Designing Embodied Data Storytelling for Cultural Heritage in Virtual Reality

Yuxuan Guo*
Shanghai Jian Qiao University

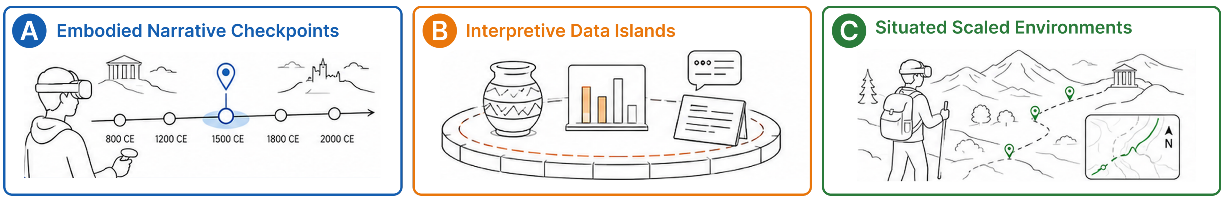

Figure 1: Overview of the proposed Temporal–Event–Geographic mapping method: (A) embodied narrative checkpoints, (B) interpretive data islands, and (C) situated scaled environments.

Abstract

Virtual reality is widely used for cultural heritage, but many systems prioritize spatial reconstruction over the interpretation of heterogeneous historical evidence. We present a generative design method that classifies heritage data by narrative role—temporal, event-based, and geographic—and maps these roles to embodied narrative checkpoints, interpretive data islands, and situated scaled environments. We instantiate the method in a Silk Road VR prototype and conduct a preliminary study with six participants. Participants reported that user-triggered progression clarified historical sequences, and object-based analogies made event-related information more memorable. They also relied on interpretive and orientation cues to connect embodied actions with historical meaning. These findings suggest that embodied data storytelling is most effective when interaction is paired with concise interpretive support.

Index Terms: Virtual reality, cultural heritage, data storytelling.

1 Introduction

Virtual reality (VR) has become an important medium for cultural heritage, enabling visitors to enter reconstructed sites, inspect artifacts, and access places that are distant, lost, fragile, or otherwise inaccessible. Many VR heritage systems focus on spatial reconstruction, artifact presentation, or virtual museum exploration. These approaches can provide strong presence, visual engagement, and access to reconstructed heritage environments. However, previous work also suggests that immersive cultural heritage systems still face challenges in interpretation, authenticity, and the communication of historical context [2]. This is important because cultural heritage is often supported by heterogeneous forms of evidence, including dates, routes, events, biographies, object circulation, environmental conditions, quantitative records, and uncertain interpretations [12]. Therefore, designing VR heritage experiences requires more than reconstructing places or displaying artifacts. It also requires methods to transform heritage evidence into interpretable narrative experiences.

This evidential nature of cultural heritage connects VR heritage design with data-driven storytelling. Narrative visualization has developed mature techniques for communicating evidence on screens, such as timelines, maps, charts, annotations, and guided story structures [11, 1]. Immersive analytics suggests that VR can support spatial engagement, bodily interaction, and presence in data interpretation [3]. However, turning all heritage data into 3D objects may reduce data legibility and make exact relationships harder to understand. This tension is especially important for non-expert audiences, who may benefit from immersive experience but still need clear explanations, comparisons, and narrative guidance.

To address this challenge, we propose a generative method that classifies cultural heritage evidence by its narrative role rather than its media format. The method distinguishes temporal, event-based, and geographic evidence and maps them to corresponding embodied narrative forms, as summarized in Fig. 1. We instantiate the method in a standalone Silk Road VR prototype and conduct a preliminary study with six participants to examine interpretation, interaction, navigation, and embodied understanding.

This work makes three contributions. First, it offers an evidence-based data storytelling perspective for designing VR experiences of cultural heritage beyond reconstruction-centered presentation. Second, it proposes a design method as a generative framework to translate heterogeneous heritage evidence into narrative VR interactions. Third, it provides a prototype-based case study and preliminary user feedback that illustrate how embodied interaction can support cultural heritage interpretation when paired with clear narrative structure.

2 Related Work

2.1 Cultural Heritage in Virtual Reality

VR has been widely used in cultural heritage to support virtual museums, historical reconstruction, artifact presentation, and access to distant or inaccessible sites. However, cultural heritage VR still faces an important interpretive challenge. Rahaman and Tan argued that many digital heritage projects are overly focused on visual products or technical processes, while paying insufficient attention to interpretation and end-user understanding [10]. Windhager et al. further showed that cultural heritage collections form complex information spaces involving heterogeneous data, such as objects, temporal structures, spatial relations, and cultural contexts [12].

Recent research has increasingly moved beyond reconstruction toward embodied and contextual storytelling. Yu et al. employed narrative analysis and contextual storytelling to communicate the historical context surrounding intangible cultural heritage in VR [14]. From a broader design perspective, Polydorou synthesized three immersive productions into a methodology that emphasizes active participation, multisensory engagement, frag-

*e-mail: yuxuanguo@msn.com

mented narratives, and the use of technology to expand narrative space [9]. These studies demonstrate a shift from presenting heritage objects toward involving audiences in situated processes of interpretation and meaning-making. However, they primarily provide case-specific designs or principles for structuring an immersive experience as a whole. They offer limited guidance on how heterogeneous evidence within the same heritage narrative should be differentiated and translated into distinct representational and interaction forms. Our work addresses this gap by classifying heritage evidence according to its narrative role rather than its media type.

2.2 Embodied Data Storytelling

Recent work in immersive analytics and immersive data storytelling [8] extends data stories into spatial environments. Immersive visual data stories use presence and embodied engagement while retaining the communicative goals of data storytelling [3], while data physicalization and visceralization translate abstract quantities into object-like forms, bodily scale, and spatial experience [4, 5]. Recent systems further demonstrate these possibilities. TimeTables supports embodied exploration of spatio-temporal data [15], Spaces to Think shows how immersive workspaces enable spatial strategies during sensemaking [7], and 3DStoryline organizes temporal, spatial, and narrative relationships within VR [13].

These studies indicate that immersive space and embodied interaction can make complex information more concrete and support spatial organization. However, existing work mainly focuses on analytical datasets, individual visualization structures, or general learning environments. It provides limited guidance on how heterogeneous cultural heritage evidence should be mapped to differentiated but coherent narrative forms. Our method addresses this gap by classifying evidence according to its narrative role and integrating embodied representation with interpretive and orientation support.

3 Design Method

We classify heritage evidence according to its narrative function. Evidence may primarily organize historical sequence, contextualize a bounded event or theme, or communicate spatial relationships. These roles imply different representational and interaction requirements in VR, leading to the three mappings shown in Fig. 1.

3.1 Embodied Narrative Checkpoints

Temporal data include biographies, historical stages, dynastic changes, war processes, journeys, and other materials that depend on sequence, progression, and causality. In screen-based data storytelling, such data are often represented through timelines, animated maps, or scrollytelling structures. In VR, however, passive playback may reduce users' sense of agency, while unrestricted exploration may interrupt the intended historical sequence. Prior work on narrative visualization suggests that effective data stories often balance author-driven sequencing with reader-driven interaction [11, 6].

To balance authorial control and user agency, we map temporal data into embodied narrative checkpoints. A narrative checkpoint refers to a key moment in the temporal sequence that users actively trigger through a simple and low-cost interaction. Its purpose is not to introduce complex gameplay, but to allow users to participate in the pacing of the story. Each checkpoint can advance the temporal sequence, activate narration, change visual states, or reveal new evidence. In this way, the design preserves the authorial structure of historical progression while giving users a sense of actively moving the story forward. The main design principle is that temporal data should not merely be displayed as a static timeline; it should become an embodied rhythm of user-triggered progression.

3.2 Interpretive Data Islands

Event-based data concerns historical events or cultural themes that involve multiple forms of evidence, including people, objects, locations, quantities, causes, and consequences. Their primary narrative function is to communicate contextual relationships within a bounded topic. To organize such materials in VR, we propose the concept of interpretive data islands.

An interpretive data island is a spatially bounded unit that assembles related evidence around a specific event or theme. It can include models, object analogies, physicalized data representations, visualizations, captions, and narration. This approach builds on data physicalization, which suggests that data can become more accessible when represented through object-like or spatial forms [4]. By allowing users to approach, inspect, compare, and listen within a focused space, the data island structure helps connect visual richness with interpretive purpose.

The main risk is that object-rich scenes may become visually attractive but interpretively weak. An effective data island must therefore make the relationships among objects, quantities, and historical meaning explicit.

3.3 Situated Scaled Environments

Geographic data include routes, cities, mountains, coordinates, distance, elevation, terrain barriers, and spatial relationships. Such data are often represented through 2D maps, but cultural heritage stories may require users to understand how geography shapes movement, exchange, risk, and encounter. Immersive analytics suggests that spatial environments can support data understanding by allowing users to engage with data through bodily movement and spatial perception [3]. We therefore map geographic data into situated scaled environments.

A situated scaled environment is an abstracted and scaled representation of a heritage geography that users can enter and explore. Within this environment, key places can be marked by culturally recognizable landmarks, while routes and terrain provide spatial structure. To prevent disorientation, a real-time mini-map can serve as an orientation anchor, showing the user's current position, visited areas, and route direction. For data such as elevation, embodied actions such as climbing, moving upward, or comparing vertical forms can make height differences experientially meaningful, aligning with prior work on VR's potential to visceralize scale and physical quantity [5].

The main risk is that users may experience the environment as a landscape without understanding its geographic structure. Situated scaled environments should therefore connect first-person exploration with an overview of routes, locations, and relative scale.

4 User Study

To examine how the proposed method works in practice, we implemented a VR prototype and conducted a preliminary user study. The goal of the study was to gather early feedback on whether the method can support interpretable, embodied, and narratively coherent cultural heritage data experiences in VR. The study also aimed to identify design issues related to interaction, onboarding, navigation, motion comfort, and the relationship between embodied experience and data interpretation.

4.1 Prototype

We developed a VR prototype (Fig. 2) for Quest headsets using the Silk Road as a case study. The case was selected because it combines temporal development, event-centered evidence, trade objects, routes, landmarks, terrain, and quantitative geographic information. The prototype was implemented in Unity and supports teleportation, object manipulation, controller-based interaction, and climbing. It consists of an introductory scene and three content

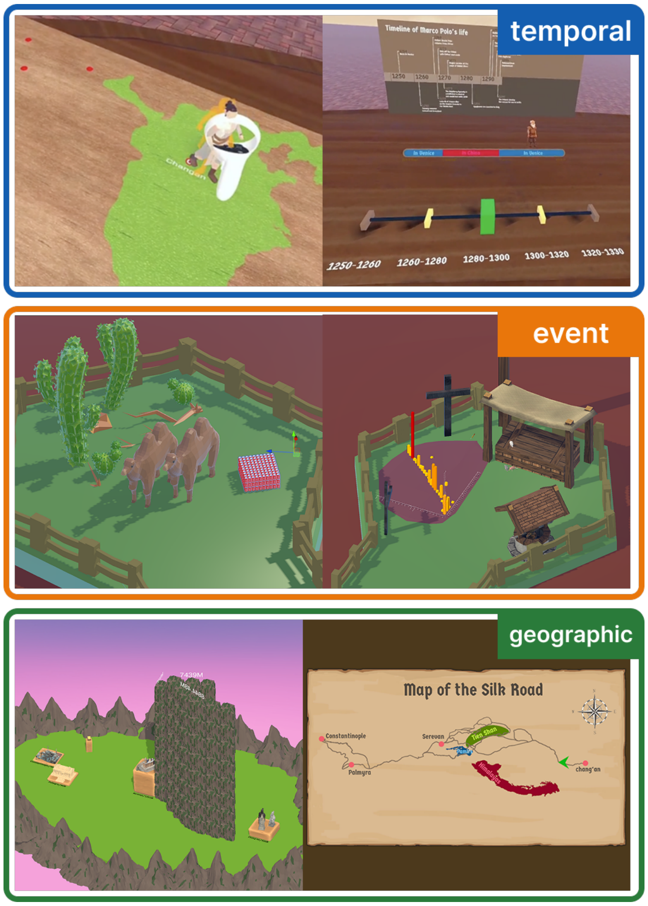

Figure 2: The screenshots of the Silk Road VR prototype implementing the three components of the proposed method: temporal progression, event-based interpretation, and geographic exploration.

components corresponding to the proposed design methods. The introductory scene helps users become familiar with basic controller operations and provides access to the three main components.

The **temporal component** presents historical development and related narrative paths through interactive maps, staged narration, and user-triggered checkpoints. While preserving the chronological order of historical events, this component allows users to actively advance the narrative, with scene content changing in response to different times. The **event component** organizes trade-, transportation-, and epidemic-related content into interpretive data units. These units combine object models, physicalized data analogies, visualized quantities, and narration to explain specific historical themes. Users can interact with objects within each unit, while voice-over narration and text panels provide contextual explanation. The **geographic component** transforms routes, cities, mountains, and elevation data into a scaled environment. Users can identify the relative locations and elevation differences of key cities through landmark models, use a mini-map for orientation, and experience the elevation differences among three major mountain ranges along the Silk Road through climbing.

4.2 Participants and Procedure

We recruited six participants (four females and two males), aged between 22 and 33 years. Participants had diverse educational backgrounds, and four had prior experience using VR headsets. None of the participants had professional expertise in Silk Road history.

Each session followed four steps. First, participants received a brief introduction to the study purpose and basic VR controls. Second, they entered the introductory scene, where they could become familiar with the interaction mechanism and select different content components. Third, participants experienced the temporal, event-based, and geographic components of the prototype. They were encouraged to explore the scenes naturally while paying attention to the historical narratives, data representations, and interaction mechanisms. Finally, after the VR experience, participants completed a semi-structured interview based on their experience and reflections.

The interview collected both background information and reflective feedback. It asked participants to evaluate their understanding of the Silk Road's history, geographic information, and trade dynamics after the experience. Additional questions focused on perceived interactivity, ease of navigation, confusing moments, changes in understanding, the most engaging parts of the experience, and suggestions for improvement. This combination of questions was intended to assess not only general user experience, but also whether the prototype helped users connect interaction with heritage data interpretation.

4.3 Results

Overall, participants responded positively to the prototype. They reported that the combination of scenes, narration, and interaction helped them develop a more concrete understanding of the Silk Road's historical events, trade activities, and geographic information. These comments suggest that the proposed method can support cultural heritage data storytelling when each interaction is clearly connected to a narrative or data function. However, the results also show that immersion and interaction alone are insufficient; users need explicit guidance to understand what each action, object, or spatial representation means.

For the temporal component, participants found that interactive maps and user-triggered progression helped organize the historical sequence. This dimension provided a clear narrative structure, enabling users to actively progress through different historical stages rather than relying solely on passive playback. Meanwhile, the checkpoint-based presentation also helped users understand the temporal span between stages more intuitively. As one participant explained, "triggering the next stage by myself made me pay more attention to the order of events." Another participant commented that the changing maps and narration, "It helped connect the different stages of the person's life, and the timeline format allowed me to see the duration of each key stage clearly." This supports the design goal of balancing author-guided sequencing with user participation.

For the event component, participants found the thematic spaces and object analogies memorable. The use of models and physicalized data objects helped connect abstract information with recognizable forms. This was especially useful for quantities or historical relationships that might otherwise remain distant from everyday experience. For example, one participant commented on the object-based comparison: "The corresponding number of cola bottle models in the scene helped me clearly understand just how much water a camel can drink at once." At the same time, participants also needed stronger explanatory anchors to understand why particular objects, quantities, or scenes were grouped together. One participant explained, "Some objects were interesting, but I still needed more text or voice explanation to understand why they were placed together". This suggests that interpretive data islands should not only contain engaging objects, but should also include captions, legends, voice-over, or prompts that explicitly explain the relationship between the objects and the historical evidence.

For the geographic component, climbing and moving through the scaled terrain were among the most engaging parts of the experience. Participants responded positively to the mountain-climbing interaction because it transformed elevation data into a bodily comparison. One participant stated, "Climbing mountains allowed me to perceive the height difference much more intuitively than simply seeing a number, and it left a deep impression on me". This supports the idea that VR is particularly suitable for representing data related to scale, height, distance, and movement. However, participants also needed orientation support to understand their position within the larger route. The mini-map helped users reconnect first-person exploration with an overview of the Silk Road geography. As one participant noted, "The mini-map helped me understand my orientation and location, while the landmark models gave me a stronger perception of the elevation differences among multiple cities." This is consistent with prior work on worlds-in-miniature techniques, which use scaled overviews to support navigation and spatial awareness in virtual environments. The result suggests that geographic data in VR should combine embodied exploration with orientation aids, rather than relying on immersion alone.

The study also revealed several cross-cutting design issues. First, interaction should be treated as an interpretive mechanism. Participants often remembered actions such as climbing or manipulating objects, but these actions supported understanding only when their data meaning was clear. This finding is consistent with prior work suggesting that interaction in immersive storytelling should be connected to learning and narrative goals rather than treated as an isolated feature. Second, embodied data experiences require explanatory anchors. Mini-maps, captions, scales, narration, and highlighted values help users translate bodily experience back into data interpretation.

Finally, participants suggested that future versions could include more dynamic environmental responses synchronized with narration. Although the prototype included interactive objects and data representations, some environmental elements remained relatively static. One participant suggested that "the environment could change more with the voice-over, so the story feels more alive." Another participant wanted "clearer hints about what to do next and what changed after an interaction." These comments point to the need for more responsive animation, clearer interaction cues, and stronger coordination between narration and environmental change.

5 Discussion and Conclusion

This work presents a method for embodied data storytelling in cultural heritage VR. The method does not prescribe a fixed set of scenes or interactions. Instead, it helps designers derive interaction strategies from the narrative role of heritage evidence. Temporal data can be transformed into embodied narrative checkpoints, event-based data into interpretive data islands, and geographic data into situated scaled environments. In this way, the method supports a shift from asking "what content should be shown in VR?" to "what role does this evidence play in the heritage narrative, and how should users experience it?"

The Silk Road prototype suggests that this approach can connect narrative structure, spatial experience, and data interpretation. User feedback indicates that interactive maps helped organize historical progression, object-based analogies made event-related information more memorable, and climbing-based interaction made elevation differences easier to perceive. However, the study also shows that embodiment alone is not sufficient. Users still need explanatory anchors, such as narration, captions, labels, mini-maps, and highlighted values, to understand what an interaction or spatial representation means.

This work has several limitations. First, the prototype is based on a single case study. Although the Silk Road is well suited to the method because it includes several kinds of evidence, other cultural heritage topics may require different adaptations. Second, the user study was preliminary and involved only six participants.

Future research should conduct larger and more systematic evaluations. Useful measures may include comprehension tasks, delayed recall, spatial orientation, engagement, cognitive load, presence, and usability. A controlled study could compare the proposed VR data story with a screen-based data story, a conventional virtual museum, or a reconstruction-centered VR experience. Such comparisons would help clarify whether the benefits come from immersion itself, from explicit narrative-data mapping, or from the combination of both.

References

[1] B. Bach, M. Stefaner, J. Boy, S. Drucker, L. Bartram, J. Wood, P. Ciuccarelli, Y. Engelhardt, U. Koppen, and B. Tversky. *Narrative Design Patterns for Data-Driven Storytelling*, chap. 5, pp. 107–133. Taylor Francis, March 2018. 1

[2] M. K. Bekele, R. Pierdicca, E. Frontoni, E. S. Malinverni, and J. Gain. A survey of augmented, virtual, and mixed reality for cultural heritage. *J. Comput. Cult. Herit.*, 11(2), Mar. 2018. doi: 10.1145/3145534 1

[3] P. Isenberg, B. Lee, H. Qu, and M. Cordeil. *Immersive Visual Data Stories*, pp. 165–184. Springer International Publishing, Cham, 2018. doi: 10.1007/978-3-030-01388-2_6 1, 2

[4] Y. Jansen, P. Dragicevic, P. Isenberg, J. Alexander, A. Karnik, J. Kildal, S. Subramanian, and K. Hornbæk. Opportunities and challenges for data physicalization. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, p. 3227–3236. Association for Computing Machinery, New York, NY, USA, 2015. doi: 10.1145/2702123.2702180 2

[5] B. Lee, D. Brown, B. Lee, C. Hurter, S. Drucker, and T. Dwyer. Data Visceralization: Enabling Deeper Understanding of Data Using Virtual Reality . *IEEE Transactions on Visualization & Computer Graphics*, 27(02):1095–1105, Feb. 2021. doi: 10.1109/TVCG.2020.3030435 2

[6] B. Lee, N. H. Riche, P. Isenberg, and S. Carpendale. More than telling a story: Transforming data into visually shared stories. *IEEE Computer Graphics and Applications*, 35(5):84–90, 2015. doi: 10.1109/MCG.2015.99 2

[7] L. Lisle, K. Davidson, L. Pavanatto, I. A. Tahmid, C. North, and D. A. Bowman. Spaces to think: A comparison of small, large, and immersive displays for the sensemaking process. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 1084–1093, 2023. doi: 10.1109/ISMAR59233.2023.00125 2

[8] J. Méndez, W. Luo, R. Rzayev, W. Büschel, and R. Dachselt. Immersive data-driven storytelling: Scoping an emerging field through the lenses of research, journalism, and games. *IEEE Transactions on Visualization and Computer Graphics*, 31(10):6839–6851, 2025. doi: 10.1109/TVCG.2025.3531138 2

[9] D. Polydorou. Immersive storytelling experiences: a design methodology. *Digital Creativity*, 35(4):301–320, 2024. doi: 10.1080/14626268.2024.2389886 2

[10] H. Rahaman and B.-K. Tan. Interpreting digital heritage: A conceptual model with end-users' perspective. *International Journal of Architectural Computing*, 9(1):99–113, 2011. doi: 10.1260/1478-0771.9.1.99 1

[11] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139–1148, 2010. doi: 10.1109/TVCG.2010.179 1, 2

[12] F. Windhager, P. Federico, G. Schreder, K. Glinka, M. Dörk, S. Miksch, and E. Mayr. Visualization of cultural heritage collection data: State of the art and future challenges. *IEEE Transactions on Visualization and Computer Graphics*, 25(6):2311–2330, 2019. doi: 10.1109/TVCG.2018.2830759 1

[13] H. Yao, L. Zhao, B. Chen, K. Li, H.-N. Liang, and L. Yu. 3DStoryline: Immersive visual storytelling. *Journal of Visualization*, 28(3):681–697, 2025. doi: 10.1007/s12650-025-01058-5 2

[14] J. Yu, Z. Wang, Y. Cao, H. Cui, and W. Zeng. Centennial drama reimagined: An immersive experience of intangible cultural heritage through contextual storytelling in virtual reality. *J. Comput. Cult. Herit.*, 18(1), Feb. 2025. doi: 10.1145/3705613 1

[15] Y. Zhang, B. Ens, K. A. Satriadi, A. Prouzeau, and S. Goodwin. Timetables: Embodied exploration of immersive spatio-temporal data. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 599–605, 2022. doi: 10.1109/VR51125.2022.00080 2