"Being Eroded, Piece by Piece": Enhancing Engagement and Storytelling in Cultural Heritage Dissemination by Exhibiting GenAI Co-Creation Artifacts

Kexue Fu
City University of Hong Kong
Hong Kong SAR, China
kexuefu2-c@my.cityu.edu.hk

Ruishan Wu
City University of Hong Kong
Hong Kong SAR, China
ruishanwu2-c@my.cityu.edu.hk

Yuying Tang
Tsinghua University
Beijing, China
tyy21@mails.tsinghua.edu.cn

Yixin Chen
University of Washington
Seattle, United States
yixin7@uw.edu

Bowen Liu
Central Academy of Fine Arts
Beijing, China
nfksq@hotmail.com

RAY LC*
City University of Hong Kong
Hong Kong SAR, China
LC@raylc.org

ABSTRACT

Cultural Heritage is not just about tangible artifacts; it also includes intangible elements such as personal memories, community ties, and envisioned futures. Traditional museums and archives often emphasize physical items like architectural pieces and photos, while overlooking people's personal and emotional connections to cultural heritage. To illustrate the personal connections people have with cultural heritage sites, we designed an exhibition that displayed images created by participants, which represent their perspectives and future visions of cultural heritage sites. The exhibition's images, generated through GenAI, helped participants narratively describe cultural heritage locations, allowing them to express their visions of future threats like over-tourism and climate change on these sites. Contrary to constraints, co-creating with Generative AI associates participants with personal memories of cultural heritage, stimulating personal narratives and promoting deep reflection on cultural heritage preservation. The dissemination strategies we designed illustrate the use of GenAI to empower the expression of matters of cultural value beyond the physical.

CCS CONCEPTS

• **Human-centered computing → Collaborative and social computing design and evaluation methods**.

KEYWORDS

Generative AI, Cultural Heritage, Dissemination Strategy, Exhibition Design

*Correspondences should be addressed to LC@raylc.org.

DIS '24, July 01–05, 2024, IT University of Copenhagen, Denmark

ACM ISBN 979-8-4007-0583-0/24/07
https://doi.org/10.1145/3643834.3660711

ACM Reference Format:
Kexue Fu, Ruishan Wu, Yuying Tang, Yixin Chen, Bowen Liu, and RAY LC. 2024. "Being Eroded, Piece by Piece": Enhancing Engagement and Storytelling in Cultural Heritage Dissemination by Exhibiting GenAI Co-Creation Artifacts. In *Designing Interactive Systems Conference (DIS '24), July 01–05, 2024, IT University of Copenhagen, Denmark.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3643834.3660711

1 INTRODUCTION

Traditionally, museums and archives have focused on preserving physical cultural heritage elements, such as architectural works, artworks, and photographs [1, 3]. With the introduction of interactive new media forms such as augmented reality (AR) [57] and eye and gesture interaction technologies [22], they help encourage navigation of cultural heritage sites and enhance the learning experience of cultural heritage. However, cultural heritage encompasses more than just the tangible artifacts that we can see and touch. cultural heritage also includes non-material aspects about how people engage with these sites, their personal perception of what is happening to these sites, and how they imagine these places to be in the future. This intangible information informs us about community ties and the stories community members tell each other [59].

As posited by the Authorized Heritage Discourse (AHD) theory, heritage transcends physicality, not just as an object, place, or event, but as a cultural and social narrative process [58], so it is essential to protect these intangible narratives as part of the preservation and dissemination process [22, 26, 41]. Given the increasing threats to historic buildings and monuments from various unavoidable factors such as climate change and war [27, 32, 65], preserving the intangible narratives and knowledge associated with these architectural heritages is particularly crucial. This preservation of stories, traditions, and practices also enables the transmission of these understandings and narratives to future generations as a form of enduring preservation [36, 37].

When visitors tour architectural landmarks or museums, their interaction with these attractions often remains superficial, with a primary focus on the physical aspects of the heritage. This approach overlooks the intricate, intangible connections between people and the potential future challenges these cultural heritage sites may face. From the exhibition side, while exhibitions devoted to cultural heritage aim to provide immersive experiences [36, 37], it can

be challenging for visitors to tangibly engage with the intangible knowledge and reflect upon the future of these cultural legacies.

However, participants in cultural heritage storytelling are generally not skilled in narrating personal stories and lack the appropriate tools for effective expression [57]. This highlights the importance of lowering entry barriers and encouraging sustained personal storytelling to enrich the diversity of perspectives [64]. Furthermore, understanding participants' perceptions and connections to cultural heritage is essential; without this, we cannot impart profound information to visitors.

Bridging the gap between fleeting visits and meaningful engagement with cultural heritage, technological advancements of Generative AI (GenAI) offer a unique solution. GenAI allows those with little technical expertise to create images by using natural language. GenAI's ability to recognize patterns in existing data and generate new and unexpected content empowers novice users to gain inspiration by directly prompting GenAI models [24, 33]. Given that the data fed into the AI represents a wide spectrum of viewpoints [53], it has the potential to shape individual narratives and offers unique visual perspectives during the generative process, thereby encouraging deeper reflection [16].

While GenAI can generate inspiring ideas through an interactive methodology, it also possesses the capability to vividly convey ideas through imagery [16]. In cultural heritage dissemination, image prowess to convey messages more powerfully than words aligns perfectly with the need to provoke visitor introspection and engagement [48, 67]. Consequently, we employed text-to-image GenAI tools, enabling those with ties and expertise in cultural heritage within their practice communities to craft visual narratives. This approach not only showcases the process and the resultant visuals to the visitors during the exhibition but also invites them to look beyond cultural heritage's physical elements, encouraging a deeper appreciation of its human significance. This leads us to formulate the following three research questions:

RQ1: *How to design a co-creation process and exhibition that displays participant-generated images from GenAI conveying personal narratives?*

RQ2: *How does using GenAI in narrating and envisioning cultural heritage impact participants' understanding and perception of the heritage site?*

RQ3: *What are the common behavioral patterns and challenges that participants face when co-creating with GenAI?*

To answer these research questions, we asked participants to use Midjourney, an accessible text-to-image generation tool, to compose prompts that reflected their understanding and connection to cultural heritage. The generative process is iteratively refined by each participant to promote deeper exploration and expression of their connections to cultural heritage. We then demonstrated an exhibition design based on the materials from the interview process, including the images generated and the interview insights, using a mobile AR application linked to a projected environment and a robotic drawing system.

Our contributions are twofold: (1) In the context of cultural heritage preservation and dissemination, we explored using GenAI as a tool for expression and personal narratives. We observed the patterns of behaviors, tapping and summarizing ideas, and inspiration of participants that emerged from the iterative generative process. (2) We have revealed the constraints in the process of GenAI functioning as a co-creation tool for cultural heritage storytelling and proposed corresponding design considerations. (3) Instead of showcasing physical fragments and photographs, we designed and prototyped an exhibition strategy to present individuals' intangible and personal perspectives on places they are connected to, using innovative exhibition design techniques.

2 RELATED WORK

2.1 Intangible Value of Cultural Heritage

Museums and archives traditionally prioritize preserving tangible aspects of cultural heritage, such as architectural models, architectural images, and physical objects like clothes and tools [1, 3]. With the advent of new and interactive media, fresh opportunities for cultural heritage exhibitions arise [15, 25]. This is exemplified in projects like the Pure Land project, which merges high-resolution digital archaeological data with immersive, interactive displays, offering unique engagement with sites like the Mogao Caves, a World Heritage site [37]. Similarly, Place Hampi provides an interactive 3D immersive experience of the entire site, allowing viewers to navigate virtually and explore 3D models of Hindu deities. AR has also been used to enhance this by potentially displaying historical structures, environments, or additional information alongside current real-world environments [36]. These immersive approaches aim to increase visitor engagement and participation in cultural heritage exhibitions while expanding the scope beyond traditional preservation methods.

Yet, the essence of cultural heritage goes beyond its physical attributes to encompass intangible significance. Despite the increasing prevalence, the exhibition of these intangible heritage aspects remains relatively scarce in innovative design practices [22]. A few exhibitions have started to weave personal reflections and emotions about heritage into their narratives. For example, a collaborative audio-visual exhibition featuring four enthusiastic participants from the tenth UNESCO Youth Forum creates a participatory space for intercultural dialogue. It presents new narratives of young people actively preserving and passing on their dynamic living traditions for present and future generations [4]. While providing a good way for participants to tell their stories, how visitors engage with the exhibition to gain access to this intangible information has not been well designed. Furthermore, previous Museum studies have also indicated that exhibition visitors appreciate the opportunity to engage with both the exhibition participants and other visitors [17, 18]. This also prompts us to consider how to use technology to foster dynamic relationships between participants and visitors when exhibiting the intangible value of cultural heritage. Refocusing on design, a prior study has begun exploring how exhibition design can effectively highlight the intangible values of cultural heritage [22]. By initially classifying tangible interaction case studies, this study identifies two design strategies that utilize tangible interaction to enhance visitors' engagement with the intangible values of cultural heritage. The first strategy involves embedding meaning, where sensor-equipped objects are created to convey the significance of intangible cultural heritage values through their physical presence. An exemplary illustration of this

strategy is found in artist Joreige's "Object de Guerre 1-4" (2000-2006), an installation that powerfully encapsulates the intangible heritage of the Lebanese War through participant-selected objects, accompanied by personal war narratives displayed within glass cabinets, each activated upon a visitor's approach [35]. This approach resonates with us as it allows visitors to intimately witness participants' narratives through participants' chosen objects, fostering a sense of "public intimacy". Embodying meaning is the second strategy proposed in this study [22]. It suggested utilizing gestures to convey significance and trigger digital interpretations of artwork. An example is the Drinking Symposium installation at the Allard Pierson Museum, where visitors can interact with virtual characters and replicas of Greek artifacts, fostering engagement with cultural heritage [60]. This concept encourages the exploration of interactive and participatory methods to appreciate intangible values [42]. Furthermore, while there are discussions on storytelling intangible cultural heritage in exhibitions, the existing practice often overlooks the narrative potential of tangible cultural heritage, such as buildings and sites [49]. These physical spaces carry our memories and experiences, representing an untapped domain for exploration and experimentation within exhibition contexts.

There's a growing interest in HCI communities to facilitate storytelling about architectural cultural heritage and how to enhance the relationship between local architectural cultural heritage sites and participants. For example, one project involved individuals in creating augmented stories centered around outdoor cultural heritage sites, showcasing AR's potential in bringing personal stories associated with these locations to life [57]. Another study introduced Cultural Probes, a storytelling method that employs physical artifacts like postcards and maps to elicit personal reflections and connect individuals to local heritage [26]. Additionally, a prior study introduced Lost State College (LSC), a mobile app exploring community perceptions and interactions with local cultural heritage. LSC offers photos and descriptions of local heritage alongside user-generated social data like visits, likes, comments, and photos. Participants utilized these social features to share information, enhancing accessibility and fostering diverse connections to local history [29]. These studies utilized text and imagery to forge an emotional bond between participants and local cultural heritage. It aligns with prior research demonstrating imagery's effectiveness in conveying meaning across various contexts, including tourism, identity formation, and media representations [67].

Drawing from previous studies, we're motivated to explore the crafting of a process that encourages participants to share their abstract thoughts and insights on local cultural heritage. Additionally, we aim to devise an exhibition format that effectively showcases these intangible aspects of cultural heritage. Moreover, we aspire to create an interactive experience for exhibition visitors, leveraging mediums such as text and imagery to stimulate reflection on the significance and trajectory of cultural heritage.

2.2 Expressing and Creating with Generative AI

GenAI tools like GPT[1], PaLM 2[2], MidJourney[3], and Stable Diffusion [55], has gained increased integration into our lives. These tools can create human-like text and images from natural language prompts, excel in domains like image generation [14, 54, 56] product design [16, 33, 43], and fashion design [34, 68]. These applications of GenAI have underscored its use in supporting creativity in human-AI co-creation tasks [50, 56, 62]. For example, ReelFramer [66], a GenAI tool, assists journalists in creating social media reels by converting news articles into scripts and storyboards. In this collaborative process, AI rapidly generates ideas and prototypes while humans oversee each step, ensuring content coherence and appropriateness.

Compared to text-to-text generation tools, text-to-image generators are convenient in ideation processes because they can quickly produce images [53], which enables real-time visual communication beyond the vagueness of text output [40]. This visualization ability of AI image generator has been used for architectural design [51], visually-guided open-ended story-creation [71], and art proposal design [16]. The interaction between AI image generation tools and humans can be depicted as a two-way search, visually illustrated through the feedback exchanged, facilitating participants' conceptualizations and expression [16]. Earlier research introduced crea.blender SDG, an AI image generation game, revealing AI images' capacity to enable users to depict feelings of both anxiety and hope for the future. This study underscores that user-generated AI images effectively convey these sentiments to audiences beyond their creators [54], emphasizing the communicative power of text-to-image generation tools in expressing participants' intentions.

Another advantage of AI image generation tools is that they can assist in image reference searches, expanding ideas, and drawing inspiration from unconventional, out-of-the-box examples [40]. This ability of AI tools can provide fresh perspectives, fueling creativity [16]. For instance, in a speculative design workshop where AI was used to create utopian structures, researchers observed that the unexpectedness of AI-generate images led creators in new, valuable directions for ideation. The difference between the participants' expected outputs and the AI-generated images facilitated new insights [24]. Likewise, a previous study examined AI uncertainty through a phenomenological lens, highlighting the significance of data noise and model variance in creative design. This research proposes that embracing uncertainty in decision-making can serve as a valuable design element, providing opportunities to develop artifacts and scenarios that capitalize on these distinctive characteristics [8]. The unpredictability introduced by AI can be understood as unintended ambiguity and imperfections in AI-generated images, which can stimulate deeper interpretation and the discovery of new insights among participants [16].

Drawing from insights in the literature, we posit that text-to-image tools facilitate the presentation of inspiring concepts through interactive processes and expressive visualization. These tools empower non-experts to visually articulate their insights, connections, and relationships with artifacts, moving beyond mere description

[1] https://openai.com/gpt-4
[2] https://ai.google/discover/palm2/
[3] https://www.midjourney.com/

Exp. No.	Gender	Age	Heritage Type	Threat Type
1	Male	24	The Forbidden City	The Forbidden City is on fire.
2	Male	23	The Forbidden City	The increased population due to tourism leads to the oxidation of the Forbidden City.
3	Female	49	The Forbidden City	High tourist numbers contribute to wear and tear of the Forbidden City
4	Female	21	Beijing Drum Tower	Inflexible commercialization without adapting to local conditions
5	Male	24	The Hanging Temple in China	Structural aging and other reasons caused the Hanging Temple to be damaged or even collapsed.
6	Female	23	Milan Cathedral	Milan Cathedral collapses due to hidden structural issues.
7	Female	27	The Forbidden City	Excessive development of tourism in the Forbidden City
8	Female	27	The Colosseum of Rome	The Colosseum of Rome was hit and submerged by the torrential rain
9	Female	24	Historical Centre (Old Town) of Tallinn	Historical Centre (Old Town) of Tallinn fall off over time or because of the humidity/over-tourism of Historical Centre (Old Town) of Tallinn
10	Female	24	Lijiang Old Town	Over-tourism and an influx of visitors are harming the local cultural ambiance of Lijiang Old Town.
11	Female	23	The British Museum	The British Museum is on fire
12	Female	25	The Temple of Heaven in Beijing	The Temple of Heaven in Beijing was completely submerged by the sudden flood
13	Male	20	Dayan Pagoda	Dayan Pagoda in Xi'an suffered from over-tourism
14	Female	32	The Mogao Grottoes	The Mogao Grottoes change after the wind was eroded by sand
15	Female	30	State Theatre (North Point)	State Theatre (North Point) in Hong Kong is on fire

Table 1: Summary of cultural heritage threat choices of participants with demographics.

to tangible, prototypable outcomes that others can comprehend and share. In the realm of cultural heritage dissemination, this ability of pictures to "speak more than a thousand words" is exactly the trigger needed for visitor reflection and engagement [48, 67]. Hence, we employed text-to-image GenAI tools, enabling participants to craft images deeply connected to local cultural heritage sites within their communities of practice. The images, alongside their creation process, were intended for exhibition dissemination. This approach sought to move beyond the tangible aspects of cultural heritage, bridging the connection from participants to exhibition visitors.

3 METHODS OF CO-CREATION PROCESS

We interviewed and observed participants engaged in the process of using GenAI to explore personal storytelling related to cultural heritage. This process was conducted online using virtual meeting platforms (Zoom and Tencent). Each session lasted approximately one hour and included one participant and two researchers. One researcher was responsible for conducting the semi-structured interviews. At the same time, due to some participants lacking accessibility to the GenAI tool, another researcher was tasked with copying the participants' prompts sent via the online meeting chatbox while sharing the real-time output of the GenAI tool.

We chose Midjourney as our text-to-image generation tool because of its user-friendly interface and flexibility, making it ideal for participants new to GenAI when expressing cultural heritage narratives. Midjourney's natural language interaction and flexible prompt requirements also simplify its usage [13, 23, 52].

3.1 Participants

We recruited 15 participants (11 females, 4 males) whose ages ranged from 20 to 49 years, with an average of 26 years. 10 participants resided in Asia, while 5 resided in Europe. No participant had extensive previous experience with any GenAI tools, including Midjourney, except P10, who previously used ChatGPT on two occasions. A recruitment notice has been shared within the WeChat community, inviting individuals passionate about cultural heritage preservation to participate in research exploring the intersection of cultural heritage and GenAI. The participants we recruited were Chinese speakers who could speak English (detailed demographic information in Table1). All participants possessed at least an undergraduate degree. Specifically, 10 participants held master's degrees, while 1 participant held a doctorate degree. Of the total, 11 were students, and 4 were employed. Notably, Participants 1, 5, 6, and 14 came from fields closely associated with cultural heritage protection, including urban and rural development, architecture, and heritage preservation.

Since Midjourney only accepts English prompts, as our participants are not all native English speakers, we allowed participants to utilize translators for editing prompts, enabling them to accurately convey the meaning of the content expressed initially in Chinese into English. Two researchers are also responsible for the accuracy and grammatical integrity of the translations to maintain fluency without altering the original intent of their prompts.

Ethical approval for all research procedures was obtained from the university's Institutional Review Board (IRB), and informed consent was obtained from all participants. The use of anonymized data was ensured to protect participants' privacy. Additionally,

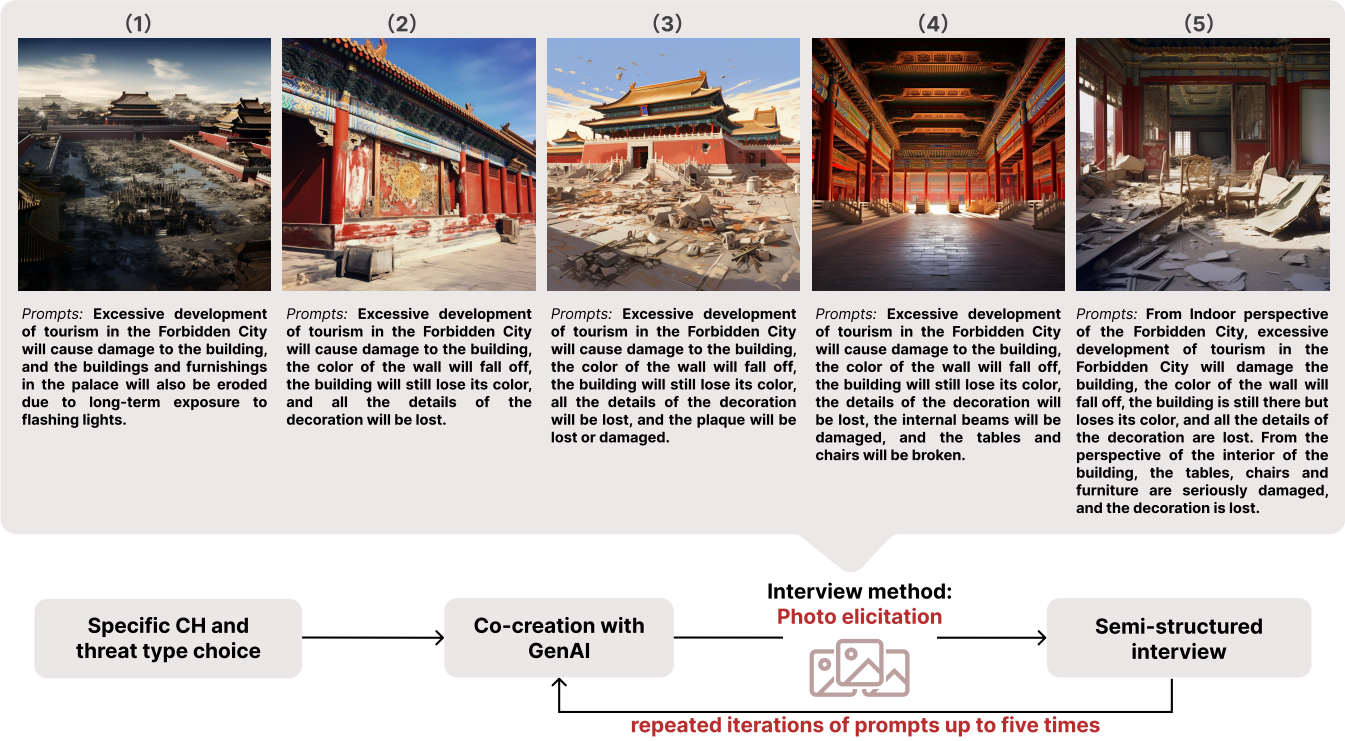


Figure 1: The workflow of the co-creation process and examples from P7, showcasing the co-creation process with GenAI, featuring images generated from start to finish along with their respective prompts.

participants received a USD $10 compensation upon completion of the interview.

3.2 Co-creation Process Procedure Design

We requested participants to choose a local cultural heritage site with which they feel a strong personal connection. This could be a site located in their residence or one they have visited, especially one that left a lasting impression on them. Our approach is grounded in the belief that personal connections allow participants to develop profound insights into physical sites and their own relationships with them. Building on the success of strategies like mobile apps and AR in enhancing cultural heritage storytelling and strengthening connections with local heritage sites [26, 29], we investigate how GenAI, as a technical tool, can further enrich these connections. Given the pressing challenges facing architectural cultural heritage sites, such as destruction, inadequate maintenance, over-tourism, and climate change [2, 61, 63], participants were prompted to consider the importance of cultural heritage preservation during the task [5]. They were tasked with imagining potential future threats to their chosen cultural heritage sites, drawing from their own experiences, and describing these scenarios in natural language to serve as prompts for the image generation process. To facilitate this process, we provided examples of common threats, including tsunamis, floods, extreme temperatures, over-commercialization, and conflict, as documented in previous literature [12]. This approach aimed to engage participants by generating images that depict potential future scenarios of cultural heritage under threat, fostering speculative reflection on their future.

To ensure that participants' envisioned scenarios align with the AI-generated images, we permitted up to five rounds of prompt modifications. Previous research has demonstrated that people's perceptions may change after viewing AI-generated images, potentially diverging from their initial prompts [24]. This iterative process acknowledges the potential for evolving expectations and prompts participants to refine their prompts based on reflections from previous image generations. Before each subsequent generation, participants are encouraged to evaluate any disparities between their envisioned scenarios and the AI-generated images from the previous round, prompting them to adjust their prompts accordingly.

After each image generation, we conducted semi-structured interviews with participants. Utilizing the photo-elicitation method [30, 31], we leveraged the unique ability of photographs to evoke information, emotions, and memories through their specific representations. This involved using AI-generated images to prompt reactions and responses from participants in each round. Initially, we asked participants to describe their observations from the generated images. Depending on their initial responses, we asked questions about any disparities between the images and their expectations. We also asked for suggestions on how subsequent prompts could better align with their vision. Furthermore, we encouraged participants to share thoughts or ideas about the cultural heritage depicted in the images or their interpretation of cultural heritage

Figure 2: In this generation process, the modification involves changing and enriching details and narrative descriptions. Text displaying the prompts provided by participants and the corresponding AI-generated images. Blue denotes expressions of details modified during the generation process, while red signifies the narratives of added details.

preservation. Following the entire generation process, participants were asked about the most challenging aspects of their interaction with GenAI and the strategies they employed to overcome these challenges.

We applied grounded theory to qualitatively analyze the interview data and the corresponding image and prompt data in the co-creation process with GenAI [21] using the following procedure: (1) Two authors read interview transcripts multiple times to become familiar with the data; (2) Two authors identified thematic topics and common features in the collected data together; (3) Two authors conducted a comprehensive review of the thematic topic to delineate sub-themes. They further examined the qualitative findings from the interviews and the corresponding images and prompts sequence to identify any correlations between them; (4) One author meticulously refined and labeled all the subthemes and presented the findings concerning our research questions. Additionally, they utilized corresponding AI-generated images and prompts to supplement the qualitative data, providing further elucidation.

4 RESULTS OF CO-CREATION PROCESS

4.1 Reviving Memories and Expanding Cultural Heritage Narratives through Interaction with GenAI

In this section, our analysis revealed a fascinating interplay between GenAI co-creation and participants' deep-seated memories of cultural heritage sites. This dynamic engagement prompted participants to iteratively refine their cultural heritage descriptions, driven by past experiences and personal connections. In the process, they corrected inaccuracies and infused their narratives with rich,

Figure 3: (Above) Images of the ancient city in Tallinn, as imagined by P9, indicating a change in the perspective of looking at the cultural heritage site. Text illustrating the prompt used by P7, which was changed to unveil the indoor views of the Forbidden City. (Below) The changing of styles by P12 makes the presentation of cultural heritage specific.

beyond-the-surface elements, delving into the essence of cultural heritage. This exploratory journey through diverse perspectives underscored GenAI's potential in unveiling the layers of intangible understanding tied to cultural heritage, as participants wove personal stories and reflections into their visualizations.

4.1.1 Detail-Oriented Narrative Refinement of Cultural Heritage. The first behavioral pattern we observed was the participants noticing details of the heritage in the GenAI-created images that do not match their memories. For example, P8 iteratively modified the prompts to include specific information like *"just a bit of the ruins peeking out of the water"*, to change the point in the images where it depicted stones instead of ruins, which is inconsistent with the scenes in his recollection (**Figure.2 (1)**). Other cases of modifying the details in the images also included adjusting structural and environmental information, such as road widths (P5, P9, P6) and wall colors (P4).

Despite the participants adjusting details, the images generated by GenAI often included elements beyond the initial prompts, such as the representation and expression of the environment, which might not have been present in the participants' original descriptions. This motivated them to revise their prompts further to adjust their vision for future scenarios of cultural heritage that differed from the GenAI output (**Figure.2 (2)**). For example, as P4 said,

> *P4: "While I initially only envisioned the appearance of the Drum Tower alone, the image included scenes around the Drum Tower. Consequently, my subsequent prompts included more descriptions of the surroundings and added prompts like "Surrounded by many modern shops. Brand new roads and transportation facilities.*

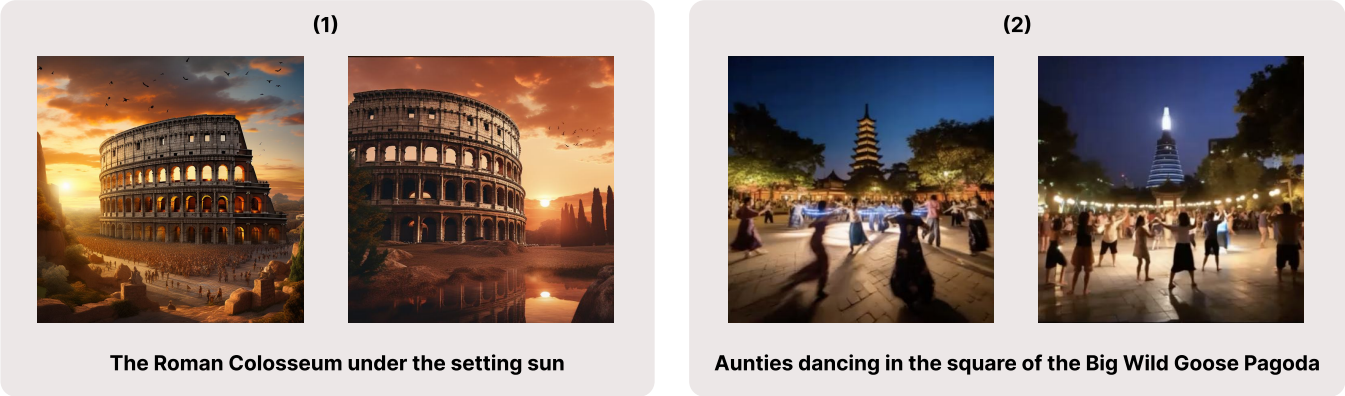


Figure 4: (1) The AI-generated image evoked memories of P8's past visit to the Roman Colosseum. (2) The AI-generated image sparked P13's recollections of the public dancing in the square before the Dayan Pagoda.

> *At the same time, I see people in the Drum Tower in the AI-generated images, but in reality, people aren't allowed inside the Drum Tower, so I want to add 'no people.'".*

Participants gave further specifications of fine details previously not considered by them after observing the images generated by GenAI. For instance, while envisioning a scenario of the Forbidden City suffering wear and tear due to excessive tourism, P3 expanded on the description with detailed specification of colors and environments of cultural heritage (**Figure.2(3)**). Similarly, P14 further elaborated on the style and specific details of the deterioration of the statues in the Mogao Grottoes (**Figure.2(4)**).

This process highlights GenAI's role in enhancing narrative precision and underscores its potential in capturing the nuanced, often intangible, connections individuals have with cultural heritage. Each modification served as a testament to the participants' desire to represent their intangible understanding of these heritage sites accurately.

4.1.2 Expanding Perspectives on Cultural Heritage Narratives. Beyond refining details, participants expanded cultural heritage narratives by adopting diverse perspectives, enriching the cultural heritage storytelling fabric. This included altering viewpoints and styles to capture the multifaceted essence of cultural heritage sites. For example, P7 expressed that in addition to generating images of the exterior of the Palace Museum, she also wanted to see scenes from inside the Palace Museum. Therefore, she added 'From the indoor perspective' to her prompts to explore more cultural heritage narratives. Similarly, P9 explored the ancient city of Tallinn, moving from street views to bird-view views of the city (**Figure.3**). She expressed her motivation as follows,

> *P7: "After seeing the AI-generated images, I thought AI could present me with scenes I couldn't see during my real-life visit. Because the real-life ancient city of Tallinn is near the sea, the scene must have been beautiful and spectacular. Still, when I visited there, I didn't have a drone to see the scene, so I wanted to see what the bird-view ancient city of Tallinn looked like with GenAI's capability."*

In addition to perspective changes, participants also varied the style of the images generated by prompting. P12 felt that the non-realistic depiction of the Temple of Heaven during a flood failed to portray the cultural heritage accurately. Therefore, she introduced "realist style" prompts to alter the image's aesthetic approach.

These narrative expansions reflect a deeper engagement with cultural heritage facilitated by GenAI, where participants explored unseen or unconsidered aspects of cultural heritage, revealing the technology's potential to broaden our understanding and appreciation of cultural heritage through new, imaginative visions.

4.1.3 Triggering Personal Storytelling. After viewing the GenAI-created images, participants were inspired to connect the scenes in the images to their own heritage-related experiences, leading them to share personal stories about heritage. For example, after P8 saw the image of the flooded Colosseum of Rome generated by Midjourney, she recalled being impressed by the sight of the Colosseum at sunset during her visit,

> *P8: "Seeing GenAI's depiction of the Roman Colosseum flooded at sunset reminded me of my visit there. The grandeur of the Colosseum, bathed in the evening sun as portrayed by GenAI, echoed my experience of leaving the site at dusk, a moment that profoundly impacted me. This experience and the GenAI scene also reinforced my belief in the resilience and possibility of rebuilding after disasters."*(**Figure.4**(1))

P13 exhibited a similar reaction. Upon viewing the GenAI-generated scene of the Dayan Pagoda at night, he was reminded of past experiences of seeing a crowd dancing in the square in front of the pagoda,

> *P13: "The AI-generated image brings to mind my previous nighttime walks past the Dayan Pagoda square. The scene is lively with many people dancing to music blaring from loudspeakers, a noise that, each time, seems to erode the pagoda's inherent tranquility."*(**Figure.4**(2))

In a similar vein, P5 reminisced about a previous trip to the Hanging Temple, where crowds of people journeyed along the narrow cliff-side pathway together. This pattern of behavior is also mirrored in

Figure 5: Unexpected results in AI-generated images allow participants to rethink heritage conservation and give participants new inspirations. (1) The AI-generated images diverged from P4's anticipations, depicting the environments of cultural heritage sites with a style echoing the Drum Tower. This led participants to reflect on whether this approach represents a viable method for commercialization. (2)(3) The images contained details that diverged from the participants' expectations, with these deviations from reality prompting reflections on strategies for cultural heritage preservation. (4)(5) The scene in the AI-generated images, which are different from the actual scene, makes P2 and P11 reflect on cultural heritage preservation.

P9's account of a shopping outing at a souvenir store in the quaint old town of Tallinn.

These personal narratives enrich participants' memory of cultural heritage and highlight the potential of GenAI in fostering a deeper, more personal understanding of cultural heritage.

4.2 Insights and Reflections about Cultural Heritage Protection through Interaction with GenAI

Furthermore, interactions with GenAI have proven to be a rich source of insights and reflections on cultural heritage protection, triggering significant emotional impact among participants. Through the lens of GenAI, individuals could visualize and reconsider aspects of heritage preservation in ways that traditional methods had not allowed, revealing the nuanced and often overlooked dimensions of cultural conservation.

4.2.1 Unexpected Results Spark Reinterpretation and Reflection. The power of GenAI to surprise and challenge preconceptions was evident in its generation of images containing details and scenarios beyond participants' expectations. For instance, when examining a GenAI-generated image, P14 noticed a representation of the Mogao Grottoes experiencing erosion, accompanied by an oasis in front of

the caves. However, as no natural oases surround the real Mogao Grottoes, she omitted this detail from her prompts. Nonetheless, upon observing it, she suggested that planting trees around the grottoes could serve as a viable conservation strategy (**Figure.5**(3)). Similarly, P4, with her local understanding of Beijing's Drum Tower and concerns about its over-commercialization, finds inspiration in a GenAI-generated image of this site,

> *P4: "The thoughtful integration of elements like lanterns and umbrellas in the streets, matching the style of the Drum Tower in the AI-generated image. It starkly contrasts the haphazard commercialization when I visit Drum Tower, as indicated in prompts "messy modern shops of different styles". However, I think it suggests that in real-life commercial developments, especially in areas surrounding cultural heritage sites, efforts should be made to ensure that elements like umbrellas and shops are in harmony with the cultural heritage site's aesthetics and ambiance, thereby maintaining a balance between modern commerce and historical preservation."*(**Figure.5**(1))

During iterative generative prompting, P4 also saw an AI-generated image that depicted the Drum Tower adorned with neon lights, a detail not specified in the prompts. This led P4 to comment:

> *P4: "Currently, the Drum Tower doesn't have such decorations in real life, but some ancient buildings are already outfitted with neon lights. I believe if the Drum Tower were to be decorated in this way in reality, it would certainly be a form of destruction."*(**Figure.5**(2))

This response reflects a concern for maintaining the historical authenticity of cultural heritage sites amidst modernization efforts after seeing the unexpected results from GenAI.

In addition to unexpected details, GenAI-created images also contained unforeseen scenarios that prompted reflection among participants. For example, P7 encountered a scene depicting the Forbidden City obscured by trees while generating images. For her, this image evoked a sense of isolation in the Forbidden City, conveying a lonely atmosphere. It led her to reflect that although over-tourism should be avoided, cultural heritage sites being ignored is not an ideal state either, highlighting the need to find a balance between overwhelming tourism and complete neglect (**Figure.5**(4)). Similarly, while generating images of the streets of the old town of Tallinn, P9 saw a close-up of souvenir shops:

> *P9: "In the AI-generated images, the items in these shops differ from what is typically sold in Tallinn's souvenir stores; they resemble the generic merchandise found in overly commercialized European souvenir shops. However, this doesn't seem like a good approach to commercialization. Each old town has its unique character, and I believe the products sold should reflect the local distinctiveness."*(**Figure.5**(5))

These instances highlight GenAI's distinct capacity to stimulate contemplation and ignite inventive approaches to heritage preservation through the presentation of visual narratives that transcend the conventional perception of cultural heritage sites.

4.2.2 The Role of AI-Generated Images in Raising Public Awareness. Beyond unexpected scenarios and details, after viewing cultural heritage scenes generated by GenAI, most participants expressed shock by the depicted cultural heritage under threat scenarios in AI-generated images. They expressed concerns about the future of sites they hold personal connections to, highlighting GenAI's ability to raise public awareness visually. For instance, P11, who had never experienced a fire, was impressed by the image of the British Museum engulfed in flames. She noted that the AI-generated images had a strong visual impact when considering the tragic destruction of invaluable artifacts and the grim possibility that such heritage sites might only exist in memory one day. P7, upon viewing the Forbidden City overwhelmed by excessive tourism, commented:

> *P7: "The images of crowds engulfing the Forbidden City struck me, vividly portraying cultural heritage being devoured like a vast cake by endless swarms of ants. It felt as if the Forbidden City was being eroded, piece by piece, by the relentless flow of visitors."*

These scenarios of cultural heritage under threat generated by GenAI resonated with participants, demonstrating the potential of GenAI in raising public awareness about the challenges facing cultural heritage.

4.3 Cognitive Engagement with Constraints of the Generative Process

In this section, we highlight the challenges participants face in co-creating with GenAI, including issues with identifying keywords for prompting, lack of consistency, the need for adherence to real-world logic, and cultural biases in generated images.

4.3.1 Challenges in Identifying Keywords within Prompts. Five participants described the necessity of prompt modification due to GenAI's limited understanding of "specific keywords" and its lack of effectiveness in executing precise prompt instructions.

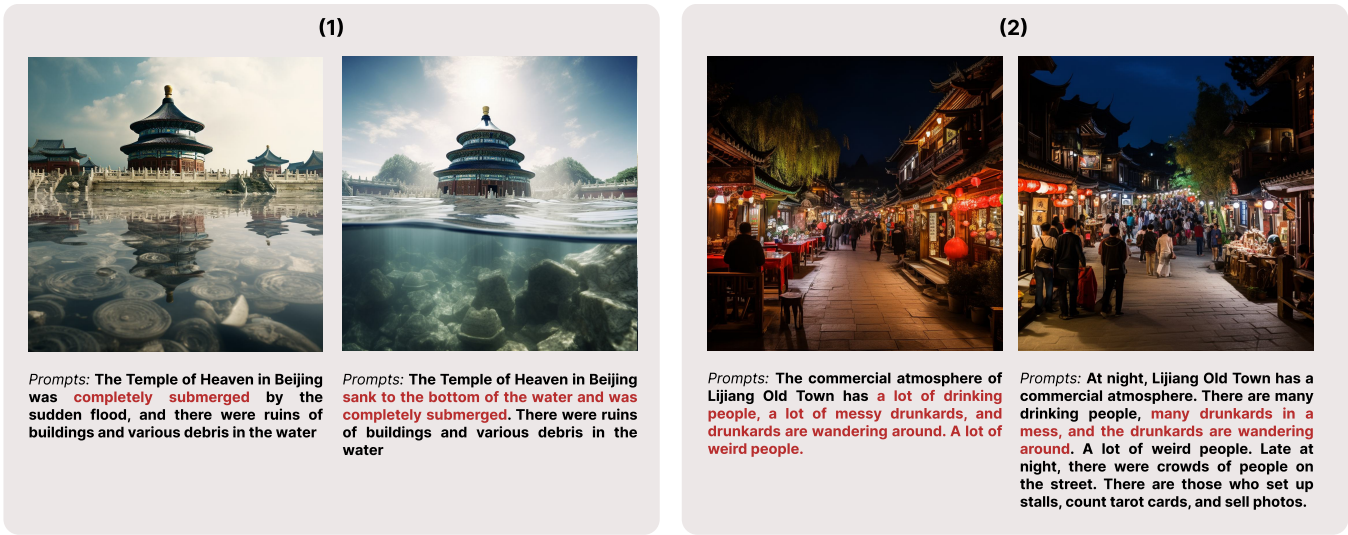
> *P5: "I added some prompts 'blend modern elements into the ancient building' to incorporate modern elements into the tower. However, Midjourney interpreted it as a direct attempt to modernize the cultural heritage through reconstruction. It seems that "blend elements" is particularly difficult for Midjourney to understand."*

Participants noted that some words used in the prompts are not reflected in the images generated. Moreover, for some cultural heritage sites such as the Dayan Pagoda, Midjourney is not able to present the form of the shape of the architectural cultural heritage correctly:

> *P13: "I've generated the appearance of the Dayan Pagoda several times but found that it doesn't quite match the actual form of the pagoda itself. Sometimes, the main body is correct, but the eaves and the number of floors are inaccurate. I guess this might be because Midjourney fails to recognize the keyword "Dayan Pagoda" and only identifies the word "Pagoda" instead."*(**Figure.7**(3))

The inability to recognize keywords is one of the main reasons why participants modify prompts, and a few participants modify them repeatedly and are still not too satisfied with the effect of the images. For instance, when P12 wanted to generate a scene of

Figure 6: Enhancing participants' awareness of cultural heritage protection through AI-generated images of threatened sites.



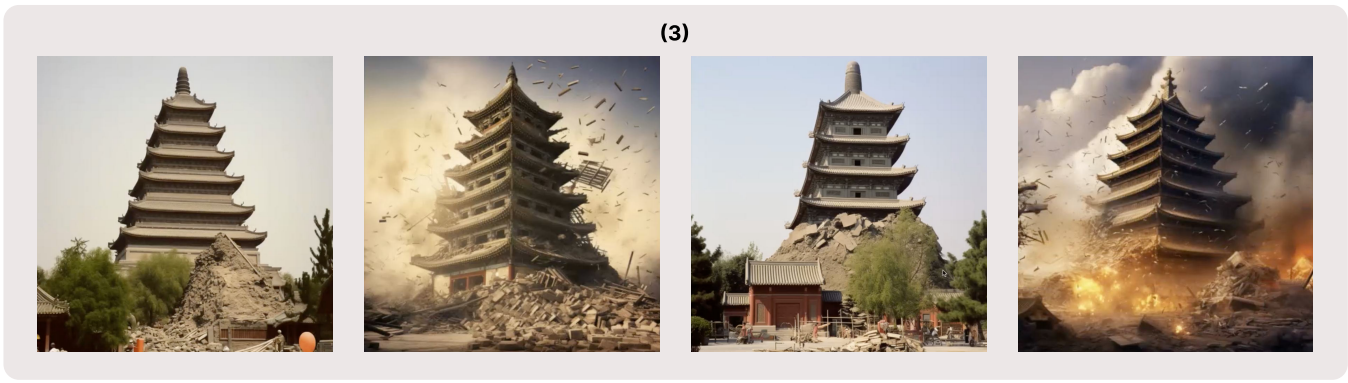


Figure 7: (1) When generating the scene of the Temple of Heaven being flooded, the phrase "completely submerged" is not well recognized. Still, after modifying the keywords (the red text is the modified keyword), the result is satisfactory to P12. (2) Even after expanding the description and repeatedly modifying it, the image generated by AI does not reflect the participants' description (in red text). (3) The morphology of the Dayan Pagoda generated by P13 changed during each generation, but none of them accurately reflected the true morphology of the Dayan Pagoda.

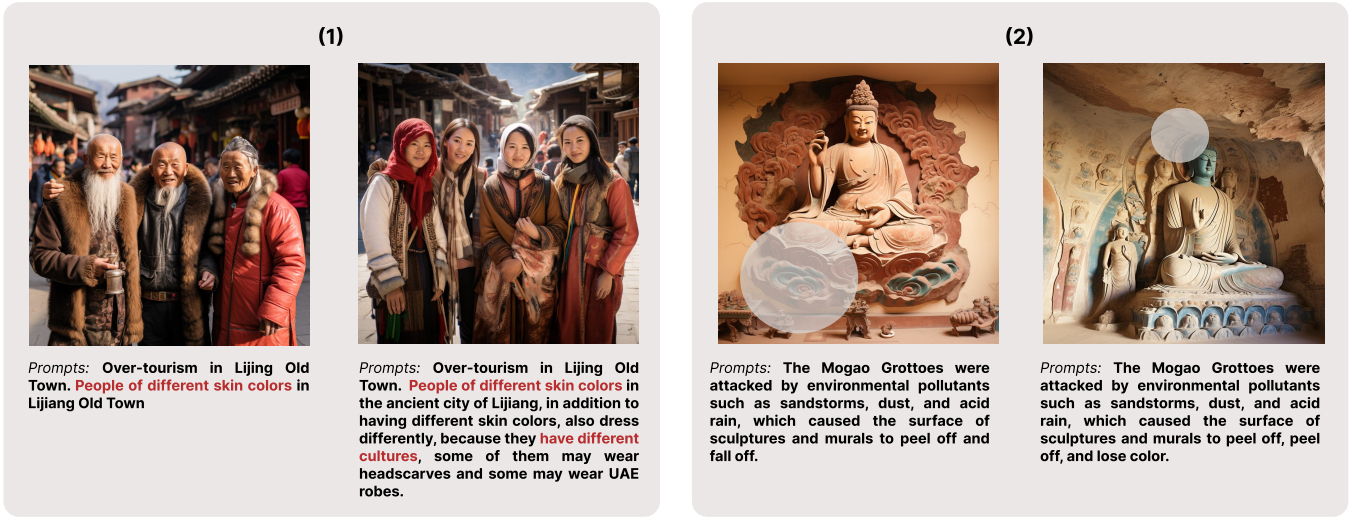


Figure 8: (1) The face of the generated character is considered to have stereotypes by P10; (2) The sculptures of the Mogao Caves in Dunhuang are considered by P14 to incorporate elements from other cultures, which is misinformation.

the Temple of Heaven being completely submerged in water, she found that the term "completely submerged" did not convey the desired depiction of the Temple of Heaven being fully underwater. However, the situation changed when she changed the description to "sank to the bottom of the water and was completely submerged (**Figure.7**(1)). A similar case is that P10 repeatedly modified prompts to generate a scene of the ancient town of Lijiang with many drunkards in the streets. Still, the images produced depicted the town in its usual state (**Figure.7**(2)). This left P14 quite disappointed. This demonstrates that for beginners using AI image generation tools, the inability to correctly use prompts and provide keywords that the generation tools can recognize may affect their experience.

4.3.2 Inconsistency in Conveying Participants' Intentions. We found that the ability of AI-generated images to convey participants' intentions in the image sequences was lacking, leading to strategic prompting to adjust for the lack of consistency.

> *P6: "In the set of images in the last round, poles were supporting the Hanging Temple. However, those poles are absent in the images generated in this round, making the temple seem to float halfway up the mountain. So, I had to adjust the prompts to return the previous architectural structure shown in the earlier steps, which is correct. Sometimes, even with the same prompts, it's difficult to reproduce specific details correctly from previously generated images."*

P3 also mentioned this when she prompted regarding the degree of wear and tear in the Forbidden City. Even though the prompts to describe the extent of the damage have not been changed, sometimes the extent of cultural heritage damage is considered by participants to be too out of line or too little compared with expectations (**Figure.1**(3)).

4.3.3 Resolving Contradictions to Real-World Logic. Participants reported that the logical relationships among depicted elements in GenAI-generated images contradicted real-world logic. Upon encountering this, participants attempted to change prompts to resolve such contradictions and ensure alignment with the logical progression of real-world scenarios.

> *P14: "In the AI-generated image, the depiction of the Mogao Grottoes fading while still retaining a smooth surface appears illogical. In actuality, wind and sand erosion would alter the texture, making it impossible for just the color to fade. This discrepancy prompted me to enhance the scene by adjusting details such as material and color to create a more realistic portrayal."*

This was also expressed in P13's attempt to generate a scene of the Dayan Pagoda suffering from an earthquake. The AI-generated scene depicted the lower part of the pagoda collapsing severely while the upper part remained intact, which contradicts the logic of how earthquakes affect buildings in reality. Similarly, in the scene generated by P2 of the Forbidden City buildings being oxidized, one image showed some areas in ruins. In contrast, others were very neat and smooth, lacking consistency in the depiction. This dissatisfaction with the results also spurred them to further iterations in the generative process.

4.3.4 Bias brought by GenAI tools. In a few cases during the generation process, participants also expressed concerns that the images generated by GenAI might be biased. For example, P10 found that while attempting to generate scenes of diverse people in Lijiang Old Town, no matter how she modified the prompts, the faces in the images were consistently stereotypes of Asian features, as perceived by her:

> *P10: "My vision is a diverse mix of tourists from different racial backgrounds. However, the AI predominantly generated images with stereotypical yellow-skinned Chinese characters, failing to reflect the diversity I prompted. And I think AI seems to represent Asian*

characters with stereotypical Western perceptions, portraying them with exaggerated wrinkles."(**Figure.8**(1))

A similar case is that P14 pointed out an issue with the representation of the Buddha figures in the Dunhuang Mogao Grottoes. According to P14, the AI-generated images blended elements from various cultural interpretations of Buddha, incorporating details that are not characteristic of the Dunhuang style but are instead derived from other religious cultures (**Figure.8**(2)). From her point of view, this misrepresentation risks distorting users' understanding of specific cultural artifacts and could inadvertently contribute to the spread of cultural stereotypes.

5 EXHIBITION DESIGN

To express the intangible relationships, stories, perceptions, and connections of people with cultural heritage instead of their physical forms, we also explored designing and prototyping two strategies to exhibit the GenAI process outcomes from participants stated in the previous section to visitors. One strategy we designed involves a physical drawing robot to sketch out a hand-drawn version of the previously generated images so that visitors can see a tangible manifestation of the GenAI process that they can post on the wall or take home with them. The other strategy uses projection on a 3D model of physical landscapes to show the environment that cultural heritage sites occupy. In both strategies, the drawn images and the 3D landscape contain AR markers that allow visitors to see the previously Midjourney-generated images associated with the particular cultural heritage created by the participants. Thus, visitors connect with the participants' personal histories and future perceptions regarding cultural heritage, expanding beyond historical knowledge to understand people's relationships with and expectations about the future of these cultural heritage sites. In short, the participants produced knowledge concerning the cultural heritage they know in the GenAI process, and the visitors took in the knowledge they created during the exhibition. This also creates a more dynamic and engaging relationship with participants and visitors to the exhibition [17, 18].

In the first design strategy using the drawing robot, we sought to create a tangible interface that would engage visitors and show a connection with the participants' GenAI creation process through the physical drawing process. Previous work has shown that the robot drawing system can create emotional expression through its movements [44], thus providing the potential to engage emotionally with visitors who watch the GenAI-co-expressed vision being drawn live during the exhibition. We reasoned that the physical creation of the sketch versions of images from the GenAI process could also engage the visitors to see an intangible process made tangible. To this end, we applied the Axidraw drawing robot. The images to sketch are line-drawing versions of cultural heritage sites that correspond to cultural heritage previously imagined by participants in the GenAI co-creation process. These images were processed into Axidraw-compatible SVG format for the robot to interpret and execute. We excluded images that participants identified as potentially causing bias and stereotypes (4.3.4). When visitors approach the Axidraw, they activate the drawing process of Axidraw through a tactile interaction method by touching.

Axidraw's output serves a dual purpose: it is both an artifact of the visitor's interaction and a marker for AR recognition. By scanning these Axidraw sketches with a mobile AR application, visitors can view a series of images (from beginning to end) corresponding to the sketches generated by the participants in the co-creation process. Additionally, they can browse through these images by dragging on the app interface. We provided a QR code to download the AR software to their phones. Also, we provided two Android devices in the gallery itself, ensuring the accessibility of the interaction for engaging with the content.

In the second strategy, we created a landscape with AR markers for different cultural heritage sites that lead to the same AR images shown in an application indicated previously. In our design, a 3D-printed model of chosen cultural heritage sites described by the participants in the co-creation process was manufactured off-site and shipped to the exhibition space. This 3D model of cultural heritage sites was polished and sanded to ensure a flawless surface for the projection. A colorful projection was shown on top of the white surface of the 3D model to mimic the landscape. At the exhibition, the visitors can interact with the physical cultural heritage model by scanning the AR markers found next to it. Scanning the AR markers with the app shows the images created by the participants in the co-creation process with AI.

Our exhibition design rationale is based on two design strategies: delivering the intangible values of cultural heritage by embedding and embodying meaning [22]. Embedding meaning is manifested through sensorized objects such as the drawing robot and 3D models, which physically represent cultural narratives and allow direct interaction with the intangible aspects of heritage. Concurrently, embodying meaning is expressed as visitors engage in meaningful actions, like operating the robot or scanning the models. These gestures do more than trigger digital content; they enhance visitors' connection to the cultural significance and intangible values of the heritage displayed.

6 DISCUSSION

6.1 Enhancing Expressive Narratives through Co-creation with GenAI

GenAI is emerging as a powerful co-creative medium, sparking personal narratives and expressions among participants about cultural heritage. It visualizes these narratives in the form of images, thereby serving as a tool for knowledge dissemination. In the co-creation process, participants initially struggle with abstract ideas and disjointed memories, finding it difficult to convey their desired themes. However, the iterative prompting process showed how individuals can fine-tune and clarify their concepts through visual feedback [16, 40]. Moreover, while envisioning the future of cultural heritage, participants consistently link scenes of cultural heritage with their memories. This is evident in section 4.1, where participants add more prompts to ensure the details in the images align with their memories. For instance, participants change the stone's description, the roads' width, etc., in the prompts to match their memory. This aligns with previous literature that suggests humans tend to think about the future based on assumptions about the past [28]. The powerful visualization capabilities of text-to-image generation tools precisely aid participants in better associating their

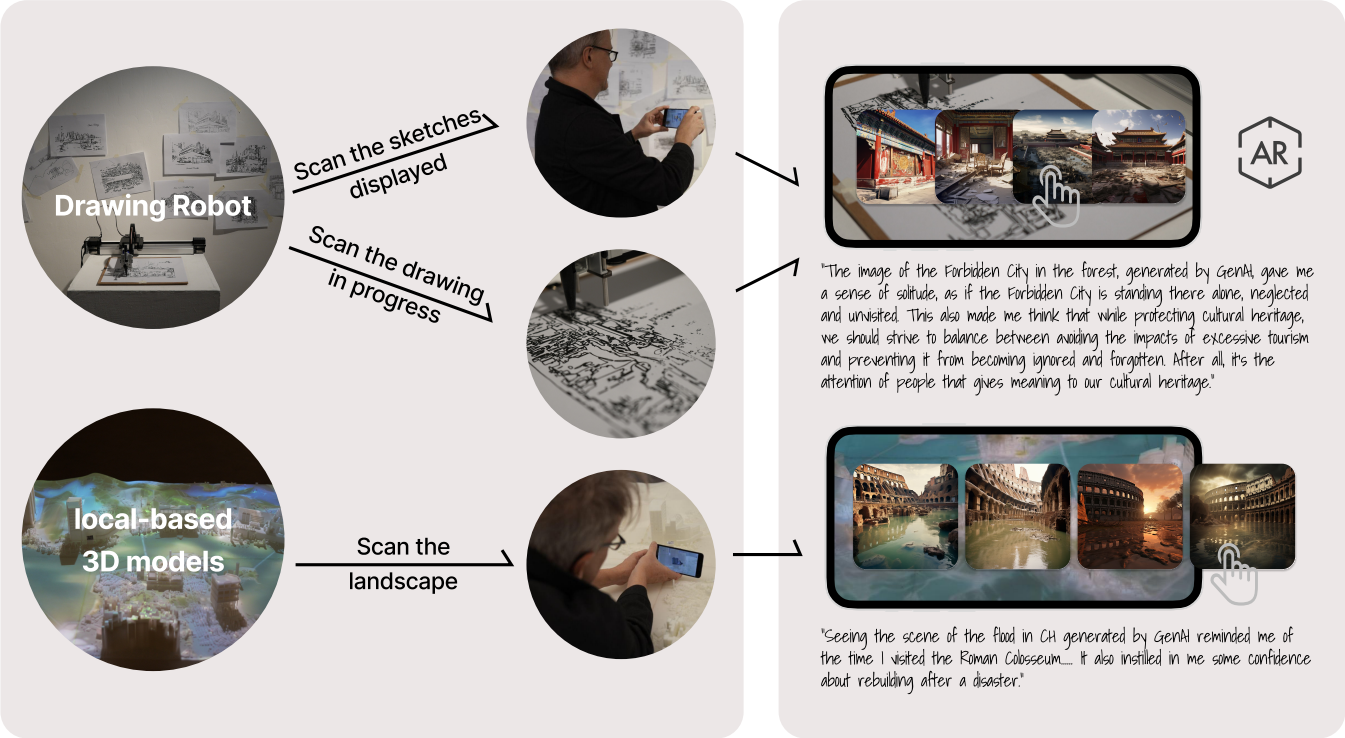


Figure 9: In the demonstration, two types of makers are available for visitors to scan through a mobile AR app: Axidraw sketches and AR tags on local-based 3D models. After scanning, visitors will see the generative sequences co-created by workshop participants and GenAI corresponding to the sketches or the 3D models of cultural heritage sites. By clicking on each image, visitors can also view the participants' key comments.

allocations and experiences with cultural heritage, resulting in more nuanced thematic expressions. After viewing the AI-generated images, which reveal aspects beyond her initial focus solely on the cultural heritage, P4 offered more intricate descriptions of the Drum Tower's surroundings. These elements can potentially guide the prompt's creator towards new, ideation-rich directions, making the narratives richer [24]. This highlights GenAI's ability to aid in self-expression and innovative storytelling in areas that demand narrative and expression, like writing novels [69], creating metaphors for science writing [39], and creating design proposals [16] etc. These findings have implications for the design of image synthesis systems for creativity support.

6.2 Innovating and Reflecting on Cultural Heritage through Co-creation with GenAI

We also observe that GenAI not only enhances individual creativity but also fosters reflection and coming up with innovative strategies for cultural heritage protection. This is accomplished by visualizing unforeseen future scenarios where cultural heritage is threatened, stimulating participants' speculation. Many of these reflections and strategies originate from the unexpected results within the images of cultural heritage under threat. For instance, the surrounding stores are adorned in a style harmonious with the Drum Tower, neon lights absent in reality, and the Forbidden City ensconced in a forest—these are the unexpected elements GenAI infuses into its generated images, surpassing the confines of the initial prompts. These differences from reality drive participants to compare and reflect on the discrepancies between the imagined scenarios in AI-generated images and actuality, leading to new understandings and ideas. This is also shown in previous research that the unexpected differences between the prompt and the generated image's interpretation yielded new insight for and excitement from participants [24]. Furthermore, previous studies have indicated that low-friction and provocative AI-generated images are more likely to spark positive lateral thinking due to their ability to open up greater imaginative space and possibilities [24, 33].

Another reason participants developed new ideas about cultural heritage preservation is due to our designed co-creation process, which places them speculatively envision future scenarios of cultural heritage under threat. Moreover, this co-creation process with GenAI has the potential to raise awareness of cultural heritage protection (4.2.2). This suggests that one promising application of text-to-image generation technology is aiding speculative design [45–47]. Unlike traditional speculative approaches, such as design fiction and co-design [70], GenAI offers a visually intuitive and empowering platform for exploring alternative worlds through iterative prompting, thereby provoking reflections on the issues under consideration [20]. Furthermore, integrating personal narratives with these visual expressions could capture public interest

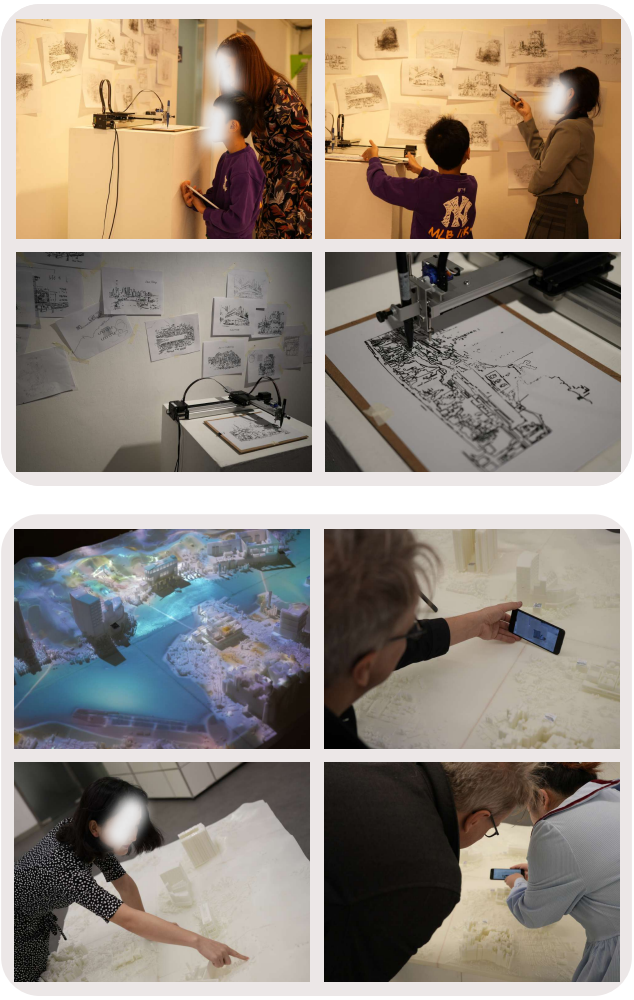


Figure 10: Images from the exhibition display prototypes of the design for AR-based interaction with participant-generated images: (Above) Utilizing an Axidraw robot to sketch images based on line-drawing versions of cultural heritage images. The images drawn also serve as markers for the AR interaction, revealing the full-colored versions of images created by participants in the co-creation process. (Below) Employing landscape-based AR interaction to showcase participant-generated images of cultural heritage futures via AR markers on a 3D printed landscape. The markers are identical to those used in the Axidraw drawings and similarly trigger the display of participant-created images on an Android device.

in urgent matters like climate change and heritage conservation. These generated images may serve as compelling visuals for public events [54], illustrating how personal perspectives on global issues can be conveyed through imaginative and striking imagery, thereby elevating awareness of these socially significant concerns.

6.3 The Double-Edged Sword of GenAI Co-Creation

In the co-creation process, we observed several challenges novice participants face when co-creation with GenAI (4.3). A notable challenge is that they find that GenAI sometimes fails to identify keywords in prompts, exacerbated by novices not always crafting clear, complete, or understandable prompts for GenAI. This often leads to ambiguities or incompleteness in the interactions, highlighting limitations in GenAI's capacity to comprehend contextual knowledge.

Furthermore, issues have been raised regarding the contradiction of GenAI-generated images with real-world logic. This may stem from GenAI's apparent lack of understanding of physical laws, causality, or the logical relationships between objects, as it relies on statistical patterns in data rather than real-world logic [6]. This observation is consistent with the concern that AI-generated images may present new challenges for fact-checkers [38]. While rarely mentioned by previous work, this issue is particularly pronounced given our study's focus on using GenAI tools to replicate real-world scenarios, underscoring our concern over the limitations of employing GenAI in contexts that can adversely affect user experience. The inconsistent logic and the lack of consistency in the outcomes of images generated from the same prompts underscore the inconsistency in GenAI's image creation process. This randomness often hampers the ability to accurately replicate specific details or visual elements.

In addition to this inconsistency, another issue of interacting with text-to-image GenAI tools is the potential for misinformation and bias [10]. Ruha Benjamin's "Race After Technology" exposes the hidden biases in algorithms, revealing their potential to perpetuate harm, particularly in predictive policing, risk assessment tools, and facial recognition technology. Despite being perceived as impartial, these technologies can actually deepen racial inequalities due to biased design and implementation processes, amplifying their negative effects on marginalized communities [9]. The bias introduced by GenAI as a new technology was also mentioned in our results (4.3), which showed that P13 believes that the generated character exhibits stereotypical Eastern facial features, and P14 thinks that the generated Buddha statue incorporates elements from other styles. This can be explained by previous work regarding how training data for generative models may contain historical and cultural biases, potentially reinforcing existing inequities and stereotypes rather than enabling users to envision new possibilities [24]. This suggests that future use of GenAI tools in cultural heritage dissemination must be particularly vigilant in ensuring that the cultural knowledge expressed is qualified with proper assumptions about how the data is used to generate the artifacts. It also suggests that cultural heritage information about past and future records needs to be carefully vetted for inaccuracies and biases when using GenAI.

The challenges encountered in the co-creation experience with GenAI included imperfections that introduced ambiguity diverging from the anticipated outcomes derived from participant prompts. However, this deviation from expectations also stimulated participants' creativity, encouraging innovative ideas and expanding their initial narratives. Therefore, the influence of GenAI on the user

experience presents both benefits and drawbacks, highlighting the necessity for improved design strategies in GenAI co-creation processes to take advantage of inspiration while reducing limitations of bias and inconsistency, aiming to nurture creativity while enhancing the user experience, particularly in contexts that demand creative storytelling with GenAI support. Drawing from our findings, we advocate for these specific design considerations:

(1) **Adaptive Image Editing:** In the iteration process of generating images with GenAI, it becomes challenging to revert to a previous step for modifications when considering the inconsistencies of the results produced by the AI. This predicament raises the question of whether we can edit specific sections of an already-generated image during creation. Ideally, this would involve incorporating elements from prior satisfactory images into the current creation or pinpointing particular regions for refinement via prompt input.
(2) **Predictive and Customizable Prompts:** Novice users found it challenging to ensure every keyword in their prompts is effectively reflected in GenAI. Thus, we propose the need for GenAI to support predictive adjustments and automatic modifications before generation. Allowing the system to parse a user's input and create text prompts automatically could enhance this process. Previous research suggests that a bottom-up approach would enable users to manually refine prompt improvements by selecting from pre-defined parameters manually, thus offering more granular control over the text prompt [19]. Additionally, providing more selectable options for prompt editing near the input box, such as style choices, could inform users of the various ways to exert more control over the generated images.
(3) **Mitigating Bias with Ethical Modals:** Regarding potential biases that text-to-image generation tools may introduce, it is encouraged to utilize more complete and unbiased training sets that do not contain historical biases and cultural practices. Ethical modal [7] can also be employed to ensure that the data generated is fair to people from diverse cultural backgrounds. This is especially important when images from text-to-image generation tools are disseminated to the public because large public engagements can potentially spread erroneous and biased information if unchecked. Furthermore, ensuring greater transparency and regulation of algorithms is important to prevent them from replicating or exacerbating existing social biases [9].
(4) **Improving Idea Capture:** It appears that participants can generate creative ideas when co-creating with GenAI. However, existing GenAI tools do not adequately facilitate participants in capturing these insights and inspirations for later reference. This becomes particularly crucial in fields related to art and design [16, 40], where reliance on GenAI for inspiration is prevalent. This also urges us to investigate methods of incorporating this feature into the co-creation process with AI. This integration would help capture and reflectively process the diverse and integrated ideas that emerge during the interaction. For example, one solution can be an integrated dashboard within the GenAI interface that allows users to capture, visualize, and organize their creative inspirations and ideas in real time. Others can be an idea evolution tracker, a feature that tracks the evolution of ideas over time, showing how initial concepts inspired by GenAI evolve into final artworks or designs.

6.4 Limitations and Future Work

In exhibition design, we've proposed an interactive approach grounded in 3D modeling and Axidraw-based AR scanning to facilitate tangible engagement. However, we've only taken preliminary steps in this direction. In the future, we're eager to deepen our exploration and iteration on leveraging materials acquired through our co-creation process for exhibition design. In our final formal exhibition, we aim to gather visitor feedback through interviews to draw the design implications of exhibition design and show the cultural heritage's intangible value. Furthermore, we want to explore how to elevate visitors from passive receivers of information to proactive creators of a sense of place within the exhibition design [11]. For instance, we're considering strategies to involve visitors in contributing their stories, drawings, or reflections on cultural heritage sites to other visitors, whether through digital media platforms, boards, or physical walls. This initiative aims to weave together a vibrant tapestry of the site's history and future aspirations.

Our participants hail from Chinese-speaking communities and are interested in cultural heritage, though they are not specialists in the field. Consequently, their insights may lack technical and historical accuracy depth, particularly in specialized discussions or interpretations. Additionally, our participant group may not adequately represent the broader public's views, including those indifferent or opposed to cultural heritage preservation. Moreover, since cultural heritage is closely linked to specific cultural contexts, given that our participants are exclusively from Chinese-speaking communities, our results might not be generalizable to other cultural or linguistic groups. Therefore, future research should explore a variety of user groups, including those with varying levels of expertise, from different communities, countries, and cultural backgrounds, and with diverse attitudes towards cultural heritage. Even though our participants can speak English, it is not their native language. Despite the accuracy of translation software, participants may still face barriers to expressing complex ideas.

Our current focus has been primarily on participants' imaginative projections and thoughts regarding the future of cultural heritage. It's also worth expanding our exploration to understand how GenAI can function as a tool in narrating the past and present of cultural heritage. For instance, when participants utilize GenAI to reconstruct past cultural heritage, we can investigate the narratives they create and the memories they evoke. Moreover, we can explore how participants generate scenarios related to familiar and unfamiliar aspects of cultural heritage. This examination may uncover whether they envision less in unfamiliar scenarios or articulate more in familiar ones, shedding light on potential differences. Furthermore, since our study has primarily involved individual narratives, there's a compelling need to investigate how communities engage in dialogues about their perceptions of cultural heritage through a collaborative prompting process with AI.

7 CONCLUSION

Our research delved into the dual aspects of leveraging GenAI for both narrating and propagating the intangible knowledge encapsulated in cultural heritage. Initially, we set up a co-creation framework with GenAI, enabling participants to project future cultural heritage scenarios. Through the aid of visual prompts generated by GenAI, this initiative fostered enhanced storytelling and personal bonds with heritage, encouraging a more profound and extensive comprehension of cultural heritage. The serendipitous insights gained from the GenAI conversions introduced reflective discussions on the preservation of heritage. We also outlined GenAI's constraints in facilitating storytelling and suggested targeted design strategies in response. Expanding our exploration, we designed how images and stories crafted in partnership with GenAI could be instrumental in curating an exhibition of cultural heritage intangible knowledge dissemination utilizing the Axidraw robot and AR. Such integration in our study aims to safeguard and enhance our grasp of cultural narratives for posterity, illustrating a forward-thinking approach to heritage dissemination.

ACKNOWLEDGMENTS

The authors would like to thank the reviewers for their insightful feedback. This work is supported by the Theme-based Research Scheme (T45-205/21-N) of the Hong Kong University Grants Committee (CityU Project No: 8779030) and Innovation and Technology Commission (ITS/326/21FP) of the Government of the Hong Kong Special Administrative Region (CityU Project No: 9449018). We also thank Yihang Zuo, Yihuan Chen, and Hongni Ye for supporting this project.

REFERENCES

[1] 2022. Inseparable Ties: Cohesion as Told by Hong Kong Historic Buildings" exhibition opens. https://www.info.gov.hk/gia/general/202211/09/P2022110900138.htm. Accessed: 2024-01-20.
[2] 2023. https://pragmatika.media/en/khramy-mystetstv-istoriia-rozvytku-ukrainskykh-opernykh-teatriv/
[3] 2023. We are Living Heritage - Photo exhibition for the 20th anniversary of the Convention for the Safeguarding of the Intangible Cultural Heritage. https://ich.unesco.org/en/we-are-living-heritage-photo-exhibition-2023-01331. Accessed: 2024-01-20.
[4] 2024. Tell your living heritage story - Interviews with participants at the UNESCO Youth Forum (audio-visual exhibition). https://ich.unesco.org/en/tell-your-living-heritage-story-2018-01002. Accessed: 2024-01-20.
[5] Dickson Adom. 2016. Inclusion of Local People and Their Cultural Practices in Biodiversity Conservation: Lessons from Successful Nations. *American Journal of Environmental Protection* 4 (2016), 67–78. https://api.semanticscholar.org/CorpusID:157586522
[6] Ajay Bandi, Pydi Venkata Satya Ramesh Adapa, and Yudu Eswar Vinay Pratap Kumar Kuchi. 2023. The power of generative ai: A review of requirements, models, input–output formats, evaluation metrics, and challenges. *Future Internet* 15, 8 (2023), 260.
[7] Hritik Bansal, Da Yin, Masoud Monajatipoor, and Kai-Wei Chang. 2022. How well can text-to-image generative models understand ethical natural language interventions? *arXiv preprint arXiv:2210.15230* (2022).
[8] Jesse Josua Benjamin, Arne Berger, Nick Merrill, and James Pierce. 2021. Machine learning uncertainty as a design material: a post-phenomenological inquiry. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–14.
[9] Ruha Benjamin. 2019. *Race after technology: Abolitionist tools for the new Jim code*. John Wiley & Sons.
[10] Charlotte Bird, Eddie Ungless, and Atoosa Kasirzadeh. 2023. Typology of risks of generative text-to-image models. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*. 396–410.
[11] Kirsten Boehner, Jennifer Thom-Santelli, Angela Zoss, Geri Gay, Justin S Hall, and Tucker Barrett. 2005. Imprints of place: creative expressions of the museum experience. In *CHI'05 extended abstracts on Human factors in computing systems*. 1220–1223.
[12] Lee Bosher, Dowon Kim, Takeyuki Okubo, Ksenia Chmutina, and Rohit Jigyasu. 2020. Dealing with multiple hazards and threats on cultural heritage sites: an assessment of 80 case studies. *Disaster Prevention and Management: An International Journal* 29, 1 (2020), 109–128.
[13] Bowwe. 2023. *Guide to Mid-Journey Prompts*. https://bowwe.com/en/blog/guide-to-midjourney-prompts Accessed: 2023-09-11.
[14] Vivien Cabannes, Thomas Kerdreux, Louis Thiry, Tina Campana, and Charly Ferrandes. 2019. Dialog on a canvas with a machine. *arXiv preprint arXiv:1910.04386* (2019).
[15] Jiaxun Cao, Qingyang He, Zhuo Wang, RAY LC, and Xin Tong. 2023. DreamVR: curating an interactive exhibition in social VR through an autobiographical design study. In *Proceedings of the 2023 CHI conference on human factors in computing systems*. 1–18.
[16] Li-Yuan Chiou, Peng-Kai Hung, Rung-Huei Liang, and Chun-Teng Wang. 2023. Designing with AI: An Exploration of Co-Ideation with Image Generators. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 1941–1954.
[17] Nuno Correia, Tarquínio Mota, Rui Nóbrega, Luís Silva, and Andreia Almeida. 2010. A multi-touch tabletop for robust multimedia interaction in museums. In *ACM International Conference on Interactive Tabletops and Surfaces*. 117–120.
[18] Dan Cosley, Joel Lewenstein, Andrew Herman, Jenna Holloway, Jonathan Baxter, Saeko Nomura, Kirsten Boehner, and Geri Gay. 2008. ArtLinks: fostering social awareness and reflection in museums. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 403–412.
[19] Hai Dang, Lukas Mecke, Florian Lehmann, Sven Goller, and Daniel Buschek. 2022. How to prompt? Opportunities and challenges of zero-and few-shot learning for human-AI interaction in creative applications of generative models. *arXiv preprint arXiv:2209.01390* (2022).
[20] Anthony Dunne and Fiona Raby. 2013. *Speculative everything: design, fiction, and social dreaming*. MIT press.
[21] Ciarán Dunne. 2011. The place of the literature review in grounded theory research. *International journal of social research methodology* 14, 2 (2011), 111–124.
[22] Daniele Duranti, Davide Spallazzo, Raffaella Trocchianesi, et al. 2016. Tangible interaction in museums and temporary exhibitions: embedding and embodying the intangible values of cultural heritage. In *SYSTEMS & DESIGN BEYOND PROCESSES AND THINKING 2016. PROCEEDINGS*. EDITORIAL UNIVERTITAT POLITÈCNICA DE VALÈNCIA, 160–171.
[23] Prompt Engineering. 2023. *Midjourney vs DALL-E 2: Same Prompt, Different Output*. https://www.promptengineering.org/midjourney-vs-dall-e-2-same-prompt-different-output/ Accessed: 2023-09-11.

[24] Ziv Epstein, Hope Schroeder, and Dava Newman. 2022. When happy accidents spark creativity: Bringing Collaborative Speculation to life with generative AI. *arXiv preprint arXiv:2206.00533* (2022).

[25] Kexue Fu, Yixin Chen, Jiaxun Cao, Xin Tong, and RAY LC. 2023. " I Am a Mirror Dweller": Probing the Unique Strategies Users Take to Communicate in the Context of Mirrors in Social Virtual Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.

[26] Bill Gaver, Tony Dunne, and Elena Pacenti. 1999. Design: cultural probes. *interactions* 6, 1 (1999), 21–29.

[27] National Geographic. 2019. *Many historic icons face the same threats as Notre Dame Cathedral.* https://www.nationalgeographic.com/culture/article/many-historic-icons-face-same-threats-notre-dame-cathedral

[28] Maurice Halbwachs. 1992. On collective memory University of Chicago Press. *Chicago IL* (1992).

[29] Kyungsik Han, Patrick C Shih, Mary Beth Rosson, and John M Carroll. 2014. Enhancing community awareness of and participation in local heritage with a mobile application. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. 1144–1155.

[30] Douglas Harper. 1986. Meaning and work: A study in photo elicitation. *Current sociology* 34, 3 (1986), 24–46.

[31] Douglas Harper. 2002. Talking about pictures: A case for photo elicitation. *Visual studies* 17, 1 (2002), 13–26.

[32] Think Global Heritage. 2017. *The Disappearing Chinese Ancient Buildings.* https://thinkglobalheritage.wordpress.com/2017/11/25/the-disappearing-chinese-ancient-buildings/

[33] Matthew K Hong, Shabnam Hakimi, Yan-Ying Chen, Heishiro Toyoda, Charlene Wu, and Matt Klenk. 2023. Generative AI for Product Design: Getting the Right Design and the Design Right. *arXiv preprint arXiv:2306.01217* (2023).

[34] Youngseung Jeon, Seungwan Jin, Patrick C Shih, and Kyungsik Han. 2021. FashionQ: an ai-driven creativity support tool for facilitating ideation in fashion design. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–18.

[35] Lamia Joreige. 2024. Objects of War No. 4. https://www.tate.org.uk/art/artworks/joreige-objects-of-war-no-4-t13250. Accessed: 2024-01-20.

[36] Jae-Shin Kang. 2017. Application method of cultural heritage contents exhibition combining augmented reality technology. *Journal of the Korea Convergence Society* 8, 5 (2017), 137–143.

[37] Sarah Kenderdine, Leith KY Chan, and Jeffrey Shaw. 2014. Pure land: futures for embodied museography. *Journal on Computing and Cultural Heritage (JOCCH)* 7, 2 (2014), 1–15.

[38] G. Khan. 2023. Will AI-generated images create a new crisis for fact-checkers? Experts are not so sure. Reuters Institute. Volume 11.

[39] Jeongyeon Kim, Sangho Suh, Lydia B Chilton, and Haijun Xia. 2023. Metaphorian: Leveraging Large Language Models to Support Extended Metaphor Creation for Science Writing. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 115–135.

[40] Hyung-Kwon Ko, Gwanmo Park, Hyeon Jeon, Jaemin Jo, Juho Kim, and Jinwook Seo. 2023. Large-scale text-to-image generation models for visual artists' creative works. In *Proceedings of the 28th International Conference on Intelligent User Interfaces*. 919–933.

[41] RAY LC. 2023. HUMAN ENOUGH: A Space for Reconstructions of AI visions in Speculative Climate Futures. In *Proceedings of the 15th Conference on Creativity and Cognition*. 217–222.

[42] RAY LC, Sijia Liu, and Qiaosheng Lyu. 2023. IN/ACTive: A Distance-Technology-Mediated Stage for Performer-Audience Telepresence and Environmental Control. In *Proceedings of the 31st ACM International Conference on Multimedia*. 6989–6997.

[43] Yu-Hsu Lee and Chun-Yao Chiu. 2023. The Impact of AI Text-to-Image Generator on Product Styling Design. In *International Conference on Human-Computer Interaction*. Springer, 502–515.

[44] Yanheng Li, Lin Luoying, Xinyan Li, Yaxuan Mao, and Ray Lc. 2023. " Nice to meet you!" Expressing Emotions with Movement Gestures and Textual Content in Automatic Handwriting Robots. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 71–75.

[45] Lauren Lin and Duri Long. 2023. Generative AI Futures: A Speculative Design Exploration. In *Proceedings of the 15th Conference on Creativity and Cognition*. 380–383.

[46] HUGH XUECHEN LIU, YUXUAN HUANG, and JUSSI HOLOPAINEN. 2023. How to use Generative AI as a design material for future human-computer (co-) creation? (2023).

[47] Hugh Xuechen LIU, Yuxuan HUANG, and Jussi Pekka HOLOPAINEN. 2023. Call for Critical and Speculative Design in Human-Computer Co-creativity: An Overview Study. In *14th International Conference on Computational Creativity (ICCC'23)*.

[48] Ying Liu. 2020. Evaluating visitor experience of digital interpretation and presentation technologies at cultural heritage sites: a case study of the old town, Zuoying. *Built Heritage* 4, 1 (2020), 14.

[49] Qazi Azizul Mowla. 2004. Memory association in place making: understanding an urban space. *Memory* 9 (2004), 52–54.

[50] Michael Muller, Lydia B Chilton, Anna Kantosalo, Charles Patrick Martin, and Greg Walsh. 2022. GenAICHI: generative AI and HCI. In *CHI conference on human factors in computing systems extended abstracts*. 1–7.

[51] Ville Paananen, Jonas Oppenlaender, and Aku Visuri. 2023. Using Text-to-Image Generation for Architectural Design Ideation. *arXiv preprint arXiv:2304.10182* (2023).

[52] Niche Pursuits. 2023. *Midjourney Review.* https://www.nichepursuits.com/midjourney-review/ Accessed: 2023-09-11.

[53] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. 2021. Learning transferable visual models from natural language supervision. In *International conference on machine learning*. PMLR, 8748–8763.

[54] Janet Rafner, Steven Langsford, Arthur Hjorth, Miroslav Gajdacz, Lotte Philipsen, Sebastian Risi, Joel Simon, and Jacob Sherson. 2021. Utopian or Dystopian?: using a ML-assisted image generation game to empower the general public to envision the future. In *Creativity and Cognition*. 1–5.

[55] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. 2022. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 10684–10695.

[56] Othman Sbai, Mohamed Elhoseiny, Antoine Bordes, Yann LeCun, and Camille Couprie. 2018. Design: Design inspiration from generative networks. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*. 0–0.

[57] Jae-Eun Shin and Woontack Woo. 2023. How Space is Told: Linking Trajectory, Narrative, and Intent in Augmented Reality Storytelling for Cultural Heritage Sites. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–14.

[58] Laurajane Smith. 2015. Intangible Heritage: A challenge to the authorised heritage discourse? *Revista d'etnologia de Catalunya* 40 (2015), 133–142.

[59] Laurajane Smith. 2020. *Emotional heritage: Visitor engagement at museums and heritage sites.* Routledge.

[60] Waag Society. 2015. https://www.mesch-project.eu/apm-feint-video/. https://www.mesch-project.eu/apm-feint-video/. Accessed on 23 April 2024.

[61] Dirk H. R. Spennemann and Kristy Graham. 2007. The importance of heritage preservation in natural disaster situations. *International Journal of Risk Assessment and Management* 7 (2007), 993–1001. https://api.semanticscholar.org/CorpusID:111082212

[62] Yuqian Sun, Ying Xu, Chenhang Cheng, Yihua Li, Chang Hee Lee, and Ali Asadipour. 2023. Explore the Future Earth with Wander 2.0: AI Chatbot Driven By Knowledge-base Story Generation and Text-to-image Model. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–5.

[63] Jřrí Toman. 2017. The Protection of Cultural Property in Time of Armed Conflict and the Convention on the Means of Prohibiting and Preventing the Illicit Import, Export and Transfer of Ownership of Cultural Property (Paris, 14 November 1970). https://api.semanticscholar.org/CorpusID:169892215

[64] Violeta Tsenova, Gavin Wood, Andrea Dolfini, Annie Tindley, and David Kirk. 2020. Un-authorised view: leveraging volunteer expertise in heritage. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.

[65] UNESCO. 2017. *UNESCO Director-General condemns destruction of Tetrapylon and severe damage to Theatre in Palmyra.* https://en.unesco.org/news/unesco-director-general-condemns-destruction-tetrapylon-and-severe-damage-theatre-palmyra

[66] Sitong Wang, Samia Menon, Tao Long, Keren Henderson, Dingzeyu Li, Kevin Crowston, Mark Hansen, Jeffrey V Nickerson, and Lydia B Chilton. 2023. ReelFramer: Co-creating News Reels on Social Media with Generative AI. *arXiv preprint arXiv:2304.09653* (2023).

[67] Emma Waterton and Steve Watson. 2010. *Culture, heritage and representation: Perspectives on visuality and the past.* Ashgate Publishing, Ltd.

[68] Di Wu, Zhiwang Yu, Nan Ma, Jianan Jiang, Yuetian Wang, Guixiang Zhou, Hanhui Deng, and Yi Li. 2023. StyleMe: Towards Intelligent Fashion Generation with Designer Style. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.

[69] Daijin Yang, Yanpeng Zhou, Zhiyuan Zhang, Toby Jia-Jun Li, and Ray LC. 2022. AI as an Active Writer: Interaction strategies with generated text in human-AI collaborative fiction writing. In *Joint Proceedings of the ACM IUI Workshops*, Vol. 10. CEUR-WS Team.

[70] Shichao Zhao. 2022. Creating Futuristic Heritage Experiences: An Exploratory Co-Design Study through Design Fiction. *Sustainability* 14, 19 (2022), 12027.

[71] Wanrong Zhu, An Yan, Yujie Lu, Wenda Xu, Xin Eric Wang, Miguel Eckstein, and William Yang Wang. 2022. Visualize Before You Write: Imagination-Guided Open-Ended Text Generation. *arXiv preprint arXiv:2210.03765* (2022).