discussed above). No explicit definition of cultural values was provided—allowing us to observe which value-related elements, if any, would emerge spontaneously in the generated game concept. Rather than tasking the model to invent a new heritage scenario, we constrained it to an existing artwork to isolate how values emerge within a predefined cultural framework. The outputs from this phase will serve as a baseline for comparison in a subsequent iteration of the experiment, in which clearly articulated values—derived from the analysis of the current output—will be made explicit to the model. This approach not only tests the creative affordances of generative AI but also informs broader reflections on how cultural values can be encoded—and decoded—within design tools.

3.1. Methodology

Our methodology involves the use of ChatGPT-4o to generate a concept document for a serious game based on the *Poème Électronique*, using structured prompts informed by the workflow proposed in the Co.Lab framework (Phase 1). Once the initial concept is produced, a series of follow-up prompts are issued to ChatGPT to clarify specific design elements relevant to the development of a working prototype, which is then implemented in Unity 6 (Phase 2). Finally, the resulting prototype is evaluated by three expert reviewers, who are asked to reflect on both the values represented within the game and those that could be potentially incorporated into the design process. The following sections detail the experimental design and its three phases.

Our experimental design is structured around role-based interactions between ChatGPT and the authors, following the approach proposed by Tyni et al. [21]. To scaffold the design process, we adopted the use of *personas*—fictional professional profiles—to guide the model's behavior. In both phases of the experiment, interactions with ChatGPT were consistently initiated using the role-based instruction: *"Act as a professional game designer working within the Co.Lab framework for serious games. Each prompt will restrict your task to defining components of the Co.Lab framework."*.

While Phase 1 focuses on generating core design components such as mechanics, interactions, and feedback systems, Phase 2 extends the model's role to cover aesthetic and artistic decisions, aligning more closely with the responsibilities of an art director. This role shift is necessary to support the implementation of a playable prototype and ensure visual coherence with the game's conceptual foundations.

3.1.1. Phase 1 – Concept Generation

The first phase of the study focused on generating a complete game concept by systematically applying the Co.Lab framework through a sequence of structured prompts. Rather than engaging in open-ended dialogue, we adopted a modular prompt–response format. To standardize the interaction, every prompt includes the following core components:

- **Step:** The specific Co.Lab component to be generated (e.g., Context, Game Universe, Learning Goals). It defines the design focus for the current prompt.
- **Example:** A short sample output taken from another project or scenario, used to illustrate the expected structure, level of detail, and tone of the response.
- **Definition:** The official Co.Lab description of the selected component. It clarifies the purpose and scope of the element to be designed and often explains how it contributes to the overall framework.
- **Guidelines:** A list of targeted instructions that restrict or direct the content of the response. These focus on specific sub-dimensions to include or exclude, helping to reduce ambiguity and guide the AI toward a useful and implementable output.
- **Key Questions:** A set of focused, directive questions used to prompt reflection on relevant aspects of the design element. These questions are aligned with the guidelines and help structure the AI's reasoning.
- **Given Knowledge:** Optional, project-specific background information and constraints that the model should take into account. These details replace the need for assumptions and allow the model to produce a grounded and context-aware output.

To prevent memory contamination across prompts, we used ChatGPT in temporary session mode, which clears all contextual memory at the end of each session. All prompts and responses were systematically recorded[1] to ensure transparency and allow for future replication. It is worth noting that the "Given Knowledge" field was used only in the initial "Context" step, where the authors provided the taxonomy of sound fragments derived from the VEP project, along with a few other design constraints.

Each Co.Lab component was generated once and re-prompted only to clarify underspecified items; no multiple creative iterations were run. The only Co.Lab component intentionally omitted for this experiment was the "Pedagogical Scenario," as suggested by the framework's authors, since the game is not intended for use within a formal educational setting. Once all components were completed, we generated six illustrative storyboards to simulate the user experience.

3.1.2. Phase 2 - Technical Prototyping

The second phase focuses on the development of a functional prototype using Unity, as suggested by ChatGPT itself during the concept phase. The goal is to implement a minimal version of the game capable of demonstrating core mechanics and allowing expert users to experience and evaluate the design. This phase follows an approach similar to that described by Anjum et al. [20], in which a video game is developed entirely through iterative prompting with ChatGPT, focusing on implementation-oriented aspects of design. To support consistency across interactions, we employ ChatGPT's "Project" feature, which allows shared knowledge to persist across multiple chats. The knowledge base provided to the model consists of the concept document generated in Phase 1 and the VEP sounds table, ensuring continuity and fidelity to the original design vision.

While the Co.Lab framework enables structured concept generation, its design abstractions are not always sufficient for implementation: the authors suggest continuous iterative steps that polish the document and help the subsequent development [11]. To bridge this gap, we introduce a process we term *gap mapping*. Gap mapping refers to the systematic identification of missing technical or

[1]Available at: https://docs.google.com/presentation/d/1OXLZMKMFrdeSeLnaUV4Ow_X7kkEjT87bKPWqiH_0KBI/edit?usp=sharing

behavioral specifications within the outputs of each Co.Lab component. For each section of the concept document, we evaluated whether the information provided was sufficient to inform implementation within the Unity engine. When information was missing, ambiguous, or underspecified, a clarification prompt was generated and submitted to ChatGPT. For example, if the Game Universe section defined the tone and setting but omitted camera position, a specific prompt was issued: *"Where is the player camera placed?"*.

All clarification questions were retrospectively mapped to their corresponding Co.Lab components and compiled into a checklist (Appendix Table 2), which can be reused in future projects to systematically identify implementation-critical gaps.

3.1.3. Phase 3 – Expert Evaluation

In the final phase, the prototype is evaluated by three domain experts (scholars in musicology, sound design, and heritage) with specialized knowledge of the *Poème Électronique*. The three experts had no prior involvement in the development of the prototype. The evaluation is conducted through a structured feedback session, during which each expert individually interacts with the prototype, and then, as a group discussion, three guiding questions are answered:

- What aspects of the original work are meaningfully represented in the game?
- What aspects of the original work are missing or underdeveloped?
- Where and how do you perceive the presence of cultural heritage value?

These open-ended prompts encourage reflection not only on fidelity to the original heritage item but also on the interpretive richness and communicative potential of the prototype. Future evaluations could complement open-ended questions with Likert-scale ratings to quantify perceived fidelity and value representation.

4. Results

4.1. Analysis of the Concept Document

The prototype game, titled *Edgar à Gogo*, is structured as a short, interactive experience designed to familiarize the player with the sonic world of Varèse's *Poème Électronique*. The player is tasked to explore "the world of sound" through four distinct levels. The first level focuses on exploring, unlocking, and recognizing sounds that are represented as physical solid floating objects in an empty space; the second one is about manipulating (changing pitch, volume, panning, and reverb) one or more sounds (unlocked from the previous level); the third one is about composing a personal version of the *Poème* (always with the unlocked sounds); finally, the fourth one is a turn-based multiplayer version of the third level (with sharing and social features).

Cognitively, the game aims to enhance players' understanding of key features of the musical piece, including sound texture, spatialization, and experimental electronic composition techniques. Players are expected to grasp how sound is organized in space and how different sonic elements interact within a tridimensional environment. On a practical level, the game enables users to interact with and manipulate sonic fragments through digital tools, guiding them toward the creation of short compositions and the resolution of sound-based puzzles. Finally, the game fosters interpersonal and affective skills by encouraging collaborative decision-making (particularly in shared physical settings in the multiplayer level), developing critical and creative listening habits, and deepening appreciation for the artistic and historical significance of experimental electronic music.

Mechanically, the game combines basic point-and-click interaction with simple puzzle solving and auditory recognition. For example, players may be tasked with classifying sounds, navigating through a sound-driven environment, or matching sound fragments to a given target sound. Feedback is provided through color-coded cues and progression through levels. Although the game lacks explicit scoring or time-based challenges, it encourages attentive listening and curiosity-driven exploration.

4.2. Analysis of the Technical Specifications

During this phase, several features outlined in the concept document were further developed, addressing ambiguities and high-level descriptions that had remained unresolved in the previous stage. For instance, the progression logic between levels was unclear, as was the number and nature of the puzzles present in Level 1. Likewise, some fundamental aspects of player interaction—such as camera placement and parameters or the mechanics of locomotion and interaction—were not specified in the original concept and required explicit clarification.

The concept document envisioned a sound space populated by physical objects, each associated with a distinct sound fragment from the *Poème Électronique*. As previously stated, to minimize authorial decisions, even the selection of sound fragments was delegated to ChatGPT. This choice led the model to make both aesthetic and semantic assumptions based on the onomatopoeic names of the sounds. From a dataset of 142 extracted segments, ChatGPT selected 12 and defined their visual representation and interactive behavior. For example, the sound "WBlock" was represented as a jagged tetrahedron with rapid vibrations—choices that likely stem from the short duration and percussive nature implied by the sound's name (probably interpreted as wooden blocks). Finally, the model also introduced a distinction between "dormant" and "active" sounds: the former only emit audio upon player interaction, while the latter loop continuously from the start of the game (this was a clarification on the concept of "unlocking" sounds).

In Image 1, a comparison is shown between the storyboard (created in Phase 1) and the prototype developed in the current phase. As illustrated, the prototype adopts a first-person camera perspective—evidenced by the mouse pointer at the center of the image—with a horizontal and frontal view of the objects, which contrasts with the perspective used in the storyboard. The sound fragments (represented by the solid objects) are color-coded and exhibit distinctive shapes, whereas the storyboard only featured basic forms such as cubes, spheres, and a pyramid.

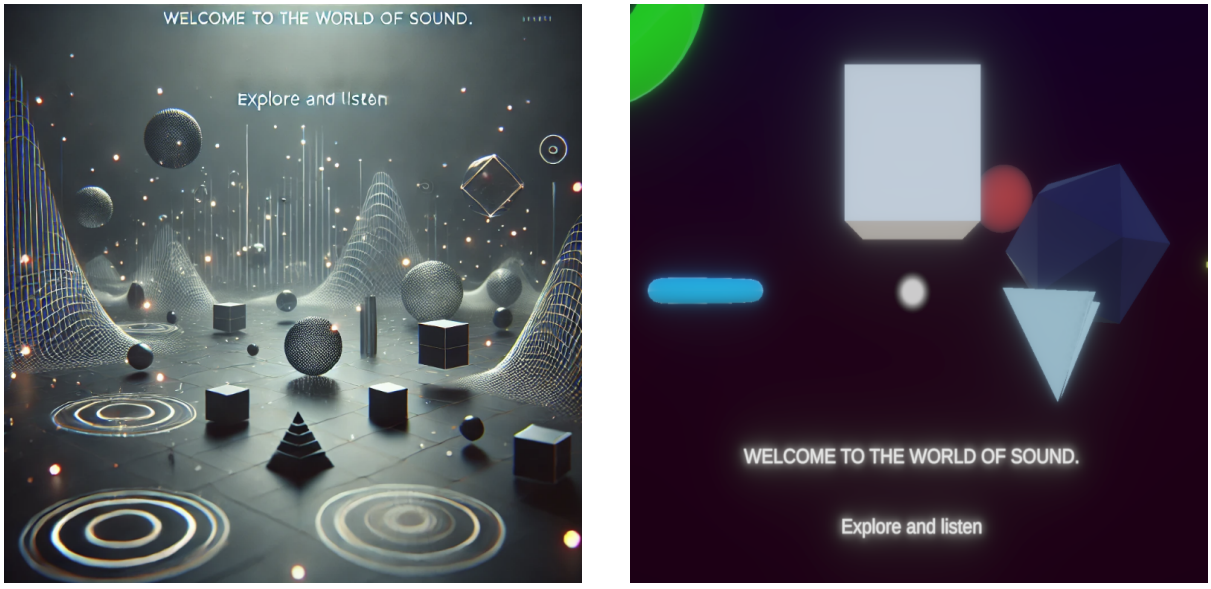


Figure 1: A storyboard and its implementation in the prototype.

4.3. Expert Feedback

The evaluation sessions with domain experts surfaced both affirmations of fidelity to the original work and constructive critiques pointing to missing or underdeveloped cultural dimensions. The experts recognized that the prototype effectively reflects certain compositional features of the *Poème Électronique*, such as its fragmentation into discrete sound units and the use of spatial diffusion. However, several

limitations emerged, which together form a valuable foundation for defining cultural heritage values to be explicitly incorporated in future iterations.

First, experts stressed the need to distinguish between different representational layers of the original work: the original sounds and their gamified treatment. A recurring request was the inclusion of a "reset" and "compare" function, allowing users to restore each sound (after the manipulation) to its unaltered form and also compare the altered version to the original sound. Without such mechanisms, the exploratory and compositional aspects risk conflating player interventions with the original sonic identity.

Second, experts highlighted the importance of temporal context in understanding the cultural value of the piece. The *Poème Électronique* is not only a collection of sounds but also a carefully sequenced narrative. The prototype currently treats sound fragments as isolated entities, neglecting their placement in Varèse's compositional structure. Suggestions included visualizing adjacent or repeated fragments to reflect their actual ordering and recurrence.

Third, the metaphor of a "soundspace"—a 3D abstract environment populated by interactive sound objects—was received with skepticism. Experts found the environment visually evocative of the Philips Pavilion, particularly in its dark color palette and pulsing lights. However, they also criticized its lack of a clear organizational logic. The density and proximity of sound sources generated auditory clutter, impeding close listening and analysis. The experts proposed alternative spatial arrangements based on perceptual criteria (e.g., duration, timbre, category).

A particularly salient suggestion was to anchor the game design in historical metaphors, such as simulating a 1950s tape editing studio, to evoke the production conditions of the original work. This would allow for richer engagement with the material and greater contextualization of its cultural significance.

Finally, in the composition phase, experts asked for greater visual and semantic coherence between in-game elements; for example, ensuring that the visual appearance of fragments in the composition timeline matches their in-environment representations. They also requested clearer playback indicators, such as highlighting the fragment currently being played and showing its remaining duration.

Collectively, these insights reveal a set of emergent but weakly supported values in the current prototype, including (1) preservation of original sonic identity, (2) sequencing and temporal context, (3) intelligibility and clarity of spatialization, and a proposal of (4) historical fidelity to the production context. These values—while latent—were not sufficiently scaffolded by the game's current mechanics and interface. They will serve as guiding constraints in a future iteration where values are explicitly modeled in the design pipeline.

5. Discussion

This study investigated whether cultural heritage values could emerge in a design process where those values were not explicitly defined but instead mediated through a structural framework (Co.Lab) and a generative AI model (ChatGPT). The findings suggest that some values did emerge, but in fragmented, weak, or ambiguous forms. These included the identity of individual sound fragments, the idea of original sequencing, and the spatial or material characteristics of the *Poème Électronique*. However, their emergence was highly contingent on the AI's interpretive behavior and not structurally guaranteed by the design process.

We introduce the notion of *value emergence* to describe this phenomenon: values can emerge through the interplay of instructional content, interaction logic, and representational choices, even in the absence of formal modeling. In our case, the Co.Lab framework served as a neutral scaffold, and ChatGPT operated using associations encoded in its training data. These were sufficient to produce culturally suggestive design components—but not to ensure their coherence, accuracy, or pedagogical depth.

For example, ChatGPT interpreted the filenames of sound fragments as semantic cues, mapping them onto visual and spatial attributes such as color, shape, and position. Percussive-sounding names led to angular forms; lighter names were linked to smaller, higher-positioned objects. This reflects a genuine

but limited effort to construct communicative coherence. The AI attempted to fill semantic gaps but did so with insufficient cultural grounding.

Crucially, it was the absence of deeper cultural values—such as the sequencing logic of the original composition or the symbolic context of its historical performance—that prompted the most insightful expert feedback. Experts were able to articulate what was missing precisely because the prototype made those gaps visible. In this way, the game acted not only as a playable artifact but also as a discursive probe, capable of eliciting reflection and value definition through its limitations.

This leads to an important design insight: value emergence is not equivalent to value embodiment. While some aspects of cultural meaning may arise spontaneously through inference or association, others—especially those related to historical context, symbolic intent, or sequencing—require intentional modeling. AI, when unprompted, may offer a coherent structure, but not necessarily a culturally faithful one.

More broadly, this experiment demonstrates that value absence can be as informative as presence. A game that fails to convey certain values can still function as a powerful starting point for expert dialogue and co-design. Rather than assuming that cultural fidelity must be fully built in from the start, we might also design for revelation through absence—using generative tools not only to prototype content but also to diagnose which values need to be made explicit in future iterations.

In sum, this experiment demonstrates that value emergence—while conceptually intriguing—is insufficient for ensuring cultural fidelity in serious games. Instead, we argue for a hybrid, iterative approach in which generative tools serve as catalysts for expert reflection, enabling the systematic articulation of values to be embedded in future design cycles.

6. Conclusions and Future Works

This study explored whether and how cultural heritage values can emerge in the design of a serious game developed using generative AI without explicit value modeling. By combining the Co.Lab design framework with ChatGPT's generative capabilities, we produced a playable prototype of *Edgar à GoGo*, a game based on Varèse's *Poème Électronique*. Through a three-phase process—concept generation, technical prototyping, and expert evaluation—we analyzed how values emerged, were omitted, or were implicitly approximated by the AI.

Our findings indicate that while some values (e.g., the identity of individual sound fragments, the immersive structure of the original performance) did emerge in limited or fragmented forms, other crucial aspects (e.g., historical sequencing, symbolic intent, and production context) remained absent. These gaps, however, proved generative: they allowed experts to articulate what was missing and thereby identify a candidate set of values to be explicitly modeled in future iterations.

We argue that value emergence—the spontaneous emergence of cultural values without formal instruction—is possible but insufficient for values-sensitive design. It offers a starting point, not an endpoint. For heritage games, where cultural fidelity and interpretive depth are essential, values must be deliberately embedded in the design process.

Generative AI, when paired with structured frameworks like Co.Lab, can function as both a creative partner and a diagnostic tool. It provides a neutral scaffold through which assumptions become visible, gaps can be located, and expert dialogue can be initiated. The resulting game is not only an artifact but also a prompt for co-design—a platform for surfacing and negotiating the cultural values to be mapped.

In future work, we will test this approach in a second iteration of *Edgar à GoGo*, where the values identified through expert feedback will be made explicit and formally introduced into the Co.Lab components. This will allow us to compare value emergence and value embodiment directly and to refine a replicable, hybrid workflow that integrates AI, design frameworks, and expert evaluation in the service of cultural heritage communication. Future replications will explore prompt variability and model robustness across iterations. While this study relied on ChatGPT-4o, alternative models could yield distinct interpretations of cultural values due to differences in training data and instruction tuning.

Declaration on Generative AI

The game concept document and the subsequent technical questions were realized with ChatGPT 4o, which acted as a game designer for the whole experiment.

References

[1] S. Deterding, D. Dixon, R. Khaled, 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, pp. 9–15. doi:10.1145/2181037.2181040.

[2] A. Dewhirst, R. Laugharne, R. Shankar, Therapeutic use of serious games in mental health: scoping review, BJPsych Open 8 (2022) e37. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC8867878/. doi:10.1192/BJO.2022.4.

[3] D. A. Plecher, C. Eichhorn, A. Köhler, G. Klinker, Oppidum - A Serious-AR-Game About Celtic Life and History, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), volume 11899 LNCS, Springer, 2019, pp. 550–559. URL: http://link.springer.com/10.1007/978-3-030-34350-7_53. doi:10.1007/978-3-030-34350-7{_}53.

[4] A. Calvo-Morata, C. Alonso-Fernández, M. Freire, I. Martínez-Ortiz, B. Fernández-Manjón, Serious games to prevent and detect bullying and cyberbullying: A systematic serious games and literature review, Computers & Education 157 (2020) 103958. URL: https://www.sciencedirect.com/science/article/abs/pii/S0360131520301561. doi:10.1016/J.COMPEDU.2020.103958.

[5] F. M. Williams-Bell, B. Kapralos, A. Hogue, B. M. Murphy, E. J. Weckman, Using serious games and virtual simulation for training in the fire service: A review, Fire Technology 51 (2015) 553–584. URL: https://link.springer.com/article/10.1007/s10694-014-0398-1. doi:10.1007/S10694-014-0398-1/METRICS.

[6] E. Bonacini, S. C. Giaccone, Gamification and cultural institutions in cultural heritage promotion: a successful example from italy, Cultural Trends 31 (2022) 3–22. URL: https://www.tandfonline.com/doi/full/10.1080/09548963.2021.1910490. doi:10.1080/09548963.2021.1910490.

[7] M. Mortara, C. E. Catalano, F. Bellotti, G. Fiucci, M. Houry-Panchetti, P. Petridis, Learning cultural heritage by serious games, Journal of Cultural Heritage 15 (2014) 318–325.

[8] L. H. Fredheim, M. Khalaf, The significance of values: heritage value typologies re-examined, International Journal of Heritage Studies 22 (2016) 466–481. doi:10.1080/13527258.2016.1171247.

[9] V. Murtas, V. Lombardo, Rediscovering the past: Serious games for archaeology, IEEE Transactions on Games (2025) 1–15. doi:10.1109/TG.2025.3577572.

[10] L. Hanes, R. Stone, A model of heritage content to support the design and analysis of video games for history education, Journal of Computers in Education 6 (2019) 587–612. URL: http://link.springer.com/10.1007/s40692-018-0120-2. doi:10.1007/s40692-018-0120-2.

[11] D. Jaccard, L. Suppan, E. Sanchez, A. Huguenin, M. Laurent, The co.LAB Generic Framework for Collaborative Design of Serious Games: Development Study, JMIR Serious Games 9 (2021) e28674. URL: https://games.jmir.org/2021/3/e28674. doi:10.2196/28674.

[12] A. Antoniou, G. Lepouras, S. Bampatzia, H. Almpanoudi, An approach for serious game development for cultural heritage: Case study for an archaeological site and museum, Journal on Computing and Cultural Heritage 6 (2013). doi:10.1145/2532630.2532633.

[13] R. Andreoli, A. Corolla, A. Faggiano, D. Malandrino, D. Pirozzi, M. Ranaldi, G. Santangelo, V. Scarano, A framework to design, develop, and evaluate immersive and collaborative serious games in cultural heritage, J. Comput. Cult. Herit. 11 (2017). URL: https://doi.org/10.1145/3064644. doi:10.1145/3064644.

[14] D. Jaccard, L. Suppan, F. Bielser, Contribution of the co.LAB framework to the collaborative design of serious games: Mixed methods validation study, JMIR Serious Games 9 (2021). doi:10.2196/33144.

[15] R. Dobson, J. Fitch, K. Tazelaar, A. Valle, V. Lombardo, Varèse's poème Électronique regained: Evidence from the vep project, International Computer Music Conference Proceedings 2005 (2005). URL: http://hdl.handle.net/2027/spo.bbp2372.2005.164.

[16] V. Šisler, Contested memories of war in czechoslovakia 38-89: Assassination: Designing a serious game on contemporary history, Game Studies 16 (2016). URL: https://gamestudies.org/1602/articles/sisler.

[17] S. Goud, Value Communication for Cultural Heritage: Operational Workflow for Digital Environments, Phd thesis, University of Turin, Turin, Italy, 2023. Available on https://hdl.handle.net/2318/1952441.

[18] P. Sweetser, Large Language Models and Video Games: A Preliminary Scoping Review, in: ACM Conversational User Interfaces 2024, ACM, New York, NY, USA, 2024, pp. 1–8. URL: https://dl.acm.org/doi/10.1145/3640794.3665582. doi:10.1145/3640794.3665582.

[19] R. Gallotta, G. Todd, M. Zammit, S. Earle, A. Liapis, J. Togelius, G. N. Yannakakis, Large Language Models and Games: A Survey and Roadmap, IEEE Transactions on Games (2024) 1–18. URL: https://ieeexplore.ieee.org/document/10680313/. doi:10.1109/TG.2024.3461510.

[20] A. Anjum, Y. Li, N. Law, M. Charity, J. Togelius, The Ink Splotch Effect: A Case Study on ChatGPT as a Co-Creative Game Designer, in: Proceedings of the 19th International Conference on the Foundations of Digital Games, ACM, New York, NY, USA, 2024, pp. 1–15. URL: https://dl.acm.org/doi/10.1145/3649921.3650010. doi:10.1145/3649921.3650010.

[21] J. Tyni, A. Turunen, J. Kahila, R. Bednarik, M. Tedre, Can ChatGPT Match the Experts? A Feedback Comparison for Serious Game Development, International Journal of Serious Games 11 (2024) 87–106. URL: https://journal.seriousgamessociety.org/index.php/IJSG/article/view/744. doi:10.17083/ijsg.v11i2.744.

[22] P. L. Lanzi, D. Loiacono, ChatGPT and Other Large Language Models as Evolutionary Engines for Online Interactive Collaborative Game Design, in: Proceedings of the Genetic and Evolutionary Computation Conference, ACM, New York, NY, USA, 2023, pp. 1383–1390. URL: https://dl.acm.org/doi/10.1145/3583131.3590351. doi:10.1145/3583131.3590351.

[23] V. Lombardo, A. Valle, J. Fitch, K. Tazelaar, S. Weinzierl, W. Borczyk, A virtual-reality reconstruction of poème électronique based on philological research, Computer Music Journal 33 (2009) 24–47. URL: https://doi.org/10.1162/comj.2009.33.2.24. doi:10.1162/comj.2009.33.2.24. arXiv:https://direct.mit.edu/comj/article-pdf/33/2/24/1855306/comj.2009.33.2.24.pdf.

[24] L. Izzo, Edgard Varèse's Poème Électronique: From the Sketches to the Sound Spatialization, Computer Music Journal 47 (2023) 5–28. URL: https://direct.mit.edu/comj/article/47/4/5/127576/Edgard-Varese-s-Poeme-Electronique-From-the. doi:https://doi.org/10.1162/COMJ_a_00700.

Appendix

Co.Lab Step	Question
Game Mechanics	What does it mean to activate a fragment?
Game Mechanics	Which and how many sound fragments are present in the scene?
Game Mechanics	Where and how can the player move?
Game Mechanics / Interactions	Puzzle Matching - What sounds are associated with the three buttons? What is the target to identify?
Game Mechanics / Interactions	Puzzle Matching - What happens when the user clicks on the corresponding hub?
Game Mechanics / Interactions	How do the objects behave when clicked (in proximity)?
Game Mechanics / Interactions / Goals and Rules	What sounds are available in the composition interface? Can they be modified?
Game Structure / Game Mechanics	How does the player proceed from level 1 to level 2?
Game Structure / Game Universe	Is the environment shared across levels?
Game Universe / Interfaces and UX	What does the abstract environment (soundspace) look like?
Game Universe / Interfaces and UX	Are the fragments solid objects? What do they look like?
Game Universe / Interfaces and UX	Where is the camera positioned?
Game Universe / Interfaces and UX / Game Mechanics	Where are the fragments positioned in the scene? Which are active and which are dormant?
Goals and Rules	What is the player expected to do in the second level?
Goals and Rules	What is the player expected to do in the third level?
Goals and Rules	What is the role of the puzzle hubs in level 2?
Goals and Rules / Incentives	Puzzle Matching - Which new fragments are unlocked?
Interactions / Goals and Rules	Puzzle Spatialization - What is the response and feedback logic (attempts, errors, success)?
Interfaces and UX	Can the camera rotate freely vertically and horizontally?
Interfaces and UX	What does the mouse cursor look like?
Interfaces and UX	What is the scale of the fragments?
Interfaces and UX	What are the rotation parameters of the fragments?
Interfaces and UX	What spatialization parameters do the fragments have?
Interfaces and UX	Puzzle Spatialization - What does the UI look like? + LAYOUT
Interfaces and UX	What is the appearance and position of the portal to go from one level to another?
Interfaces and UX	What is the layout of the composition interface?
Interfaces and UX / Goals and Rules	Can multiple fragments be inserted into the same slot?
Interfaces and UX / Goals and Rules	Is the composition interface fullscreen? Can the user still explore the environment?
Interfaces and UX / Interactions	List all feedback elements (conveyance)
Interfaces and UX / Interactions	Do all puzzles trigger an overlay UI in which the puzzle is played?
Interfaces and UX / Interactions / Goals and Rules	Puzzle Matching - What happens to the hub after completion? Both in failure and success?

Table 2
Clarification questions mapped to Co.Lab steps.