"I Recall the Past": Exploring How People Collaborate with Generative AI to Create Cultural Heritage Narratives

ZHITING HE, Communication University of China, China
JIAYI SU, Guangzhou Academy of Fine Arts, China
LI CHEN, City University of Hong Kong, China
TIANQI WANG, Guangdong University of Technology, China
RAY LC*, City University of Hong Kong, China

Visitors to cultural heritage sites often encounter official information, while local people's unofficial stories remain invisible. To explore expression of local narratives, we conducted a workshop with 20 participants utilizing Generative AI (GenAI) to support visual narratives, asking them to use Stable Diffusion to create images of familiar cultural heritage sites, as well as images of unfamiliar ones for comparison. The results revealed three narrative strategies and highlighted GenAI's strengths in illuminating, amplifying, and reinterpreting personal narratives. However, GenAI showed limitations in meeting detailed requirements, portraying cultural features, and avoiding bias, which were particularly pronounced with unfamiliar sites due to participants' lack of local knowledge. To address these challenges, we recommend providing detailed explanations, prompt engineering, and fine-tuning AI models to reduce uncertainties, using objective references to mitigate inaccuracies from participants' inability to recognize errors or misconceptions, and curating datasets to train AI models capable of accurately portraying cultural features.

CCS Concepts: • **Human-centered computing** → *Empirical studies in collaborative and social computing*.

Additional Key Words and Phrases: cultural heritage, familiarity, narrative, generative AI

1 INTRODUCTION

People visit cultural heritage sites for diverse reasons, engaging in a meaningful dialogue with the history, culture, and humanities intertwined with them. The tangible and intangible elements of cultural heritage that they can see, hear, and touch collectively shape a unique cultural ambiance, offering valuable experiences and knowledge. Within these experiences and knowledge, the social values linked to the historic environment stand out as a crucial characteristic of cultural heritage [27]. This encompasses not only historical value as defined by heritage professionals but also includes memory, oral history, symbolism, and cultural practices associated with the

---

*Correspondences should be addressed to LC@raylc.org

Authors' addresses: Zhiting He, zhiting@cuc.edu.cn, Communication University of China, Beijing, China; Jiayi Su, etnentryno.1@gmail.com, Guangzhou Academy of Fine Arts, Guangzhou, China; Li Chen, lchen488-c@my.cityu.edu.hk, City University of Hong Kong, Hong Kong, China; Tianqi Wang, tttttiq@outlook.com, Guangdong University of Technology, Guangzhou, China; RAY LC, LC@raylc.org, City University of Hong Kong, Hong Kong SAR, China.

---

historic environment [27]. However, visitors to cultural heritage sites are often presented with information related to heritage history, significant events, and important status provided by official historical institutions, known as Authorized Heritage Discourse (AHD) [65]. In contrast, unofficial memories, emotions, and stories related to local people are often transient and invisible. Therefore, heritage studies underscore the importance of presenting the valuable perspectives and stories of communities and individuals familiar with the place, rather than relying solely on official narratives [22, 45, 66].

However, the challenge in collecting community stories is that, although individuals familiar with heritage sites may possess expertise regarding their cultural heritage [10, 11], they are not necessarily narrative experts and lack the means and tools to participate in familiar cultural heritage storytelling. Therefore, we aim to explore the potential of generative AI technology in promoting individuals' participatory activities in the contexts of culture heritage [20] The openness and open-source nature of some generative AI tools [7, 68], along with their auxiliary applications in the creative field, make these AI tools have the potential to lower the barriers to narrative creation, enabling more people to access cultural narratives. This, in turn, facilitates sustained personal storytelling and amplifies diverse voices within the cultural narrative space [65]. Moreover, considering that the data supplied to the AI encompasses a diverse range of perspectives, it can influence individuals' narratives and provide a distinctive visual perspective throughout the generation process, thus fostering inspiration and reflection [7]. However, generative AI also comes with limitations, such as the potential for cultural bias that may mislead individuals unfamiliar with the cultural heritage. Therefore, we aim to explore the ways, advantages, and challenges of using generative AI to assist individuals in cultural narrative creation, and propose possible solutions.

Therefore, the research questions that we aim to answer are:

RQ1: How do participants use narrative strategies when collaboratively creating narratives about familiar cultural heritage with generative AI?

RQ2: What are the strengths and limitations of generative AI in supporting the creation of personal narratives about familiar cultural heritage?

RQ3: How does the creation of personal narratives about unfamiliar cultural heritage, without prior knowledge and experience, differ from that of narratives related to familiar cultural heritage?

To pursue these research objectives, we conducted workshops with 20 participants. During the workshops, participants used Stable Diffusion to generate a series of images related to familiar cultural heritages, such as those from their hometowns or locations where they spent considerable time [56], to create a cultural narrative. In subsequent semi-structured interviews, participants shared their creative themes, insights, and challenges encountered during the collaborative creation process with AI. We also asked them to create narratives about unfamiliar cultural heritages to compare and explore how the level of familiarity with heritage sites influenced their creations.

Our results indicated that (1) participants adopted three distinct narrative strategies, ranging from wholly depicting objective heritage sites to entirely expressing subjective interpretations, with some participants changing their strategies midway through the process; (2) AI assisted in illuminating, amplifying, and reinterpreting personal narrative creation; (3) the cultural limitations, cultural bias and uncertainty of AI had adverse effects on personal narrative; and (4) participants were more prone to overlooking mistakes and biases within output images about unfamiliar cultural heritage. These findings shed light on potential methods, strengths, and limitations in using generative AI to create personal narratives of familiar cultural heritage, thereby demonstrating its potential to facilitate individuals' access to cultural narratives.

2 RELATED WORK

2.1 Cultural Heritage Narrative

Cultural heritage narratives can enhance individuals' comprehension and documentation of cultural heritage [18, 19, 40, 46, 51, 60]. Personal narratives about cultural heritage have the capacity to promote social cohesion by addressing the marginalization and undervaluation of public participation in heritage archives [3, 19, 45, 66]. These individual accounts can further enrich the understanding and perception of cultural heritage for others [17, 51], fostering a more profound resonance within the community and broader contexts [1, 2, 9, 52, 65, 76].

A familiar place refers to individuals' hometowns as well as locations where they have spent considerable time and consequently feel familiar and have extensive previous knowledge of the setting [56]. Familiar places are highly valued in familiar tourism. For example, Clarke and Bowen's investigation into the behaviors and attachments of familiar tourists claimed that locals have extensive knowledge of their area, enabling them to participate in tourism co-creation and become valuable providers of tourism resources [10]. They also mentioned that the community has a historical, storytelling perspective on familiar places [11]. Hung et al. applied VR to enable people to revisit well-known places, illustrating how modern technology can expand and enrich the cultural experience of familiar tourists [24]. These studies reveal the valuable knowledge and understanding that familiar visitors have of places. We believe that locals' familiarity with their area also encompasses an understanding of local cultural heritage, which can provide valuable individual perspectives to cultural narratives.

Narrative is defined as a structured form that people use to organize stories, events, and facts related to their experiences, memories, or understanding of the world, thereby rendering these elements meaningful [49]. Empowering familiar groups to narrate their cultural heritage stories leverages this potential of narratives. Tsenova et al. defined narrative as encompassing multiple stories, offering a broad perspective through which people make sense of their surroundings, with each story focusing on individual characters or personal experiences [66]. Nevertheless, the challenge is evident when individuals residing in heritage sites, despite possessing knowledge about their cultural heritage and personal experiences, lack proficiency in narrative techniques. Consequently, they might encounter difficulties in utilizing tools to effectively convey their stories. Therefore, it becomes imperative to minimize entry barriers, facilitating sustained personal storytelling and thereby amplifying a diverse range of voices within this domain [65].

Recent studies underscore the need for employing technology to facilitate participatory activities that enhance community engagement with cultural heritage. Do et al. demonstrated this through a digital photo display system that actively involved a rural community in sharing and preserving their local heritage [12]. Similarly, Echavarria et al. engaged younger demographics by using augmented reality to transform local heritage sites into interactive learning environments [14]. Giglitto et al. highlighted the power of digital tools in enabling marginalized groups to participate more fully in cultural heritage activities, thereby promoting inclusivity [20]. Ciolfi et al. further emphasized the need for collaborative platforms that bring together professionals and community members in heritage conservation efforts [8]. Collectively, these studies reflect a strong communal interest in leveraging technology to boost engagement and ensure broad participation in cultural heritage preservation.

Technological advancements significantly enhance how communities and individuals narrate cultural heritage. Kambunga et al. proposed that participatory design and interactive exhibitions allow for the excavation of personal narratives within cultural contexts, engaging communities in the preservation and dissemination of cultural heritage [29]. Similarly, the work by Giglitto et al. emphasizes the role of digital technologies in facilitating community participation in cultural

heritage, particularly for marginalized groups such as migrants and refugees, enhancing social inclusion through tailored digital tools that respect cultural nuances [20]. Shin and Woo applied augmented reality (AR) storytelling at cultural heritage sites to create site-specific narratives that enhance the historical and cultural context for visitors, making the narrative experience immersive and localized [61]. Nisi et al. claimed that using digital narrative tools facilitates the sharing of immigrant experiences and cultural heritage, promoting social inclusion and enhancing public dialogue through story-driven digital tools [52]. AI technologies further support these endeavors by enabling sophisticated narrative analytics and personalized content generation, thus enriching the narrative landscape of cultural heritage [15, 72]. Bernstein et al. proposed that generative AI can significantly enhance individual access to narrative creation by lowering the technical entry threshold, thus democratizing the tools necessary for cultural expression [5]. Similarly, Muller et al. claimed that AI technology enables even those without deep technical or creative training to craft complex, creative content, enriching cultural narratives with diverse perspectives [50]. As a result, generative AI has the potential to allow a broader community to contribute their stories and experiences, thereby enriching the cultural narrative in profound ways.

In this study, we aim to involve individuals in collaborative efforts with generative AI to create personal narratives about cultural heritage, thus exploring the potential of generative AI in enhancing individual access to cultural narratives.

2.2 Supporting Cultural Heritage Narratives with Generative AI

Generative AI uses existing media as data to produce new media in various formats such as images, texts, and videos [50]. Generative AI has exerted influence on users in diverse domains [13, 28, 44, 55, 63, 74]. often making creative tasks more accessible by democratizing the tools necessary for creation and engagement [57, 70, 71]. In the realm of social media, Lyu et al. applied generative AI to reduce content creation costs and break down skill barriers, thereby augmenting human creativity and enabling broader participation in content generation, as evidenced in the use of generative tools by YouTubers [47]. Louie et al. applied these technologies to empower untrained users to engage in complex creative endeavors such as music composition, traditionally requiring specific skills and training [43]. Moreover, generative AI has shown promise in art and design [6, 30, 32, 39], such as fostering 3D modeling [41, 75], supporting creative artwork-making [33–35], and aiding in more interactive and inclusive design processes [7, 36, 42, 58, 68]. Guo et al. applied AI tools like DALL-E and Stable Diffusion to facilitate the participation of individuals in artwork creation by lowering technical barriers and increasing productivity [23], fostering broad community participation [21]. This extends to design, where research team developed the BrainFax platform integrated with communication tools to enhance collaborative design efforts, making visual creation accessible to designers irrespective of their visualization skills [67]. Generative AI also plays a transformative role in cultural engagement. Trichopoulos used GPT-4 as an interactive guide in museums, enhancing the visitor experience through tailored narratives and recommendations based on user preferences, thus making cultural education more immersive and personalized [64]. These examples underscore how generative AI can affect community and cultural participation by making sophisticated tools accessible and fostering creative expression.

While generative AI unlocks tremendous potential in creativity and cultural engagement, it also presents significant challenges, particularly concerning ethics, accessibility, and user understanding. Verheijden and Funk claimed that technical limitations arise in visualizing complex or abstract prompts, revealing the constraints of AI in accurately rendering specific functionalities [67]. They also emphasized that ethical concerns are prevalent, with issues like data provenance and the potential reproduction of biases in AI-generated images raising alarms about the integrity of AI models [67]. Lyu et al. pointed out that accessibility challenges are highlighted by the ambiguity

and limited explainability of AI systems, which can alienate non-expert users and erode trust in these technologies [47]. Guo et al. highlighted that the regulation of online creativity also faces threats from AI's capability to produce low-quality and potentially infringing content, triggering conflicts within creative communities [21]. Moreover, Muller et al. proposed that the unpredictable nature of AI-generated content complicates user control, questioning the effectiveness of current user interface paradigms and necessitating a reevaluation of interaction design [50]. The shift from user-driven to model-driven control underscores the need for designing safer and more effective interactions tailored to generative AI [69]. Mahdavi Goloujeh pointed out that ethical dilemmas extend to the sharing of prompts and creative ownership in community settings, where the balance between inspiration and originality must be carefully managed to avoid cultural and ethical missteps [48]. In cultural contexts, Trichopoulos claimed that the deployment of large language models demands meticulous implementation and continuous evaluation to maintain cultural integrity and ensure inclusivity [64]. Addressing these multifaceted challenges is crucial for fostering an environment where generative AI can be both innovative and ethically sound, enhancing its benefits while mitigating risks.

Therefore, it is crucial to explore the potential of generative AI in enabling individuals to better participate in cultural narratives. AI technology holds the promise of enhancing human creativity through co-creation, as seen in design and music, where it supports the creation of high-quality outputs and empowers users with greater control and self-efficacy [67]. Generative AI also facilitates richer engagement in cultural heritage, providing personalized and informative experiences that can transform visitor interactions [64]. However, AI also has limitations, including the need for careful integration to avoid overshadowing human creativity and ensuring that ethical standards and cultural integrity are maintained. By addressing these challenges, we can maximize the benefits of generative AI while providing a more inclusive and creative community engagement.

3 METHODS

3.1 Workshop Tasks and Procedure

We conducted workshops to investigate human-AI co-creation of personal narratives related to cultural heritage. Each workshop was held virtually using an online meeting tool and lasted around 1.5 to 2 hours, involving one participant and one researcher. The workshop included a pre-study interview, a training session, two distinct image generation processes of familiar and unfamiliar cultural heritage. Each generation process was followed by a post-study semi-structured interview (see Fig.1). Stable Diffusion [59] was chosen as the image generation tool because, compared to other generative AI tools such as MidJourney, it is open-source, making it accessible to more people. Additionally, it offers more parameter options and models, allowing participants to flexibly manipulate the style and content of generated images to meet their requirements. We utilized version 1.3.2 of the local version of Stable Diffusion, along with the v5-PrtRE version of the Anything Model. All procedures followed institutional guidelines IRB for human subject testing.

3.1.1 Introduction to the Task. After signing a consent form, participants were given the opportunity to use Stable Diffusion to complete two image creation tasks: one related to familiar cultural heritage and another related to unfamiliar cultural heritage. For the familiar cultural heritage task, participants were asked to first briefly recall and describe their impressions and personal experiences with the familiar cultural heritage to evoke their memories, and then create an image about this familiar cultural heritage, with no restrictions on the theme. For the unfamiliar cultural heritage task, participants were instructed to choose one of the ten provided cultural heritage sites they were unfamiliar with. We showed them around three pictures and a brief introduction to give them a general idea of the place. After that, they were asked to create an image about this
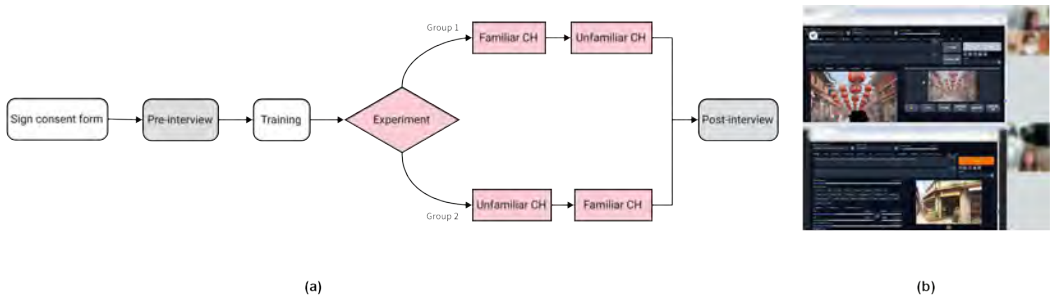
Fig. 1. (a) is procedure diagram, (b) is the screenshot of the online workshop meeting.

unfamiliar cultural heritage based on their own understanding and guesses, with no restrictions on the theme. They could continue generating images until they were satisfied with the result.

We chose Hong Kong as the unfamiliar cultural heritage site due to our familiarity with the locations, which allowed us to identify overlooked details in the images generated by participants. Additionally, we presented 10 distinct cultural heritage sites for participants to choose from, ensuring that each site was selected almost equally to minimize differences in their influence on the generated results, thereby avoiding any bias resulting from specific site characteristics.

3.1.2 Specific Process. Participants remotely controlled our computer to generate images, ensuring uniformity in equipment usage among all participants. Following the introduction, we presented a video tutorial illustrating the utilization of three features in Stable Diffusion for image generation and adjustment—specifically, text-to-image and inpainting. However, participants were not confined to these three features. We standardized the initial settings of Stable Diffusion to the "anything" model, "DPM++ 2M Karras" sampler, "Automatic" VAE model, and a size of 768*512 for each generation process.

Throughout the generation process, participants could write prompts in any way they preferred, whether using complete sentences or single words, and had the freedom to employ any feature of Stable Diffusion to create and adjust images until they deemed the images to meet their expectations. After participants wrote Chinese prompts, researchers translated them into English using ChatGPT and then input them into Stable Diffusion. We made sure that the translation process did not add new words or meanings; researchers instructed ChatGPT to translate the prompts directly, preserving the original Chinese meaning and phrasing. All researchers used the same command-line template to input the prompts into ChatGPT for translation during the experiment. Researchers did not provide any guidance for the content creation, offering only technical support. Each participant completed two counterbalanced image generation processes for both familiar and unfamiliar cultural heritage.

After each generation process, participants were invited to participate in a semi-structured interview. During the interview, participants were prompted to introduce the themes of the generated images and share their comments. Additionally, they were asked to describe and discuss the process of their collaboration with generative AI in creating the narrative.

3.1.3 Data Recording. During the process, we obtained participants' consent to record videos of the Stable Diffusion interface and the online meeting for the entire workshop. Participants were also required to copy the image and its related information and paste it into an online shared file. This served both as a means for them to track their output images and prompts during the process and for us to collect this information for further analysis.

3.2 Data Analysis

We employed thematic analysis to identify potential principal themes from the interview data. In analyzing the participants' interview data, we reviewed related prompts, generated images, and real-life photos to verify and understand the details they mentioned during the interviews, including how they altered their prompts, evaluated the generated images, and compared the images to actual cultural heritages. Three researchers conducted a round of independent open coding [62] and discussed the resulting themes. The primary codes in the first round of coding encompassed AI's presentation of specific cultural features, new elements introduced by AI in narratives, participants' emotions in the narratives, describing specific scenarios in narratives, adapting narrative strategies based on AI, etc. We then went back to the data to conduct two additional rounds of focused coding [62], ultimately achieving agreement on the finalized themes. All quotes used in the paper were translated into English by the lead author and checked by the co-authors.

Furthermore, we organized the images generated by participants into sequences based on familiar and unfamiliar cultural heritages to explore differences in the quantity of images generated, the extent to which participants altered image elements, and other aspects.

3.3 Participants

A total of 20 individuals (Table 1) were recruited from various regions across China through online social platforms, with an age range of 18 to 35 years (M = 26.5, SD = 4.85). The participant group comprised 15 females and 5 males. Participants needed to create images of a familiar cultural heritage and another of an unfamiliar one. Therefore, our recruitment criteria required applicants to be very familiar with a cultural heritage site, such as a place in their hometown or a place where they had spent considerable time, and to have a good understanding of its appearance, culture, and other relevant aspects. Additionally, from the list of unfamiliar cultural heritage sites we provided, they needed to select at least one place they had never visited or knew very little about for creating images. We excluded participants who did not satisfy these criteria. Out of the 20 participants, only 2 had frequently used generative AI tools before; the rest had either never used them or had only used them a few times. Thus, most participants had limited experience using generative AI tools before the experiment. Regarding familiar cultural heritages, 13 out of 20 participants are local residents familiar with their cultural heritage, which they frequently visited. 1 participant lived in the city where their cultural heritage is located for several years. The remaining 6 participants have visited the familiar cultural heritage sites several times and possess considerable knowledge about them. Regarding unfamiliar cultural heritages, 19 out of 20 participants had never been to the unfamiliar cultural heritage sites we provided before the experiment; only P10 had visited the Star Ferry once, but she mentioned that she had merely passed by and had not entered, so she was not familiar with the place. They relied solely on pictures and words to learn about the unfamiliar places. Participants were compensated with a Stable Diffusion handbook containing feature demonstrations. Prior to participation, informed consent was obtained from all individuals. Additionally, all collected data was anonymized with the participants' permission to ensure their privacy.

4 RESULTS

4.1 Strategies Participants Adopted When Creating a Narrative Related to Familiar Cultural Heritage in Collaboration with Generative AI

The results of thematic analysis indicated that, when collaborating with Stable Diffusion to generate images of familiar cultural heritage, participants adopted similar creation strategies. These findings

Table 1. Demographic Information of Participants

ID	Gender	Age	Education Background	Familiarity with generative AI tools	Region of Residence	Familiar Cultural Heritage	Familiarity	Unfamiliar Cultural Heritage	Familiarity
P1	Female	18-25	Humanities and Social Sciences	Used a few times	Jiangsu Province	Ming City Wall	Local resident	Fruit Market	Never visited
P2	Male	18-25	Sciences and Engineering	Never used	Guangdong Province	Da Fen Oil Painting Village	Local resident	Star Ferry	Never visited
P3	Female	18-25	Sciences and Engineering	Never used	Guangdong Province	Monument to the Autumn Harvest Uprising	Local resident	Fruit Market	Never visited
P4	Female	26-35	Sciences and Engineering	Never used	Guangdong Province	Dunhuang	Visited several times	Fruit Market	Never visited
P5	Male	26-35	Humanities and Social Sciences	Frequently uses	Beijing	The Forbidden City	Local resident	Mid-level Central	Never visited
P6	Female	26-35	Business	Used a few times	Guangdong Province	Nantou Ancient Town	Local resident	Fruit Market	Never visited
P7	Female	18-25	Humanities and Social Sciences	Used a few times	Shandong Province	Haiyun Temple Tangqiu Festival	Local resident	Kowloon Walled City	Never visited
P8	Male	18-25	Sciences and Engineering	Never used	Guangdong Province	Shantou Small Park	Local resident	Lui Seng Chun	Never visited
P9	Female	18-25	Humanities and Social Sciences	Used a few times	Zhejiang Province	Grand Buddha Temple	lived for a few years	State Theatre	Never visited
P10	Female	18-25	Humanities and Social Sciences	Used a few times	Guangdong Province	Chang'an Road	Local resident	Star Ferry	Visited once
P11	Female	18-25	Humanities and Social Sciences	Used a few times	United Kingdom	York Minster	Visited several times	State Theatre	Never visited
P12	Female	18-25	Humanities and Social Sciences	Used a few times	Guangdong Province	Chaozhou Shadow Play	Visited several times	Ding Ding	Never visited
P13	Male	18-25	Humanities and Social Sciences	Frequently uses	Guangdong Province	Guanyin Mountain	Local resident	Kowloon Walled City	Never visited
P14	Female	18-25	Sciences and Engineering	Used a few times	Guangdong Province	Nanjing Massacre Memorial Hall	Visited several times	Central Market	Never visited
P15	Female	26-35	Humanities and Social Sciences	Never used	Jiangsu Province	Bell Tower	Local resident	Lui Seng Chun	Never visited
P16	Male	26-35	Humanities and Social Sciences	Never used	Shanghai	Suzhou Gardens	Visited several times	Ding Ding	Never visited
P17	Female	18-25	Humanities and Social Sciences	Used a few times	Shanxi Province	Shanxi University Sculptures	Local resident	Yau Ma Tei	Never visited
P18	Female	18-25	Media and Communication	Never used	Shandong Province	Laoshan Taoist Temple	Visited several times	Yick Cheong Building	Never visited
P19	Female	18-25	Humanities and Social Sciences	Never used	Henan Province	Earthen Wall Grass House	Local resident	Yick Cheong Building	Never visited
P20	Female	26-35	Humanities and Social Sciences	Never used	Guangdong Province	Yuyin Shanfang under Canton Tower	Local resident	Yau Ma Tei	Never visited

present the strategies participants tended to adopt when expressing their personal narratives of familiar cultural heritage through AI collaboration.

4.1.1 Three creation strategies for generating images of familiar cultural heritages. Based on participants' descriptions of their creative themes in the interviews, we found that their approaches varied: some participants tended to objectively recreate the actual state and characteristics of cultural heritages, some added their own memories, experiences, and emotional connections to depict a scenario within the cultural heritage, and others used a creative approach to make interesting creations based on the cultural heritage. We categorized these three creative approaches as (1) depicting realistic scenes, (2) recreating memorable scenarios, and (3) expressing exploratory creation. We then placed them on a spectrum from objectively depicting heritage sites to subjectively expressing interpretations to indicate whether participants' creative attitudes were more objective or subjective (see fig.2).

Participants who aimed to depict realistic scenes typically described specific details of the scene, such as the visual angle (P4), the appearance of the architecture (P2, P5), and the environmental characteristics (P6, P10, P13, P19). They compared the details of the output images with the real scenes that they had observed, and corrected any inaccuracies to make the images more closely resemble what they remembered seeing.

Those participants intending to recreate memorable scenarios often infused their emotions and memories into the images, focusing on impressions (P1, P7, P9, P11, P12) or events (P8, P14, P15, P17) related to cultural heritages. In order to convey their feeling associated with the heritages, participants adjusted color schemes, sizes and shapes of components, composition, and more. The goal was not necessarily to replicate the real scene but to capture their impressions. For instance, P9 explained, "*What I wanted to show was how the Dafo Temple looked to me when I first saw it. It's this old Buddhist building standing tall in the middle of a busy modern city. I wanted to capture how the old and new mix and clash.*". Participants often placed a small figure in the scene as a symbol of themselves to anchor their past selves in the narrative. As P7 noted, "*I added a character in the scene, hoping to depict myself in that past scenario.*"

Participants expressing exploratory creation pursued beyond the cultural heritages, presenting fantastical and imaginative ideas primarily rooted in their personal creativity rather than reality or memory (P3, P16, P18, P20). For example, P20 told us: "*What I initially envisioned was a cyclic pattern, where everything undergoes a continuous cycle of transformation.*" They embraced bold attempts, viewing the imaginative results generated by AI as creative inspiration, as expressed by

Fig. 2. Three creation approaches used by participants, each corresponding to the images they ultimately generated. These three approaches encompass depicting realistic scenes, recreating memorable scenarios, and expressing exploratory creation.

P20: "*I think one good thing about AI is that it can stimulate your imagination; it can come up with things that people wouldn't think of. So, as I proceed, I hope it can be more daring.*"

4.1.2   Participants may change their creative strategies after gaining some understanding of the AI.
Some participants, after seeing AI-generated images, would have a general understanding of the

possible creative results, thus changing their creative methods to align with their perception of the AI's characteristics (P3, P17, P20). For example, P20 changed her creation strategy from "depicting realistic scenes" to "expressing exploratory creation" during the process (see fig.3). She initially added a surreal element to a prompt that originally described an objective scene by including the phrase "*inside a massive eye*," and the resulting image transformed from a realistic style to a more dreamlike one. She then further refined the description of the massive eye and added more mystical terms such as "*dream*," "*surreal*," and "*psychedelic*." This gradually shaped the image into the desired composition, with a circular area resembling an eye at the center, surrounded by buildings. She told us,

> "*At first, I just wanted to present the lake, ancient buildings, and the Canton Tower, but when I found it difficult to achieve, I changed my approach. I chose one element that could be extended and added some of my favorite elements. I think the good thing about AI is that it can stimulate your imagination, coming up with things that people might not think of. So, the further I went, the more I hoped it could be bolder. As a result, the later works almost resemble the universe, not just limited to my imagination.*"
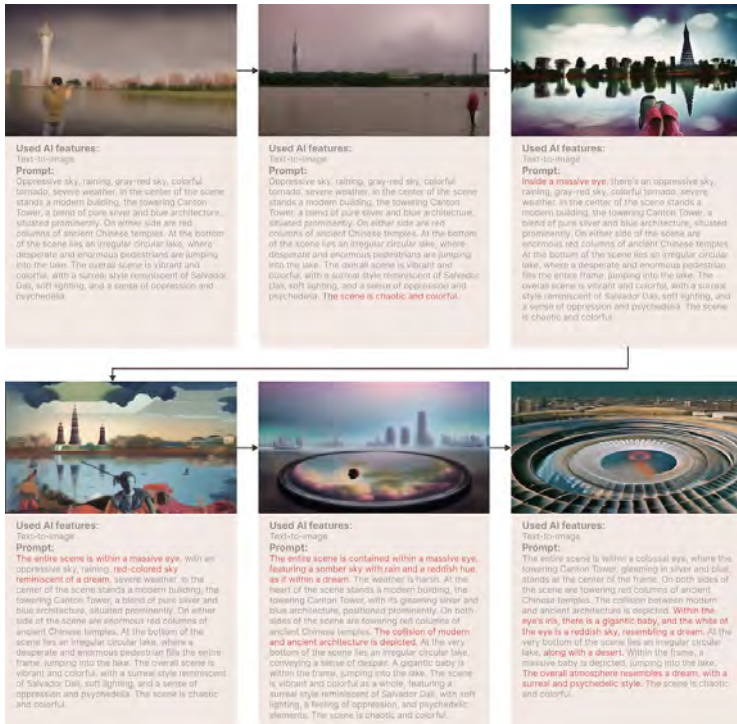

Fig. 3. P20's generated images and prompts. She gradually incorporated many surreal elements into her prompts.

4.2 Generative AI's Strengths in Assisting Personal Narrative Creation About Familiar Cultural Heritage

Through thematic analysis, we identified three types of assistance that AI provides to personal narratives: illuminating personal narratives, amplifying personal narratives, and reinterpreting personal narratives.

4.2.1 Illuminating Personal Narratives. AI can assist participants in constructing their personal narratives by iteratively generating images and gradually awakening their memories. Since our participants are experts in the cultural heritage they are familiar with, they were able to evaluate and correct inaccuracies in the generated images. However, they were not expert in storytelling, thus they initially did not have a clear anticipation of the visual outcome to be generated, only some scattered ideas (P1, P9, P11, P14). However, the rapid presentation of AI-generated results enabled participants to continuously experiment with new keyword combinations until achieving their desired effect. For example, P11 initially focused on describing a building but gradually incorporated diverse narrative elements, such as the Bible and priests, to enrich the storytelling aspects of the image (see Fig. 4). As she added more detailed descriptions in the prompts, such as "*priest in a red robe*" and "*a Bible on the table,*" the images evolved from merely showcasing the church building to depicting a narrative scene with a priest praying. Although the image did not perfectly capture details from the prompt, such as "*three worshippers,*" it still effectively enhanced the overall storytelling aspect that P11 intended. Another example is P1, who used the inpainting feature to continuously adjust the traffic volumes on the road (see Fig. 5). She initially used a brush to erase the road area and then input "*cars,*" but no cars appeared in the image. She then tried "*cars on the road*" to depict cars driving on the road, which successfully generated cars. Afterwards, she added "*bus*" to the prompt, but the resulting image had too much traffic, making it feel crowded. Ultimately, she chose the second generated result. This process illustrates how stable diffusion quickly provided P1 with various results, helping her find the desired visual effect.

Some participants' memory of the cultural heritage is vague, making it challenging for them to recall specific details when generating images initially without any references. In the process, The AI-generated images concretely portrayed the blurred images in participants' minds, enabling them to compare these visuals with their own memories. Interestingly, some of the AI's randomly generated unexpected elements occasionally prompt them to recall additional details and emotions associated with the cultural heritage (P1, P14, P17, P19). For example, P17 told us that the generated image truly evoked her childhood fear of that sculpture (see fig.6):

> "*At first, I said it was 'really, really scary' when I saw it as a child, but now, it's just three words for me. However, when I was a child, those three words felt like the world was about to end. During the generation process, for instance, I added the keyword 'thrilling.' As a result, the imagery became more abstract and distorted. Many bizarre and grotesque figures emerged in the picture, and I really felt that those three words, 'really, really scary,' came alive.*"

4.2.2 Amplifying Personal Narratives. AI amplify personal narratives by introducing supplementary narrative elements or suggesting alternative visual representation, which, after a selection process, eventually become integrated into the final image. The generated images incorporated additional elements not initially specified by the participants in their keywords, occasionally unexpectedly enhancing the original narrative (P6, P10, P18). For example, in the image of the Qilou (a traditional arcade buildings found in Southern China) generated by P10, AI introduced numerous shops, products, and customers. While this depiction differed from her existing impression of the Qilou scene, it significantly enriched her narrative (see fig.7). P10 told us:
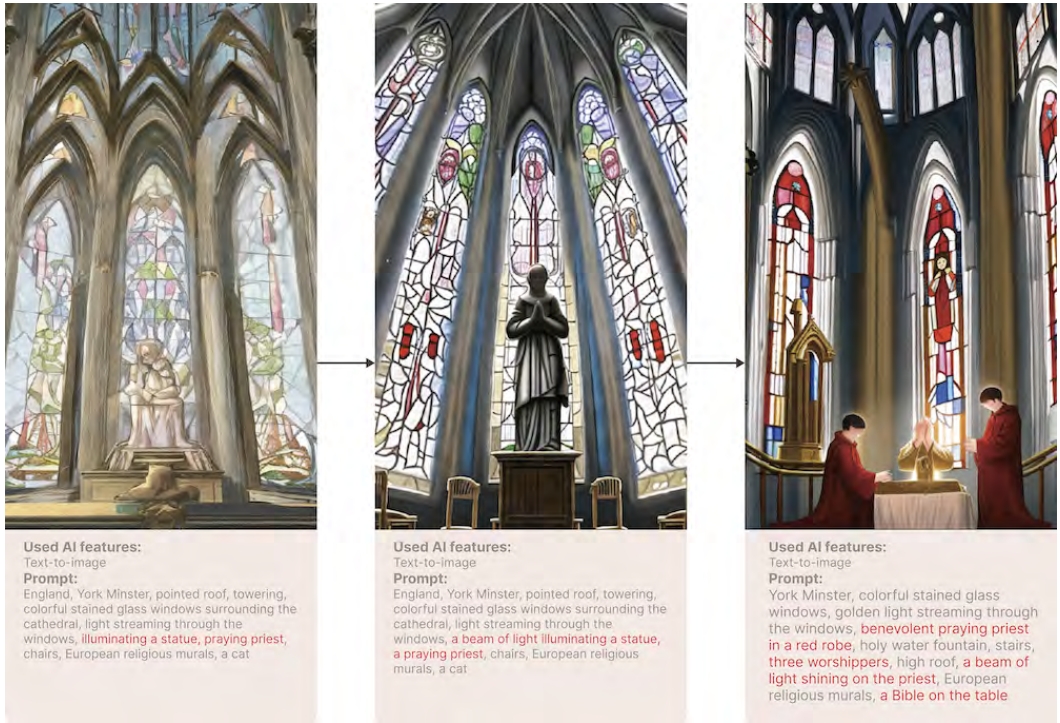
Used AI features:
Text-to-image
Prompt:
England, York Minster, pointed roof, towering,
colorful stained glass windows surrounding the
cathedral, light streaming through the
windows, illuminating a statue, praying priest,
chairs, European religious murals, a cat

Used AI features:
Text-to-image
Prompt:
England, York Minster, pointed roof, towering,
colorful stained glass windows surrounding the
cathedral, light streaming through the
windows, a beam of light illuminating a statue,
a praying priest, chairs, European religious
murals, a cat

Used AI features:
Text-to-image
Prompt:
York Minster, colorful stained glass
windows, golden light streaming through
the windows, benevolent praying priest
in a red robe, holy water fountain, stairs,
three worshippers, high roof, a beam of
light shining on the priest, European
religious murals, a Bible on the table

Fig. 4. P11's generated images and prompts evolved from her initial goal of depicting a building to gradually incorporating various narrative elements.

Used AI features:
Text-to-image
Prompt:
Nanjing, a Ming Dynasty city wall, the
city wall surrounded by trees, people
walking on the city wall, vehicles
passing underneath the city wall, blue
sky and white clouds, overall bright
color tone, sunlight.

Used AI features:
Inpainting
Prompt:
cars

Ultimate

Used AI features:
Inpainting
Prompt:
cars on the road

Used AI features:
Inpainting
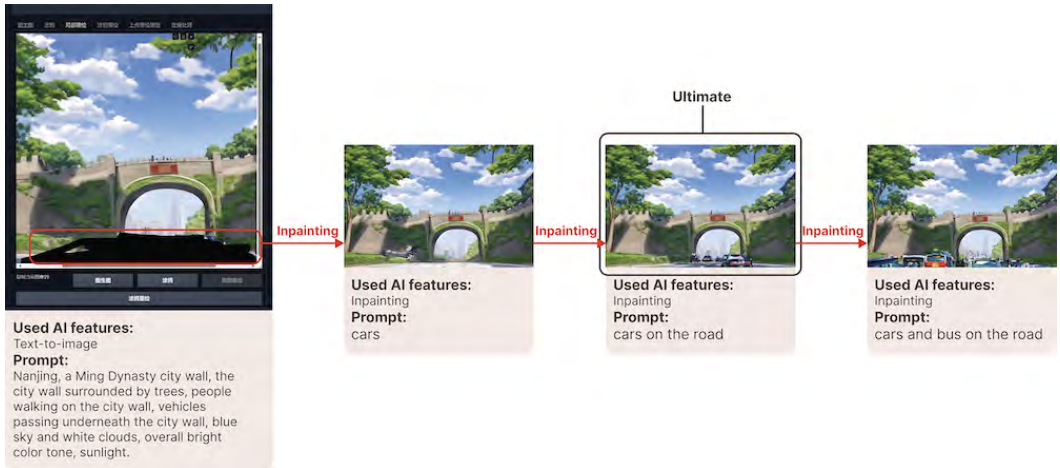**Prompt:**
cars and bus on the road

Fig. 5. P1's inpainting interface and the generated images. She tested a number of scenarios with different traffic volumes on the road and finally choose one from them.
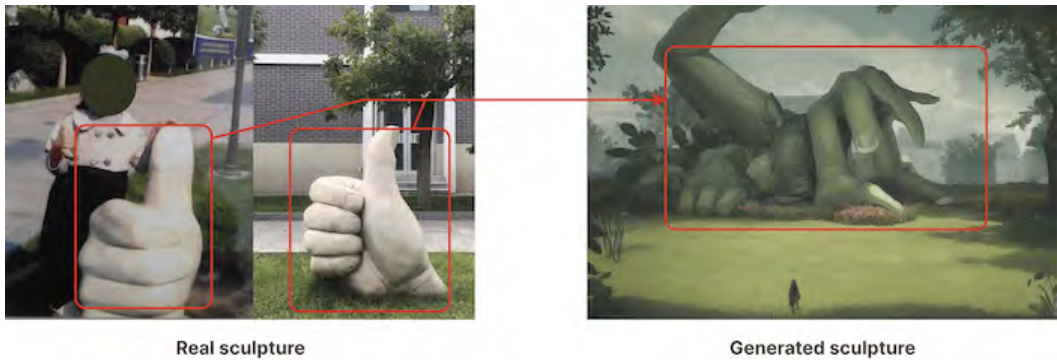
Fig. 6. The photograph of the authentic sculpture (provided by P17) is compared with the image generated by P17.

"*The generated images are somewhat different from my memories; they appear more vivid, and the entire scene seems crowded. In reality, the pathway wasn't as crowded. It gives the impression that there are many more shops and their products inside, making me feel happy and content walking through it. Because it used to be lively here, and now it feels deserted.*"

Fig. 7. The photograph of the authentic Qilou scene (sourced from the internet) and the image of Qilou generated by P10.

Additionally, AI might present an image with visual representations different from what the participant initially intended, but it more effectively captures their main idea (P1, P13, P15, P16, P17). For instance, P15 experimented with various descriptions in the prompt, such as "*a crowd of people riding bicycles*" and "*a father riding a bicycle carrying a little girl.*" Most of the images were depicted from a third-person perspective. However, one randomly generated image, which appeared to be from a first-person perspective, reminded her of how she used to see the clock tower while sitting on the back of her father's bicycle. She felt that this unexpected image better expressed her idea (see Fig. 8):

"*I've been struggling with depicting a scene of someone riding a bicycle. Initially, I wanted to show it from a third-person perspective. Later, I found that the current generated perspective actually resembles what I intended. It's as if I'm still sitting on the back seat of that bicycle, looking back at the scene of the building.*"
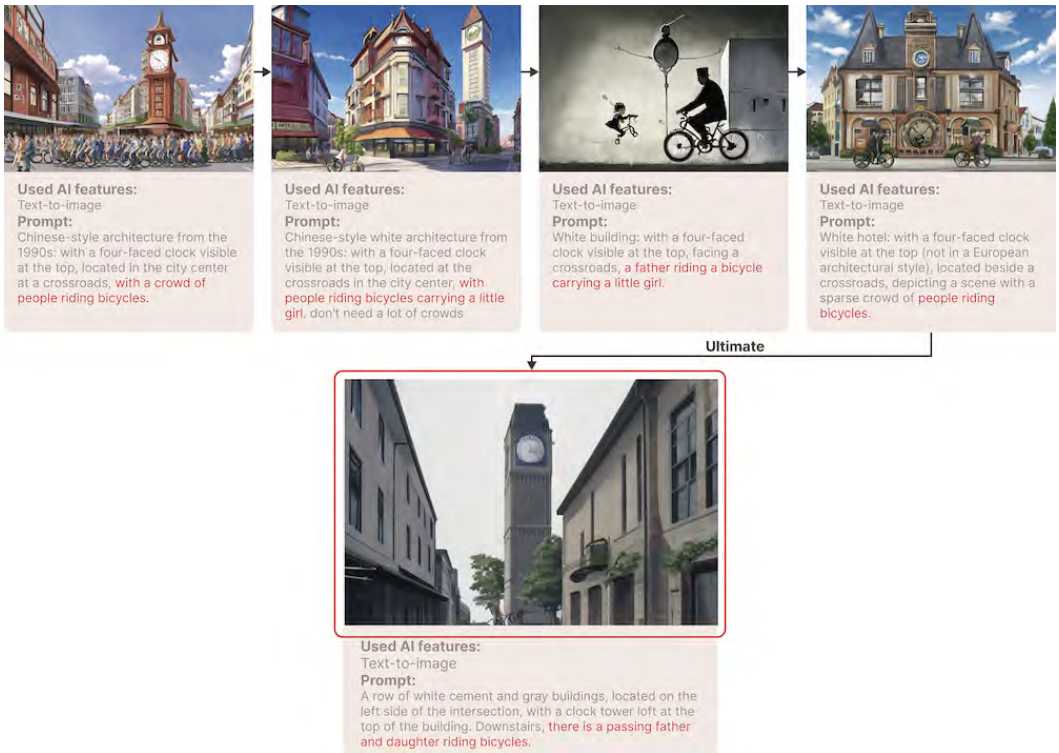
Fig. 8. The images of the clock tower scene generated by P15, she chose one that happened to be generated, resembling a first-person perspective.

4.2.3 Reinterpreting Personal Narratives. Some of the unconventional outcomes produced by AI appear thought-provoking to certain participants, prompting them to reconsider their narratives and, in turn, come up with new interpretation (P13, P18, P20). For instance, when AI misunderstood "*red moon*" and generated a "*red eyeball*" (see fig.9) it sparked new thoughts for P18:

> "*Usually, when we look at this painting, we perceive the scene from our own perspective. However, with an eyeball appearing on it, it feels like the scene is staring back at us, creating a shift in perspective.*"

The image of the deity generated by AI is very different from the one P13 had in mind (see fig.10), but he finds it interesting, he told us:

> "*What's kind of cool is how the AI shows the deity in different ways. Normally, we all have a fixed idea of what gods look like. They have a set appearance in our minds. But the AI comes up with different images, like how their face or clothes look. This made me wonder, why do we always picture deities in a specific way?*"

Overall, these findings indicated that AI can assist participants to varying degrees in completing their personal narratives about cultural heritage. It can help them clarify their narrative thoughts, enabling them to recall more details related to cultural heritage. AI can also suggest supplementary narrative elements and present ideas to help participants more comprehensively express their themes. Further, the uncertainty introduced by AI may even alter participants' original understanding, leading to the creation of new and more meaningful narratives.
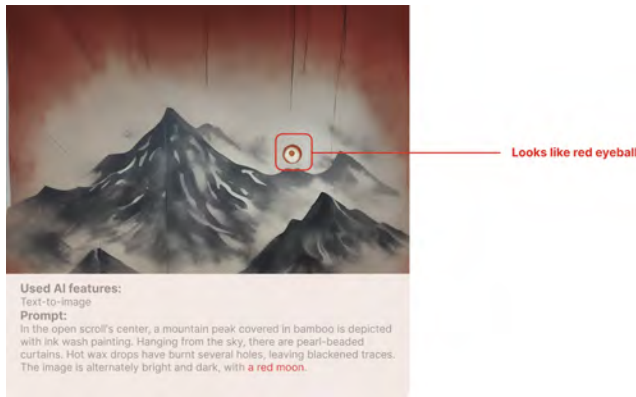
Fig. 9. In a image generated by P18, the red moon appears to resemble a red eyeball.

Fig. 10. The photograph of the deity (sourced from the internet) and the image of deity generated by P13.

4.3 Generative AI's Limitations in Personal Narrative Creation About Familiar Cultural Heritage

Generative AI presents limitations in familiar cultural heritage narrative creation, including limitations in meeting detailed participant requirements, struggles in depicting specific cultural features, and cultural bias in favor of Western styles.

4.3.1 Generative AI's Limitation in Meeting Detailed Requirements for Familiar Cultural Heritage Narrative. The images generated by AI at times fall short of meeting participants' detailed requirements (P3, P15). Consequently, certain elements in the scenes may lack relevance to both the cultural heritage and the narrative intended by the participants, potentially resulting in an incomplete representation of their storytelling and leading to misunderstandings. For instance, when P3 attempted to describe a scene with five soldiers standing beside the monument, the AI consistently portrayed the soldiers standing on top of the monument (see fig.11). As P3 said:

> "*Even though I wrote that all five people are standing beside it, AI always insists on placing them on top of the monument, but originally, there wouldn't be anyone standing on it. It merges the people with the monument, thinking that I meant people are also part of the monument.*"

However, some participants became aware of this limitations of AI during the generation process, prompting them to modify their expectations and creative strategies (P5, P7, P11, P17, P20). For instance, P7 mentioned, "*After the first round of generating, I pretty much got the vibe, so I moved*
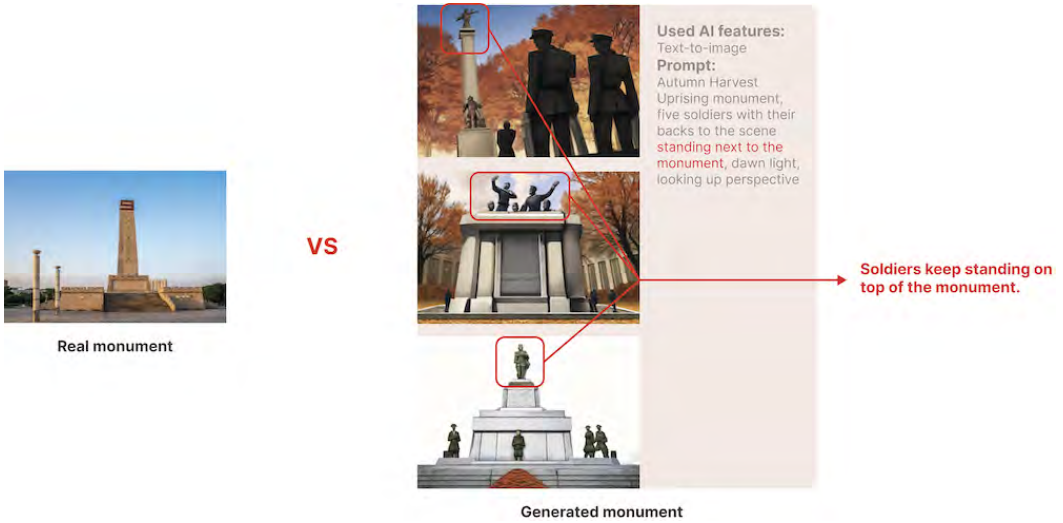
Fig. 11. The photograph of the monument (sourced from the internet) and P3's AI-generated image and prompt for the same monument.

on to the next step without making too many adjustment. Also, when I'm adjusting, I don't set my expectations too high." Recognizing both the weaknesses and strengths of AI, participants discovered that adjusting their generation strategies sometimes yielded unexpectedly positive outcomes. For instance, P17, after gaining a better understanding of AI's strengths, began incorporating abstract descriptions into her prompts, such as "*an overall atmosphere of suspense,*" to make the images more surreal and uncanny (see Fig.12). P17 told us:

> "*At first, I'd get pretty frustrated. I mean, the inputs I gave were so spot-on, and then it gives me these completely different images. Later on, I figured since it can't nail down a specific image, why not take a more abstract approach myself? I tried describing things in a different way, turning the real-world look into more of an imagined or abstract feel. As we gradually got in sync, I started feeling like it's a kind of relationship adjustment between me and the machine. Now, I'm quite satisfied with the current image. If we keep refining, there might even be a chance to generate an even better result.*"
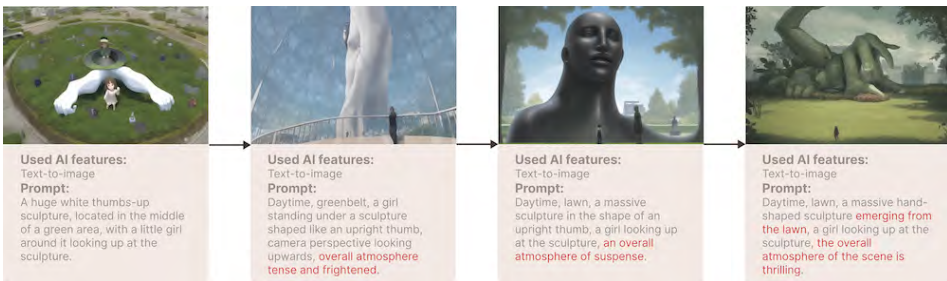

Fig. 12. The AI-generated images by P17 and the accompanying input prompts. She gradually added more abstract descriptions within the prompts.

4.3.2   Generative AI's Limitation in Presenting Specific Cultural Features. Cultural heritages encompass unique cultural forms, and participants were particularly keen on accentuating those distinctively culturally significant features in the images. However, AI faced challenges in comprehending or portraying those specific cultural images due to a lack of relevant datasets. This inadequacy led participants to feel that the images lacked crucial cultural elements (P7, P13, P16, P17, P20). For instance, P7 attempted to use various words to describe the image of "Tangqiu" (a traditional Chinese food of sugar balls). She tried terms like "*sugar ball,*" "*gobstopper,*" and "*jawbreaker,*" but AI consistently failed to depict the distinctive appearance of Tangqiu as envisioned by her (see fig.13). P7 told us:

> "*The most significant issue is with the key term 'Tangqiu'. When describing the Tangqiu gathering, it's essential to highlight the main aspect of 'Tangqiu'. However, by the end of the generation process, the image of 'Tangqiu' still did not come across as well as I had hoped in my mind.*"

Fig. 13. The photograph of "Tangqiu" (sourced from the internet), along with the AI-generated images and prompts created by P7 for "Tangqiu".

4.3.3   Generative AI's Cultural Bias. We have observed that AI tends to generate buildings with a predominantly Western style, often diverging significantly from the traditional architectural aesthetics found in China. For example, P15 noted that the buildings in the generated images mostly lean towards a European style, lacking the sense of historical periods and local characteristics inherent in Chinese architecture (see fig.14). She commented,

> "*The pictures we generated along the way mostly look like European-style buildings, which is totally different from what I remember. I recall there being modern buildings with some*

era-specific and local characteristics. Although these buildings in my hometown doesn't have that ancient history spanning dynasties, it has witnessed changes over the past thirty years or so."
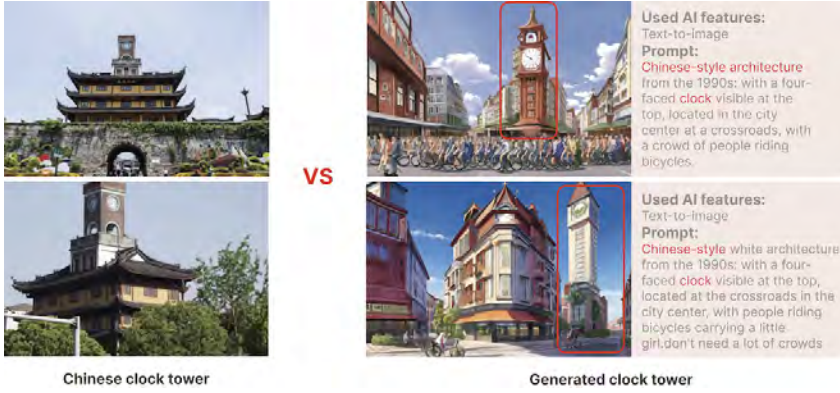

Fig. 14. The photograph of the Chinese clock tower (sourced from the internet) and the AI-generated clock tower by P15.

In addition, AI often relies on stereotypical depictions of Chinese buildings and characters. For instance, when P7 described a unique and charming Chinese temple street for the "Tangqiu" festival, the resulting architecture in image (see fig.15) gave her the impression that "*The buildings it's creating lean more towards a flashy, colorful vibe, not quite that serene charm you'd find in a temple. It's more like something you'd see in a Chinatown.*" Moreover, extending beyond architecture, when P7 tried to use the inpainting feature to generate an image of a young girl placed on the street, the AI defaulted to creating a girl with a Western appearance when she input "*a little girl.*" Even after modifying the prompt to "*Chinese little girl,*" the generated characters still retained Western features (see Fig. 16).

Fig. 15. The photograph of the real Chinese street (sourced from the internet) and P7's AI-generated image and prompt for the Chinese street. The AI-generated image specifically captures the style of a Western Chinatown.
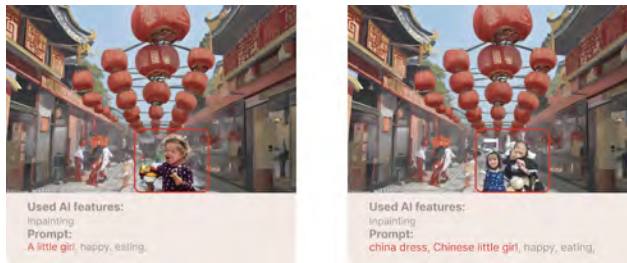
Fig. 16. In the images generated by P7, the young girls display characteristics typical of Western children.

4.4 The Impact of Lacking Local Knowledge on Using Generative AI to Create Images of Unfamiliar Cultural Heritages

To understand how people's expertise in local culture heritage influences their generative process and outcomes when collaborating with AI in storytelling, we had participants generate additional images based on provided information about a cultural heritage they were not familiar with. Our findings indicated that, owing to a lack of professional knowledge about the specific cultural heritage, participants were more inclined to accept the generated results, in contrast to their hesitancy in generating content related to familiar cultural heritage. Also, the generated images were more likely to exhibit bias, and participants' perception of the cultural heritage is more easily influenced by the generated images.

4.4.1 The Sequence of AI-Generated Images About Familiar and Unfamiliar Cultural Heritage. Figure 17 illustrates the sequence of output images generated by 20 participants, comparing the image sequences of familiar and unfamiliar cultural heritages. The red frames highlight images that were significantly altered by participants in terms of color, structure, and elements. The black arrows indicate the image sequences that contain more than twice as many images as the other set (familiar/unfamiliar cultural heritages). Within the context of familiar cultural heritages, participants consistently adjusted output images. Notably, P3, P7, P9, P15, made substantial adjustments to the generation process of familiar cultural heritages. However, only P20 made a substantial adjustments in the generation of unfamiliar cultural heritages, as evidenced by the red frame in Figure 17 . Furthermore, participants P7, P13, and P18 produced twice as many images of familiar cultural heritages compared to unfamiliar ones, as indicated by the black arrows in Figure 17 . This behavior can be attributed to their in-depth familiarity with these heritages, leading them to continuously modify details of the output images to align with their memories. In contrast, when generating images of unfamiliar cultural heritages, participants provided general descriptions and made fewer modifications to the output images. Due to a lack of knowledge about these cultural heritages, participants exhibited a higher level of acceptance towards the details generated by AI.

4.4.2 Participants' Creative Strategies for Unfamiliar Cultural Heritage. When describing unfamiliar cultural heritage, participants depicted the architectural features they see in the available materials (see fig.18). For example, P1 aimed to convey the appearance of "*a fruit market located among skyscrapers*" based on the information seen in the images. She mentioned, *"I think it matches the characteristics of this fruit market. Although subjectively, I don't like this composition, objectively, I think it expresses what it wants to express. It is a run-down wholesale market within the city of Hong Kong"* Additionally, participants often incorporate impressions of the place formed through other channels. For instance, P5 integrated the "*cyberpunk*" impression of Hong Kong formed from

Fig. 17. Sequence of output images generated by 20 participants

movies into his creation. He mentioned, "*The second picture resembles many of those scenes in sci-fi movies,. It feels like, even though it's crowded, it could be an extreme form of a cyberpunk future.*"

4.4.3 Biases in Participants' AI-Generated Images of Unfamiliar Cultural Heritage. When generating images related to unfamiliar cultural heritage, participants frequently incorporated their generalized impressions of a place (P5, P6, and P12), which may deviate from reality. As a result, participants might inadvertently introduced their own stereotypes into the images. For instance, when P5 described Central Mid-levels in Hong Kong, a place he had never visited, he infused it with "*cyberpunk*" features and believed that the generated image aligned with his perception of Hong Kong as "*crowded*" (see fig.19). He mentioned:

> "*I've been to Hong Kong a few times, and the vibe there is quite stifling, humid, and crowded. The first generated image, even though it has this cyberpunk and sci-fi feel, still captures that crowded essence. I use the term 'cyberpunk' because I've seen a lot of sci-fi movies and shows that are set in Hong Kong and Tokyo.*"

Additionally, participants' lack of familiarity with the location hinders their ability to discern potential cultural biases within the generated images, making it challenging for them to identify and modify such biases. For example, in the cultural heritage related to the fruit market site that

Fig. 18. The photograph of the Fruit Market and the Central Mid-levels in Hong Kong (sourced from the internet) and P1 and P5's AI-generated image and prompt for the same place.

Fig. 19. The photograph of the Central Mid-levels in Hong Kong (sourced from the internet and P5's AI-generated image and prompt for the same place. P5 incorporated "cyberpunk" elements into the image.

we provided, the most typical architectural feature is the rooftop adorned with old-style three-dimensional calligraphy on stone signs. However, during her creative process, P1 failed to notice that the buildings in the image lacked this distinctive cultural feature, appearing too commonplace (see fig.20). Furthermore, her perception of the place upon seeing the image was "*I think it's kind of like an urban village where young people have probably gone out to work, and there might be some elderly folks here, setting up stalls to sell fruits.*", while, in reality, this fruit market serves as a local central hub for the import of fruits from around the world. Many vendors engage in trading and transportation here, and it is not a backward urban village.

Due to their limited knowledge of the location, the generated images have the potential to influence and shape their impressions of the place, possibly introducing cultural biases. As expressed by P1, "*Although I haven't been there, I feel like when creating it, it refreshed my impression of it. Maybe the reality is not exactly like this, but my impression might be mixed with the images I generated and the images I've seen before, blending into a fusion of memories.*"

old-style three-dimensional calligraphy on stone signs
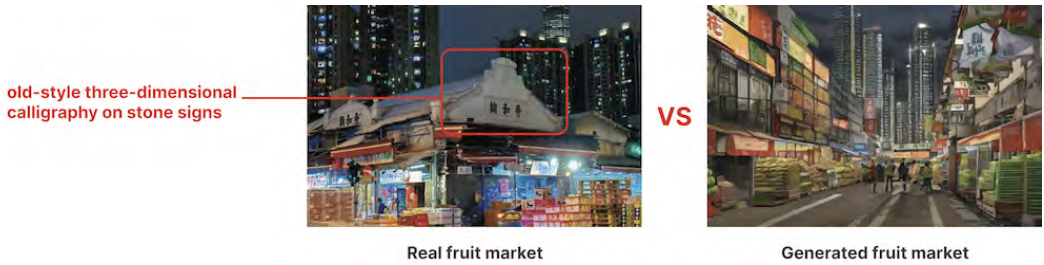
VS

Real fruit market

Generated fruit market

Fig. 20. The photograph of the fruit market (sourced from the internet) and P1's AI-generated image for the same place. The most typical architectural feature of the fruit market is the rooftop adorned with old-style three-dimensional calligraphy on stone signs.

5 DISCUSSION

5.1 The Impact of Generative AI on Personal Narrative Creation for Familiar Cultural Heritages

Our findings revealed that participants held diverse perspectives and memories related to various heritage sites, including those deemed of relatively minor historical significance by heritage professionals. The collaboration with AI facilitated participants in visually presenting their experiences, memories, and understanding of familiar cultural heritage. These personal narratives may contribute to the dynamic historical information of cultural heritage and amplifying individual voices distinct from official institutions within the space [65], offering fresh perspectives distinct from those presented in AHD [3, 29].

The creation of human-AI co-generated images involves an interactive collaboration between humans and AI. When participants attempted to express their understanding and emotions concerning cultural heritage, they often initially held only abstract concepts and fragmented memories. Generative AI can still generate complete images from vague keywords, allowing individuals to refine their ideas based on the output images [16, 41, 75]. As demonstrated in section 4.2.1 of the results, participants' memories related to cultural heritage become increasingly vivid through continuous iterations [9] and the clarity of their themes progressively improved. Additionally, given that participants were not necessarily experts in narrative and visual communication, they often struggled to effectively convey their ideas through images [76]. However, as demonstrated in Section 4.2.1, generative AI provided participants with feedback and creative insights, assisting them in refining their images [7, 71]. This helped them find better ways to express their cultural narratives through images.

However, generative AI often generates results that do not align with participants' expectations during the generation process. Some mistakes may be attributed to the inherent limitations of generative AI, but participants, unaware of these limitations beforehand, may spend a significant amount of time without achieving the desired results. A potential solution to this issue is to provide participants with information of the reasons behind each generation mistake, as suggested by Jeung et al. [26]. This might help them to gain a better understanding of the capabilities and limitations of AI in generating diverse content. Our results indicated that participants, such as P17, who possessed this understanding, can promptly adjust their creative expectations and strategies. Another approach is to employ prompt engineering to test and optimize prompts, thereby enhancing the effectiveness and controllability of Stable Diffusion [31, 38]. This enables participants to input more precise prompts and achieve desired results. Additionally, reinforcement learning from human feedback can be used to fine-tune stable diffusion [53, 54], better aligning with people's preferences

and requirements when creating cultural heritage narratives. These methods can serve as potential solutions to reduce the challenges faced by individuals unfamiliar with generative AI tools, thus increasing their accessibility to cultural narrative creation.

Additionally, it is important to note that in this workshop, we asked participants to engage in narrative through image creation. However, if they were required to supplement with other narrative elements, such as incorporating a written story, their creative results might differ. For instance, they might use new types of creative strategies, and their images might include more narrative elements instead of raw depictions.

5.2 Generative AI Should Serve as a Suggestion Provider, Not a Controller

The AI-generated outcomes contained uncertainties that introduced unexpected elements into the images. Sometimes these uncertainties provided valuable insights for participants' creations, such as the "first-person perspective scene" for P15 and the "red eyeball" for P18, but other times they were beyond the control of participants. As noted in Section 4.4.2 of the results, when creating images related to unfamiliar cultural heritage, participants lacking relevant knowledge were unable to identify mistakes related to the unfamiliar cultural heritage, which could result in AI leading the narrative process and leaving many erroneous uncertainties in the final narrative [16]. However, when participants generated images related to familiar cultural heritage, as indicated in Section 4.4.1, they were able to continuously evaluate and select AI-generated results based on their own experiences and knowledge. This demonstrates that a foundational understanding of cultural heritage[40] can help mitigate mistakes in the images to some extent.

Nonetheless, we must acknowledge that people's understanding of familiar cultural heritage can still contain subjective misconceptions or stereotypes. Therefore, to reduce inaccuracies introduced by both AI and individuals in the output images, it is beneficial to provide individuals with more objective references, such as actual photographs and official explanations of the cultural heritage. Additionally, collaborating with officials or staff in cultural heritage can help correct discrepancies between the images and objective facts.

5.3 AI Exhibits Cultural Bias While Also Offering Cultural Inspiration

To begin with, we need to contextualize this work within Chinese populations, as national identity may influence how narratives are generated. Chinese individuals generally hold a strong recognition of their culture [73] and want to maintain its independence and uniqueness within the global context. Consequently, when creating narratives, they often emphasized accurately conveying Chinese styles. For example, when P7 generated the street scene and P15 generated the clock tower, both emphasized Chinese-style features. They criticized that AI often falls short in capturing the unique features of Chinese cultural elements within the cultural heritage they aim to emphasize in their narratives. This includes AI-generated results with obvious Western rather than Chinese architectural characteristics and some that are filled with Western stereotypes of Chinese architecture. The inadequacy of AI's capabilities in this regard hinders the expression of these crucial and unique cultural aspects and exhibits stereotypes and inaccuracies in understanding identity descriptions related to specific cultures [4, 77].

Surprisingly, however, we found that people's attitudes toward these cultural biases were not entirely negative. The bias introduced by AI also offered a fresh perspective, especially for participants familiar with the cultural heritage. Engaging with the AI-enabled process allowed participants to perceive their familiar cultural heritage from a new perspective and understand diverse interpretations of the same elements from other cultures [25], such as P13 reflecting on the appearance features of deities from different cultures. This also implies that participants may have the opportunity to engage in critical reflection about their potential cultural biases towards other cultures

through the biases present in the generated images related to their own culture. We hypothesize that being able to view cultural elements from the perspective of other cultures may enhance people's awareness of cultural diversity [68].

To address this challenge while enhancing its benefits, it is advisable to gather information about the unique cultural features of specific cultural heritages from communities or official sources. This ensures the curation of appropriate datasets for training more accurate models [37]. Additionally, employing techniques such as prompt engineering and fine-tuning can enable the expression of different cultural characteristics more accurately through prompts. Continuously integrating individual narratives into the model can also help diversify the datasets, ensuring they reflect a broader range of voices, rather than being primarily shaped by influential or popular content driven by power.

5.4 Limitations

We have identified several limitations in this study. Firstly, we acknowledge that professional and cultural backgrounds may influence participants' use of narrative approaches, an aspect that we have not extensively explored. Future research can delve into the analysis of narrative approaches among individuals from diverse backgrounds, such as Eastern and Western backgrounds, to comprehend variations in AI narrative strategies across different cultural contexts. Additionally, our participants were primarily aged between 18-35 years and most had good proficiency with electronic devices. However, older individuals, who have rich narratives about cultural heritage, may not be able to learn how to use generative AI tools as quickly as the current participants. Future research could focus on how older age groups use AI tools to tell stories about cultural heritage, the strategies they employ, and the difficulties they encounter.

Secondly, the workshop format may not have fully allowed participants to openly discuss their familiar cultural heritage of interest. Utilizing alternative research methods, such as a series of workshops that provide more time for participants other functions of Stable Diffusion and refine their personal narratives, or incorporating gamified approaches to modify their collaborative interactions with generative AI, may yield different findings.

Thirdly, our choice of Stable Diffusion as the AI generation tool may have influence on the creation process. Other generative AI tools, like Midjourney, could potentially have a different impact on participants' creation processes. For instance, Midjourney might be more user-friendly for novice users compared to Stable Diffusion, potentially altering their creation workflow. Additionally, using a multilingual dataset such as LAION to allow participants to input in their native language, as well as selecting a model tuned for landscapes to depict tangible cultural heritage, could also influence participants' creations.

Lastly, the unfamiliar cultural heritages provided to the participants were all from Hong Kong, which might have influenced their creation process due to their previous impressions of the region. Future research could consider diversifying the selection of unfamiliar cultural heritages to explore how this impacts participants' creation strategies.

6 CONCLUSION

We conducted workshops with 20 participants to explore their creative strategies using generative AI for creating narratives about familiar cultural heritages, as well as to evaluate both the strengths and limitations of generative AI in this context. We also examined how participants' familiarity with cultural sites influenced their creation by comparing the results of familiar and unfamiliar cultural heritages. The results showed three distinct narrative strategies adopted by participants. While generative AI proved beneficial in illuminating, amplifying, and reinterpreting personal

narratives, its inherent cultural biases and uncertainties also posed challenges, particularly when participants lacked specific knowledge about the cultural heritages involved.

To address uncertainties and errors in AI-generated content, we recommend providing participants with detailed explanations for each generation mistake, utilizing prompt engineering, and applying reinforcement learning from human feedback to fine-tune Stable Diffusion. Moreover, to mitigate potential inaccuracies stemming from participants' inability to recognize AI-generated errors or subjective misconceptions, it is crucial to supplement their creative process with objective references such as authentic photographs and official descriptions of cultural heritages. Additionally, to enhance AI's ability to accurately capture and portray unique cultural elements, we propose the development of curated datasets tailored to specific cultural contexts for training more precise AI models.

Regarding the content of creation, workshop participants were asked to create representations of tangible cultural heritage. Future research may explore whether people achieve different results and creative strategies when using AI to describe intangible heritage-related themes that have a greater process component to the data material. Additionally, the participants' narrative expressions were limited to images. If we allow them to describe a story to assist in the narration of the images, their creation strategies might change.

REFERENCES

[1] Genevieve Alelis. 2013. Exhibiting emotion: using digital technologies to discover emotional connections. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*. Association for Computing Machinery, New York, NY, USA, 1017–1022. https://doi.org/10.1145/2468356.2468538

[2] Kate Allen, Nic Hollinworth, Faustina Hwang, Andy Minnion, Gosia Kwiatkowska, Ticky Lowe, and Nick Weldin. 2013. Interactive sensory objects for improving access to heritage. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*. Association for Computing Machinery, New York, NY, USA, 2899–2902. https://doi.org/10.1145/2468356.2479569

[3] Paulo Bala, Pedro Sanches, Vanessa Cesário, Sarah Leão, Catarina Rodrigues, Nuno Jardim Nunes, and Valentina Nisi. 2023. Towards Critical Heritage in the wild: Analysing Discomfort through Collaborative Autoethnography. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–19. https://doi.org/10.1145/3544548.3581274

[4] Cynthia L. Bennett, Cole Gleason, Morgan Klaus Scheuerman, Jeffrey P. Bigham, Anhong Guo, and Alexandra To. 2021. "It's Complicated": Negotiating Accessibility and (Mis)Representation in Image Descriptions of Race, Gender, and Disability. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–19. https://doi.org/10.1145/3411764.3445498

[5] Michael S. Bernstein, Joon Sung Park, Meredith Ringel Morris, Saleema Amershi, Lydia B Chilton, and Mitchell L. Gordon. 2023. Architecting Novel Interactions with Generative AI Models. In *Adjunct Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. ACM, San Francisco CA USA, 1–3. https://doi.org/10.1145/3586182.3617431

[6] Baptiste Caramiaux and Sarah Fdili Alaoui. 2022. "Explorers of Unknown Planets": Practices and Politics of Artificial Intelligence in Visual Arts. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 477:1–477:24. https://doi.org/10.1145/3555578

[7] 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 (DIS '23)*. Association for Computing Machinery, New York, NY, USA, 1941–1954. https://doi.org/10.1145/3563657.3596001

[8] Luigina Ciolfi, Areti Damala, Eva Hornecker, Monika Lechner, Laura Maye, and Daniela Petrelli. 2015. Cultural heritage communities: technologies and challenges. In *Proceedings of the 7th International Conference on Communities and Technologies*. ACM, Limerick Ireland, 149–152. https://doi.org/10.1145/2768545.2768560

[9] Caroline Claisse, Daniela Petrelli, Luigina Ciolfi, Nick Dulake, Mark T. Marshall, and Abigail C. Durrant. 2020. Crafting Critical Heritage Discourses into Interactive Exhibition Design. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376689

[10] Jackie Clarke and David Bowen. 2018. Familiar tourists, their behaviours and place attachments: an empirical framework. *Tourism Recreation Research* 43, 4 (Oct. 2018), 417–431. https://doi.org/10.1080/02508281.2018.1498158

[11] Jackie Clarke and David Bowen. 2021. Repeat tourists and familiar place formation: Conversion, inheritance and discovery. *Journal of Destination Marketing & Management* 20 (June 2021), 100605. https://doi.org/10.1016/j.jdmm.2021.100605

[12] Trien V. Do, Keith Cheverst, and Nick Taylor. 2015. Content analysis of a rural community's interaction with its cultural heritage through a longitudinal display deployment. In *Proceedings of the 2015 British HCI Conference*. ACM, Lincoln Lincolnshire United Kingdom, 46–55. https://doi.org/10.1145/2783446.2783567

[13] Stefania Druga, Fee Lia Christoph, and Amy J Ko. 2022. Family as a Third Space for AI Literacies: How do children and parents learn about AI together?. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3491102.3502031

[14] Karina Rodriguez Echavarria, Myrsini Samaroudi, Laurie Dibble, Edward Silverton, and Sophie Dixon. 2022. Creative Experiences for Engaging Communities with Cultural Heritage through Place-based Narratives. *Journal on Computing and Cultural Heritage* 15, 2 (June 2022), 1–19. https://doi.org/10.1145/3479007

[15] 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 *Proceedings of the 2024 ACM Designing Interactive Systems Conference*. 2833–2850.

[16] Jie Gao, Yuchen Guo, Toby Jia-Jun Li, and Simon Tangi Perrault. 2023. CollabCoder: A GPT-Powered WorkFlow for Collaborative Qualitative Analysis. In *Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing (CSCW '23 Companion)*. Association for Computing Machinery, New York, NY, USA, 354–357. https://doi.org/10.1145/3584931.3607500

[17] Elisa Giaccardi, Elizabeth Churchill, and Sophia Liu. 2012. Heritage matters: designing for current and future values through digital and social technologies. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems (CHI EA '12)*. Association for Computing Machinery, New York, NY, USA, 2783–2786. https://doi.org/10.1145/2212776.2212720

[18] Rebecca Gibson, Camilla May Babbage, Hanne Gesine Wagner, Dominic James Price, Sarah Martindale, Neil H Chadborn, Sachiyo Ito-Jaeger, Dimitrios Paris Darzentas, Helena Webb, Rachel Jacobs, Ayça Atabey, Boriana Koleva, Martin Flintham, Heidi Winklhofer, Lachlan D Urquhart, and Elvira Perez Vallejos. 2023. "It's not just for the Past but it's for the Here and Now": Gift-Giver Perspectives on the Memory Machine to Gift Digital Memories. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference (DIS '23)*. Association for Computing Machinery, New York, NY, USA, 987–1001. https://doi.org/10.1145/3563657.3595962

[19] Danilo Giglitto. 2015. Enhancing Community Heritage Empowerment through Wiki Software. In *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing*. ACM, Vancouver BC Canada, 85–88. https://doi.org/10.1145/2685553.2699327

[20] Danilo Giglitto, Luigina Ciolfi, Caroline Claisse, and Eleanor Lockley. 2019. Bridging cultural heritage and communities through digital technologies: Understanding perspectives and challenges. In *Proceedings of the 9th International Conference on Communities & Technologies-Transforming Communities*. 81–91.

[21] Qingyu Guo, Kangyu Yuan, Changyang He, Zhenhui Peng, and Xiaojuan Ma. 2024. Exploring the Evolvement of Artwork Descriptions in Online Creative Community under the Surge of Generative AI: A Case Study of DeviantArt. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–7. https://doi.org/10.1145/3613905.3650851

[22] 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 (CSCW '14)*. Association for Computing Machinery, New York, NY, USA, 1144–1155. https://doi.org/10.1145/2531602.2531640

[23] Yuanning Han, Ziyi Qiu, Jiale Cheng, and RAY LC. 2024. When Teams Embrace AI: Human Collaboration Strategies in Generative Prompting in a Creative Design Task. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3613904.3642133

[24] Peng-Kai Hung, Rung-Huei Liang, Shih-Yu Ma, and Bo-Wen Kong. 2024. Exploring the Experience of Traveling to Familiar Places in VR: An Empirical Study Using Google Earth VR. *International Journal of Human–Computer Interaction* 40, 2 (Jan. 2024), 255–277. https://doi.org/10.1080/10447318.2022.2114141 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10447318.2022.2114141.

[25] Myounghoon Jeon, Rebecca Fiebrink, Ernest A. Edmonds, and Damith Herath. 2019. From rituals to magic: Interactive art and HCI of the past, present, and future. *International Journal of Human-Computer Studies* 131 (Nov. 2019), 108–119. https://doi.org/10.1016/j.ijhcs.2019.06.005

[26] Jun Li Jeung and Janet Yi-Ching Huang. 2023. Correct Me If I Am Wrong: Exploring How AI Outputs Affect User Perception and Trust. In *Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing (CSCW '23 Companion)*. Association for Computing Machinery, New York, NY, USA, 323–327. https://doi.org/10.1145/3584931.3606997

[27] Siân Jones. 2017. Wrestling with the Social Value of Heritage: Problems, Dilemmas and Opportunities. *Journal of Community Archaeology & Heritage* 4, 1 (Jan. 2017), 21–37. https://doi.org/10.1080/20518196.2016.1193996

[28] Purnima Kamath, Zhuoyao Li, Chitralekha Gupta, Kokil Jaidka, Suranga Nanayakkara, and Lonce Wyse. 2023. Evaluating Descriptive Quality of AI-Generated Audio Using Image-Schemas. In *Proceedings of the 28th International Conference on Intelligent User Interfaces (IUI '23)*. Association for Computing Machinery, New York, NY, USA, 621–632. https://doi.org/10.1145/3581641.3584083

[29] Asnath Paula Kambunga, Heike Winschiers-Theophilus, and Rachel Charlotte Smith. 2020. Participatory Memory Making: Creating Postcolonial Dialogic Engagements with Namibian Youth. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference (DIS '20)*. Association for Computing Machinery, New York, NY, USA, 785–797. https://doi.org/10.1145/3357236.3395441

[30] Eunseo Kim, Jeongmin Hong, Hyuna Lee, and Minsam Ko. 2022. Colorbo: Envisioned Mandala Coloringthrough Human-AI Collaboration. In *27th International Conference on Intelligent User Interfaces (IUI '22)*. Association for Computing Machinery, New York, NY, USA, 15–26. https://doi.org/10.1145/3490099.3511135

[31] Jiwon Kim, Jiwon Kang, Taeeun Kim, Hayeon Song, and Jinyoung Han. 2023. AlphaDAPR: An AI-based Explainable Expert Support System for Art Therapy. In *Proceedings of the 28th International Conference on Intelligent User Interfaces (IUI '23)*. Association for Computing Machinery, New York, NY, USA, 19–31. https://doi.org/10.1145/3581641.3584087

[32] Tomas Lawton, Francisco J Ibarrola, Dan Ventura, and Kazjon Grace. 2023. Drawing with Reframer: Emergence and Control in Co-Creative AI. In *Proceedings of the 28th International Conference on Intelligent User Interfaces (IUI '23)*. Association for Computing Machinery, New York, NY, USA, 264–277. https://doi.org/10.1145/3581641.3584095

[33] RAY LC. 2022. Imitations of Immortality: Learning from Human Imitative Examples in Transformer Poetry Generation. In *Proceedings of the 10th International Conference on Digital and Interactive Arts (ARTECH '21)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3483529.3483537

[34] 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 (C&C '23)*. Association for Computing Machinery, New York, NY, USA, 217–222. https://doi.org/10.1145/3591196.3593341

[35] RAY LC. 2023. TOGETHER ENOUGH: Collaborative Constructions of Adaptations to Climate Futures. In *Companion Publication of the 2023 ACM Designing Interactive Systems Conference (DIS '23 Companion)*. Association for Computing Machinery, New York, NY, USA, 55–59. https://doi.org/10.1145/3563703.3596805

[36] RAY LC and Yuying Tang. 2024. Speculative Design with Generative AI: Applying Stable Diffusion and ChatGPT to imagining climate change futures. In *Proceedings of the 11th International Conference on Digital and Interactive Arts (ARTECH '23)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3632776.3632827

[37] Tianyi Li and Francisco Iacobelli. 2023. Purposeful AI. In *Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing (CSCW '23 Companion)*. Association for Computing Machinery, New York, NY, USA, 563–565. https://doi.org/10.1145/3584931.3606954

[38] Youjia Li, Jianjun Shi, and Zheng Zhang. 2024. An Approach for Rapid Source Code Development Based on ChatGPT and Prompt Engineering. *IEEE Access* 12 (2024), 53074–53087. https://doi.org/10.1109/ACCESS.2024.3385682

[39] Long Ling, Xinyi Chen, Ruoyu Wen, Toby Jia-Jun Li, and RAY LC. 2024. Sketchar: Supporting Character Design and Illustration Prototyping Using Generative AI. *Proc. ACM Hum.-Comput. Interact.* 8, CHI PLAY (Oct. 2024), 337:1–337:28. https://doi.org/10.1145/3677102

[40] Guanhong Liu, Xianghua(Sharon) Ding, Jinghe Cai, Weiyun Wang, Xinyue Wang, Yuting Diao, Jin Chen, Tianyu Yu, Haiqing Xu, and Haipeng Mi. 2023. Digital Making for Inheritance and Enlivening Intangible Cultural Heritage: A Case of Hairy Monkey Handicrafts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3544548.3581539

[41] Vivian Liu, Jo Vermeulen, George Fitzmaurice, and Justin Matejka. 2023. 3DALL-E: Integrating Text-to-Image AI in 3D Design Workflows. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference (DIS '23)*. Association for Computing Machinery, New York, NY, USA, 1955–1977. https://doi.org/10.1145/3563657.3596098

[42] Duri Long, Mikhail Jacob, and Brian Magerko. 2019. Designing Co-Creative AI for Public Spaces. In *Proceedings of the 2019 Conference on Creativity and Cognition (C&C '19)*. Association for Computing Machinery, New York, NY, USA, 271–284. https://doi.org/10.1145/3325480.3325504

[43] Ryan Louie, Andy Coenen, Cheng Zhi Huang, Michael Terry, and Carrie J. Cai. 2020. Novice-AI Music Co-Creation via AI-Steering Tools for Deep Generative Models. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. https://doi.org/10.1145/3313831.3376739

[44] Ryan Louie, Jesse Engel, and Cheng-Zhi Anna Huang. 2022. Expressive Communication: Evaluating Developments in Generative Models and Steering Interfaces for Music Creation. In *27th International Conference on Intelligent User Interfaces (IUI '22)*. Association for Computing Machinery, New York, NY, USA, 405–417. https://doi.org/10.1145/3490099.3511159

[45] Alex Jiahong Lu, Shruti Sannon, Cameron Moy, Savana Brewer, Jaye Green, Kisha N Jackson, Daivon Reeder, Camaria Wafer, Mark S. Ackerman, and Tawanna R Dillahunt. 2023. Participatory Noticing through Photovoice: Engaging Arts- and Community-Based Approaches in Design Research. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. ACM, Pittsburgh PA USA, 2489–2508. https://doi.org/10.1145/3563657.3596041

[46] Fei Lu, Feng Tian, Yingying Jiang, Xiang Cao, Wencan Luo, Guang Li, Xiaolong Zhang, Guozhong Dai, and Hongan Wang. 2011. ShadowStory: creative and collaborative digital storytelling inspired by cultural heritage. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 1919–1928. https://doi.org/10.1145/1978942.1979221

[47] Yao Lyu, He Zhang, Shuo Niu, and Jie Cai. 2024. A Preliminary Exploration of YouTubers' Use of Generative-AI in Content Creation. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–7. https://doi.org/10.1145/3613905.3651057

[48] Atefeh Mahdavi Goloujeh, Anne Sullivan, and Brian Magerko. 2024. The Social Construction of Generative AI Prompts. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–7. https://doi.org/10.1145/3613905.3650947

[49] Irini A. Malegiannaki, Thanasis Daradoumis, and Symeon Retalis. 2020. Teaching Cultural Heritage through a Narrative-based Game. *Journal on Computing and Cultural Heritage* 13, 4 (Dec. 2020), 1–28. https://doi.org/10.1145/3414833

[50] Michael Muller, Anna Kantosalo, Mary Lou Maher, Charles Patrick Martin, and Greg Walsh. 2024. GenAICHI 2024: Generative AI and HCI at CHI 2024. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–7. https://doi.org/10.1145/3613905.3636294

[51] Reese Muntean, Alissa N. Antle, Brendan Matkin, Kate Hennessy, Susan Rowley, and Jordan Wilson. 2017. Designing Cultural Values into Interaction. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 6062–6074. https://doi.org/10.1145/3025453.3025908

[52] Valentina Nisi, Paulo Bala, Hollie Bostock, Vanessa Cesário, and Nuno Jardim Nunes. 2023. "Before gentrification, we claim for habitation": Eliciting Values and Assets through Cultural Heritage Storytelling. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference (DIS '23)*. Association for Computing Machinery, New York, NY, USA, 2423–2436. https://doi.org/10.1145/3563657.3596124

[53] Khalid K. Osman, Florian Barbaro, and Andy Skumanich. 2023. Leveraging Social Media Analysis for Effective Water Management. In *Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing (CSCW '23 Companion)*. Association for Computing Machinery, New York, NY, USA, 289–292. https://doi.org/10.1145/3584931.3607003

[54] Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, and others. 2022. Training language models to follow instructions with human feedback. *Advances in neural information processing systems* 35 (2022), 27730–27744.

[55] Pat Pataranutaporn, Valdemar Danry, Lancelot Blanchard, Lavanay Thakral, Naoki Ohsugi, Pattie Maes, and Misha Sra. 2023. Living Memories: AI-Generated Characters as Digital Mementos. In *Proceedings of the 28th International Conference on Intelligent User Interfaces (IUI '23)*. Association for Computing Machinery, New York, NY, USA, 889–901. https://doi.org/10.1145/3581641.3584065

[56] Philip L. Pearce. 2012. The experience of visiting home and familiar places. *Annals of Tourism Research* 39, 2 (April 2012), 1024–1047. https://doi.org/10.1016/j.annals.2011.11.018

[57] Han Qiao, Vivian Liu, and Lydia Chilton. 2022. Initial Images: Using Image Prompts to Improve Subject Representation in Multimodal AI Generated Art. In *Proceedings of the 14th Conference on Creativity and Cognition (C&C '22)*. Association for Computing Machinery, New York, NY, USA, 15–28. https://doi.org/10.1145/3527927.3532792

[58] Jeba Rezwana and Mary Lou Maher. 2022. Understanding User Perceptions, Collaborative Experience and User Engagement in Different Human-AI Interaction Designs for Co-Creative Systems. In *Proceedings of the 14th Conference on Creativity and Cognition (C&C '22)*. Association for Computing Machinery, New York, NY, USA, 38–48. https://doi.org/10.1145/3527927.3532789

[59] 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.

[60] Marie-Monique Schaper and Narcis Pares. 2022. Enhancing Students' Social and Emotional Learning in Educational Virtual Heritage through Projective Augmented Reality. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems (CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3491101.3503551

[61] 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 (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3544548.3581414

[62] Klaas-Jan Stol, Paul Ralph, and Brian Fitzgerald. 2016. Grounded theory in software engineering research: a critical review and guidelines. In *Proceedings of the 38th International conference on software engineering*. 120–131.

[63] Minhyang (Mia) Suh, Emily Youngblom, Michael Terry, and Carrie J Cai. 2021. AI as Social Glue: Uncovering the Roles of Deep Generative AI during Social Music Composition. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3411764.3445219

[64] Georgios Trichopoulos. 2023. Large Language Models for Cultural Heritage. In *Proceedings of the 2nd International Conference of the ACM Greek SIGCHI Chapter*. ACM, Athens Greece, 1–5. https://doi.org/10.1145/3609987.3610018

[65] 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 (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376558

[66] Violeta Tsenova, Gavin Wood, and David Kirk. 2023. Loci Stories: Exploring Design for Polyvocality. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference (DIS '23)*. Association for Computing Machinery, New York, NY, USA, 15–30. https://doi.org/10.1145/3563657.3596114

[67] Mathias Peter Verheijden and Mathias Funk. 2023. Collaborative Diffusion: Boosting Designerly Co-Creation with Generative AI. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg Germany, 1–8. https://doi.org/10.1145/3544549.3585680

[68] Greg Walsh and Eric Wronsky. 2019. AI + Co-Design: Developing a Novel Computer-supported Approach to Inclusive Design. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing (CSCW '19)*. Association for Computing Machinery, New York, NY, USA, 408–412. https://doi.org/10.1145/3311957.3359456

[69] Justin D. Weisz, Jessica He, Michael Muller, Gabriela Hoefer, Rachel Miles, and Werner Geyer. 2024. Design Principles for Generative AI Applications. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–22. https://doi.org/10.1145/3613904.3642466

[70] Justin D. Weisz, Michael Muller, Stephanie Houde, John Richards, Steven I. Ross, Fernando Martinez, Mayank Agarwal, and Kartik Talamadupula. 2021. Perfection Not Required? Human-AI Partnerships in Code Translation. In *26th International Conference on Intelligent User Interfaces (IUI '21)*. Association for Computing Machinery, New York, NY, USA, 402–412. https://doi.org/10.1145/3397481.3450656

[71] Justin D. Weisz, Michael Muller, Steven I. Ross, Fernando Martinez, Stephanie Houde, Mayank Agarwal, Kartik Talamadupula, and John T. Richards. 2022. Better Together? An Evaluation of AI-Supported Code Translation. In *27th International Conference on Intelligent User Interfaces (IUI '22)*. Association for Computing Machinery, New York, NY, USA, 369–391. https://doi.org/10.1145/3490099.3511157

[72] Ruishan Wu, Chunlei Gong, Li Chen, Jiayi Su, and LC Ray. 2024. The Present in the Future is the Past: Applying Generative AI to Visualize and Imagine Cultural Heritage Sites in Both Augmented and Physical Reality. In *29th International Symposium on Electronic Art (ISEA 2024)*.

[73] Shuqin Xu. 2018. Cultivating national identity with traditional culture: China's experiences and paradoxes. *Discourse: Studies in the Cultural Politics of Education* 39, 4 (2018), 615–628.

[74] 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 IUI 2022 Workshops: APEx-UI, HAI-GEN, HEALTHI, HUMANIZE, TExSS, SOCIALIZE*. CEUR-WS Team, 56–65. https://scholars.cityu.edu.hk/en/publications/publication(d901f5a2-0600-422f-b588-db5a59871961).html

[75] Chengzhi Zhang, Weijie Wang, Paul Pangaro, Nikolas Martelaro, and Daragh Byrne. 2023. Generative Image AI Using Design Sketches as input: Opportunities and Challenges. In *Proceedings of the 15th Conference on Creativity and Cognition (C&C '23)*. Association for Computing Machinery, New York, NY, USA, 254–261. https://doi.org/10.1145/3591196.3596820

[76] Shichao Zhao. 2023. Involving British-Chinese Immigrants in Participatory Action Research: Lessons Learnt from the Field. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference (DIS '23)*. Association for Computing Machinery, New York, NY, USA, 45–60. https://doi.org/10.1145/3563657.3596107

[77] Jianlong Zhou, Heimo Müller, Andreas Holzinger, and Fang Chen. 2023. Ethical ChatGPT: Concerns, Challenges, and Commandments. https://doi.org/10.48550/ARXIV.2305.10646 Version Number: 1.

A INTERVIEW QUESTIONS

A.1 Pre-study interview question

Where are you from?

Can you share something about [specific topic or place] or related to your personal experiences? / Can you imagine some Hong Kong CH that you have never visited before?

How familiar are you with [specific region or area]? Can you provide some information about [specific topic or place]? Also, do you know why this place is called [chosen CH]?

What were your initial thoughts when you first entered these places?

What emotional connections do you have with these places?

A.2 Post-study interview question

Why do you have these thoughts? Can you introduce your final work?

What are the differences between talking about or imagining CH and generating CH in SD?

Please evaluate the final generated work. Which image do you like the most, and why? How does it relate to your thoughts or imagination?

During the generation process, did you have any new feelings or thoughts about CH?

After the generation, do you have any other feelings or thoughts?

Why did you write this prompt? What kind of result do you hope to get from SD?

When you first tried to generate images of CH, what came to your mind initially?

What surprises did SD bring you during the generation process?

What difficulties did you encounter during the generation process? What do you wish SD could do?

Do you think the results matched your expectations?

Why did you choose these specific [topic] words? (Is it because these places have certain characteristics in your memory?)

What else do you think the entire generation process brought to you?

What are the main differences between these two generation processes?