

SiMUSation: An Interactive Visitor Experience Simulation Framework to Support Museum Exhibition Design

Huanchen Wang
Southern University of Science and Technology
Shenzhen, Guangdong, China
City University of Hong Kong
Hong Kong, China
wanghc2022@mail.sustech.edu.cn

Qiuming Chen
Southern University of Science and Technology
Shenzhen, Guangdong, China
12432735@mail.sustech.edu.cn

Zhonghao Ji
Southern University of Science and Technology
Shenzhen, Guangdong, China
12432740@mail.sustech.edu.cn

Ruqi Sun
Southern University of Science and Technology
Shenzhen, Guangdong, China
sunrq2024@mail.sustech.edu.cn

Zhichao Lu
City University of Hong Kong
Hong Kong, China
zhichao.lu@cityu.edu.hk

Yuxin Ma*
Southern University of Science and Technology
Shenzhen, Guangdong, China
mayx@sustech.edu.cn

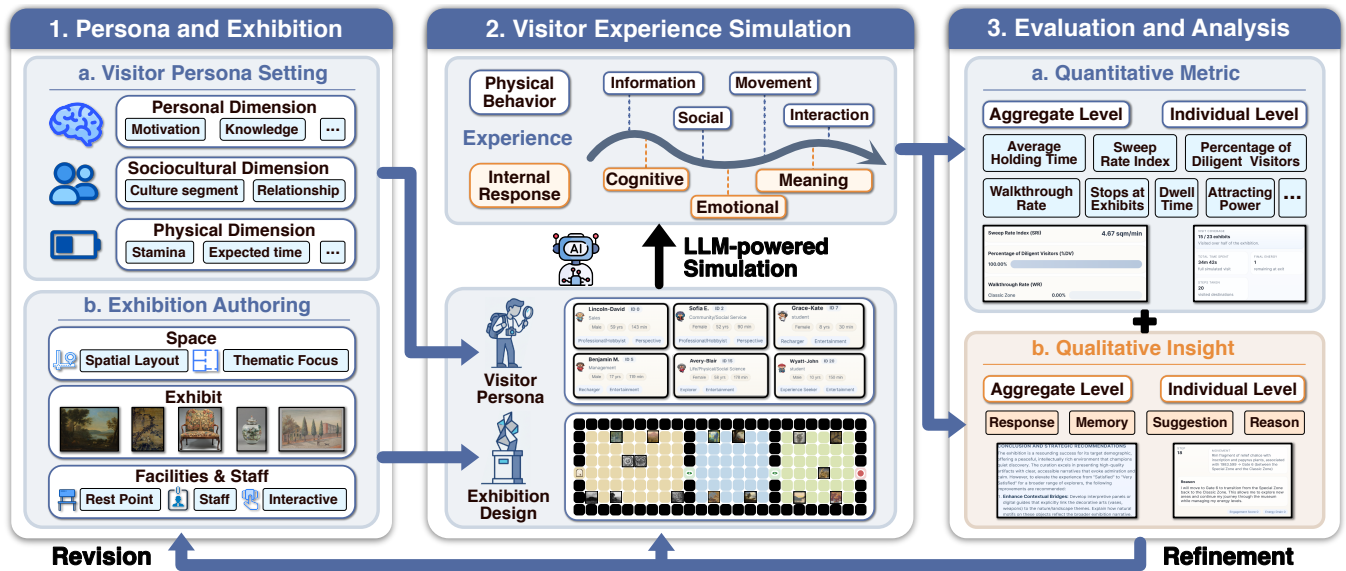


Figure 1: The SiMUSation framework consists of: 1) Persona & Exhibition, supporting the diverse and plausible visitor personas building and the interactive authoring of exhibition space and narrative; 2) Visitor Experience Simulation, which simulates dynamic experiences from the defined personas and exhibition environment, with feedback represented as physical behaviors and internal responses; and 3) Evaluation & Analysis, providing multi-faceted analysis to inform early-stage design iteration.

Abstract

Understanding how diverse audiences engage with narratives and content is central to exhibition design, yet designers often rely on intuition. Existing experience evaluation methods are typically retrospective, costly, and offer limited access to visitors' internal states,

hindering early-stage iterative refinement. Rather than relying only on post-implementation evaluation with real visitors, we explore LLM-driven persona simulation as a reference for early-stage design. Following this idea, we present SiMUSation, an interactive framework designed to support early-stage exhibition design. SiMUSation models diverse visitor personas and simulates their exhibition experiences through a dual-layer representation that couples observable behaviors, such as movement and gaze, with corresponding internal responses, such as confusion and narrative engagement. Designers can steer simulations, inspect feedback from simulated visits, and iteratively revise layouts, content, and narrative flow to further examine how changes reshape visitor experience. We implemented a prototype and evaluated it through a user study (N=12), showing

*Corresponding author.



This work is licensed under a Creative Commons Attribution 4.0 International License.
UIST '26, Detroit, MI, USA

© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2856-3/2026/11
<https://doi.org/10.1145/3830398.3830560>

that SiMUSation provides insights for reflection and refinement in early-stage exhibition design. Our findings further highlight the potential of persona-driven simulation to support audience-informed evaluation and iterative decision-making across design tasks.

CCS Concepts

• **Human-centered computing** → **Interactive systems and tools**.

Keywords

LLMs, audience simulation, museum exhibition design, persona, creativity support tool

ACM Reference Format:

Huanchen Wang, Qiuming Chen, Zhonghao Ji, Ruqi Sun, Zhichao Lu, and Yuxin Ma. 2026. SiMUSation: An Interactive Visitor Experience Simulation Framework to Support Museum Exhibition Design. In *The 39th Annual ACM Symposium on User Interface Software and Technology (UIST '26), November 02–05, 2026, Detroit, MI, USA*. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3830398.3830560>

1 Introduction

Understanding how a museum exhibition’s narrative is perceived, interpreted, and reconstructed by its audience is central to exhibition design and evaluation [21, 28, 76]. Yet designers still rely largely on professional intuition, making design outcomes susceptible to individual perspectives and biases [22, 95]. Because the effects of design decisions often remain unclear until implementation or public opening, evaluation typically depends on costly observation and visitor feedback collected after the fact [22, 37, 82, 102]. These constraints limit early-stage assessment and iterative refinement.

Prior work has explored digital approaches for capturing and analyzing personalized visitor experiences in museums [4, 45, 101], but these approaches remain limited: behavioral data cannot fully account for visitors’ internal experiences [20, 24, 52, 105], and user studies are difficult to scale and often lack representativeness [74, 77]. Recent advances in large language models (LLMs), especially in LLM-based role-playing [14, 84], suggest a promising alternative. By constructing coherent and diverse personas, LLMs can generate complex human behaviors and responses across contexts, as demonstrated in domains such as social simulation, games, and writing [34, 59, 70, 96]. However, both traditional digital techniques and recent LLM-based museum works have primarily focused on visitor-facing support, including navigation and guidance [4, 53], VR/AR-based interactive experiences [64, 101, 104], and LLM-based role and interaction design [9, 13, 46, 67, 89]. In contrast, support for curators and exhibition designers during early-stage design remains underexplored.

Motivated by these gaps and informed by a formative study (N=7), we present SiMUSation, an interactive framework for early-stage museum exhibition design that organizes the workflow as an authoring-simulating-iterating loop. SiMUSation integrates LLM-powered visitor simulation into the exhibition design workflow, linking physical behavior with internal responses to provide interpretable representations of visitor experience feedback and support reflection and design iteration. We implemented SiMUSation as a prototype and evaluated its usability and effectiveness through a

user study (N=12). The results suggest that SiMUSation can provide actionable insights into visitor experience and support design refinement by helping designers identify where issues arise through behavioral traces and why through simulated internal responses. Based on these findings, we discuss implications for LLM-powered agent systems in user-related design tasks. Our contributions are:

- We conducted a formative study that identified challenges in early-stage exhibition design workflows and derived design goals.
- We present SiMUSation, an interactive framework that embeds LLM-powered visitor experience simulation into the exhibition design workflow for iterative refinement.
- We developed an interactive system and evaluated the framework’s usability and effectiveness in supporting design refinement through a user study.

2 Background and Related Work

2.1 Visitor Experience in Museums

In museums, visitors are not merely viewers of displayed objects but active participants in meaning-making and experience construction [17, 22, 88]. Accordingly, museums are widely understood as places of experience, where people come not only to learn, but also to enjoy leisure and socialize with others [22, 49, 88]. Designing engaging visitor experiences has therefore become a central concern in contemporary museum practice, as museums seek to create experiences that are meaningful, inclusive, and memorable and to sustain public engagement across diverse audiences [47, 81].

Substantial research has examined how museum experiences are formed and interpreted [22, 40, 50, 69]. The visitor experience can be modeled as multi-phased and context-dependent, spanning before, during, and after the visit [22]. It thus frames museum visiting as a dynamic and situated experience rather than a purely observational activity. Prior studies further suggest that such experiences are shaped through the interaction of personal, social, and physical contexts, together with visitors’ feelings and memories [21, 22, 69]. At the same time, this complexity makes it difficult to capture and assess the visitor experience holistically, which in turn limits efforts to evaluate and refine exhibition experiences [22, 35, 49].

Digital methods have been increasingly applied to analyze museum visitor experience [30, 41, 82]. Prior work often uses sensing technologies, such as Bluetooth [105], cameras [24, 38], and complex multi-device systems [5, 52, 75], to collect physical behavioral data and identify patterns in related visitor measures that support evidence-based assessment. However, visitor experience is inherently complex, and behavioral data alone cannot fully capture visitors’ interpretations, emotions, and reflections [1]. To access these experiential dimensions more directly, researchers often use surveys and interviews to examine what visitors learn, think, and feel [7, 68, 74, 77]. Yet these methods are difficult to scale and labor-intensive to conduct and analyze [30]. Therefore, these limitations hinder institutions and designers from systematically analyzing visitor experience for exhibition evaluation and design improvement.

Digital technologies have been widely used to enhance museum experience [45, 65]. Prior work has supported visitor navigation, guidance, and information access through positioning and mobile systems [4, 78, 94] and robotic guides [36, 51]. Related studies have also explored tangible and embodied interactions, enabling visitors

to engage with exhibitions [101]. Building on these approaches, AR and VR link physical and virtual exhibits and spaces to support immersive exploration [15, 80, 85, 104]. This line of work further extends to virtual museums, which examine how exhibitions can be designed and experienced in fully digital environments, including screen-based and panoramic settings [11, 57, 64]. Building on prior work, we shift attention from visitors themselves to curators and designers, positioning visitor experience as a lens integrated into the exhibition design workflow at an early stage.

2.2 LLMs for Museum Exhibition

Recent advances in LLMs, including awareness, emotion, and personality, have enabled more human-like responses and stronger adherence to personas and roles [42, 43, 56, 79, 84]. Building on this progress, recent studies have developed LLM role-playing approaches that use persona descriptions and role profiles to shape agent behavior in different tasks and interaction contexts, with some approaches representing both internal thoughts and physical actions in settings [14, 91, 93, 97]. As a result, LLMs are increasingly used in role-playing applications across domains such as game design [59, 96], writing [2, 34, 73], programming [39, 109], business [63, 100], and social simulation [3, 70], where they can represent diverse beliefs, perspectives, and tendencies to generate personalized and situated content and decisions for various user-centered task-solving [14, 32, 55]. In parallel, growing human-AI co-creation research has examined how to understand and interact with LLM-powered personas and roles to support several creative tasks [32], including ideation [16, 61], graphic design [86], writing [71, 90], and game design [48]. These developments suggest that visitor experience simulation as a reference may also support early-stage museum exhibition design.

In museum and exhibition contexts, recent LLM-based work has primarily focused on enhancing visitor-facing personalized interaction and guidance [9]. One main direction uses LLM-powered guides, such as docent agents and robotic guides, to provide explanation and exploration support [4, 67]. Another direction introduces interactive exhibition-related artifacts and characters, such as historical figures, to enrich engagement and interpretation [46, 98]. Related studies also explore the role of LLM and how LLM with multiple perspectives can help visitors sequence content, understand knowledge, and engage with the exhibition in more adaptive ways [13, 58, 89]. However, little work has investigated a more comprehensive framework for visitor experience simulation and examined how such simulation can directly support exhibition evaluation and design iteration. This gap motivates our framework, which enables rich visitor experience simulation and experience-driven, iterative editing to support a closed design loop.

3 Formative Study

To identify challenges in early-stage exhibition design, we conducted a formative study with 7 participants (aged 20–50 years, $M=30.86$; 3 females and 4 males) recruited via campus forums and promotional referrals, including museum-exhibition-related domain practitioners, researchers, and students (Table 1). Each participant joined in a 90-minute Zoom session and received 100 CNY (approximately USD 14) in compensation. The study aimed to examine their current design practices and how they incorporate visitor

experience into design thinking. In each session, participants first completed a brief painting exhibition design task based on a floor plan, during which they developed a narrative, marked intended visitor routes, and identified areas of uncertainty (Appx. A.1). This was followed by a semi-structured interview about their experiences and challenges in conventional design workflows spanning the stages of planning, evaluating, and refining exhibition designs, as well as their expectations and preferences for future support.

3.1 Findings

3.1.1 F1: Difficulty in Anticipating and Considering Diverse Visitor Experiences. During early-stage exhibition planning, all participants emphasized the importance of adopting a visitor perspective by considering movement paths, narrative comprehension, and different audience groups. At the same time, they noted that analyzing visitor experience requires accounting for many nuanced and intersecting factors, including prior knowledge (P1, P2, P4, P6), interest (P3), sociocultural background (P3, P6, P7), motivation (P3, P4, P5, P6), and physical constraints such as stamina (P3, P7) and expected visit duration (P5). The complexity of these factors made it difficult for participants to anticipate diverse responses and translate them into design decisions. Several participants (P1–P5) further noted that, without concrete ways to examine how different visitor groups might experience an exhibition, they often relied on their own professional experience, which could introduce personal biases into early-stage exhibition design planning.

3.1.2 F2: Lack of Real-time and Multi-faceted Evaluation of Experience. Most participants also emphasized that effective evaluation and design refinement should be grounded in visitor experience. However, they noted that existing evaluation methods are mostly retrospective, which are conducted only after an exhibition is installed or opened (P3, P7). Without feedback before deployment, identifying and correcting design flaws can become costly. All participants argued that current approaches do not adequately capture the dynamic and complex nature of visitor experience. Several participants (P1, P3, P6) further noted that although observable behaviors such as movement paths and dwell times can be recorded, they provide limited access to visitors' internal responses, including narrative understanding and emotional engagement. P6 and P7 both stressed that visitor experience should not be assessed by a single standard, since different audiences vary in their visiting contexts. Therefore, participants expressed a strong need for evaluation methods that can provide timely, multi-faceted, and audience-sensitive feedback to support visitor understanding and evaluation.

3.1.3 F3: Limitations of Static Methods for Exhibition Authoring and Iteration. Many participants described revision and refinement as difficult parts of early-stage exhibition design. They often rely on static formats such as document outlines (P1, P6) and rough sketches (P3, P4, P6, P7), which made changes cumbersome and fragmented. These workflows also provided limited support for informed revision, making iteration largely trial and error. Several participants therefore wanted more practical feedback to support revision, including both quantitative summaries and qualitative comments (P1, P5). Participants (P1–P4, P6–P7) also noted that, while not ideal, 2D canvas authoring was sufficient for early-stage design. When discussing potential authoring support, participants

consistently emphasized the importance of preserving creative control. Rather than expecting tools to generate finalized solutions, they preferred tools that offer references and suggestions while leaving design decisions to the designer.

3.2 Design Goals

To address these challenges identified in the findings, we derived three design goals for an LLM-powered framework that simulates complex and diverse visitor experiences to support early-stage exhibition design, refinement, and revision:

3.2.1 DG1: Support Reliable and Diverse Visitor Personas Building. This goal supports the construction of heterogeneous visitor personas, helping designers move beyond subjective assumptions (F1). The framework allows users to define intersecting and multidimensional factors grounded in prior literature and our formative study findings, enabling early-stage design to better account for diverse and plausible visitor experiences.

3.2.2 DG2: Provide Dynamic Simulation and Multi-Granular Feedback. To address the retrospective conventional evaluation and the "black box" nature of visitor experience (F2), the framework simulates exhibition visits based on diverse visitor personas (F1) and provides a visual and interpretable visit process by connecting observable behaviors with internal responses over time. In addition, the framework provides multi-granular feedback, combining quantitative metrics with qualitative insights at both the aggregate and individual levels, to support timely and multi-faceted analysis. Such feedback can help designers move beyond intuition and identify potential design issues for refinement and revision (F2, F3).

3.2.3 DG3: Facilitate Insight-Driven Iteration via Interactive Authoring. To facilitate more flexible, experience-based iteration during early-stage design (F2, F3), the framework seamlessly integrates simulation feedback into an interactive authoring environment. Rather than automated solutions, it provides suggestions and clues as reference, allowing designers to iteratively refine layout and narrative plans while maintaining creative agency (F3).

4 SiMUSation

Drawing on the derived design goals, we proposed SiMUSation (Figure 1), a framework with persona-driven visitor experience simulation to support early-stage museum exhibition design, and developed a prototype based on it. The framework comprises four main components: visitor persona agent construction (DG1, DG2), exhibition spatial and narrative authoring (DG2, DG3), visitor experience simulation and representation (DG2), and multi-granular feedback for insight-driven iteration (DG2, DG3). These components support persona building, dynamic simulation, interpretable feedback, and interactive design refinement, helping designers understand visitor experience and revise exhibition plans. In this section, we detail SiMUSation's framework (Figure 1) and prototype (Figure 2).

4.1 Visitor Persona Agents Construction

4.1.1 Structuring Personas via the Contextual Model: To support persona construction that reflects real-world visitor diversity (DG1), we ground SiMUSation in Falk and Dierking's Contextual Model of Learning [22], which frames museum experience through the interaction of Personal, Sociocultural, and Physical contexts across

the phases before, during, and after a visit (DG2). In this component, we use the model to structure visitor attributes that can be specified prior to the visit. Accordingly, the persona agent captures pre-visit characteristics from the Personal, Sociocultural, and Physical dimensions on the visitor side (Figure 1.1.a). Contextual factors that emerge during the visit, such as the exhibition environment and its interaction with visitors, are incorporated later in the experience simulation pipeline described in subsection 4.3. Specifically, SiMUSation structures persona agents through three dimensions:

1. The Personal Dimension: This dimension captures individual characteristics that shape the unique needs and assets a visitor brings to a visit, as well as the self-fulfillment they seek, which influences their visiting experience. The specific attributes include:

Identity-related motivation: To reflect the diverse primary reasons for visiting in a structured way, we incorporate Falk's framework [21], which argues identity-related motivation largely shapes visitor experience and categorizes visitors as *Explorers*, *Facilitators*, *Professionals/Hobbyists*, *Experience Seekers*, or *Rechargers*.

Prior knowledge: The visitor's familiarity with the exhibition or related domain expertise before the visit (e.g., related exhibits, history, background).

Prior experience: Relevant previous museum-going experiences that may shape expectations and comparison.

Interests: Specific themes, topics, or types of exhibits that the visitor is particularly drawn to.

2. The Sociocultural Dimension: This dimension captures the social and cultural background through which a visitor engages with an exhibition. It includes both a broader cultural background and an immediate social setting. The specific attributes include:

Cultural segment: An audience segmentation system widely used in culture and heritage organizations, characterized by cultural background and social identity, and situating individual perspectives within a broader social and cultural context [66].

Relationship: whether the visitor attends alone or with others, such as friends or family.

3. The Physical Dimension: Within persona construction, this dimension focuses on visitor-side physical conditions that can be specified before the visit, rather than on exhibition-side physical factors. These attributes capture physical factors that may impact the experience. The specific attributes include:

Expected visiting duration: The approximate amount of time the visitor plans or is willing to spend in the exhibition.

Stamina: The visitor's general physical endurance for moving through and attending to an exhibition over time [29].

These attributes provide a structured basis for representing visitor diversity in SiMUSation. However, directly exposing all of them as open-ended configuration variables would make persona construction difficult to manage in practice. To retain descriptive richness while supporting interpretability and scalability, we ground persona construction in audience survey data, as described below.

4.1.2 Grounding Persona from Audience Surveys. While the attribute set outlined above enables comprehensive visitor characterization, exposing users to too many variables can increase cognitive burden and affect the scalability and reliability of persona construction. To address this challenge, prior market research suggests that a two-dimensional (2×2) framework can offer a structured

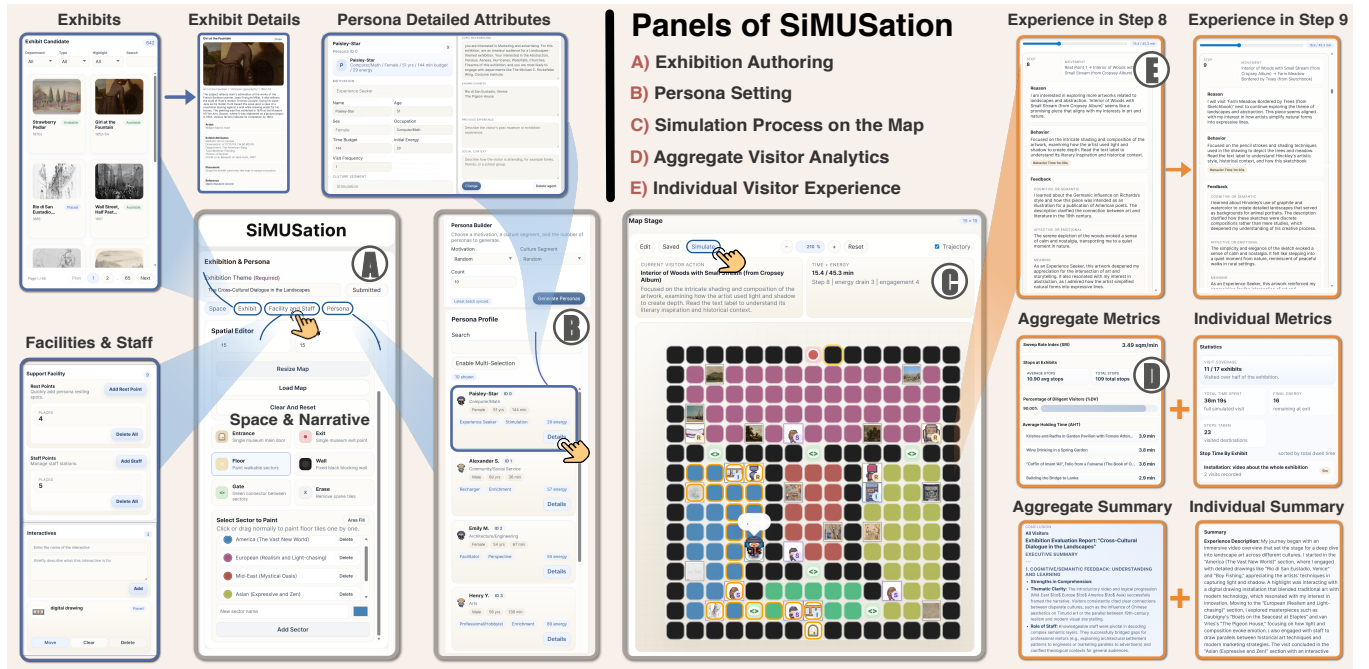


Figure 2: The prototype interface supports an authoring-simulating-iterating workflow through five panels: A) Exhibition Authoring Panel for editing exhibition space, exhibits, facilities, and staff; B) Persona Setting Panel for building and editing visitor personas; C) Simulation Process on the Exhibition Map for switching between editing and simulation modes and visualizing visitor trajectories and interactions; D) Aggregate Visitor Analytics Panel for group-level summaries and behavioral metrics; and E) Individual Visitor Experience Panel for step-by-step inspection of each visitor’s behaviors and responses.

yet flexible way to understand target audiences [62, 72]. Building on this rationale, SiMUSation adopts an approach centered on two anchor dimensions: **Identity-related motivation** and **Cultural segment**. We selected these two attributes because formative study participants highlighted them as particularly important for shaping the visitor experience, while prior work suggests that identity-related motivation largely influences the museum experience [21], and cultural segments offer a practical representation widely adopted in cultural and heritage contexts [66]. Thus, these two attributes inform visitors’ exhibition expectations through both personal orientation and sociocultural positioning.

On this basis, the designer can first specify the two anchor dimensions and the number of personas to define the target audience. The remaining attributes are then populated by sampling empirical distributions derived from culture- and museum-related audience surveys (see supplemental materials). Using survey-based distributions instead of LLM generation reduces hallucinated or weakly supported attribute combinations while keeping persona construction practical, interpretable, and grounded, while preserving within-group variation. Designers can also manually refine personas for target audiences or edge cases using the attributes above.

4.2 Exhibition Spatial and Narrative Authoring

Early-stage exhibition design is an inherently iterative process that requires flexible, experience-driven adjustments. To support this process (DG3), SiMUSation provides an interactive authoring environment that enables designers to rapidly create and revise exhibition plans. The system also translates these plans into a computational representation that LLM-powered agents can perceive

and reason about, enabling dynamic simulation of visitor experience within this environment (DG2).

Based on F3, we abstracted the authoring canvas into a 2D grid-based canvas with topology, reducing the LLM’s reasoning overhead while preserving the critical exhibition contexts that influence the visitor experience: routing paths, spatial connectivity, and encounters with narratives and exhibition content. The authoring process is organized into three functional layers that provide the simulation with exhibition-related context (Figure 1.1.b). The first layer defines the space, including walls, entrances, exits, and passage points, as well as floor areas that indicate functional and narrative sectors. All these elements establish the exhibition’s accessible boundaries and provide the spatial and semantic context for visitor movement. The second layer consists of exhibits placed within these defined zones. Each exhibit is represented through multi-modal information, including images, content and background descriptions, and other attributes such as artist and medium. The third layer includes facilities and staff, such as interactive installations, resting areas, and staffs, which represent environmental and operational factors that may influence visitor experience. After authoring, SiMUSation converts the configured layout into a JSON representation, translating the exhibition plan into a visiting simulation.

4.3 Visitor Experience Simulation and Representation

Building on the persona structure introduced above, SiMUSation models visitor experience as an evolving process shaped by the interaction of personal, sociocultural, and physical contexts throughout the visit. Rather than treating personas as static profiles, the

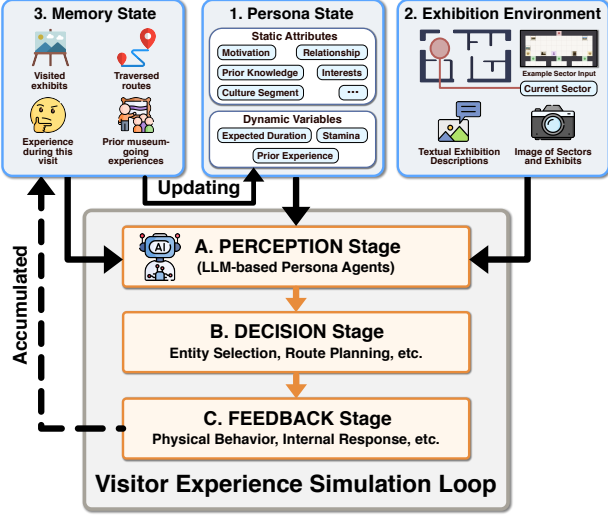


Figure 3: Perception-Decision-Feedback simulation loop in SiMUSation, where each LLM-driven visitor agent updates its behavior and responses over time based on persona state, situated environmental input, and memory state.

framework uses these factors as simulation inputs to account for how LLM-driven visitors attend to exhibits, move through the exhibition space, and respond to what they encounter. To achieve this, we design the experience simulation as a multi-modal Perception-Decision-Feedback loop (Figure 3) and translate the simulation process into a dual-layer representation, Figure 1.2 (DG2).

At each step of the simulation loop, the LLM-based persona agent is conditioned on three core input streams.

1. Visitor persona state (Figure 3.1) provides the basis for decision making. These include the persona attributes defined in subsection 4.1 (i.e., **Identity-related motivation**, **Prior knowledge**, **Interests**, **Culture segment**, **Relationship**), which affect expectations, exhibit selection, and route planning. They also include dynamic state variables such as **Stamina** [29] and **Expected visiting duration**, which may influence how long visitors continue exploring, how selectively they browse, and whether they seek rest. In addition, **Prior experience** can be derived from prior simulated visiting experience to model returning visitors.

2. Situated environmental input (Figure 3.2) provides the locally available context for perception. SiMUSation limits each visitor’s sight to the current sector, including nearby exhibits, facilities, and staff, rather than assuming access to the full exhibition at once. This input is provided through both a layout image of the current sector and structured textual descriptions of accessible entities. When the visitor chooses to inspect a specific exhibit, the agent further receives exhibit-level multi-modal input, including the exhibit image and its descriptive metadata.

3. Memory state (Figure 3.3) maintains information accumulated across the visit, such as visited exhibits, traversed routes, and experience during this visit, while also incorporating relevant **Prior experience**. It supports continuity across steps and enables more coherent sequential behavior and interpretation over time.

Given these inputs, the simulation proceeds iteratively through perception, decision, and feedback. In the perception phase, the visitor receives multi-modal information from the local environment,

along with its persona-related state and memory. In the decision phase, the model determines the next action based on these combined inputs, such as moving toward another exhibit, interacting with an installation, asking staff for directions or interpretation, or resting in a designated area. In the feedback phase, the consequences of that action update the visitor’s state, memory, and subsequent access to information. For example, resting may reduce fatigue, while interaction with exhibits, facilities, or staff may alter later movement, attention, and response. These updated states then form part of the input to the next simulation step.

To make the resulting simulation trace useful for exhibition design, SiMUSation represents visitor experience feedback through two aligned layers: observable physical behavior and internal response [22, 25, 44, 54, 99, 108]. This representation draws on prior work and our formative study findings, with further details in the supplemental material. **The first layer** captures visible behaviors that designers could reasonably infer or observe in real settings [54], including movement and navigation behaviors, information-access and attention behaviors such as viewing an exhibit, interaction and engagement behaviors such as taking notes or sketching, and social behaviors such as speaking with companions. All these behaviors make the simulated visit legible as a trajectory through the exhibition space; **The second layer** captures the visitor’s moment-to-moment internal response during the visit. Accordingly, SiMUSation represents internal response across three closely related dimensions: cognitive and comprehension response (e.g., whether the visitor understands the content or can connect it to prior knowledge), affective response (e.g., feelings of interest, confusion, surprise, boredom, or emotional resonance), and meaning- and value-related response [22, 44]. The latter includes ways in which the visit may contribute to personal, intellectual, social, or physical well-being, such as a sense of identity, reflection, or restoration.

Both layers allow SiMUSation to represent visitor experience not only as a sequence of actions, but also as an interpretable process of situated reasoning and response. For each simulated step, SiMUSation can therefore output both what the visitor does and how the visit is internally experienced, helping designers examine how specific spatial, curatorial, and interpretive choices may influence different audiences throughout the visit.

4.4 Multi-Granular Feedback for Insight-Driven Iteration

To support insight-driven iteration, SiMUSation follows an overview-to-detail analysis logic [87] and provides multi-granular quantitative and qualitative feedback from individual visitors to aggregated groups (DG2). This structure allows designers to move between broad patterns and specific cases, while comparing how visitor experience varies across different audiences in the exhibition, thus further supporting iterative refinement and revision (DG3).

From each simulated visit, SiMUSation first extracts individual-level behavioral traces, including visited exhibits and sectors, dwell time, stops, and routes. Based on these traces, it computes behavioral evaluation metrics [101], such as Average Holding Time and Walkthrough Rate (Table 3), to assess attention, circulation, and engagement patterns across personas and visitor groups. In parallel, SiMUSation summarizes each visitor’s qualitative feedback from the full visit trace, including overall impressions, suggestions,

and satisfaction (as detailed in the supplemental material). These persona-grounded summaries can then be aggregated further for selected visitor groups or the full simulated population, enabling comparisons of shared and divergent responses across audiences.

4.5 Prototype

Based on the SiMUSation framework, we developed a working prototype of an LLM-powered interactive system that simulates various visitor personas in a virtual environment as a reference for early-stage exhibition design (Figure 2).

4.5.1 Implementation Details. The SiMUSation prototype is developed as a web-based application. The front-end interface is built using Vite and Vue.js to support responsive and interactive visualizations. The back-end relies on FastAPI to manage system requests and LangChain to orchestrate the iterative perception-decision-feedback simulation loops for the LLM-based agents. We utilize Qwen3.5-Plus [92] as our core LLM for its vision-language understanding and spatial reasoning capabilities, enabling it to process multi-modal inputs such as textual descriptions and exhibit images.

The simulation loop is mainly driven by three LLM-based agents. The **Visitor Agent** governs each simulated persona through two chained sub-stages and one summary stage: a *Navigation Decision* step that decides which entity in environment to approach next based on the visitor’s profile, current position, and prior experience of before and during visiting, and an *Experience Feedback* step that generates the visitor’s concrete behavior, responses, and inner attribution upon arriving at the chosen exhibit. When ending of the visiting, the agent will *Summary* the experience. The **Staff Agent** simulates docent or guide interactions when a visitor persona initiates a question, producing context-aware dialogue grounded in the exhibit’s curatorial narrative. Finally, the **Analyst Agent** synthesizes each visitor’s full trajectory and feedback into a structured post-visit summary that highlights key experiential patterns and potential design issues. All agents receive structured output formats to ensure machine-parseable responses that can be directly rendered by the front-end. Further implementation details and prompts are provided in the supplemental material.

4.5.2 Interface. The prototype interface consists mainly of five panels (Figure 2): Exhibition Authoring Panel; Persona Setting Panel; Simulation Process on the Exhibition Map; Aggregate Visitor Analytics Panel; Individual Visitor Experience Panel to support an iterative authoring-simulating-iterating workflow, enabling designers to easily test and refine their design plans:

Exhibition Authoring Panel (DG3): This panel supports authoring for the exhibition. Users first specify the exhibition theme, and then switch among three authoring views corresponding to the layers in subsection 4.2: *space view*, *exhibit view*, and *facilities and staff view* (Figure 2.A). They are organized around a shared Map Panel, allowing users to configure the exhibition context for simulation.

In the *space view*, users can upload a floor plan or edit the map directly by placing entrances, exits, floors, gates/passages, and walls. Designers can name each sector for narrative or functional roles and assign floor colors to distinguish areas. In the *exhibit view*, users can browse exhibits by filtering with department, type, and highlight tags, or by search. Clicking an exhibit card reveals its

detailed information, and selected exhibits can be placed onto the Map Panel. In the *facilities and staff view*, users can place rest points and staff locations, and configure the content of interactive facilities before positioning them on the Map Panel. These views support the structured authoring of the spatial, semantic, and operational context required for visitor simulation.

Persona Setting Panel (DG1, DG2): This panel includes a *persona building view* and a *persona profiles view* (Figure 2.B). In the persona building view, users can generate a specified number of personas based on selections from the two anchors in subsection 4.1.2, or generate them fully at random. The persona profiles view displays each persona’s basic information. Clicking a persona card reveals the detailed attributes described in subsection 4.1.1, which users can further edit. Users can also delete personas in batches.

Simulation Process on the Exhibition Map (DG2, DG3): This panel supports both design-plan editing and the visualization of the visiting simulation process on the exhibition map (Figure 2.C). Users can switch between *edit* and *simulate* modes through the controls at the top of the panel. In *simulate* mode, the panel visualizes each visitor’s movement trajectory and interactions with entities in the environment. It also displays the current simulation step with the corresponding behavior, time, and stamina state.

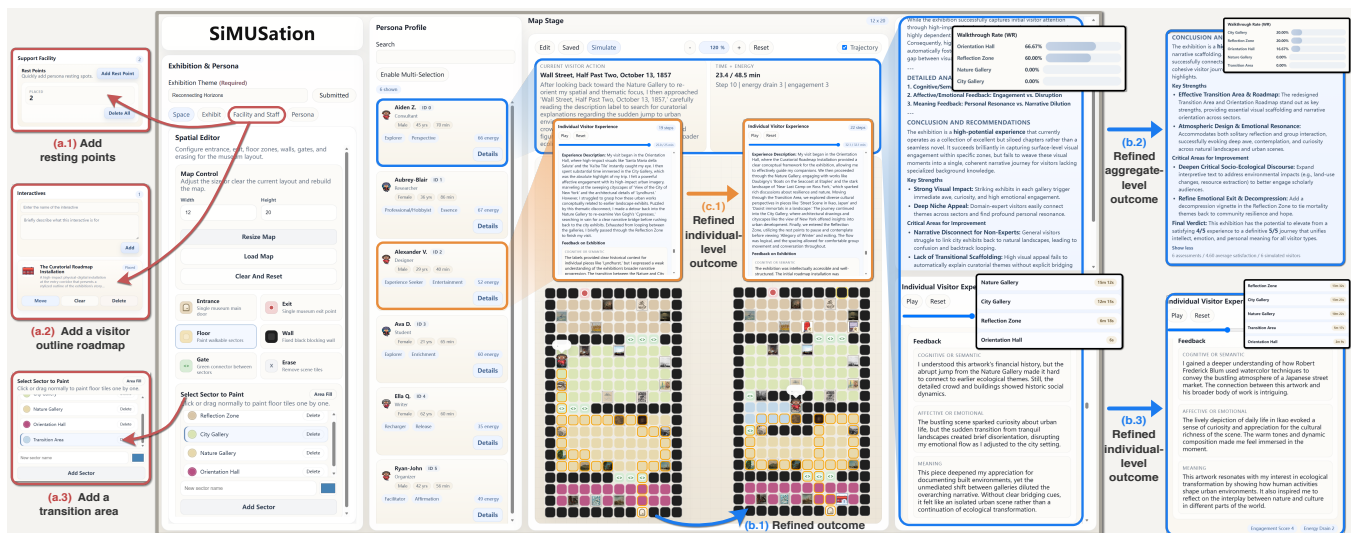
Aggregate Visitor Analytics Panel (DG2, DG3): Following an overview-to-detail analysis mantra, this panel provides aggregated summaries and behavioral metrics for selected visitor groups or the full simulated population (Figure 2.D). Users can select groups through a drop-down menu based on the two persona anchors. The panel presents group-level summaries of visitor satisfaction, overall feedback, perceived strengths, and suggestions, with evaluation metrics including Visitor Count (VC), Average Holding Time (AHT), Attracting Power (AP), Sweep Rate Index (SRI), Percentage of Diligent Visitors (%DV), Walkthrough Rate (WR), Stops at Exhibits, and Dwell Time (DT), in Table 3. These outputs help users examine overall exhibition performance across different audiences.

Individual Visitor Experience Panel (DG2, DG3): In the individual panel, this panel presents each visitor’s step-by-step experience during the visit with a final summary of the overall experience (Figure 2.E). It also reports individual behavioral measures, such as visited sectors and time spent, stopped exhibits and dwell time, and total visit duration. Users can pause the simulation shown in the Map Panel or drag the timeline to inspect what the visitor was doing at a specific step. These two panels support both cross-group analysis and case-level inspection of the simulated visitor experience, enabling multi-granular analysis to inform iteration.

4.6 User Scenario

To illustrate how SiMUSation supports design reasoning, we describe an early-stage exhibition scenario. Designer Lily is developing a temporary exhibition on how urban development reshapes human relationships with the natural environment. The goal is to connect scenic imagery, ecological observation, and urban experience through a coherent visitor journey. Lily has a preliminary concept and candidate exhibits, but is unsure whether the arrangement can convey the intended narrative to different visitors.

Lily divides the exhibition into four sectors: an *Orientation Hall*, a *Nature Gallery*, a *City Gallery*, and a *Reflection Zone* linking the two perspectives. Exhibits and pathways are then arranged on the



map panel to guide visitors from natural observation to urban development and personal reflection (Figure 4). Lily then defines a small set of personas to represent likely audience diversity, such as an *Explorer* interested in ecology and local geography, and an *Experience Seeker* visiting casually for visual enjoyment. Using SiMUSation's persona construction, Lily reviews the generated profiles and makes minor adjustments to match the expected audience.

To understand why, Lily inspects individual simulated visits. *Alexander* (ID 2) spends substantial time in the *City Gallery* and shows strong affective engagement with urban imagery, but expresses confusion and fatigue about how these works relate to earlier landscape exhibits (Figure 4.c.1). *Aiden* (ID 0) follows the intended route more closely, yet feels the shift from the *Nature Gallery* to the *City Gallery* is abrupt (Figure 4.b.3).

Based on these findings, Lily revises the plan by adding an outline map into the *Orientation Hall*, adding a *Transition Area* between the *Nature Gallery* and *City Gallery*, and placing a resting point near the entrance to the *Reflection Zone* (Figure 4.a). Based on the refinement, the feedback of the second simulation shows these changes improve narrative flow continuity and sustain engagement across these audience groups (Figure 4.b, 4.c).

5 Evaluation

RQ2: How does the iterative simulation-feedback framework support the early-stage exhibition design?

5.1 Methods

Task and Data. We asked participants to design an exhibition around the theme *cross-cultural dialogue in landscapes*. To support this task, we curated a set of open-access information on 600+ artworks related to landscape representation across different cultures from The Metropolitan Museum of Art collection¹. We obtained the artworks’ metadata and images through the museum’s official API² and converted them into the format required by SiMUSation. Given a predefined exhibition space, participants were asked to select around 20 exhibits, arrange them in the layout, and construct a coherent narrative for the exhibition.

Procedure. Participants were asked to familiarize themselves with the exhibition theme and dataset one week before the study. On the study day, each session began with a 15-minute introduction to the study background and relevant domain knowledge, followed by a 30-minute system walkthrough and free exploration to familiarize with the functions and interactions. Participants then completed a

¹<https://www.metmuseum.org/art/collection>

²<https://metmuseum.github.io/>

50-minute design task, progressing from initial exhibition authoring to simulation, iteration, and refinement. Afterward, they spent 10 minutes completing questionnaires and 15 minutes in a semi-structured interview. We logged each participant's initial designs (v1, authored without simulation) and final designs (v2, revised with simulation), and invited four experts to review and score the results. For each participant, the experts evaluated the two designs in pairs. To support blind review and reduce bias, we shuffled both the order of the design pairs across participants and the placement order of the initial and final designs within each pair. These mappings were known only to the research team during scoring.

Measurements. To assess users' perceptions of SiMUSation, we evaluated its usability using the System Usability Scale (SUS) [8]. We also measured perceived reliability and early-stage design support using 7-point Likert-scale items. The measurements were as follows. For reliability and acceptance of persona simulation:

- **Persona differentiation:** whether the framework captured differences in exhibition experience across different personas;
- **Behavioral plausibility:** whether the simulated behavior aligned with expected visitor behavior in real exhibition settings;
- **Response realism:** whether simulated internal responses, such as emotions and narrative understanding, appeared natural and consistent with visitor personas;
- **Experience explainability:** whether the simulation helped participants understand the reasons behind visitors' feedback.

For the effectiveness of the SiMUSation, we assessed:

- **Design insight:** whether the simulation revealed design blind spots or potential problems overlooked in initial planning;
- **Decision support:** whether outputs such as trajectories and individual feedback provided useful support for revising layout or narrative structure;
- **Design confidence:** whether using SiMUSation increased confidence in the design compared with relying only on intuition;
- **Iteration effort:** whether SiMUSation enabled easy and low-cost testing of different layouts and exhibit combinations;
- **Workflow optimization:** whether integrating editing and experience simulation improved early-stage exhibition planning.

For expert evaluation, each pair of designs from the same participant was reviewed separately using 7-point Likert-scale ratings on the following criteria: 1) **Narrative coherence:** whether the cross-cultural theme was presented through a clear and reasonable narrative logic; 2) **Spatial organization:** whether spatial exhibit order and circulation were natural and well organized; 3) **Facility placement suitability:** whether the locations of facilities and staff were suitable in the exhibition layout.

5.2 Results

Based on questionnaire responses, post-study interviews, and expert evaluations of participants' design outcomes, the results indicate that SiMUSation was positively received in terms of both usability and perceived usefulness for early-stage exhibition design. The SUS analysis yielded a mean score of 79.38 (SD = 7.78), suggesting a good level of usability. Beyond overall usability, the findings further suggest that participants generally found the LLM-driven personas plausible and their responses interpretable. Participants also considered integrating simulation with feedback into the workflow as

Expert Scores for the Initial and Revised Designs in Box Plot

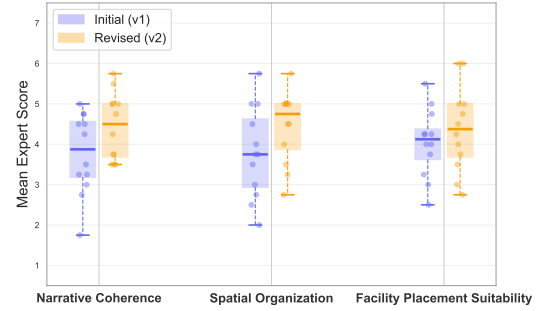


Figure 5: Four experts' ratings of the design from twelve participants between the initial and revised designs.

helpful for identifying design issues and supporting refinement. These perceptions were further reflected in the expert assessments, in which revised designs generally received higher scores than the initial versions on narrative coherence, spatial organization, and facility placement suitability (Figure 5). The inter-rater agreement among experts was $ICC(2,1) = 0.83$ for v1 and $ICC(2,1) = 0.77$ for v2.

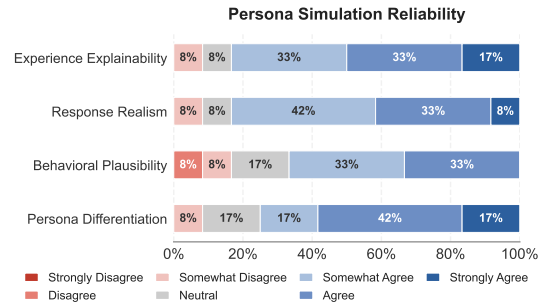


Figure 6: Twelve participants' ratings of the questionnaire related to the reliability of SiMUSation.

5.2.1 Plausibility and Explainability of LLM-Driven Personas. Participants generally regarded the simulated personas as credible proxies for exhibition audiences, and the quantitative results showed consistently positive evaluations across all four measures (Figure 6). For **persona differentiation** ($M=5.42$, $SD=1.19$), participants perceived the generated visitors as meaningfully distinct rather than interchangeable. This perception was not only based on demographic variation, but also on exhibition-related dimensions such as prior experience and interests. For example, U1 noted that motivations and cultural segments were core dimensions, allowing them to generate personas in batches and then further customize extreme or boundary cases one by one. In this way, audience consideration did not rely solely on personal intuition.

For **behavioral plausibility** ($M=4.75$, $SD=1.23$), the overall evaluation remained positive. Participants generally considered behaviors sufficiently reasonable for judging visitor flow and attention distribution in early-stage planning. U12 commented that paths and dwell time were relatively objective in exhibition settings and therefore aligned well with real design concerns. At the same time, some participants pointed out limits to realism. For instance, U4 noted that other behaviors, such as interaction behaviors, need greater richness from real visitors. Notably, these concerns were typically framed as limitations in modeling complex human behavior rather

than as a rejection of the approach itself. As U8 suggested, “even if some aspects of real-world experience are somehow simplified, the current framework remains useful for early-stage design.”

Participants gave more positive evaluations to **response realism** ($M=5.25$, $SD=1.01$), and especially to **experience explainability** ($M=5.42$, $SD=1.11$). The findings suggest that participants valued SiMUSation because it showed not only what a persona did, but also why each decision was made. Several participants emphasized that the combination of trajectory visualization, internal responses, and step-by-step reasoning made the simulated experience easier to understand and, in turn, more credible. As U9 said, “I found a repeated viewing pattern initially seemed confusing, but became reasonable after considering the persona’s profile as an energy explorer and stepwise feedback.” This indicates that explainability directly influenced whether participants regarded a persona as trustworthy. In this respect, the internal responses were particularly important. Participants appreciated SiMUSation’s ability to display immediate reactions across multiple dimensions and relate them to persona-specific states. U11, for instance, noted that they read the internal responses most frequently and found them the most useful, particularly for assessing whether the persona’s thoughts aligned with the ideas they intended to convey. Rather than producing generic comments on the exhibition, these responses were often interpreted as grounded in each persona’s interests and background. Despite these strengths, participants also noted risks related to LLM sycophancy and hallucination. U7 and U4 found the aggregate-level summary overly positive and insufficiently critical, even when some layouts were intentionally confusing for visitors. U2 also questioned whether real visitors’ motivations would ever be so explicit.

Overall, participants’ trust in the personas depended less on whether every action fully reproduced real human behavior, and more on whether the simulated experience was interpretable, insightful, consistent with the persona, and supported by a traceable causal chain. This perceived credibility was also important for later use of SiMUSation, as participants were willing to treat simulated visitor experiences as actionable evidence precisely because they could inspect both movement trajectories and internal responses.

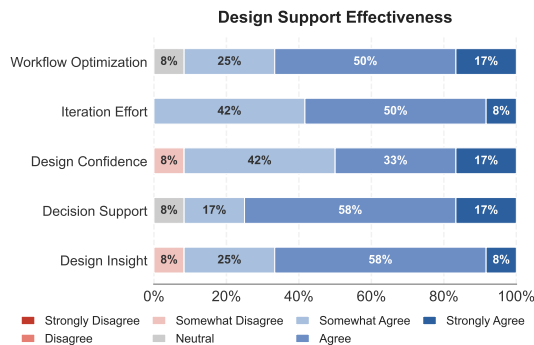


Figure 7: Twelve participants’ ratings of the questionnaire related to the effectiveness of SiMUSation.

5.2.2 Support for Early-Stage Exhibition Design through Iterative Simulation and Feedback. The results also suggest that participants found SiMUSation useful for early-stage exhibition design, especially in revealing overlooked issues, supporting revision, reducing

iteration effort, and increasing confidence in decision-making (Figure 7). This indicates that SiMUSation is not only a simulation tool, but also an iterative design aid for early-stage design.

For **design insight** ($M=5.58$, $SD=0.95$), participants reported that the simulation helped expose design blind spots that were not apparent in their initial plans. U3 noted from the visitor trajectories that, although the cultural sections had been arranged according to their real-world geographic positions, many personas mainly attended to the central and rear spaces while overlooking the side areas, suggesting that this spatial logic was not immediately clear and that stronger guidance was needed. Similarly, U4 reflected after reading the feedback that the overview area, which they had initially designed around a highlighted exhibit to serve as a strong hook, was not effective enough in helping visitors both grasp the overall exhibition structure and transition into the main theme.

For **decision support** ($M=5.83$, $SD=0.80$), SiMUSation supports design revision by combining aggregated and individual-level evidence across both quantitative and qualitative forms. At the macro level, U9 noted that visualized metrics made overall trends and attention hotspots easy to identify. At the micro level, U12 emphasized that individual behaviors and thoughts helped them better take the visitor’s perspective to review the design. This combination, in which high-level patterns indicated where problems occurred, and persona-level evidence helped explain why, also strengthened participants’ overall **design confidence** ($M=5.50$, $SD=1.04$). U7 suggested that because the simulated persona behavior was rational and explainable, it increased their confidence for designs rather than personal intuition. This finding is also consistent with the role of explainability discussed in subsection 5.2.1 in supporting understanding and decision-making.

The results for **iteration effort** ($M=5.67$, $SD=0.62$) and **workflow optimization** ($M=5.75$, $SD=0.83$) further indicate clear advantages. Compared with traditional approaches such as writing outlines, sketching, or manually building 3D models, SiMUSation supported a tightly integrated and relatively low-cost edit-simulate-analyze-revise loop. U11 noted that, because layout editing and simulation feedback were integrated into the same environment, designers could avoid labor-intensive manual drafting. This aligns with U2’s observation that the tool made both layout modification and simulation results visible and editable.

Expert evaluations of the initial and revised designs provided complement whether support early-stage design. The results showed that revised designs mainly received higher scores than the initial versions across all three evaluation dimensions. Among these, the improvements in **narrative coherence** (Δ , $M=0.67$, $SD=0.98$) and **spatial organization** (Δ , $M=0.65$, $SD=0.97$) were especially notable. These findings are consistent with participants’ reports that simulation visualizations and step-by-step feedback helped them identify and address imbalanced circulation, movement bottlenecks, and neglected areas, thereby improving the coherence and integration of the exhibition content and space.

Overall, the main value of SiMUSation in early-stage exhibition design lies in its support for a practical, visitor-centered iterative loop. It helped designers identify design problems, explore alternatives at relatively low cost, and understand diverse visitor experiences for decision making, thereby translating these insights into improvements in spatial organization and narrative coherence.

6 Discussion

Through the design and development of SiMUSation, informed by findings from the user study, we derive implications that extend beyond the framework itself for the design of LLM-supported tools in culture- and space-related creative tasks, such as exhibition design and planning. We also reflect on the current limitations of the framework and outline directions for future work.

6.1 Design Implications

6.1.1 From Generator to Experiential Reflection. Many AI tools for design still operate according to their end-to-end nature, producing relatively complete outputs directly from user prompts. While this approach can improve efficiency, heavy reliance on automated generation may weaken user experience and ownership [10, 19], and may also constrain human creativity and critical reflection [18, 33]. Prior research on human-AI co-creation has examined both the role of AI [27, 107] and the workflow across tasks [83, 106], in which users express intent while AI collects and understands it, then contributes to content creation, followed by user review, decision, and revision. Our findings suggest that this model may be less suitable for exhibition design, where subjectivity, cultural interpretation, and narrative construction are central. Participants did not primarily expect AI to generate or modify layouts. Instead, they valued AI as a support for experiential reflection, enabling them to examine their ideas from different visitor perspectives and to provide feedback on the design. What mattered was not automated production, but the system’s ability to offer diverse, plausible, and reflective responses that supported later refinement. This also suggests that, for other user-centered creative tasks, AI-supported systems may benefit from shifting from generation-driven to feedback-driven support. In this role, AI acts less as a designer and more as a critical reviewer that supports reflection while preserving user agency.

6.1.2 Towards Situated Agents for Design. Conventional AI-based design assistants are strong in semantic understanding and reasoning but often remain detached from physical constraints [60, 103]. For design tasks involving space and movement, AI support may benefit from going beyond prompt-response generation. We suggest considering embodied or situated modeling in agent design, where perception, decision-making, and feedback are continuously integrated with internal states, environmental conditions, and dynamic memory [26, 31, 70]. In this way, behavior is shaped by both the present context and prior interactions over time. Such stateful agents provide more realistic feedback for user-centered design, helping designers surface issues in narrative flow, spatial sequence, and experiential continuity that static generation cannot capture.

6.1.3 Explainability and Trustworthy AI. Our findings show that explainability supported transparency and understanding in the use of AI personas. By making visible why a persona behaved as it did, SiMUSation allowed designers to inspect both the observable process and the underlying rationale. This helped participants interpret the simulated experience from the visitor’s perspective and relate it to their own design intentions. However, while such explainability builds immediate trust [12, 23], fostering grounded, lasting trustworthiness [6, 23] requires recognizing that explainability should not be conflated with real-world accuracy. Although SiMUSation provides traceable rationales for reflection, these are

model-generated accounts rather than ground-truth evidence of human cognition. To prevent over-reliance on simulated feedback, future AI design aids must move beyond merely explaining behaviors to explicitly communicating AI risks and uncertainties.

6.2 Limitations and Future Work

Our work has several limitations that point to directions for future research. First, SiMUSation currently relies on a 2D representation, which simplifies some properties of real exhibition environments. Extending the system to 3D simulation environments may support richer spatial perception (e.g., 3D sightlines, lighting) and more immersive forms of evaluation. Second, although persona generation is grounded in empirical data, LLM-powered agents may reflect clichés, sycophancy, or overly positive judgments. This may reduce the system’s ability to provide sufficiently critical feedback in edge cases. To mitigate this, future work could explore fine-tuned models and broader visitor identity settings to improve the rigor and diversity of simulated responses. It could also introduce more robust principles and computational models beyond the current dual-layer representation to better align the simulation with the logic of human perception. Third, real exhibitions are not only individual experiences but also social settings. The current framework mainly models individual visiting while incorporating basic social relationships, such as peer influence. It does not yet capture richer dynamic multi-agent interactions, such as crowd formation. Modeling such processes may enhance ecological validity and reveal emergent spatial-use patterns. Finally, the current pipeline requires re-running the simulation after even minor changes. Incremental simulation methods that update only locally affected agent states to better support iterative author-simulate-iterate workflows.

7 Conclusion

In this study, we present SiMUSation, an interactive framework that uses LLM-driven visitor personas to simulate visiting experiences and support early-stage museum exhibition design. SiMUSation combines diverse visitor profiles with authored exhibition contexts to model visitor experiences, represented through physical behaviors and internal responses within the exhibition. Through these simulated personas, designers can inspect how visits unfold, obtain multi-faceted feedback from diverse audience perspectives, and iteratively refine spatial and narrative plans before physical implementation. The user study suggests that the virtual personas can serve as a reference for exhibition audiences and help users identify overlooked design blind spots, thereby enriching and supporting the early-stage exhibition design process. SiMUSation also shows the potential of LLM-powered simulation for audience-informed design and evaluation tasks.

Acknowledgments

We sincerely thank the participants in our formative study and user study for sharing their thoughts and suggestions on SiMUSation, and the four experts for their careful ratings and feedback. We are also grateful to all reviewers for their valuable insights and feedback. We acknowledge the partial use of LLMs to assist in the writing process. The LLMs were employed as a tool for polishing the manuscript to enhance the clarity and quality of the text.

References

- [1] Michela Addis, Valentina Copat, and Cecilia Martorana. 2024. Museum Experience and Its Impact on Visitor Reactions. *Journal of Philanthropy and Marketing* 29, 1 (2024), e1826. doi:10.1002/nvsm.1826
- [2] Harsh Agrawal, Aditya Mishra, Manish Gupta, and Mausam. 2023. Multimodal Persona Based Generation of Comic Dialogs. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki (Eds.). ACL, Toronto, Canada, 14150–14164. doi:10.18653/v1/2023.acl-long.791
- [3] Molood Ale Ebrahim Dehkordi, Jonas Lechner, Amineh Ghorbani, Igor Nikolic, Emile Chappin, and Paulien Herder. 2023. Using Machine Learning for Agent Specifications in Agent-based Models and Simulations: A Critical Review and Guidelines. *Journal of Artificial Societies and Social Simulation* 26, 1 (2023). doi:10.18564/jasss.5016
- [4] Moneerah Almeshari, John Dowell, and Julianne Nyhan. 2021. Museum Mobile Guide Preferences of Different Visitor Personas. *Journal on Computing and Cultural Heritage* 14, 1, Article 9 (Dec. 2021), 13 pages. doi:10.1145/3423186
- [5] Liliana Ardisson, Tsvi Kuflik, and Daniela Petrelli. 2012. Personalization in Cultural Heritage: the Road Travelled and the One Ahead. *User Modeling and User-adapted Interaction* 22, 1 (2012), 73–99. doi:10.1007/s11257-011-9104-x
- [6] Zahra Atf and Peter R. Lewis. 2026. Is Trust Correlated With Explainability in AI? A Meta-Analysis. *IEEE Transactions on Technology and Society* 7, 1 (2026), 70–77. doi:10.1109/TTS.2025.3558448
- [7] Alessandro Bollo and Luca Dal Pozzolo. 2005. Analysis of Visitor Behaviour inside the Museum: An Empirical Study. In *Proceedings of the 8th international conference on arts and cultural management*, Vol. 2. Citeseer, 1–13.
- [8] John Brooke et al. 1996. SUS-A Quick and Dirty Usability Scale. Usability Evaluation in Industry 189, 194 (1996), 4–7. *This page intentionally left blank* (1996).
- [9] Fan Bu, Zhan Li, Yu Jiang, and Ziyao Liu. 2026. Reframing Museum Intelligence: A Survey of Large Language Models in Museums. *Authorea Preprints* (2026). doi:10.22541/au.177188048.80779493/v2
- [10] Daniel Buschek, Lukas Mecke, Florian Lehmann, and Hai Dang. 2021. Nine Potential Pitfalls when Designing Human-AI Co-Creative Systems. arXiv:2104.00358 [cs.HC] <https://arxiv.org/abs/2104.00358>
- [11] Andrew Chalmers, Faisal Zaman, and Taehyun Rhee. 2024. Avatar360: Emulating 6-DoF Perception in 360° Panoramas through Avatar-Assisted Navigation. In *2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*. IEEE, 630–638. doi:10.1109/VR58804.2024.00083
- [12] Bhanu chander, Chinju John, Lekha Warriar, and Kumaravelan Gopalakrishnan. 2025. Toward Trustworthy Artificial Intelligence (TAI) in the Context of Explainability and Robustness. *Comput. Surveys* 57, 6, Article 144 (Feb. 2025), 49 pages. doi:10.1145/3675392
- [13] Bingqing Chen, Ruoyu Wen, Shufang Tan, and Yue Li. 2025. Exploring User Preferences for Museum Guides: The Role of Chatbots in Shaping Interactive Experiences. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI EA '25). ACM, New York, NY, USA, Article 256, 8 pages. doi:10.1145/3706599.3720067
- [14] Jiangjie Chen, Xintao Wang, Rui Xu, Siyu Yuan, Yikai Zhang, Wei Shi, Jian Xie, Shuang Li, Ruihan Yang, Tinghui Zhu, Aili Chen, Nianqi Li, Lida Chen, Caiyu Hu, Siye Wu, Scott Ren, Ziquan Fu, and Yanghua Xiao. 2024. From Persona to Personalization: A Survey on Role-Playing Language Agents. *Transactions on Machine Learning Research* (2024). <https://openreview.net/forum?id=xrO70E8UIZ>
- [15] Anping Cheng, Dongming Ma, Younghwan Pan, and Hao Qian. 2024. Enhancing Museum Visiting Experience: Investigating the Relationships Between Augmented Reality Quality, Immersion, and TAM Using PLS-SEM. *International Journal of Human-Computer Interaction* 40, 17 (2024), 4521–4532. doi:10.1080/10447318.2023.2227832
- [16] Yoonseo Choi, Eun Jeong Kang, Seulgi Choi, Min Kyung Lee, and Juho Kim. 2025. Proxona: Supporting Creators' Sensemaking and Ideation with LLM-Powered Audience Personas. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '25). ACM, New York, NY, USA, Article 149, 32 pages. doi:10.1145/3706598.3714034
- [17] Sophia Diamantopoulou, Eva Insulander, and Fredrik Lindstrand. 2012. Making Meaning in Museum Exhibitions: Design, Agency and (Re-) Representation. *Designs for learning* 5, 1-2 (2012), 11–28. doi:10.2478/dfl-2014-0002
- [18] Anil R. Doshi and Oliver P. Hauser. 2024. Generative AI Enhances Individual Creativity but Reduces the Collective Diversity of Novel Content. *Science Advances* 10, 28 (July 2024). doi:10.1126/sciadv.adn5290
- [19] Fiona Draxler, Anna Werner, Florian Lehmann, Matthias Hoppe, Albrecht Schmidt, Daniel Buschek, and Robin Welsch. 2024. The AI Ghostwriter Effect: When Users do not Perceive Ownership of AI-Generated Text but Self-Declare as Authors. *ACM Transactions on Computer-Human Interaction* 31, 2, Article 25 (Feb. 2024), 40 pages. doi:10.1145/3637875
- [20] Andrew Emerson, Nathan Henderson, Jonathan Rowe, Wookhee Min, Seung Lee, James Minogue, and James Lester. 2020. Early Prediction of Visitor Engagement in Science Museums with Multimodal Learning Analytics. In *Proceedings of the 2020 International Conference on Multimodal Interaction (Virtual Event, Netherlands) (ICMI '20)*. ACM, New York, NY, USA, 107–116. doi:10.1145/3382507.3418890
- [21] John H. Falk. 2016. *Identity and the Museum Visitor Experience*. Routledge. doi:10.4324/9781315427058
- [22] John H. Falk and Lynn D. Dierking. 2016. *The Museum Experience Revisited*. Routledge. doi:10.4324/9781315417851
- [23] Andrea Ferrario and Michele Loi. 2022. How Explainability Contributes to Trust in AI. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (Seoul, Republic of Korea) (FAccT '22). ACM, New York, NY, USA, 1457–1466. doi:10.1145/3531146.3533202
- [24] Alessio Ferraro, Carla Limongelli, Mauro Mezzini, and Giuseppe Sansonetti. 2022. Using Deep Learning for Collecting Data about Museum Visitor Behavior. *Applied Sciences* 12, 2 (2022). doi:10.3390/app12020533
- [25] Group for Education in Museums. 2026. *Generic Learning Outcomes Framework 2026*. <https://gem.org.uk/resource/generic-learning-outcomes-framework-2025/>
- [26] Pascale Fung, Yoram Bachrach, Asli Celikyilmaz, Kamalika Chaudhuri, Delong Chen, Willy Chung, Emmanuel Dupoux, Hongyu Gong, Hervé Jégou, Alessandro Lazaric, Arjun Majumdar, Andrea Madotto, Franziska Meier, Florian Metzger, Louis-Philippe Morency, Théo Moutakanni, Juan Pino, Basile Tervet, Joseph Tighe, Paden Tomasello, and Jitendra Malik. 2025. Embodied AI Agents: Modeling the World. arXiv:2506.22355 [cs.AI] <https://arxiv.org/abs/2506.22355>
- [27] Jie Gao, Simret Araya Gebreegziabher, Kenny Tsu Wei Choo, Toby Jia-Jun Li, Simon Tangi Perrault, and Thomas W Malone. 2024. A Taxonomy for Human-LLM Interaction Modes: An Initial Exploration. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI EA '24). ACM, New York, NY, USA, Article 24, 11 pages. doi:10.1145/3613905.3650786
- [28] Adrian George. 2015. *The Curator's Handbook: Museums, Commercial Galleries, Independent Spaces*. Thames & Hudson.
- [29] Benjamin Ives Gilman. 1916. Museum Fatigue. *The Scientific Monthly* 2, 1 (1916), 62–74.
- [30] Amy Grack Nelson and Sarah Cohn. 2015. Data Collection Methods for Evaluating Museum Programs and Exhibitions. *Journal of Museum Education* 40, 1 (2015), 27–36. doi:10.1080/10598650.2015.11510830
- [31] Xudong Guo, Kaixuan Huang, Jiale Liu, Wenhui Fan, Natalia Vélez, Qingyun Wu, Huazheng Wang, Thomas L. Griffiths, and Mengdi Wang. 2026. Embodied LLM Agents Learn to Cooperate in Organized Teams. *IEEE Transactions on Computational Social Systems* (2026), 1–17. doi:10.1109/TCSS.2025.3637527
- [32] Juhye Ha, Hyeon Jeon, Daeun Han, Jinwook Seo, and Changhoon Oh. 2024. CloChat: Understanding How People Customize, Interact, and Experience Personas in Large Language Models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). ACM, New York, NY, USA, Article 305, 24 pages. doi:10.1145/3613904.3642472
- [33] Sabrina Habib, Thomas Vogel, Xiao Anli, and Evelyn Thorne. 2024. How Does Generative Artificial Intelligence Impact Student Creativity? *Journal of Creativity* 34, 1 (2024), 100072. doi:10.1016/j.jyoc.2023.100072
- [34] Senyu Han, Lu Chen, Li-Min Lin, Zhengshan Xu, and Kai Yu. 2024. IBSEN: Director-Actor Agent Collaboration for Controllable and Interactive Drama Script Generation. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. ACL, Bangkok, Thailand, 1607–1619. doi:10.18653/v1/2024.acl-long.88
- [35] George E Hein. 2002. *Learning in the Museum*. routledge.
- [36] Mehdi Hellou, JongYoon Lim, Norina Gasteiger, Minsu Jang, and Ho Seok Ahn. 2022. Technical Methods for Social Robots in Museum Settings: An Overview of the Literature. *International Journal of Social Robotics* 14, 8 (2022), 1767–1786. doi:10.1007/s12369-022-00904-y
- [37] Bill Hillier and Kali Tzortzi. 2006. Space Syntax: the Language of Museum Space. In *A Companion to Museum Studies*. Wiley Online Library, 282–301. doi:10.1002/9780470996836
- [38] Sukjoo Hong, Taeha Yi, Joosun Yun, and Ji-Hyun Lee. 2021. Visitor-artwork Network Analysis Using Object Detection with Image-retrieval Technique. *Advanced Engineering Informatics* 48 (2021), 101307. doi:10.1016/j.aei.2021.101307
- [39] Sirui Hong, Mingchen Zhuge, Jonathan Chen, Xiaowu Zheng, Yuheng Cheng, Jinlin Wang, Ceyao Zhang, Zili Wang, Steven Ka Shing Yau, Zijuan Lin, Liyang Zhou, Chenyu Ran, Lingfeng Xiao, Chenglin Wu, and Jürgen Schmidhuber. 2024. MetaGPT: Meta Programming for A Multi-Agent Collaborative Framework. In *The Twelfth International Conference on Learning Representations*. <https://openreview.net/forum?id=VtmBAGCN7o>
- [40] Eilean Hooper-Greenhill. 2006. Studying Visitors. In *A Companion to Museum Studies*. Wiley Online Library, 362–376. doi:10.1002/9780470996836.ch22
- [41] Eva Hornecker and Luigina Ciolfi. 2022. *Human-Computer Interactions in Museums*. Springer Nature.
- [42] Jen-tse Huang, Man Ho Lam, Eric John Li, Shujie Ren, Wenxuan Wang, Wenxiang Jiao, Zhaopeng Tu, and Michael R Lyu. 2024. Apathetic or Empathetic? Evaluating LLMs' Emotional Alignments with Humans. *Advances in Neural Information Processing Systems* 37 (2024), 97053–97087.

- [43] Jen-tse Huang, Wenxuan Wang, Eric John Li, Man Ho Lam, Shujie Ren, Youliang Yuan, Wenxiang Jiao, Zhaopeng Tu, and Michael Lyu. 2024. On the Humanity of Conversational AI: Evaluating the Psychological Portrayal of LLMs. In *The Twelfth International Conference on Learning Representations*.
- [44] Institute for Learning Innovation. 2025. *Measuring What Matters: Museums, Well-Being, and the Real Value of Experience*. <https://www.instituteforlearninginnovation.org/measuring-what-matters-museums-well-being-and-the-real-value-of-experience/>
- [45] Rosen Ivanov and Victoria Velkova. 2025. Analyzing Visitor Behavior to Enhance Personalized Experiences in Smart Museums: A Systematic Literature Review. *Computers* 14, 5 (2025). doi:10.3390/computers14050191
- [46] Zhuoling Jiang, Yipeng Qin, and Daniel J. Finnegan. 2025. 'Chattable' Avatars: Using LLMs to Power Visitor Engagement with Historical Persons. In *Proceedings of the 38th International BCS Human-Computer Interaction Conference* (Cardiff University, UK) (BCS HCI '25). BCS Learning & Development Ltd, Swindon, GBR, 91–102. doi:10.14236/ewic/BCSHCI2025.10
- [47] Ceri Jones. 2015. Enhancing our Understanding of Museum Audiences: Visitor Studies in the Twenty-first Century. *Museum & Society* 13, 4 (2015), 539–544.
- [48] Nicolai Hejlesen Jørgensen, Sarmilan Tharmabalan, İlhan Aslan, Nicolai Brodersen Hansen, and Timothy Merritt. 2026. Guiding, Not Railroad: Design and Evaluation of a Multi-Agent System for Narrative Redirection in Role-playing Games. In *Proceedings of the 31st International Conference on Intelligent User Interfaces* (Paphos, Cyprus) (IUI '26). ACM, New York, NY, USA, 1551–1601. doi:10.1145/3742413.3789218
- [49] Ellie King, M. Paul Smith, Paul F. Wilson, Janet Stott, and Mark A. Williams. 2023. Creating Meaningful Museums: A Model for Museum Exhibition User Experience. *Visitor Studies* 26, 1 (2023), 59–81. doi:10.1080/10645578.2022.2129944
- [50] Volker Kirchberg and Martin Tröndle. 2012. Experiencing Exhibitions: A Review of Studies on Visitor Experiences in Museums. *Curator: the museum journal* 55, 4 (2012), 435–452. doi:10.1111/j.2151-6952.2012.00167.xDigitalObjectIdentifier(DOI)viewmetrics
- [51] Yoshinori Kuno, Kazuhisa Sadazuka, Michie Kawashima, Keiichi Yamazaki, Akiko Yamazaki, and Hideaki Kuzuoka. 2007. Museum Guide Robot Based on Sociological Interaction Analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (CHI '07). ACM, New York, NY, USA, 1191–1194. doi:10.1145/1240624.1240804
- [52] Joel Lanir, Tsvi Kuflik, Julia Sheidin, Nisan Yavin, Kate Leiderman, and Michael Segal. 2017. Visualizing Museum Visitors' Behavior: Where Do They Go and What Do They Do There? *Personal and Ubiquitous Computing* 21, 2 (2017), 313–326. doi:10.1007/s00779-016-0994-9
- [53] Chaeun Lee, Jinwook Kim, Hyeonbeom Yi, and Woohun Lee. 2024. Viewer2Explorer: Designing a Map Interface for Spatial Navigation in Linear 360 Museum Exhibition Video. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). ACM, New York, NY, USA, Article 400, 15 pages. doi:10.1145/3613904.3642952
- [54] Martine Levasseur and Eliseo Veron. 1983. Ethnographie d'une Exposition. In *Histoires d'expo, Peuple et culture*. Bibliothèque Publique d'Information (BPI) du Centre Pompidou, 29–32. <https://shs.hal.science/halshs-01484185>
- [55] Jiayang Li, Jiarui Jiang, Shuqing Tang, Mutao Yu, Xinyuan Song, Caoyang Xue, Yusheng Su, Xinyang Tan, and Yang Shi. 2026. All Futures at Once: Supporting Speculative Design for Placemaking with Multi-Agent Social Simulation. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems* (Barcelona, Spain) (CHI '26). ACM, New York, NY, USA, Article 1075, 20 pages. doi:10.1145/3772318.3791543
- [56] Xinyu Li, Ruiyang Zhou, Zachary C Lipton, and Liu Leqi. 2024. Personalized Language Modeling from Personalized Human Feedback. arXiv:2402.05133 [cs.CL]
- [57] Yuzhi Li, Ruixuan Yang, Jiahui Zou, Haojun Xu, and Feng Tian. 2025. Human-centric Virtual Museum: Redefining the Museum Experience Through Immersive and Interactive environments. *International Journal of Human-Computer Interaction* 41, 13 (2025), 8426–8437. doi:10.1080/10447318.2024.2408861
- [58] Yongming Li, Hangyue Zhang, Andrea Yaoyun Cui, Zisong Ma, Yunpeng Song, Zhongmin Cai, and Yun Huang. 2025. EyeSee: Enhancing Art Appreciation through Anthropomorphic Interpretations from Multiple Perspectives. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '25). ACM, New York, NY, USA, Article 660, 23 pages. doi:10.1145/3706598.3714042
- [59] Seungwon Lim, Seungbeen Lee, Dongjun Min, and Youngjae Yu. 2025. Persona Dynamics: Unveiling the Impact of Persona Traits on Agents in Text-Based Games. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. ACL, Vienna, Austria, 31360–31394. doi:10.18653/v1/2025.acl-long.1515
- [60] Yang Liu, Weixing Chen, Yongjie Bai, Xiaodan Liang, Guanbin Li, Wen Gao, and Liang Lin. 2025. Aligning Cyber Space With Physical World: A Comprehensive Survey on Embodied AI. *IEEE/ASME Transactions on Mechatronics* 30, 6 (2025), 7253–7274. doi:10.1109/TMECH.2025.3574943
- [61] Yiren Liu, Pranav Sharma, Mehul Oswal, Haijun Xia, and Yun Huang. 2025. PersonaFlow: Designing LLM-Simulated Expert Perspectives for Enhanced Research Ideation. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference* (Madeira, Portugal) (DIS '25). ACM, New York, NY, USA, 506–534. doi:10.1145/3715336.3735789
- [62] Alex Lowy and Phil Hood. 2010. *The Power of the 2 x 2 Matrix: Using 2 x 2 Thinking to Solve Business Problems and Make Better Decisions*. John Wiley & Sons.
- [63] Benjamin J McCloskey, Phillip M LaCasse, and Bruce A Cox. 2024. Natural Language Processing Analysis of Online Reviews for Small Business: Extracting Insight from Small Corpora. *Annals of Operations Research* 341, 1 (2024), 295–312. doi:10.1007/s10479-023-05816-2
- [64] Christofer Meinecke, Chris Hall, and Stefan Jänicke. 2022. Towards Enhancing Virtual Museums by Contextualizing Art through Interactive Visualizations. *Journal on Computing and Cultural Heritage* 15, 4, Article 62 (Dec. 2022), 26 pages. doi:10.1145/3527619
- [65] Nurul Fathihin Mohd Noor Shah and Masitah Ghazali. 2018. A Systematic Review on Digital Technology for Enhancing User Experience in Museums. In *User Science and Engineering*, Natrah Abdullah, Wan Adilah Wan Adnan, and Marcus Foth (Eds.). Springer Singapore, Singapore, 35–46. doi:10.1007/978-981-13-1628-9_4
- [66] Morris Hargreaves McIntyre. 2021. Culture Segments. <https://mhm.co/signature-tools/culture-segments/>
- [67] Markus Brøndberg Nielsen, Mathias Øgaard Brams Niebuhr, Rasmus Kanne Mikkelsen, Rasmus Odgaard, and Ivan Nikolov. 2025. LLM-Integrated VR NPCs to Support Social Presence and Engagement in a Viking Museum. *IEEE Transactions on Games* 17, 4 (2025), 1125–1134. doi:10.1109/TG.2025.3619578
- [68] Mohd Kamal Othman, Helen Petrie, and Christopher Power. 2011. Engaging Visitors in Museums with Technology: Scales for the Measurement of Visitor and Multimedia Guide Experience. In *Human-Computer Interaction – INTERACT 2011*, Pedro Campos, Nicholas Graham, Joaquim Jorge, Nuno Nunes, Philippe Palanque, and Marco Winckler (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 92–99.
- [69] Jan Packer and Roy Ballantyne. 2016. Conceptualizing the Visitor Experience: A Review of Literature and Development of a Multifaceted Model. *Visitor Studies* 19, 2 (2016), 128–143. doi:10.1080/10645578.2016.1144023
- [70] Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. 2023. Generative Agents: Interactive Simulacra of Human Behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (San Francisco, CA, USA) (UIST '23'). ACM, New York, NY, USA, Article 2, 22 pages. doi:10.1145/3586183.3606763
- [71] Syemin Park, Soobin Park, and Youn-kyung Lim. 2026. Constella: Supporting Storywriters' Interconnected Character Creation through LLM-based Multi-Agents. *ACM Transactions on Computer-Human Interaction* (Feb. 2026). doi:10.1145/3796234 Just Accepted.
- [72] Neil Pollock and Luciana D'Adderio. 2012. Give Me a Two-by-two Matrix and I Will Create the Market: Rankings, Graphic Visualisations and Sociomateriality. *Accounting, Organizations and Society* 37, 8 (2012), 565–586.
- [73] Yiting Ran, Xintao Wang, Tian Qiu, Jiaqing Liang, Yanghua Xiao, and Deqing Yang. 2025. BOOKWORLD: From Novels to Interactive Agent Societies for Story Creation. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Wanxiang Che, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar (Eds.). Association for Computational Linguistics, Vienna, Austria, 15898–15912. doi:10.18653/v1/2025.acl-long.773
- [74] Stéphanie Rey, Céline Picard, Yanis Fatmi, Fanny Franco, Sarah Guilbert, Jérémie Manéré, Christophe Bortoloso, Mustapha Derras, Nadine Couture, and Anke M. Brock. 2020. Build Your Own Hercules: Helping Visitors Personalize their Museum Experience. In *Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction* (Sydney NSW, Australia) (TEI '20). ACM, New York, NY, USA, 495–502. doi:10.1145/3374920.3374978
- [75] C. Rocchi, O. Stock, M. Zancanaro, M. Kruppa, and A. Krüger. 2004. The Museum Visit: Generating Seamless Personalized Presentations on Multiple Devices. In *Proceedings of the 9th International Conference on Intelligent User Interfaces* (Funchal, Madeira, Portugal) (IUI '04). ACM, New York, NY, USA, 316–318. doi:10.1145/964442.964517
- [76] Carol L. Rogers. 2012. The Importance of Understanding Audiences. In *Communicating Uncertainty*. Routledge, 179–200. doi:10.4324/9781410601360
- [77] Maria Roussou, Akviri Katifori, Laia Pujol, Maria Vayanou, and Stefan J. Rennick-Egglestone. 2013. A Life of Their Own: Museum Visitor Personas Penetrating the Design Lifecycle of a Mobile Experience. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (Paris, France) (CHI EA '13). ACM, New York, NY, USA, 547–552. doi:10.1145/2468356.2468453
- [78] Irene Rubino, Jetmir Xhembulla, Andrea Martina, Andrea Bottino, and Giovanni Malnati. 2013. Musa: Using indoor positioning and navigation to enhance cultural experiences in a museum. *Sensors* 13, 12 (2013), 17445–17471. doi:10.3390/s131217445
- [79] Nino Scherrer, Claudia Shi, Amir Feder, and David Blei. 2023. Evaluating the Moral Beliefs Encoded in LLMs. In *Advances in Neural Information Processing Systems*, A. Oh, T. Naumann, A. Globerson,

- K. Saenko, M. Hardt, and S. Levine (Eds.), Vol. 36. Curran Associates, Inc., 51778–51809. https://proceedings.neurips.cc/paper_files/paper/2023/file/a2cf225ba392627529efef14dc857e22-Paper-Conference.pdf
- [80] Guy Schofield, Gareth Beale, Nicole Beale, Martin Fell, Dawn Hadley, Jonathan Hook, Damian Murphy, Julian Richards, and Lewis Thresh. 2018. Viking VR: Designing a Virtual Reality Experience for a Museum. In *Proceedings of the 2018 Designing Interactive Systems Conference* (Hong Kong, China) (DIS '18). ACM, New York, NY, USA, 805–815. doi:10.1145/3196709.3196714
- [81] Belinda Crawford Seagram, Leslie H Patten, and Christine W Lockett. 1993. Audience Research and Exhibit Development: A Framework. *Museum Management and Curatorship* 12, 1 (1993), 29–41. doi:10.1016/0260-4779(93)90004-9
- [82] Beverly Serrell. 2010. Paying Attention: The Duration and Allocation of Visitors' Time in Museum Exhibitions. *Curator: The Museum Journal* 40, 2 (2010), 108–125. doi:10.1111/j.2151-6952.1997.tb01292.x
- [83] Omar Shaikh, Shardul Sapkota, Shan Rizvi, Eric Horvitz, Joon Sung Park, Diyi Yang, and Michael S. Bernstein. 2025. Creating General User Models from Computer Use. In *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology* (Busan, Republic of Korea) (UIST '25). ACM, New York, NY, USA, Article 35, 23 pages. doi:10.1145/3746059.3747722
- [84] Murray Shanahan, Kyle McDonell, and Laria Reynolds. 2023. Role Play with Large Language Models. *Nature* 623, 7987 (2023), 493–498. doi:10.1038/s41586-023-06647-8
- [85] Maria Shehade and Theopisti Stylianou-Lambert. 2020. Virtual Reality in Museums: Exploring the Experiences of Museum Professionals. *Applied Sciences* 10, 11 (2020). doi:10.3390/app10114031
- [86] Donghoon Shin, Daniel Lee, Gary Hsieh, and Gromit Yeuk-Yin Chan. 2025. PosterMate: Audience-driven Collaborative Persona Agents for Poster Design. In *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology* (Busan, Republic of Korea) (UIST '25). ACM, New York, NY, USA, Article 201, 20 pages. doi:10.1145/3746059.3747769
- [87] Ben Shneiderman. 2003. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *The Craft of Information Visualization*, BENJAMIN B. BEDERSON and BEN SHNEIDERMAN (Eds.). Morgan Kaufmann, San Francisco, 364–371. doi:10.1016/B978-155860915-0/50046-9
- [88] Lois H Silverman. 1995. Visitor Meaning-making in Museums for a New Age. *Curator: The Museum Journal* 38, 3 (1995), 161–170. doi:10.1111/j.2151-6952.1995.tb01052.x
- [89] Mingyang Su, Chao Liu, Jingling Zhang, WU Shuang, and Mingming Fan. 2025. SimViews: An Interactive Multi-Agent System Simulating Visitor-to-Visitor Conversational Patterns to Present Diverse Perspectives of Artifacts in Virtual Museums. In *Proceedings of the 33rd ACM International Conference on Multimedia* (Dublin, Ireland) (MM '25). ACM, New York, NY, USA, 6740–6750. doi:10.1145/3746027.3754593
- [90] Shayan Talaei, Meijin Li, Kanu Grover, James Kent Hippler, Diyi Yang, and Amin Saberi. 2025. StorySage: Conversational Autobiography Writing Powered by a Multi-Agent Framework. In *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology* (Busan, Republic of Korea) (UIST '25). ACM, New York, NY, USA, Article 209, 26 pages. doi:10.1145/3746059.3747681
- [91] Yihong Tang, Kehai Chen, Muyun Yang, Zheng-Yu Niu, Jing Li, Tiejun Zhao, and Min zhang. 2025. Thinking in Character: Advancing Role-Playing Agents with Role-Aware Reasoning. In *Advances in Neural Information Processing Systems*, D. Belgrave, C. Zhang, H. Lin, R. Pascanu, P. Koniusz, M. Ghassemi, and N. Chen (Eds.), Vol. 38. Curran Associates, Inc., 117765–117799.
- [92] Qwen Team. 2026. Qwen3.5: Accelerating Productivity with Native Multimodal Agents. <https://qwen.ai/blog?id=qwen3.5>
- [93] Yu-Min Tseng, Yu-Chao Huang, Teng-Yun Hsiao, Wei-Lin Chen, Chao-Wei Huang, Yu Meng, and Yun-Nung Chen. 2024. Two Tales of Persona in LLMs: A Survey of Role-Playing and Personalization. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, Yaser Al-Onaizan, Mohit Bansal, and Yun-Nung Chen (Eds.). Association for Computational Linguistics, Miami, Florida, USA, 16612–16631. doi:10.18653/v1/2024.findings-emnlp.969
- [94] Willem Robert van Hage, Natalia Stash, Yiwen Wang, and Lora Aroyo. 2010. Finding Your Way through the Rijksmuseum with an Adaptive Mobile Museum Guide. In *The Semantic Web: Research and Applications*, Lora Aroyo, Grigoris Antoniou, Eero Hyvönen, Annette ten Teije, Heiner Stuckenschmidt, Liliana Cabral, and Tania Tudorache (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 46–59.
- [95] Peter Vergo. 1997. *The New Museology*. Reaktion Books.
- [96] Guanzhi Wang, Yuqi Xie, Yunfan Jiang, Ajay Mandekar, Chaowei Xiao, Yuke Zhu, Linxi Fan, and Anima Anandkumar. 2024. Voyager: An Open-Ended Embodied Agent with Large Language Models. *Transactions on Machine Learning Research* (2024). <https://openreview.net/forum?id=ehfRiF0R3a>
- [97] Xintao Wang, Heng Wang, Yifei Zhang, Xinfeng Yuan, Rui Xu, Jen-Tse Huang, Siyu Yuan, Haoran Guo, Jiangjie Chen, Shuchang Zhou, Wei Wang, and Yanghua Xiao. 2025. CoSER: Coordinating LLM-Based Persona Simulation of Established Roles. In *Proceedings of the 42nd International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 267)*, Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (Eds.). PMLR, 64822–64858. <https://proceedings.mlr.press/v267/wang25dk.html>
- [98] Zhan Wang, Lin-Ping Yuan, Liangwei Wang, Bingchuan Jiang, and Wei Zeng. 2024. VirtuWander: Enhancing Multi-modal Interaction for Virtual Tour Guidance through Large Language Models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). ACM, New York, NY, USA, Article 612, 20 pages. doi:10.1145/3613904.3642235
- [99] Lilian Weng. 2023. LLM-powered Autonomous Agents. [lilianweng.github.io](https://lilianweng.github.io/posts/2023-06-23-agent/) (Jun 2023). <https://lilianweng.github.io/posts/2023-06-23-agent/>
- [100] Luoxuan Weng, Yinghao Tang, Yingchaojie Feng, Zhuo Chang, Ruiqin Chen, Haozhe Feng, Chen Hou, Danqing Huang, Yang Li, Huaming Rao, Haonan Wang, Canshi Wei, Xiaofeng Yang, Yuhui Zhang, Yifeng Zheng, Xiuqi Huang, Minfeng Zhu, Yuxin Ma, Bin Cui, Peng Chen, and Wei Chen. 2025. DataLab: A Unified Platform for LLM-Powered Business Intelligence. In *2025 IEEE 41st International Conference on Data Engineering (ICDE)*. IEEE, 4346–4359. doi:10.1109/ICDE65448.2025.00326
- [101] Jonathan Wieland, Daniel I. Fink, Anke V. Reinschuessel, Jonathan Häfler, Tiare Feuchter, and Harald Reiterer. 2026. Using Digital Twins to Design and Evaluate Interactive Exhibitions: A Case Study with Handheld AR. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26)*. ACM, New York, NY, doi:10.1145/3772318.3790852
- [102] Jean D. Wineman and John Peponis. 2010. Constructing Spatial Meaning: Spatial Affordances in Museum Design. *Environment and Behavior* 42, 1 (2010), 86–109. doi:10.1177/0013916509335534
- [103] Huatao Xu, Liying Han, Qirui Yang, Mo Li, and Mani Srivastava. 2024. Penetrative AI: Making LLMs Comprehend the Physical World. In *Proceedings of the 25th International Workshop on Mobile Computing Systems and Applications* (San Diego, CA, USA) (HotMobile '24). ACM, New York, NY, USA, 1–7. doi:10.1145/3638550.3641130
- [104] Ningning Xu, Yue Li, Xingbo Wei, Letian Xie, Lingyun Yu, and Hai-Ning Liang. 2024. CubeMuseum AR: A Tangible Augmented Reality Interface for Cultural Heritage Learning and Museum Gifting. *International Journal of Human-Computer Interaction* 40, 6 (2024), 1409–1437. doi:10.1080/10447318.2023.2171350
- [105] Yuji Yoshimura, Stanislav Sobolevsky, Carlo Ratti, Fabien Girardin, Juan Pablo Carrascal, Josep Blat, and Roberta Sinatra. 2014. An Analysis of Visitors' Behavior in the Louvre Museum: A Study Using Bluetooth Data. *Environment and Planning B: Planning and Design* 41, 6 (2014), 1113–1131. doi:10.1068/b130047p
- [106] Shuning Zhang, Hui Wang, and Xin Yi. 2025. Exploring Collaboration Patterns and Strategies in Human-AI Co-creation through the Lens of Agency: A Scoping Review of the Top-tier HCI Literature. *Proceedings of the ACM on Human-Computer Interaction* 9, 7, Article CSCW413 (Oct. 2025), 43 pages. doi:10.1145/3757594
- [107] Jiayi Zhou, Renzhong Li, Junxiu Tang, Tan Tang, Haotian Li, Weiwei Cui, and Yingcai Wu. 2024. Understanding Nonlinear Collaboration between Human and AI Agents: A Co-design Framework for Creative Design. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). ACM, New York, NY, USA, Article 170, 16 pages. doi:10.1145/3613904.3642812
- [108] Xuhui Zhou, Hao Zhu, Leena Mathur, Ruohong Zhang, Haofei Yu, Zhengyang Qi, Louis-Philippe Morency, Yonatan Bisk, Daniel Fried, Graham Neubig, and Maarten Sap. 2024. SOTOPIA: Interactive Evaluation for Social Intelligence in Language Agents. In *The Twelfth International Conference on Learning Representations*. <https://openreview.net/forum?id=mM7VurbA4r>
- [109] Yueheng Zhu, Chao Liu, Xuan He, Xiaoxue Ren, Zhongxin Liu, Ruwei Pan, and Hongyu Zhang. 2026. AdaCoder: An Adaptive Planning and Multi-Agent Framework for Function-Level Code Generation. *IEEE Transactions on Software Engineering* 52, 2 (2026), 631–650. doi:10.1109/TSE.2025.3642621

A Formative Study

A.1 Brief Task

Participants were asked to complete a brief exhibition design task situated in a realistic museum setting. The scenario was set in the first-floor red zone of the Denon Wing at the Louvre³. The exhibition theme was *Renaissance painting*. To support the task, we provided each participant with the floor plan and 50 related paintings from the Louvre collection.

Participants were asked to arrange the artwork cards on the floor plan as an initial exhibition concept. The task focused on curatorial thinking rather than technical feasibility, encouraging them to use

³<https://collections.louvre.fr/>

grouping, sequence, and spatial placement to organize the selected works into one or more coherent narratives, ideas, or concepts related to the Renaissance.

Table 1: Details of participants interviewed in the formative study, including their professional backgrounds and museum or exhibition practice experience.

ID	Background	Domain Experience
P1	Cultural Exhibition Practitioner	6 years
P2	Industrial Design Student	4 years
P3	New Media Exhibition Practitioner	7 years
P4	Art & Design Student	2 years
P5	Natural History Museum Researcher	9 years
P6	History Museum Researcher	24 years
P7	Municipal Museum Researcher	11 years

B User Study

Table 2: Details of participants in the user study, including their professional backgrounds and museum or exhibition practice experience.

ID	Background	Domain Experience
U1	Exhibition Practitioner	7 years
U2	Art & Design Student	2 years
U3	Art & Design Student and Exhibition Enthusiast	3 years
U4	Art & Design Student	1 year
U5	Industrial Design Student	1 year
U6	Industrial Design Student	2 years
U7	Industrial Design Student	1 year
U8	Exhibition Designer	8 years
U9	Art & Design Student	1 year
U10	Art & Design Student	2 years
U11	Art & Design Student	2 years
U12	Exhibition Designer	9 years

C SiMUSation

Table 3: Key objective metrics for evaluating visitor experience and exhibition performance.

Measure	Definition & Formula
Visitor Count (VC)	Total number of visitors. ($VC = \text{Count of Visitors}$)
Average Holding Time (AHT)	Average time (s) visitors spend engaging with an exhibit ($AHT = \frac{\sum \text{Time Spent}}{VC}$).
Attracting Power (AP)	Proportion of visitors who approach a specific exhibit ($AP = \frac{\text{Stopped Visitors}}{VC} \times 100\%$).
Sweep Rate Index (SRI)	Rate of movement through the exhibition (Area/Time). Lower is better for engagement ($SRI = \frac{\text{Area (m}^2\text{)}}{\text{Avg. Visit Time}}$).
Percentage of Diligent Visitors (%DV)	Percentage of visitors stopping at > 50% of exhibits ($DV = \frac{\text{Count of Visitors (} \geq 50\% \text{ Stops)}}{\text{Count of Visitors}} \times 100\%$).
Walkthrough Rate (WR)	Percentage of visitors not stopping at any exhibit ($WR = \frac{\text{Count of Visitors (0 Stops)}}{VC} \times 100\%$).
Stops at Exhibits	Number of exhibits visited per visitor ($AvgStops = \frac{\sum \text{Total Stops}}{VC}$).
Dwell Time (DT)	Total time spent in the exhibition or a specific section ($DT = \text{Exit Time} - \text{Entry Time}$).