

# A Data-Driven Structured Prompt Framework for AI-Assisted Cultural Tourism Souvenir Design

Yi He<sup>1,2</sup>, Weijie Ling<sup>3</sup>, Zhonghang Bai<sup>4</sup>, Jinho Yim<sup>5\*</sup>

<sup>1</sup>Department of Smart Experience Design, Ph.D. Candidate, Graduate School of Techno Design, Kookmin University, Seoul, Korea

<sup>2</sup>School of Art, Lecturer, Hebei Agricultural University, Baoding, Hebei, China

<sup>3</sup>Department of Smart Experience Design, Ph.D., Graduate School of Techno Design, Kookmin University, Seoul, Korea

<sup>4</sup>School of Architecture and Art Design, Professor, Hebei University of Technology, Tianjin, China

<sup>5</sup>Department of Smart Experience Design, Professor, Graduate School of Techno Design, Kookmin University, Seoul, Korea

---

## Abstract

**Background** Artificial intelligence (AI) text-to-image systems are increasingly used in early-stage design, including cultural tourism souvenir development. In this context, however, designers often struggle to organize cultural references, product-related information, and visual descriptions clearly when translating design intent into prompts. This difficulty can lead to unstable outputs, theme drift, and weak alignment between intended cultural cues and generated images. This study aims to propose a structured, evidence-based method for organizing culture-related prompt content in ways that designers can specify and revise more clearly during iteration.

**Methods** We collected 2,460 prompt–image pairs from officially featured Midjourney posts related to cultural tourism souvenirs. Latent Dirichlet allocation topic modeling was applied to the prompt corpus, and a Delphi-based expert card-sorting procedure was used to consolidate the results into five design themes. Based on these themes, we developed five topic-based structured prompt templates that specify key elements such as souvenir type, cultural references, visual style, and related descriptive content. These templates were then summarized into a stepwise workflow for structured prompt construction and iterative refinement. The framework was examined in a one-day design workshop. Twenty-two designers used the structured prompts to generate 110 souvenir concepts. Ten experts rated how clearly each structural element was expressed in the outputs, and semi-structured interviews were conducted to capture designers' usage experiences.

**Results** Five cultural tourism souvenir design themes were extracted from the curated prompt–image dataset and translated into five topic-specific structured prompt templates, together with a corresponding workflow for structured prompt construction and iterative refinement. The workshop results showed clearer expression of concrete and visually grounded elements such as souvenir type, color, and visual style. Designers also reported that the templates reduced uncertainty at the beginning of prompting and made it easier to revise specific elements during iteration. At the same time, limitations remained for abstract narratives and cross-cultural semantic integration. In such cases, the generated images sometimes showed fragmented cultural cues or only superficial combinations of symbols.

**Conclusions** This study proposes a data-driven structured prompting framework for cultural tourism souvenir ideation and provides empirical evidence from expert ratings and designer interviews. The framework is most useful for organizing cultural references and supporting the specification and revision of concrete visual elements during early-stage ideation, while also clarifying its current limits in narrative meaning and cross-cultural semantic integration. These findings provide a practical reference for prompt design and human–AI collaboration in culturally oriented design tasks.

**Keywords** Cultural Tourism Souvenirs, Souvenir Design, Text-to-Image Generation, Data-Driven Workflow, Structured Prompting

---

\*Corresponding author: Jinho Yim (hci.yim@kookmin.ac.kr)

*Citation:* He, Y., Ling, W., Bai, Z., & Yim, J. (2026). A Data-Driven Structured Prompt Framework for AI-Assisted Cultural Tourism Souvenir Design. *Archives of Design Research*, 39(2), 37-75.

<http://dx.doi.org/10.15187/adr.2026.05.39.2.37>

**Received :** Feb. 23.  
2026 ; **Reviewed :** Apr.  
13. 2026 ; **Accepted :**  
Apr. 23. 2026  
**pISSN** 1226-8046  
**eISSN** 2288-2987

**Copyright :** This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted educational and non-commercial use, provided the original work is properly cited.

---

## 1. Introduction

In recent years, the application of artificial intelligence (AI) in creative design has expanded rapidly, reshaping how cultural products are conceived, developed, and evaluated (El Abed & Castro-Lopez, 2024; McComb & Jablokow, 2022). Among these developments, text-to-image generation tools—such as Midjourney—have demonstrated strong visual synthesis capabilities, producing aesthetically compelling images based on natural language prompts (Oppenlaender, 2024). In cultural tourism, design tasks often require the integration of local identity, narrative depth, and artistic appeal (Duan et al., 2023; Zhu & Rahman, 2025). Against this background, AI-driven tools have opened up new possibilities for cultural tourism souvenir design, particularly in terms of rapid visualization, stylistic exploration, and the iteration of multiple design options during the early stages of concept generation. Despite recent advances in text-to-image models, prompt writing in design practice is still largely experience-driven. Designers often craft prompts through trial and error, which can lead to unstable results, theme drift, and uneven visual quality (Don-Yehiya et al., 2023; Mahdavi Goloujeh et al., 2024). This issue is especially pronounced in cultural tourism souvenir design, because such tasks require not only visually appealing outputs, but also a relatively clear organization of cultural references, product-related information, and narrative cues in the development of design concepts (Zhan et al., 2024).

Recent prompt-engineering research suggests that structure matters: how information is organized in a prompt can substantially influence perceived image quality, controllability, and semantic alignment (Oppenlaender, 2024; Zhan et al., 2024). Yet, in domain-specific design work—particularly cultural souvenir creation—structured prompt modeling is still limited. In practice, a key open problem is how to derive usable prompt structures from large-scale, high-quality AI design practices and then examine how these structures perform in real design tasks.

To address this gap, this study asks two research questions:

1. How can a data-driven structured prompt modeling method be constructed for AI-assisted cultural tourism souvenir design to enable clearer organization of cultural references and related narrative cues, more explicit specification of concrete visual elements, and more targeted prompt revision during early-stage concept generation?
2. To what extent does the proposed method support designers in producing cultural tourism souvenir concepts, and how do designers perceive and evaluate the method in hands-on use?

---

## 2. Literature Review

### 2. 1. The Role of Generative AI in Cultural Tourism Souvenir Design

Cultural tourism is more than visiting a destination. It also involves how people interpret local identity, local stories, and symbolic narratives that are tied to place (Qiu et al., 2024; Richards, 2018). Souvenirs sit at the center of this meaning-making process because they provide a tangible way to “carry” place-related meanings. They help visitors keep an emotional and narrative connection to a destination, and they also make those meanings shareable with others after the trip (Littrell et al., 1993; Vecco, 2010; Zhu & Rahman, 2025).

In recent years, text-to-image platforms (e.g., Midjourney, DALL-E, and Stable Diffusion) have started to enter souvenir concept development. These systems can translate textual descriptions into visually rich outputs, which supports rapid prototyping and broad stylistic exploration for culturally inspired products (Bommasani et al., 2021; Ramesh et al., 2022). The immediate value is faster prototyping and easier style trials for culturally inspired products in tourism-related design tasks (Gîrbacia, 2024; Lyu et al., 2024). In practice, designers use these systems to quickly compare multiple product directions and narrative visuals during early ideation (Luther et al., 2024; Mahdavi Goloujeh et al., 2024; Rapp et al., 2025). Compared with traditional visual development workflows, the practical benefit is straightforward: it becomes easier to iterate on regional symbols, historical motifs, and aesthetic variations across multiple directions, without committing to a single visual path too early.

However, cultural design is sensitive to small semantic shifts. When prompts are vague or under-specified, generated results may drift away from the intended meaning, produce incoherent symbolism, or miss contextual accuracy (Don-Yehiya et al., 2023; Zhan et al., 2024). In souvenir design—where outputs are expected to evoke place, memory, and local culture—prompt formulation therefore acts as a central mechanism for maintaining expressive clarity and thematic relevance (Guo et al., 2024; Zhang et al., 2024). Yet in many real design settings, prompts are still written in an intuitive and ad hoc manner. What remains insufficiently examined is whether more systematic prompt construction can support clearer organization of cultural references and related narrative information, greater controllability of concrete visual elements, and more targeted prompt revision during AI-assisted early-stage concept generation. This gap underpins our decision to foreground prompt formulation: in practice, it is the main interaction medium through which designers express intent, set constraints, and refine meaning with generative models.

### 2. 2. Prompt Formats and the Need for Structured Design

In generative text-to-image systems, prompts are the primary channel through which users express visual goals, stylistic preferences and contextual cues. Because they rely on natural language rather than explicit graphical controls or parameter sliders, prompts make AI image tools approachable for non-technical users, but they also introduce ambiguity and instability—especially in culturally sensitive design contexts where symbolic precision is important (Liu & Chilton, 2022; Oppenlaender et al., 2024).

Platforms such as Midjourney and DALL-E generally accommodate two common prompting styles: keyword-dominant prompts and prompts written in natural language. Keyword-heavy formats can be effective when designers want to pin down specific visual parameters—for example, color, composition, or subject attributes—but they often require platform-specific know-how and are not always easy to read, share, or reuse in a team setting (Don-Yehiya et al., 2023; Zhong et al., 2023). Natural language prompts, by contrast, are easier to write and make it simpler to convey narrative intent; the trade-off is that outputs can become less predictable, and themes may drift across iterations when the prompt is revised repeatedly (Brade et al., 2023; Mahdavi Goloujeh et al., 2024).

Both research and practitioner experience have increasingly converged on a hybrid approach: combining a stable keyword scaffold with a lighter layer of natural language (Rapp et al., 2025; Zhan et al., 2024). In such prompts, structured keywords are used to lock in the elements that should remain consistent (e.g., product type, subject, landmark, and style), while short natural language phrases add tone, emotion, or narrative cues. This separation is useful because it clarifies what is “fixed” versus what is “flexible,” helping designers retain control without losing expressive room and reducing unintended semantic shifts during iterative generation.

This logic is particularly important in cultural tourism souvenir design, because such tasks are not simply about generating visually appealing images, but about establishing a relatively clear correspondence among product form, local cultural references, and visual expression (Duan et al., 2023; Zhu & Rahman, 2025). When these aspects are not clearly organized in the prompt, the resulting images are more likely to show theme drift, stacked cultural symbols, or an imbalanced expressive focus (Don-Yehiya et al., 2023; Zhan et al., 2024). From this perspective, the value of structured prompting lies not only in improving general generation control, but also in helping designers organize culture-related content more explicitly and make more targeted adjustments to concrete visual elements. On this basis, this study further develops and tests a structured prompting method for cultural tourism souvenir design.

### **2. 3. Applications of Latent Dirichlet Allocation (LDA) Topic Modeling in Design Language Analysis**

To identify recurrent latent themes and keyword clustering patterns in AI-generated design cases in a data-driven way, this study introduced Latent Dirichlet allocation (LDA) to model the prompt texts. Compared with relying entirely on manual induction, topic modeling makes it possible to identify latent themes and their representative terms from a larger text corpus (Blei et al., 2003; Kim, 2023). As a classic unsupervised topic modeling method, LDA represents a text collection as a probabilistic mixture of latent topics and extracts representative terms within each topic (Blei et al., 2003). Although generative AI prompts differ from conventional long-form natural language texts, they still tend to show features such as keyword density, relatively loose expression, and the juxtaposition of multiple kinds of semantic information. This makes it possible to use topic modeling to detect latent thematic cues and keyword organization patterns within them (Kumar et al., 2023; Zhai et al., 2022). On this basis, LDA was used in this study as a topic identification tool to derive

an initial topic–keyword structure from the prompt corpus and to provide a foundation for subsequent semantic interpretation and structural organization.

From a modeling perspective, each Midjourney prompt in this study is treated as a document, while the entire prompt corpus is represented as a probabilistic mixture of latent topics (Blei et al., 2003). To clarify how topic distributions and keyword generation are handled in this study, the core variables of the model are briefly defined below:

K: number of topics.

$\alpha$ : Dirichlet prior for document-topic distributions.

$\beta$ : Dirichlet prior for topic-word distributions.

$\theta_d$ : topic distribution for document d.

$\phi_k$ : word distribution for topic k.

$Z_{d,n}$ : topic assignment for the nth word in document d.

$w_{d,n}$ : the actual word at position n in document d.

In this study, each Midjourney prompt is treated as a document d and modeled with LDA, where  $\theta_d \sim \text{Dir}(\alpha)$  denotes its topic mixture. For each token position n, a topic assignment  $Z_{d,n} \sim \text{Mult}(\theta_d)$  is sampled, and the token  $w_{d,n}$  is generated from  $\phi_{Z_{d,n}} \sim \text{Dir}(\beta)$ . The resulting topics and high-frequency keywords are used to identify recurring themes in the prompt corpus and to inform the structured prompt design developed in later sections.

## 2. 4. Applying the Delphi Card Sorting Method for Structural Classification in Prompt Modeling

LDA topic modeling can identify the preliminary keyword results associated with each topic, but statistical topics do not directly correspond to the prompt components used in design practice. In cultural tourism souvenir design, these keywords often involve product attributes, cultural references, visual styles, and related descriptive information at the same time. Therefore, before moving to template construction, their semantic boundaries and category relationships still need to be further clarified. For this reason, this study introduced Delphi card sorting after LDA in order to conduct expert-based semantic organization of the topic keywords.

Card sorting allows participants to group and label information based on their own understanding, and is well suited to the early stages of classification when category boundaries are still unclear (Spencer, 2009). In particular, open card sorting makes it possible to retain the natural semantic associations among keywords without imposing a predefined classification framework. However, when card sorting is used alone, differences in individual interpretation can easily lead to unstable category boundaries (Righi et al., 2013).

Delphi is a method for building expert consensus through anonymous judgment, multiple rounds of feedback, and step-by-step revision, and is appropriate for classification problems in which boundaries are ambiguous and gradual convergence is needed (Dalkey & Helmer, 1963; Hsu & Sandford, 2007; Okoli & Pawlowski, 2004). For this reason, the present study

combined open card sorting with Delphi: the former was used for open grouping, while the latter was used to compare, refine, and stabilize category boundaries and labels. Through this process, the keyword sets identified under each LDA topic were further organized into semantic categories with clearer boundaries and more stable labels, which were then translated into prompt components that designers could directly use and revise in later stages.

Based on this procedure, the study examined the top 100 weighted keywords under each theme, which then served as the basis for the structured prompt templates developed in the following section.

### 3. Study 1: Construction of Design Methods

#### 3.1. Method

As shown in Figure 1, the method consisted of four steps. A prompt corpus for cultural tourism souvenir design was curated from Midjourney’s official website and cleaned for subsequent analysis. We used LDA topic modeling to identify five candidate design themes from the corpus. To interpret and stabilize the themes, we ran a three-round Delphi card-sorting process in which experts independently clustered representative keywords and iteratively agreed on the semantic meaning and boundaries of each theme. The resulting expert consensus was then translated into topic-specific prompt elements and their organizational logic, yielding five structured prompt models and an operable workflow to support cultural tourism souvenir design tasks.

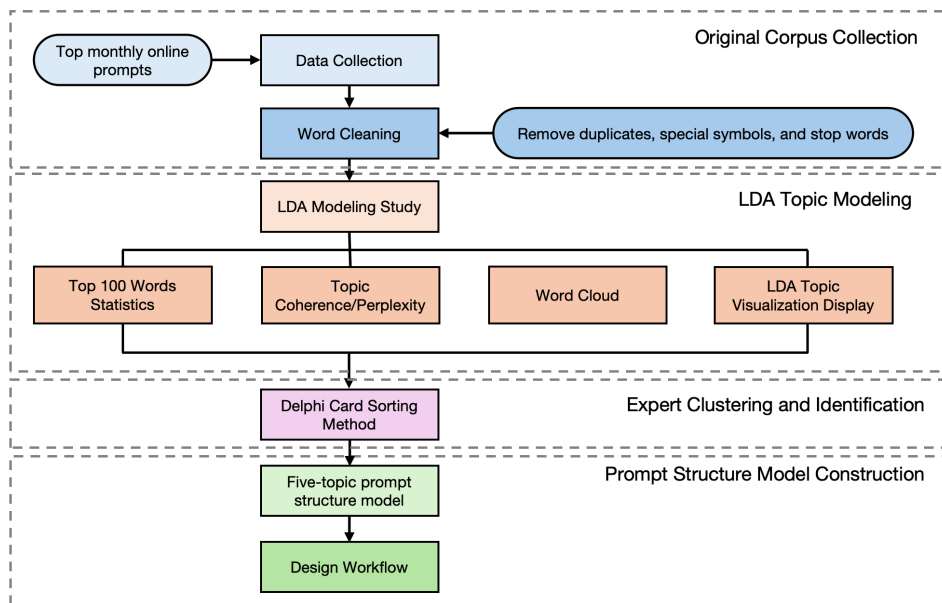


Figure 1 Research methods

### 3. 2. Data Collection

Midjourney’s official Explore page displays curated, high-engagement images together with the prompts used to generate them, and it supports browsing by day/week/month as well as keyword search (Midjourney, 2024c). We selected the Explore page as the source of prompt data because the prompt–image pairs displayed there are curated by the platform and associated with relatively high user engagement. In general, these cases represent more polished and more visible generated works within the platform, which makes them useful for examining common structural patterns and modes of expression in prompts related to cultural tourism souvenirs.

We manually compiled a prompt corpus over three months (December 2024–February 2025) by repeatedly querying Explore with the keyword “cultural travel souvenir.” For each query session, we copied prompts that corresponded to cultural tourism souvenir–related outputs with high user engagement, then merged records across sessions and removed duplicates. This process yielded 2,460 prompts. During manual compilation, the results were screened at the same time according to the aims of the study. Prompts were retained when the generated outputs showed both a clear souvenir carrier and a recognizable destination-related cultural element or tourism context. By contrast, results were excluded when they did not present a clear souvenir attribute or lacked a clear cultural tourism orientation. For example, outputs organized around landmarks, local cultural symbols, traditional crafts, or tourism scenes, and ultimately expressed as concrete souvenir forms such as badges, magnets, postcards, accessories, or lamps, were retained. In contrast, outputs that mainly appeared as general travel collages, combinations of international landmarks, conceptual display objects, or other content without a clear souvenir identity were excluded. In other words, what we removed from the dataset was not simply low-quality imagery, but content that did not match the task of cultural tourism souvenir design. To comply with platform terms and preserve the original context, we collected all prompts and visible metadata manually, without automated web crawlers (Midjourney, 2024a).

Before topic modeling, we cleaned and standardized the text to make short prompt inputs more comparable and reduce noise. Following common natural language processing (NLP) preprocessing practices (Bengfort et al., 2018; Miner et al., 2012), we performed tokenization, normalization, stop-word removal, low-frequency filtering, and vectorization. Tokenization split each prompt into word units to support downstream semantic modeling (Putra et al., 2018). We normalized the text by lowercasing and removing symbols and digits to reduce sparsity (Chai, 2023), and we removed common English stop words (e.g., “the,” “is,” “of”) using the Natural Language Toolkit (NLTK) stop-word list (Chai, 2023).

Because prompts frequently contain multi-word expressions (e.g., “Eiffel Tower,” “white background”), we also applied phrase mining after tokenization. Frequent bigrams that appeared at least twice in the corpus were merged into phrase tokens and added to the vocabulary, which helped preserve contextual meaning and reduce fragmentation in short-text modeling (Lau et al., 2013). The final corpus contained 11,277 unique tokens for subsequent analyses.

### 3. 3. Topic Modeling with LDA

This study applied LDA to identify hidden semantic structures in the prompt data. LDA is an unsupervised modeling approach well-suited for short texts with sparse labels (Steinberger & Griffiths, 2007).

We implemented the LDA model in Python (v3.9) with Gensim and fitted it to the preprocessed prompt corpus. The model was trained for 50 passes, and term frequency-inverse document frequency (TF-IDF) weighting was applied to improve topic separability. For each prompt, the model produced a topic-probability vector; we also extracted the top keywords for each topic and summarized topic prevalence across the corpus.

To select the number of topics (K), we compared candidate models using two standard metrics. Perplexity was used as an indicator of out-of-sample fit, where lower values suggest better generalization (Blei et al., 2003; Wallach et al., 2009).

Coherence Score: Reflects how semantically related top words in a topic are, aligning with human interpretation (Stevens et al., 2012).

Perplexity is calculated as (Blei et al., 2003):

$$\text{Perplexity}(D_{\text{test}}) = \exp\left(-\frac{\sum_{d=1}^M \log p(w_d)}{\sum_{d=1}^M N_d}\right) \quad (1)$$

Where  $p(w_d)$  is the likelihood of document  $d$ , and  $N_d$  is the word count in  $d$ . Lower values indicate better model fit.

Coherence is calculated as (Mimno et al., 2011):

$$\text{Coherence}(T) = \sum_{m=2}^M \sum_{l=1}^{m-1} \log\left(\frac{D(w_m, w_l) + \epsilon}{D(w_l)}\right) \quad (2)$$

Where  $T$  is a topic's keyword set,  $D(w_m, w_l)$  is the number of documents containing both  $w_m$  and  $w_l$ , and  $\epsilon$  is a smoothing constant.

To enhance interpretability, three additional indicators were applied:

Distinctiveness (Chuang et al., 2012): Measures how uniquely a word identifies a specific topic across the corpus.

$$\text{Distinctiveness}(w) = KL(P(T | w) || P(T)) \quad (3)$$

Saliency (Chuang et al., 2012): Combines word frequency with its distinctiveness to evaluate a term's prominence in a topic.

$$\text{Saliency}(w) = p(w) \cdot \text{Distinctiveness}(w) \quad (4)$$

Relevance (Sievert & Shirley, 2014): Adjusts keyword rankings by balancing term probability and distinctiveness (lift).

$$r(w, t) = \lambda \cdot \log p(w|t) + (1 - \lambda) \cdot \log \text{lift}(w, t) \quad (5)$$

Relevance is widely used in pyLDavis, a Python-based LDA visualization tool, to enhance topic interpretability.

### 3. 4. LDA Topic Selection and Visualization

Choosing a suitable topic number (K) is crucial for balancing model performance and usability. This study tested K values from 3 to 10 using perplexity and coherence metrics. Results are shown in Figure 2.

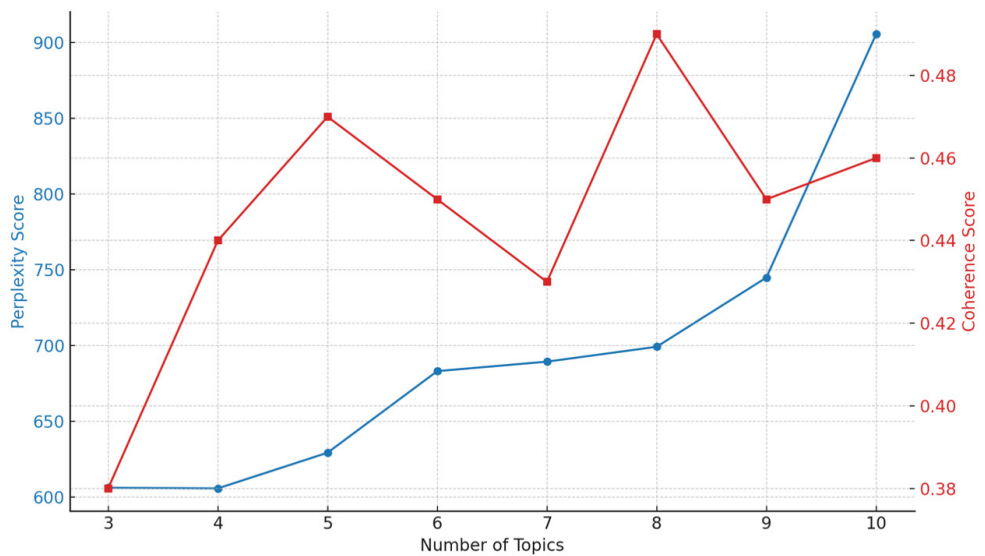


Figure 2 Trend chart of perplexity and Coherence score under the conditions of K = 3 ~ 10

K = 5 offered a favorable trade-off, yielding the lowest perplexity (629.34) and the highest coherence (0.468), which indicates relatively clear and stable topic divisions. Thus, five topics were selected for further modeling and expert classification.

To maintain interpretability for practical use, we constrained the candidate topic numbers to a small and manageable range (K = 3–10). Prior work suggests that increasing K can yield more fine-grained topic solutions, but may also produce fragmented or redundant topics and make human interpretation more difficult; therefore, topic-number selection should balance statistical fit and human interpretability rather than optimize a single metric (Chang et al., 2009; Weston et al., 2023). In contrast, very small K values (e.g., K values of 1–2) are often too coarse to align with LDA’s core assumption that documents are mixtures of multiple topics (Blei et al., 2003). Based on the above considerations and the metric trends shown in Figure 2, we ultimately selected K = 5. From the perspective of interpretive utility and downstream design use, the aim of topic selection in this study was to derive a set of themes that would remain manageable for expert classification, prompt template construction,

and subsequent workshop application. Cognitive research has shown that the number of information units people can actively handle at one time is usually limited to a relatively small range, often described as around 3–5 chunks (Cowan, 2001). Research on choice complexity has likewise shown that as the number of alternatives increases, the burden of judgment and selection also rises (Proctor & Schneider, 2018). In this sense, K = 5, while selected on the basis of statistical performance, also remained within an operationally manageable range for interpretation and design use.

For topic interpretation and reporting, we adopted two complementary visual summaries. First, word clouds (Figure 3) display the most representative keywords for each topic, where font size corresponds to relative importance.

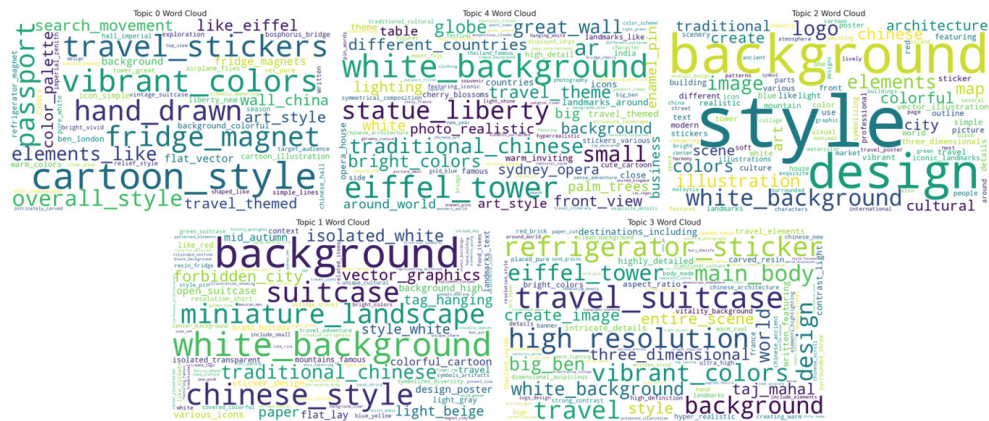
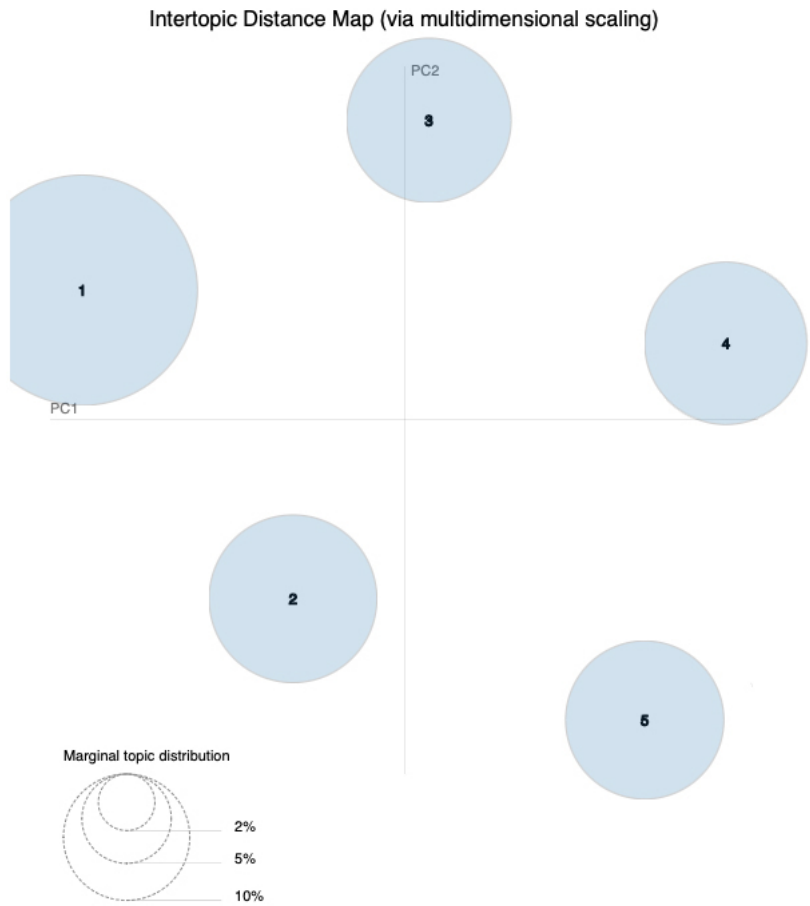


Figure 3 Word cloud diagram of keywords for 5 topics

Second, the LDA topic map (Figure 4) visualizes topic prevalence and semantic distance: each circle denotes a topic, its size indicates prevalence in the corpus, and the spacing reflects distinctiveness (Chuang et al., 2012). As shown in Figure 4, topics are well separated with minimal overlap, which supports the semantic distinctiveness of the extracted themes. Overall, the five-topic solution provides an interpretable structure for downstream prompt-structure analysis while retaining adequate model fit.



**Figure 4** Visualization map of 5 topics of LDA distribution

### 3. 5. Implementation of the Delphi Card Sorting Method

After LDA identified five preliminary topics, this study further applied a three-round Delphi card-sorting process to classify the top 100 weighted keywords under each topic. The purpose was to turn statistically derived keyword sets into structured categories with clearer semantic boundaries and more stable labels. This process followed the basic logic of Modified Delphi Card Sorting by combining open card sorting with iterative feedback and revision (Paul, 2008; Reese et al., 2018). Information on the six experts involved in the classification is presented in Table 1, and the overall procedure is shown in Figure 5. The expert panel included both university faculty and professional designers with experience in cultural tourism souvenirs and product design, and all had experience using AI image-generation tools. This allowed the keywords to be judged from multiple perspectives, including product attributes, cultural expression, visual style, and design application.

Table 1 Basic information on experts consulted

ID	Sex	Occupation/ Field of Work	Profile	Experience in Using AI
Expert 1	Male	University teacher / Industrial design	Ph.D. in industrial design, 20 years of experience in product design and development	Familiar with AI visual tools and integrates AI design practice in teaching.
Expert 2	Female	University teacher / Visual communication	Ph.D. in visual communication design, 15 years of experience in cultural product design	Familiar with AI image generation tools, frequently uses them for design prototyping and teaching
Expert 3	Female	Senior designer / Cultural tourism souvenir development	Master's degree in industrial design, over 10 years of experience in heritage souvenir design	Practical experience with AI-assisted conceptual sketching and creative brainstorming
Expert 4	Male	University teacher / Cultural heritage and art design	Ph.D. in art design, 12 years of experience in cultural and tourism product development and mentoring design competitions.	Regular user of AI visual generation tools for cultural heritage visualization
Expert 5	Male	Senior researcher / Product Design	Ph.D. in culture product design, 16 years of research in affective product design	Familiar with AI visual tools and uses them for shape language and emotion mapping studies
Expert 6	Female	University teacher / Interaction design	Ph.D. in interaction design, 14 years of teaching and practice in interactive cultural products	Familiar with AI tools for visual scenario creation and interface mock-ups

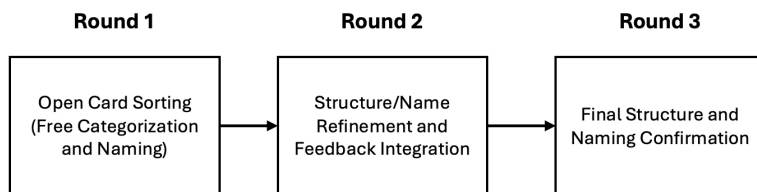


Figure 5 The Prompt classification and naming process of expert Delphi Cards Sorting

The first round focused on open exploration rather than immediate agreement. In this round, the research team used an open card-sorting approach and did not provide a fixed classification framework in advance. Instead, the keyword sets under each topic were given to the experts, who were asked to group them freely and assign their own category labels based on their understanding. This step was intended to preserve the natural semantic associations among keywords as much as possible, so that potential boundary differences between terms could emerge clearly in the initial round (Spencer, 2009).

The second round focused on comparing, organizing, and refining the initial results. The research team first reviewed the Round 1 classifications across experts to identify categories that were already showing signs of convergence, as well as keywords and boundary issues that still involved clear disagreement. On this basis, overlapping categories were merged, inconsistent labels were aligned, and an anonymized structured summary was returned to the experts. The experts then reclassified the same keywords in light of this feedback, with particular attention to semantically ambiguous terms and category boundaries (Hsu & Sandford, 2007; Okoli & Pawlowski, 2004). In this sense, Round 2 did not restart the classification process, but instead moved the category structure from an initially dispersed state toward gradual convergence while still preserving the openness of expert judgment.

The third round was used mainly to confirm keywords that had not yet fully stabilized after Round 2, rather than reopening all categories for another round of open classification. Only those terms that still showed boundary-level disagreement or retained a secondary interpretive frame were sent back to the experts for review. Categories that had already reached agreement in Round 2 were not revised again in any major way. Once the remaining differences no longer represented substantive classification discrepancies, the keyword assignment was finalized. In light of prior Delphi research on the balance between the number of rounds and participant burden, three rounds are generally sufficient to obtain a relatively stable structure while also limiting expert fatigue (Niederberger & Köberich, 2021).

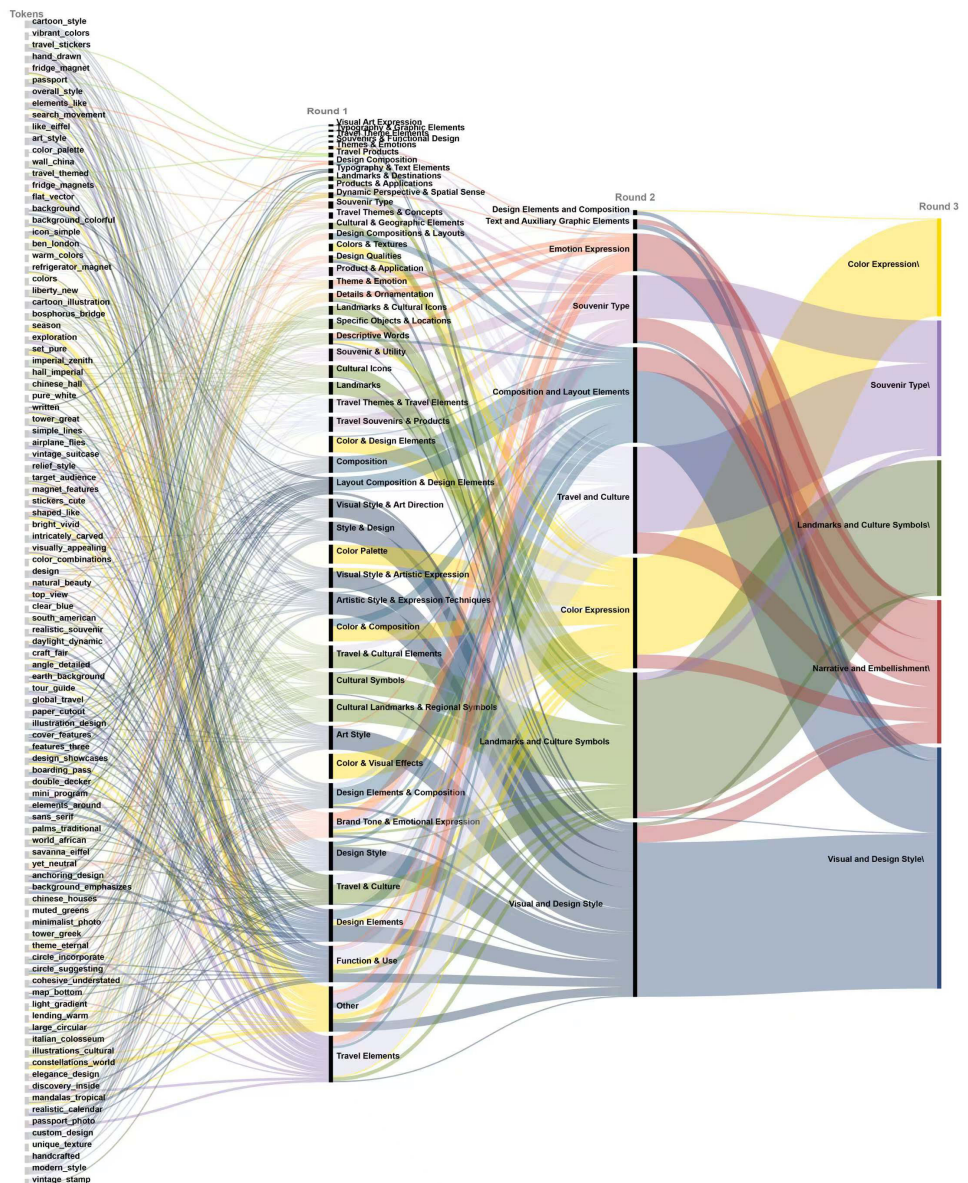


Figure 6 Three-round Delphi card-sorting clustering for Topic 0

Table 2 Representative examples of keyword reclassification and convergence across three rounds of Delphi card sorting

Stage		fridge_magnet	intricately_carved	colors
Round 1 classifications		E1: Souvenir and Utility	E1: Art Style	E1: Color Palette
		E2: Travel Souvenirs and Products	E2: Design Qualities	E2: Color and Design Elements
		E3: Travel and Culture	E3: Design Elements / Function	E3: Color and Visual Effects
		E4: Other	E4: Design Style	E4: Color and Composition
		E5: Product and Application	E5: Descriptive Words	E5: Design Elements and Composition
		E6: Souvenir Type	E6: Artistic Style and Expression Techniques	E6: Color and Visual Effects
Round 1 team summary	Main convergence areas	Souvenir/product-related: travel-cultural-related	Style/design-related; descriptive-related	Color-related; design/composition-related
	Boundary issue	Specific product identity vs. broader travel-cultural meaning	Specific craft/style treatment vs. general descriptive meaning	Specific color descriptor vs. broader design/composition meaning
	Candidate frame	Souvenir Type / Travel and Culture	Visual and Design Style / Descriptive Expression	Color Expression / Design Elements and Composition
Round 2 clustering result		E1: Souvenir Type	E1: Visual and Design Style	E1: Color Expression
		E2: Souvenir Type	E2: Visual and Design Style	E2: Color Expression
		E3: Travel and Culture	E3: Visual and Design Style	E3: Color Expression
		E4: Souvenir Type	E4: Visual and Design Style	E4: Color Expression
		E5: Souvenir Type	E5: Visual and Design Style	E5: Design Elements and Composition
		E6: Souvenir Type	E6: Visual and Design Style	E6: Color Expression
Round 2 team summary		Most interpretations converged on Souvenir Type, while Travel and Culture remained a secondary frame for final clarification	Full consensus on Visual and Design Style was achieved in Round 2	Interpretations largely converged on Color Expression, while one broader design/composition interpretation remained for final clarification
Round 3 result		Souvenir Type	Visual and Design Style	Color Expression

Through this three-round process, keyword classification gradually moved from open grouping toward a relatively stable and interpretable category structure. Using Topic 0 as an example, Figure 6 shows how keyword categories shifted and converged across the three rounds, while Table 2 presents several representative keywords to illustrate the reasoning behind these convergence paths in more detail. In Table 2, E1–E6 refer to Expert 1–Expert 6 listed in Table 1. For instance, in Round 1, fridge\_magnet was understood by some experts as a specific souvenir type, while others interpreted it as an object carrying travel-related cultural meaning. This suggests that the disagreement was not about the meaning of the term itself, but about which semantic function it should primarily serve in the later prompt structure. By contrast, the Round 1 classifications of colors were more dispersed, indicating that experts placed the term within adjacent frames such as color, visual effect, and composition. Taken together, these examples show that disagreement did not arise from arbitrary judgment, but mainly from blurred boundaries between neighboring semantic functions. By the end of Round 2, these initial disagreements were no longer expressed as open-ended divergence, but had narrowed to a small number of still unresolved classification

decisions. For example, colors had largely converged on Color Expression, though a small number of broader design- and composition-related interpretations remained. Similarly, fridge\_magnet had mostly converged on Souvenir Type, while Travel and Culture was still retained in a few cases. At the same time, not all disputes continued into Round 3. In the case of intricately\_carved, although Round 1 had produced different labels such as artistic style, design feature, and descriptive expression, the term had already converged on Visual and Design Style by Round 2, showing that for some keywords the boundary issue could be sufficiently clarified at that stage. In contrast, Round 3 dealt specifically with residual classification disagreements that had not been fully resolved after Round 2, such as those seen in fridge\_magnet and colors, and their final category labels were determined based on the results of this last review. The resulting category labels were then further organized into structured prompt components for the five topics and used in the prompt structure model construction presented in the next section.

---

## 4. Prompt Structure Model Construction

### 4. 1. Structural Components of the Five Topics and Their Cross-Topic Characteristics

Based on the LDA results, five topics were identified from Midjourney prompts related to cultural tourism souvenirs. Each topic was characterized by a set of high-weight keywords. These keyword sets were then further refined through Delphi card sorting to clarify their semantic boundaries and finalize their category labels. In other words, the five topic categories were not automatically named by LDA itself. Instead, the research team determined the topic names by interpreting the semantic center formed by the high-weight keywords within each topic and by taking into account the experts' final classification of the core word groups. Table 3 presents the thematic descriptions of the five topics together with their corresponding structured prompt components, showing the basic classification results within each topic. These structural components were not added afterward as an extra interpretive layer. Rather, they were directly grounded in the stabilized classification results and then organized into topic-specific prompt templates to provide structured guidance for designers during early-stage concept exploration and prompt construction. They can function as usable prompt units because each category already corresponds to a structural role that can be independently specified, replaced, or adjusted in prompt construction, such as souvenir type, cultural source, visual rendering, and narrative embellishment.

Table 3 Thematic descriptions and structured prompt components for the five topics

Topic and Focus	Structural Component	Description
Topic 0: Playful & Creative Souvenirs : Playful, vivid, cartoon-like souvenirs with bold visuals; suitable for cheerful and entertaining concepts.	Souvenir Types	Specifies the product form (e.g., stickers, fridge magnets) chosen for their fun and creative appeal.
	Landmarks and Cultural Elements	Iconic landmarks or cultural symbols of the destination used as textual input.
	Color Expression	Bright, lively color schemes that reflect local characteristics or playful design intent.
	Visual and Design Style	Cartoon or illustrative styles that emphasise creativity and playfulness.
Topic 1: Traditional Cultural Heritage with a Modern Twist : Blends classical cultural symbols with contemporary aesthetics to appeal to modern tastes.	Narrative and Embellishment	Descriptive terms and short phrases that enhance storytelling and cultural association.
	Souvenir Types	Specifies the product form; modern product formats (identified from high-frequency keywords) are used as carriers of traditional motifs.
	Cultural Heritage and Identity	Representative heritage elements, landmarks and regional characteristics that signal local identity.
	Visual and Design Style	Color schemes and backgrounds that balance traditional symbolism with a modern visual aesthetic.
Topic 2: Contemporary Visual & Cultural Souvenir Design: Balances minimalist contemporary aesthetics with multilayered cultural elements, aiming for clarity and cultural richness.	Narrative and Embellishment	Cultural stories and descriptive elements that add depth and interpretive richness.
	Souvenir Types and Architectural Elements	Combines souvenir forms with recognizable architectural features (e.g., skylines, facades) to anchor the visual theme.
	Local Cultural and Artistic Motifs	Uses iconic local cultural or artistic elements that frequently appeared in the prompt corpus.
	Modern Visual Foundations	Clean color palettes, simple compositions and minimal backgrounds reflecting contemporary design preferences.
Topic 3: Hyper-Realistic & Iconic Cultural Imagery: Uses hyper-realistic, photographic styles with high visual fidelity to present iconic cultural imagery and rich detail.	Primary Iconography and Souvenir Types	Specifies the souvenir form and the primary iconographic subject to be rendered (e.g., landmark / emblematic scene / artefact), so the prompt clearly defines both the carrier and the visual focus.
	Narrative and Embellishment	Adds culturally anchored descriptors that highlight symbolic meanings and contextual associations (e.g., origin, tradition, festival, historical reference) tied to the chosen iconography, supporting interpretability beyond pure appearance.
	Hyperreality Factor	Uses realism tokens as rendering constraints-resolution/detail/sharpness terms (e.g., “photorealistic,” “8K,” “fine detail”)–to push the output toward photographic fidelity and texture-rich surfaces.
	Visual and Design Style	Controls photographic look-and-feel through lighting, color tone, and background settings (e.g., studio-like lighting, neutral backdrops), ensuring the hyper-realistic subject remains visually legible and polished.
Topic 4: International Landmark & Cross-Cultural Design: Emphasises a globalised perspective, integrating diverse cultural elements and promoting a universal aesthetic for cross-regional stories.	Souvenir Types	Specifies the souvenir format as the output carrier (e.g., postcards, posters, magnets), which sets the expected layout and content density for internationally oriented themes.
	Global Landmarks and Culture Symbols	Lists internationally recognizable landmarks and culture symbols as the main visual references (e.g., skylines, monuments, national icons), providing clear, widely shared cues for cross-regional interpretation.
	Cross-Cultural Expression	Uses explicit cross-region framing (e.g., multiple countries/cities, global travel routes, intercultural encounters) to signal that the design should integrate elements from more than one cultural context rather than representing a single locale.
	Narrative and Embellishment	Adds short linking phrases that explain how the regions or cultures relate (e.g., journey, exchange, comparison), helping align disparate symbols into one coherent storyline.
	Environment and Background	Sets a neutral, readable visual context (e.g., clean background, soft lighting, simplified global backdrop) to prevent the environment from competing with multi-cultural iconography and to support clear composition.

Building on this, Table 4 reorganizes these structural components from a cross-topic perspective. Here, “Shared Component” refers to a higher-level structural role that recurs across multiple topics and serves a similar structural function. The differences are reflected in how this shared role is instantiated in different topics. At this stage, the research team did not further merge similar components or force all topics into a single unified template. The reason is that the study aimed to preserve the topic-level differences revealed by the data-driven results and the experts’ final classifications, rather than imposing formal uniformity through post hoc adjustment. A closer comparison shows that some components recur across several topics, such as Souvenir Types, Visual and Design Style, and Narrative and Embellishment, suggesting that they play relatively stable structural roles in different topics. At the same time, some components were retained only in particular topics. For example, although Color Expression belongs to the broader dimension of visual presentation, it was kept as an independent component only in Topic 0 because color-related keywords occupied a larger share of that topic, ranked more prominently, and formed a stronger semantic cluster. In the other topics, color information was present, but it did not emerge as an equally salient independent dimension. Likewise, Hyperreality Factor and Cross-Cultural Expression were retained as independent components because they were especially prominent in Topic 3 and Topic 4, respectively. It should also be noted that the recurrence of the same component label across topics does not mean that the underlying keyword content is identical. Rather, it indicates that these components play similar structural roles at a higher level. Taking Visual and Design Style as an example, Table 3 shows that this dimension in Topic 0 is characterized more strongly by cartoon and illustration-based playful rendering, whereas in Topic 1 it places greater emphasis on balancing traditional cultural elements with modern visual aesthetics. In other words, the same label points to a recurring structural role, not to exactly the same semantic content. Table 3 and Table 4 therefore serve different but connected purposes: the former presents the classification results within each topic, while the latter shows how those results relate to one another across topics.

Table 4 Cross-topic alignment of structural components derived from data-driven semantic clustering

Shared Component	General Definition	Topic 0	Topic 1	Topic 2	Topic 3	Topic 4
Souvenir Form and Carrier	Defines the souvenir form or carrier through which the design concept is materialized	Souvenir Types	Souvenir Types	Souvenir Types and Architectural Elements	Primary Iconography and Souvenir Types	Souvenir Types
Cultural Source and Symbolic Content	Identifies the cultural references, symbols, landmarks, or motifs that anchor the design concept	Landmarks and Cultural Elements	Cultural Heritage and Identity	Local Cultural and Artistic Motifs	—	Global Landmarks and Culture Symbols
Visual Rendering	Defines how the image is visually rendered, including style, color, realism, composition, and background treatment	Color Expression; Visual and Design Style	Visual and Design Style	Modern Visual Foundations	Visual and Design Style	Environment and Background
Narrative and Embellishment	Adds short descriptive or interpretive phrases that enrich mood, storytelling, or semantic emphasis in the prompt	Narrative and Embellishment	Narrative and Embellishment	Narrative and Embellishment	Narrative and Embellishment	Narrative and Embellishment
Topic-Specific Emphasis	Captures additional dimensions that become independently salient in certain topics	—	—	—	Hyperreality Factor	Cross-Cultural Expression

## 4. 2. Structured Prompt-Based Workflow for AI-Generated Design of Cultural Tourism Souvenirs

To support AI-assisted design in cultural tourism souvenir creation, this study proposes a structured, prompt-based workflow for using Midjourney (Figure 7). The workflow is built on the five prompt topics and their structural templates introduced in Section 4.1. Designers develop prompts by filling topic-specific components and then combining them into one coherent instruction. In this way, the workflow integrates keyword precision with natural-language flexibility, forming a complete semantic-to-visual generation pipeline. Design judgment enters through topic selection, topic-specific component filling, and prompt composition.

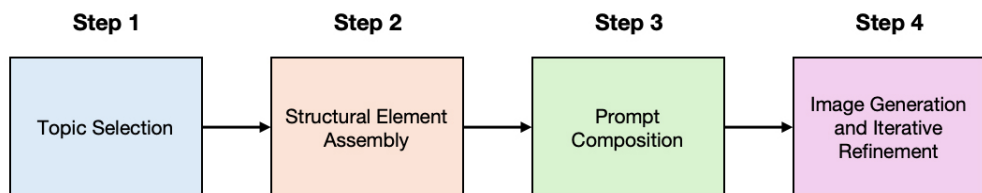


Figure 7 Workflow for structured AI-driven cultural tourism souvenir design

### Step 1: Topic Selection

Designers first choose one topic that matches the intended design direction. This choice functions as a constraint on the overall visual target and the type of cultural expression. For instance, for a Palace Museum souvenir intended to be playful, the designer would start from Topic 0: Playful & Creative Iconography Souvenirs.

### Step 2: Structural Element Assembly

Each selected topic contains its own set of 4–5 topic-specific structural components, established through the preceding LDA and Delphi card-sorting process and presented in Section 4.1. Designers populate each component with terms that reflect the destination’s cultural features and the message to be conveyed. In the Palace Museum example, one feasible set is:

- Souvenir Type: handkerchief;
- Landmarks and Cultural Elements: Chinese traditional embroidery craft; Palace Museum;
- Color Expression: Red;
- Visual and Design Style: cartoon style design;
- Narrative and Embellishment: Cat flutter butterfly.

This step helps ensure that the prompt includes the main semantic dimensions rather than relying on a few broad adjectives.

### Step 3: Prompt Composition

After selecting the element terms, designers write one prompt that combines keywords with readable natural language, which helps maintain both controllability and usability (Don-Yehiya et al., 2023; Zhong et al., 2023). Prior studies have shown that this hybrid format can support clearer prompt formulation and improved output quality (Brade et al., 2023; Zhong et al., 2023). The composition follows Midjourney’s guidance on concreteness and clarity (e.g., prioritize the main subject and remove irrelevant modifiers)(Midjourney, 2024b). A typical prompt arranges elements in a consistent order so that each component is explicit and traceable during later edits. One example is:

Red (Color) Chinese traditional embroidery craft, Palace Museum (Landmarks/Cultural elements) handkerchief (Souvenir type), embroidered with a cat–butterfly pattern (Narrative/Embellishment), cartoon style design (Visual/Design style).

### Step 4: Image Generation and Iterative Refinement

Designers input the prompt into Midjourney and review the results. When the output deviates from expectations, revisions are applied to the specific component responsible (e.g., changing the souvenir type, tightening landmark terms, or adjusting style tokens) rather than rewriting the entire sentence. This targeted revision can make iteration more focused and aligns with model-feedback-based co-creation practices (Mahdavi Goloujeh et al., 2024).

Overall, the workflow provides a practical procedure for prompt formulation and revision in cultural souvenir ideation, with each step linked to an explicit semantic decision and a corresponding editable prompt component. More importantly, the workflow does not merely reorganize the topic-specific structured prompt templates presented in Section 4.1, but also serves as the operational procedure for the subsequent workshop evaluation. In this sense, the templates developed in Study 1 were not left at the level of classification results, but were further translated into a prompt-construction process that could be directly applied and tested in design tasks.

## 5. Study 2: AI Design Generation Workshop

To examine the practical usability of the topic-based structured prompt templates and their corresponding workflow developed in Study 1 in a realistic design setting, we conducted a one-day AI design generation workshop centered on cultural tourism souvenir image creation. In other words, Study 2 was not carried out independently from Study 1, but applied and evaluated the topic templates and prompt-construction workflow established in the previous stage through actual design tasks. The workshop was organized by the research team, who explained the topic templates and operating procedure to the participants, monitored task completion, and provided necessary on-site support to ensure procedural consistency.

### 5. 1. Participants

We recruited 22 participants from China with a design background (Table 5), including senior undergraduates, graduate and doctoral students, and professional designers in product design, visual communication, and related disciplines. Eligible participants were required to have basic knowledge of cultural tourism souvenir design; have prior experience using Midjourney, and be able to develop product concepts independently.

Table 5 The description of characteristics in the sample

Item	Indicators	Frequency	Percent
Gender	Female	12	54.5
	Male	10	45.5
Educational Background	Master's	13	59.1
	Undergraduate	6	27.3
	Doctoral	3	13.6
Area of Study	Product Design	6	27.3
	Visual Communication	5	22.7
	Industrial Design	3	13.6
	Digital Media Art	3	13.6
	Interaction Design	1	4.5
	Cultural Design Related	3	13.6
	Arts and Crafts Design	1	4.5
Years of Design Study/Work Experience	2-4 years	6	27.3
	5-7 years	10	45.5
	8-10 years	5	22.7
	>10 years	1	4.5
	Total	22	100

### 5. 2. Design Task

To keep the results comparable across topics, the structured prompt-based workflow proposed in Section 4.2 was tested within a unified design context. More specifically, Topics 0-3 were all carried out under the same design task, namely the development of cultural tourism souvenirs for the Palace Museum, so that the cultural target remained consistent and the effects of different topic-based prompt templates could be compared within the same cultural context. Topic 4, by contrast, was intended for cross-cultural expression.

For this reason, while retaining Palace Museum cultural elements as a base, it introduced an international theme centered on the fusion of Eastern and Western elements, so that the applicability of the workflow in a cross-cultural design task could also be examined. The Palace Museum was used as the shared design context because it provides a highly recognizable cultural reference familiar to the participants.

### **5. 3. Procedure**

At the start of the workshop, the research team explained the structural components and the step-by-step workflow of the five topic-based prompt models. Example cases were used to show how individual prompt elements could influence the generated images.

The workshop procedure was designed around two main considerations. First, the study aimed to examine whether the five topic-based structured prompt templates developed in Study 1 could be practically used by designers in cultural tourism souvenir design tasks. Second, it aimed to observe how the different topic-based templates varied in terms of structural element expression and overall topic-level expression under a unified design context. To this end, all participants completed five topic-based tasks and followed the corresponding structured prompting procedure. This arrangement helped keep the task conditions as consistent as possible when comparing the application of different topic templates, while also reducing additional variation caused by differences in participants' prompt-writing habits, language expression, and prior prompting experience. However, this workshop did not include a free-form prompt comparison condition. Therefore, the findings at this stage cannot be taken as direct empirical proof that structured prompts are more effective than free-form prompts. Rather, they should be understood primarily as observations of how the five structured prompt templates performed under a shared condition and how they differed at the topic level.

Each participant then completed five structured prompt drafting tasks—one for each topic—and used Midjourney to generate corresponding images. After finishing each topic, participants submitted the structured prompt text they used and one final image selected as the best match to their design intent. During the workshop, participants were allowed to run multiple generation attempts for each topic within a fixed time window. When questions arose, researchers provided clarification to keep the operational procedure consistent across participants.

### **5. 4. Evaluation of the Design Method**

The evaluation addressed two aspects. First, we assessed the expression clarity of each structural element in the generated images, including how recognizable and visually identifiable these elements were in the outputs. Second, we examined participants' perceptions of using the structured prompt template for each topic.

#### **5. 4. 1. Expert Scoring**

To assess visual expression of the structural elements, ten experts with experience in cultural product design and AI image generation evaluated all submitted images. Each element was scored on a seven-point scale from 1 (completely absent) to 7 (very clearly present). Expert demographics and professional backgrounds are reported in Table 6.

To reduce fatigue, the experts completed the scoring over one week. For analysis, the scores for each element were averaged within each topic to produce an element-expression score for the corresponding prompt structure. This score was used as an indicator of how clearly the structured prompt supported the specification and visual expression of its intended elements during generation.

Table 6 Basic information on experts

Item	Indicator	Frequency	Percentage (%)
Gender	Male	4	40
	Female	6	60
Age	30–39	3	30
	40–49	4	40
	50–59	3	30
	Master's Degree	3	30
	Doctorate	7	70
Occupation	University Professor	6	60
	Designer	4	40

#### 5. 4. 2. User Interviews

To capture participants' perceptions of the topic-based prompt structures, we conducted semi-structured interviews with all 22 workshop participants after they completed the five design tasks. A structured interview guide was prepared in advance (Appendix A). It asked about (1) initial impressions and perceived differences between structured prompts and participants' usual design approach (Q5–Q6), (2) the clarity and usefulness of specific structural elements (Q7–Q8), and (3) practical issues and improvement suggestions, including technical difficulties (Q9), revisions to the template (Q10), and perceived influence on the overall design process (Q11). The final section invited reflections on potential applications of structured prompting in cultural and creative work (Q12–Q18).

Interviews took place within one to two days after the workshop to keep experiences recent. Before each session, participants were informed of the study purpose, confidentiality arrangements, and their right to withdraw. With consent, interviews were audio-recorded. Sessions were conducted either online or face-to-face and typically lasted 40–60 minutes. All interviews were conducted in Chinese to ensure clear communication and accurate expression of participants' experiences. The interviewer adjusted follow-up questions and the order of prompts when needed to clarify responses and capture task-specific examples from the workshop tasks.

We analyzed the interview data using a grounded-theory–informed qualitative content analysis supported by NVivo 15. Audio recordings were transcribed verbatim. Researchers reviewed the transcripts repeatedly and segmented the text into meaning units. Each segment was coded to reflect its central idea, and codes were iteratively compared across participants and grouped into higher-level categories and themes (the analysis workflow is summarized in Appendix C).










To strengthen consistency, coding was carried out by multiple researchers independently and then discussed in team meetings to reconcile differences and confirm the final category structure. The interview findings were used alongside the expert scoring results as complementary evidence for evaluating the structured prompt model.







## 5. 5. Results

### 5. 5. 1. AI Design Works

Over the course of a day, 22 designers produced 110 cultural souvenir designs based on the structured prompt requirements for each topic. Representative examples are provided below (Table 7).

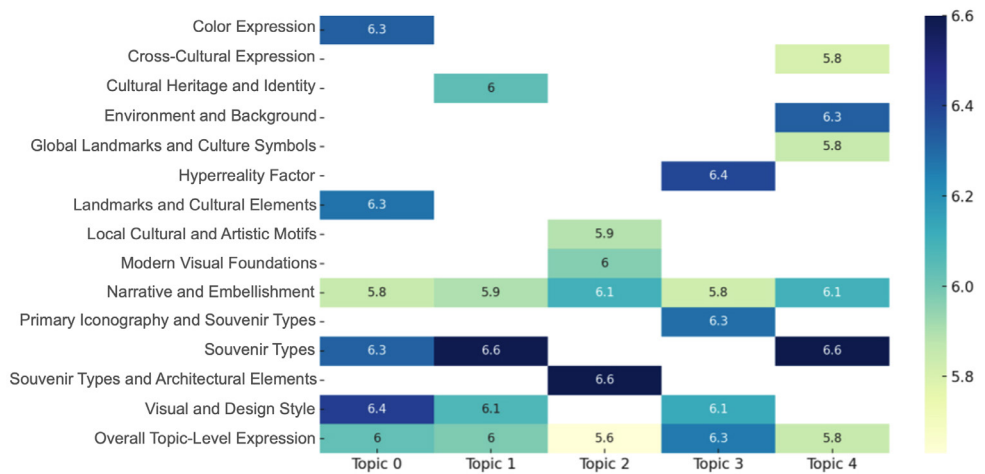
Table 7 Presents representative prompt examples and corresponding outputs from each of the five thematic prompt models

Topic	AI Image	Prompt
		An enamel brooch, cat elements of the Palace Museum and flower elements, lovely lively color scheme, modern Chinese style, embellished with diamonds, simple abstract, exquisite fashion design style, Sweet and quiet
Topic 0 Playful & Creative Souvenirs		A set of enamel travel pins showcasing landmarks of the Forbidden City, including palace roofs, traditional pavilions, China Chic color, vivid and cute style
		A set of Chinese princess doll design, Palace Museum elements, Beijing Opera costume decorative elements, traditional Chinese color matching, cartoon style, small cute and interesting, plush details
Topic 1 Traditional Cultural Heritage with a Modern Twist		A ceramic incense burner, the Palace Museum ceramic style, pure color with gold embellishing, simple minimalist style, clean background, new Chinese aesthetic, modern design, oriental sense of mystery and sacred sense
		An embroidered coin purse, architectural elements of the Palace Museum, plant embellishing, traditional Chinese color scheme, simple modern plane aesthetic, clean background, warm and peaceful atmosphere
		a Palace Museum souvenir bracelet, combine woven fabric cord with golden enamel craftsmanship. Incorporate pendants featuring imperial seals and traditional patterns to express local cultural identity, set it on a simple, solid-color background to highlight the product details, exudes a gorgeous Chinese temperament
Topic 2 Contemporary Visual & Cultural Souvenir Design		A knitted scarf printed with the Palace Museum, traditional Chinese architectural lace, back pattern, yellow color, highlighting the royal style, modern graphic art style
		folding umbrella painted on the silhouette of the Palace Museum, simple Chinese auspicious cloud decoration embellishments, tile art style, new Chinese color matching, geometric graphics, abstract plane design, simple modern fashion
		A canvas bag with the pattern of the National Palace Museum architecture, simple Chinese pattern embellishment, modern graphic art style, lively Chinese color scheme, geometric abstract graphics, fashion simple

Topic 3 Hyper-Realistic & Iconic Cultural Imagery		A crystal ball, inside is the Palace Museum under snow, super real miniature landscape, with architectural details, the red plum of the Palace Museum, snowflakes shining gold and silver light, traditional Chinese color matching, new Chinese aesthetics
		A pair of earrings is the architectural shape of the Palace Museum, super realistic architectural lines and jewelry details and light effects, red and gold color matching, new Chinese style, elegant imperial temperament
		A small lamp in the shape of the Palace Museum, super real architectural details, traditional Chinese shadow play form, a long sense of history, new Chinese design style, decorative style
Topic 4 International Landmark & Cross-Cultural Design		a set of 3D modern seals for the Beijing Palace Museum and other international landmarks, Eiffel Tower, Statue of Liberty, each seal features a different country City landmark. Show the detailed printed impression, with clear focus on the seal's face imprint. Each seal uses colorful ink in unique, vibrant hues, presented in a lively, on a clean background
		A three-dimensional paper carved table lamp composed of various world-famous landmarks, including the Palace Museum, the Forbidden City, the Temple of Heaven, the Terracotta Army, the Eiffel Tower, the Statue of Liberty pattern, cross-cultural integration, Chinese style pattern, simple abstract cartoon, placed on the desktop, clean background
		A set of modern enamel pins featuring Palace Museum in Beijing, the Eiffel Tower, the Statue of Liberty, cross-cultural international style, showing the architecture shape with bold lines and vivid red, gold, and white tones. on a tabletop, clean background

### 5. 5. 2. Expert Evaluation Results

Following the workshop, ten experts rated each design on the structural components within each topic and on an Overall Topic-Level Expression score. We first checked whether the ratings could be aggregated by examining inter-rater reliability (Cronbach's alpha) for each topic-dimension pair (Appendix B, Table B). Internal consistency ranged from  $\alpha = 0.79$  to 0.97, and most dimensions were above 0.90, suggesting high agreement across raters. In Topic 3, Hyperreality Factor ( $\alpha = 0.86$ ) and Overall Topic-Level Expression ( $\alpha = 0.79$ ) were comparatively lower than other dimensions but remained acceptable for exploratory design research.



**Figure 8** Heatmap of structured prompt element score by topic

After reliability screening, we computed mean scores for each topic and visualized them in a heatmap (Figure 8). Across topics, mean scores for structural elements were generally above 5.0 on the 7-point scale. For Topic 0, the Overall Topic-Level Expression was 6.03. Within this topic, Visual and Design Style (6.42) and Color Expression (6.33) were the highest-scoring elements, whereas Narrative and Embellishment (5.85) was lower.

For Topic 1 and Topic 2, Souvenir Types received the highest means (6.59 and 6.60), followed by culturally related elements such as Cultural Heritage and Identity (6.04) and Local Cultural and Artistic Motifs (5.89). Elements involving narrative or broader composition were slightly lower, including Narrative and Embellishment (5.90; 6.10) and Modern Visual Foundations (5.97).

For Topic 3, Hyperreality Factor (6.39) and Primary Iconography and Souvenir Types (6.29) were high, while Narrative and Embellishment (5.83) remained lower. For Topic 4, the Overall Topic-Level Expression was 5.85. Souvenir Types (6.60) and Environment and Background (6.33) scored relatively high, whereas Cross-Cultural Expression (5.81) and Global Landmarks and Culture Symbols (5.85) were lower.

At the topic level, Overall Topic-Level Expression was highest for Topic 3 (6.26), followed by Topic 0 (6.03) and Topic 1 (5.98), while Topic 2 (5.63) and Topic 4 (5.85) were lower. Overall, the scoring pattern suggests that elements describing concrete product form, visual style, color, and environment are expressed more consistently than elements requiring narrative formulation or cross-cultural semantic integration. This observation is consistent with participants' interview feedback that physical attributes were easier to control than cultural semantics.

### 5. 5. 3. Qualitative Content Analysis Results

Interviews with the 22 participants revealed recurring themes about the use of structured prompts in cultural tourism souvenir design. The themes relate to prompt-component

controllability, trust and acceptance of the tool, cultural translation and knowledge gaps, shifts in designers' working process, and theme focus during generation. Selected coding examples are provided in Appendix C, Table C.

Typical comments suggested that the template reduced uncertainty at the start of ideation by making required elements explicit. Some participants also reported that highly predefined slots could constrain exploratory thinking, particularly during early ideation when testing alternative directions.

At the element level, souvenir types, visual style, and material or craft details were frequently described as practical to specify. They were also perceived as relatively stable in the generated outputs. One participant shared, "AI understands material and craft details better than I expected, and renders faster than manual sketching." In contrast, participants described ongoing difficulty with cross-cultural symbols and complex cultural narratives. One interviewee pointed to culturally specific architectural motifs and explained, "Elements like 'roof ridge beasts' are unique to Chinese architecture. AI knows they are decorative animals on rooftops but fails to capture the intricate shapes and deeper cultural meanings."

Regarding trust in the tool, many participants linked structured prompts to efficiency and output quality. Some described the workflow as shortening detours in ideation. One participant said, "Structured prompts are like a design map, reducing detours," and another noted, "AI lets me quickly try different materials and techniques, with almost no time cost, and the results are good." However, participants also reported practical obstacles. When prompts included too many abstract semantics or multiple element combinations, controllability decreased. Outputs could become a simple stacking of elements rather than semantic integration. Several participants therefore expected more precise prompt language and a more complete local cultural knowledge base.

Participants also reported that structured prompts influenced their design thinking. Many described the approach as a "language-driven design blueprint." One designer explained, "In the past, I drew sketches first and then enriched the elements. Now I first write clear keywords, and AI generates images directly, which is very efficient." This shift was associated with faster trial-and-error cycles during ideation.

Theme control was repeatedly mentioned as important for keeping the output focused. Some participants stated, "The guidance meets expectations," and others said, "A clear definition of souvenir types prevents deviation from the intended design direction." These comments suggest that topic framing and clear carriers help reduce off-topic generations.

When discussing future development, designers were generally optimistic about structured prompting for cultural tourism souvenirs and related cultural creative work. They viewed it as suitable for "rapid prototyping and parallel development." At the same time, they noted the need for improvement in semantic understanding, cross-cultural expression, and interactive feedback. Several participants suggested providing more concrete examples and adjustable templates to reduce learning cost and improve efficiency. Others emphasized that building a

comprehensive traditional cultural knowledge base, including detailed historical background, cultural symbols, and craft details, would improve AI understanding of complex cultural semantics and enhance the cultural depth of generated outputs.

Overall, the interview themes complement the expert scoring results. Structured prompts were perceived as reliable for controlling concrete, visually grounded elements, whereas narrative formulation and cross-cultural semantic integration remained less predictable. These findings provide qualitative support for the strengths and limitations observed in the quantitative evaluation and help identify practical directions for refining the prompt structure in future work.

---

## 6. Discussion and Conclusion

### 6. 1. Five Cultural Tourism Souvenir Topics

This study derived five topic-based prompt structures for cultural tourism souvenir design through data mining and expert consensus. Their usefulness was examined in a design workshop and follow-up interviews. Together, the results indicate that the five topics span distinct design intents and provide a practical way to organize prompt writing for generative image design of souvenirs (Oppenlaender et al., 2024; Wang et al., 2024).

For Topic 0 (Playful & Creative Souvenirs), outputs were generally stable when the prompt emphasized clear visual cues such as cartoon or illustrative style and color. Designers reported that these elements were easy to adjust within the template, which reduced repeated trial-and-error during generation (Rajcic et al., 2024). At the same time, participants noted that this topic showed clearer limitations when the design goal required richer storytelling or deeper cultural interpretation. In such cases, additional refinement or post-editing was often needed to enrich the narrative content (Lyu et al., 2024).

Topic 1 (Traditional Cultural Heritage with a Modern Twist) combines heritage symbols with contemporary aesthetics. Designers described this topic as useful for keeping cultural identity visible while presenting it in a modern product form. However, participants also noted that the balance is sensitive to wording. When heritage descriptions were too general, the outputs tended to look “modern” but culturally generic. More specific cultural cues and concise narrative framing were needed to preserve recognisable heritage features within a contemporary style.

Topic 2 (Contemporary Visual & Cultural Design) supports a modern, minimalist visual language combined with cultural symbols. In our evaluation, this topic showed relatively clearer style consistency and key-symbol expression, which aligns with market preferences for clean and commercially appealing souvenir visuals. However, participants also pointed out that simplified compositions can weaken narrative depth. The resulting images may look complete in form but convey limited story or emotion (Oppenlaender, 2024).

Topic 3 (Hyper-Realistic & Iconic Cultural Imagery) showed relatively higher performance on the relevant expression dimensions when prompts stressed visual fidelity and material detail. Expert scores and interview feedback suggested that realism-related constraints and iconic subjects were generally expressed more clearly in this topic. Yet designers also described difficulties when prompts required multi-layered symbolism or complex cultural image combinations. In these cases, the model sometimes struggled to preserve both visual consistency and semantic coherence, implying a need for stronger contextual support from cultural knowledge and clearer semantic constraints within prompts (Guo et al., 2024).

Topic 4 (International Landmark & Cross-Cultural Design) encouraged broader combinations of landmarks and cross-cultural framing. Participants described this topic as flexible for composing multi-symbol concepts aimed at international audiences. However, both expert evaluation and interviews indicated that controllability varied more in this topic than in others. Designers often needed additional prompt tuning and manual adjustment to improve the logical relationship among symbols, especially when cross-linguistic cultural references were involved (Karpouzis, 2024).

Overall, the five topics provide a workable structure for decomposing souvenir prompts into explicit elements. At the same time, the results highlight a recurring limitation observed across topics. Concrete visual attributes are easier to control than narrative meaning and cross-cultural semantic integration. Future work should therefore strengthen cultural knowledge support and prompt mechanisms for narrative and intercultural coherence, while keeping the templates usable for designers in practice.

## 6. 2. Theoretical implications

This study proposes a data-driven way to structure prompt topics for cultural tourism souvenir design. Using high-quality Midjourney works as the starting point, we combined LDA topic modeling (Blei et al., 2003) with a Delphi-based card sorting procedure (Hsu & Sandford, 2007; Okoli & Pawlowski, 2004) to derive five representative topics. Compared with approaches that rely mainly on designers' intuition (Oppenlaender, 2022), this pipeline makes the topic construction process more explicit and easier to replicate in similar cultural design contexts (Oppenlaender et al., 2024).

The study also clarifies how the topic templates can be used in early-stage concept generation. We formalized a four-step procedure that includes topic selection, element assembly, prompt composition, and iterative generation. The workshop and follow-up interviews suggest that this procedure helped participants keep prompts more closely aligned with a selected design direction and made revisions more traceable during iteration.

From the perspective of organizing culture-related design content, this study suggests that structured prompting can be used to organize originally scattered cultural references and related expressive content into clearer prompt components. When such content is written into prompts in a more structured manner, designers can specify more explicitly which cultural elements are included in the prompt and how they are described. At the same time, the findings also reveal clear boundaries: structured templates were associated with clearer

expression of concrete and visually grounded attributes such as souvenir type, color, and visual style, but remained limited in supporting narrative-level expression and cross-cultural semantic integration.

This pattern can also be understood in relation to the semantic characteristics of text-to-image systems. Prior research on prompt design has shown that clearer and more concrete prompt cues generally support more stable specification and controllability in the generation process (Liu & Chilton, 2022; Oppenlaender, 2024). Studies on iterative prompting have likewise shown that, although prompt revision and editing can make the revision process more traceable, they also reveal that generated outputs depend heavily on how content is specified at the prompt level (Don-Yehiya et al., 2023; Guo et al., 2024). At the same time, research on cultural translation and cultural representativeness suggests that when generative tasks involve multilayered cultural meanings, place-based knowledge, or cross-cultural relations, the outputs are more likely to exhibit simplification, symbolic generalization, or semantic fragmentation rather than genuine narrative integration (Lyu et al., 2024; Zhang et al., 2024). From this perspective, the theoretical significance of the present study does not lie in arguing that cultural narratives should simply be reduced to visual symbols. Rather, it lies in showing that, in AI-assisted early-stage ideation, complex cultural content often needs to be decomposed into more explicit and operable prompt units before it can enter a structured human–AI interaction process. In this study, that decomposition was organized precisely through predefined slots. Predefined slots not only make it clearer which elements need to be specified first and which relations need to be expressed first, but also implicitly shape the path along which ideation develops. At the same time, this decomposition reveals an important theoretical trade-off: on the one hand, it increases the traceability and controllability of design intent and helps designers quickly focus on a given theme; on the other hand, because predefined slots delimit in advance the range of content that can be organized and combined, the ideation process is more likely to be directed toward relatively expected combinations, thereby narrowing the space for open-ended divergence and unexpected associations in the early stages of concept generation, and to some extent affecting the development of idea diversity and design originality.

Accordingly, the theoretical significance of this study lies in reconceptualizing prompts in cultural tourism souvenir design as an intermediary organizational structure that connects design intent, cultural cues, element organization, and subsequent revision. From this perspective, the study proposes a preliminary analytical and design framework for culturally oriented design tasks, offering a clearer theoretical entry point for understanding how structured prompting supports designers in organizing prompts, coordinating revision, and collaborating with generative AI during early-stage ideation. At the same time, it also provides a starting point for future research on structured prompting, designer–AI co-creation, and the organizational logic of prompt-based interaction.

### **6. 3. Practical Implications**

The proposed topic-based prompt templates serve as a practical scaffold for cultural tourism souvenir ideation with Midjourney and similar text-to-image systems. For design practice, the framework provides a simple way to organize prompts by making key elements explicit,

such as the souvenir type, cultural references, color, visual style, and related descriptive content. This makes it easier to locate which part of a prompt should be adjusted when results are off-target, and it supports more targeted iteration during concept exploration. (Pan et al., 2025; Wang et al., 2024).

For design education, the workflow can be incorporated into studio or project-based courses as a prompt-writing scaffold. Students can be asked to complete each component of a selected topic template, generate images, and then revise one component at a time to observe how changes affect the outputs. This structure helps instructors combine prompt construction training with the identification and organization of cultural references, as well as critical reflection on AI-generated outputs, rather than encouraging students to accept such outputs at face value (Garg et al., 2025; Lou, 2023). The same approach can also be used in short workshops to help beginners develop basic prompt skills and an analytical awareness of culture-oriented design tasks.

For cultural tourism communication and destination branding, the templates provide a way to organize local cultural references and commonly used visual elements into a structured prompt format for generating early-stage visual concepts. Agencies or creative teams can use the framework to build a shared prompt library that documents commonly used landmarks, motifs, and style settings for a destination. This can support consistent concept development and facilitate collaboration across designers, content creators, and local stakeholders in the early phases of souvenir and cultural-creative product development (Gonçalves et al., 2022).

#### **6. 4. Limitations and Future Research Directions**

This study still has several limitations. First, the participant sample consisted of 22 Chinese designers with design backgrounds. Their relatively similar linguistic and cultural backgrounds may have influenced how they understood and evaluated cultural symbols, place-based knowledge, and the generated outputs. In addition, the Midjourney Explore data used to construct the prompt framework were drawn from highly engaged and curated cases presented on the platform. Although such data were useful for identifying more mature patterns of prompt organization, they may also reflect platform-preferred aesthetic styles and representational tendencies, and therefore may not fully capture the broader and more diverse prompting practices found in cultural tourism souvenir design. Meanwhile, all generation tasks were conducted on a single text-to-image platform, Midjourney, so the findings were still shaped by the capabilities of that specific platform. Future research could further validate the framework across participants with different cultural backgrounds and levels of design experience, while also incorporating more diverse data sources and multiple platforms to examine its applicability in a wider range of design contexts.

Second, the workshop evaluation did not include a free-form prompt comparison condition. Therefore, the findings at this stage should not be interpreted as direct empirical evidence that structured prompting is more effective than free-form prompting. Rather, they should be understood mainly as observations of the application patterns, expression characteristics, and limitations of different topic-based templates under a shared structured-prompt condition. Future studies could introduce comparison groups to examine how

different prompting formats differ in terms of controllability, clarity of expression, and user experience.

Third, the interview findings suggested that highly predefined slots may constrain exploratory thinking to some extent and make design ideation more likely to proceed along relatively expected paths. This indicates that, while structured prompting helps maintain design focus, it may also carry a risk of design convergence. However, this study did not directly test the specific degree or scope of this convergent effect through a dedicated comparative design or quantitative measures. Future research could compare different prompting formats and different stages of use in order to better understand how structured prompting balances design focus and creative divergence, and to further explore more flexible ways of organizing templates.

Finally, the five topic-based templates proposed in this study should still be regarded as a preliminary structured framework derived from the current dataset and expert classification results. Although Table 4 already provides an initial organization of the structural components that recur across topics and their topic-specific emphases, further alignment of these repeated slots, refinement of a more general template layer, and greater consistency in how the same labels are interpreted across different topics still require additional work. Future research could therefore build on the present framework by further aligning cross-topic slots and refining a more general template layer, in order to improve the framework's consistency, reusability, and extensibility.

## 6. 5. Conclusion

This study proposes a data-driven structured prompting framework for cultural tourism souvenir design. Drawing on Midjourney prompt–image pairs, LDA topic modeling, and expert Delphi card sorting, the framework develops five topic-based prompt templates. It was further examined through a design workshop. The findings show that the framework offers clearer organizational and control advantages for the expression of concrete and visually specifiable elements, such as souvenir type, color, and visual style, while also revealing clear boundaries in narrative meaning, cultural depth, and cross-cultural semantic integration. Overall, the study provides a more organized and traceable path for human–AI collaboration in early-stage ideation for cultural tourism–oriented design tasks, and offers an initial reference point for future research on structured prompting and designer–AI co-creation.

## References

1. Bengfort, B., Bilbro, R., & Ojeda, T. (2018). *Applied text analysis with Python: Enabling language-aware data products with machine learning*. O'Reilly Media.
2. Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of machine Learning research*, 3, 993–1022.
3. Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., & Brunskill, E. (2021). *On the opportunities and risks of foundation models*. <https://arxiv.org/abs/2108.07258>

4. Brade, S., Wang, B., Sousa, M., Oore, S., & Grossman, T. (2023). Promptify: Text-to-image generation through interactive prompt exploration with large language models. *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, 1–14. <https://doi.org/10.1145/3586183.3606725>
5. Chai, C. P. (2023). Comparison of text preprocessing methods. *Natural Language Engineering*, 29(3), 509–553. <https://doi.org/10.1017/S1351324922000213>
6. Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J., & Blei, D. (2009). Reading tea leaves: How humans interpret topic models. *Advances in neural information processing systems*, 22.
7. Chuang, J., Manning, C. D., & Heer, J. (2012). Termite: Visualization techniques for assessing textual topic models. *Proceedings of the international working conference on advanced visual interfaces*, 74–77. <https://doi.org/10.1145/2254556.2254572>
8. Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and brain sciences*, 24(1), 87–114. <https://doi.org/10.1017/S0140525X01003922>
9. Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management science*, 9(3), 458–467.
10. Don-Yehiya, S., Choshen, L., & Abend, O. (2023). Human learning by model feedback: The dynamics of iterative prompting with midjourney. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 4146–4161. <https://doi.org/10.18653/v1/2023.emnlp-main.253>
11. Duan, Z. Y., Tan, S. K., Choon, S. W., & Zhang, M. Y. (2023). Crafting a place-based souvenir for sustaining cultural heritage. *Heliyon*, 9(5), Article e15761. <https://doi.org/10.1016/j.heliyon.2023.e15761>
12. El Abed, M., & Castro-Lopez, A. (2024). The impact of AI-powered technologies on aesthetic, cognitive and affective experience dimensions: a connected store experiment. *Asia Pacific Journal of Marketing and Logistics*, 36(3), 715–735. <https://doi.org/10.1108/APJML-02-2023-0109>
13. Garg, A., Soodhani, K. N., & Rajendran, R. (2025). Enhancing data analysis and programming skills through structured prompt training: The impact of generative AI in engineering education. *Computers and Education: Artificial Intelligence*, 8, 100380. <https://doi.org/10.1016/j.caeai.2025.100380>
14. Gîrbacia, F. (2024). An Analysis of Research Trends for Using Artificial Intelligence in Cultural Heritage. *Electronics*, 13(18), 3738. <https://doi.org/10.3390/electronics13183738>
15. Gonçalves, A. R., Dorsch, L. L. P., & Figueiredo, M. (2022). Digital tourism: an alternative view on cultural intangible heritage and sustainability in Tavira, Portugal. *Sustainability*, 14(5), 2912. <https://doi.org/10.3390/su14052912>
16. Guo, Y., Shao, H., Liu, C., Xu, K., & Yuan, X. (2024). Prompthis: Visualizing the process and influence of prompt editing during text-to-image creation. *IEEE Transactions on Visualization and Computer Graphics*. <https://doi.org/10.1109/TVCG.2024.3408255>
17. Hsu, C.-C., & Sandford, B. A. (2007). The Delphi technique: making sense of consensus. *Practical assessment, research, and evaluation*, 12(1). <https://doi.org/10.7275/pdz9-th90>
18. Karpouzis, K. (2024). Plato's shadows in the digital cave: Controlling cultural bias in generative AI. *Electronics*, 13(8), 1457. <https://doi.org/10.3390/electronics13081457>
19. Kim, S. Y. (2023). Investigating the effect of customer-generated content on performance in online platform-based experience goods market. *Journal of Retailing and Consumer Services*, 74, 103409. <https://doi.org/10.1016/j.jretconser.2023.103409>
20. Kumar, A., Chakraborty, S., & Bala, P. K. (2023). Text mining approach to explore determinants of grocery mobile app satisfaction using online customer reviews. *Journal of Retailing and Consumer Services*, 73, 103363. <https://doi.org/10.1016/j.jretconser.2023.103363>
21. Lau, J. H., Baldwin, T., & Newman, D. (2013). On collocations and topic models. *ACM Transactions on Speech and Language Processing (TSLP)*, 10(3), 1–14. <https://doi.org/10.1145/2483969.2483972>

22. Littrell, M. A., Anderson, L. F., & Brown, P. J. (1993). What makes a craft souvenir authentic? *Annals of tourism research*, 20(1), 197–215. [https://doi.org/10.1016/0160-7383\(93\)90118-M](https://doi.org/10.1016/0160-7383(93)90118-M)
23. Liu, V., & Chilton, L. B. (2022). Design guidelines for prompt engineering text-to-image generative models. *Proceedings of the 2022 CHI conference on human factors in computing systems*, 1–23. <https://doi.org/10.1145/3491102.3501825>
24. Lou, Y. Q. (2023). Human Creativity in the AIGC Era\*. *She Ji—the Journal of Design Economics and Innovation*, 9(4), 541–552. <https://doi.org/10.1016/j.sheji.2024.02.002>
25. Luther, T., Kimmerle, J., & Cress, U. (2024). Teaming up with an AI: Exploring human–AI collaboration in a writing scenario with ChatGPT. *AI*, 5(3), 1357–1376. <https://doi.org/10.3390/ai5030065>
26. Lyu, Y., Shi, M., Zhang, Y., & Lin, R. (2024). From image to imagination: Exploring the impact of generative AI on cultural translation in jewelry design. *Sustainability*, 16(1), 65. <https://doi.org/10.3390/su16010065>
27. Mahdavi Goloujeh, A., Sullivan, A., & Magerko, B. (2024). Is It AI or Is It Me? Understanding Users' Prompt Journey with Text-to-Image Generative AI Tools. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3613904.3642861>
28. McComb, C., & Jablokow, K. (2022). A conceptual framework for multidisciplinary design research with example application to agent-based modeling. *Design Studies*, 78, 101074. <https://doi.org/10.1016/j.destud.2021.101074>
29. Midjourney. (2024a). *Legacy features*. Retrieved February 28, 2025, from <https://docs.midjourney.com/hc/en-us/articles/33329788681101-Legacy-Features>
30. Midjourney. (2024b). *Prompt basics*. Retrieved February 28, 2025, from <https://docs.midjourney.com/hc/en-us/articles/32023408776205-Prompt-Basics>
31. Midjourney. (2024c). *Website overview*. Retrieved February 28, 2025, from <https://docs.midjourney.com/hc/en-us/articles/33329460426765-Website-Overview>
32. Mimno, D., Wallach, H., Talley, E., Leenders, M., & McCallum, A. (2011). Optimizing semantic coherence in topic models. *Proceedings of the 2011 conference on empirical methods in natural language processing*, 262–272. <https://aclanthology.org/D11-1024/>
33. Miner, G. D., Elder, J., Fast, A., Hill, T., Nisbet, R., & Delen, D. (2012). *Practical text mining and statistical analysis for non-structured text data applications*. Academic Press.
34. Niederberger, M., & Köberich, S. (2021). Coming to consensus: The Delphi technique. *European Journal of Cardiovascular Nursing*, 20(7), 692–695. <https://doi.org/10.1093/eurjcn/zvab059>
35. Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: an example, design considerations and applications. *Information & management*, 42(1), 15–29. <https://doi.org/10.1016/j.im.2003.11.002>
36. Oppenlaender, J. (2022). The creativity of text-to-image generation. *Proceedings of the 25th international academic mindtrek conference*, 192–202. <https://doi.org/10.1145/3569219.3569352>
37. Oppenlaender, J. (2024). A taxonomy of prompt modifiers for text-to-image generation. *Behaviour & Information Technology*, 43(15), 3763–3776. <https://doi.org/10.1080/0144929X.2023.2286532>
38. Oppenlaender, J., Linder, R., & Silvennoinen, J. (2024). Prompting AI art: An investigation into the creative skill of prompt engineering. *International journal of human-computer interaction*, 1–23. <https://doi.org/10.1080/10447318.2024.2431761>
39. Pan, S. M., Anwar, R. B., Awang, N. N. B., & He, Y. N. (2025). Constructing a Sustainable Evaluation Framework for AIGC Technology in Yixing Zisha Pottery: Balancing Heritage Preservation and Innovation. *Sustainability*, 17(3), Article 910. <https://doi.org/10.3390/su17030910>
40. Paul, C. L. (2008). A modified delphi approach to a new card sorting methodology. *Journal of Usability Studies*, 4(1), 7–30.
41. Proctor, R. W., & Schneider, D. W. (2018). Hick's law for choice reaction time: A review. *Quarterly Journal of Experimental Psychology*, 71(6), 1281–1299. <https://doi.org/10.1080/17470218.2017.1322622>

42. Putra, S. J., Gunawan, M. N., & Suryatno, A. (2018). Tokenization and n-gram for indexing Indonesian translation of the Quran. *2018 6th International Conference on Information and Communication Technology (IColCT)*, 158–161. <https://doi.org/10.1109/IColCT.2018.8528762>
43. Qiu, L., Rahman, A. R. A., & Dolah, M. S. b. (2024). The role of souvenirs in enhancing local cultural sustainability: A systematic literature review. *Sustainability*, *16*(10), 3893. <https://doi.org/10.3390/su16103893>
44. Rajcic, N., Llano Rodriguez, M. T., & McCormack, J. (2024). Towards a diffractive analysis of prompt-based generative ai. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 1–15. <https://doi.org/10.1145/3613904.3641971>
45. Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). *Hierarchical text-conditional image generation with CLIP latents*. arXiv. <https://arxiv.org/abs/2204.06125>
46. Rapp, A., Di Lodovico, C., Torrielli, F., & Di Caro, L. (2025). How do people experience the images created by generative artificial intelligence? An exploration of people's perceptions, appraisals, and emotions related to a Gen-AI text-to-image model and its creations. *International Journal of human-computer studies*, *193*, 103375. <https://doi.org/10.1016/j.ijhcs.2024.103375>
47. Reese, T., Segall, N., Nesbitt, P., Del Fiol, G., Waller, R., Macpherson, B. C., Tonna, J. E., & Wright, M. C. (2018). Patient information organization in the intensive care setting: expert knowledge elicitation with card sorting methods. *Journal of the American Medical Informatics Association*, *25*(8), 1026–1035. <https://doi.org/10.1093/jamia/ocy045>
48. Richards, G. (2018). Cultural tourism: A review of recent research and trends. *Journal of hospitality and tourism management*, *36*, 12–21. <https://doi.org/10.1016/j.jhtm.2018.03.005>
49. Righi, C., James, J., Beasley, M., Day, D. L., Fox, J. E., Gieber, J., Howe, C., & Ruby, L. (2013). Card sort analysis best practices. *Journal of Usability Studies*, *8*(3), 69–89.
50. Sievert, C., & Shirley, K. (2014). LDAvis: A method for visualizing and interpreting topics. *Proceedings of the workshop on interactive language learning, visualization, and interfaces*, 63–70.
51. Spencer, D. (2009). *Card sorting: Designing usable categories*. Rosenfeld Media.
52. Stevens, K., Kegelmeyer, P., Andrzejewski, D., & Buttler, D. (2012). Exploring topic coherence over many models and many topics. *Proceedings of the 2012 joint conference on empirical methods in natural language processing and computational natural language learning*, 952–961.
53. Steyvers, M., & Griffiths, T. (2007). Probabilistic topic models. In *Handbook of latent semantic analysis* (pp. 439–460). Psychology Press.
54. Vecco, M. (2010). A definition of cultural heritage: From the tangible to the intangible. *Journal of cultural heritage*, *11*(3), 321–324. <https://doi.org/10.1016/j.culher.2010.01.006>
55. Wallach, H. M., Murray, I., Salakhutdinov, R., & Mimno, D. (2009). Evaluation methods for topic models. *Proceedings of the 26th annual international conference on machine learning*, 1105–1112. <https://doi.org/10.1145/1553374.1553515>
56. Wang, Y. L., Xi, Y. C., Liu, X. X., & Gan, Y. (2024). Exploring the Dual Potential of Artificial Intelligence-Generated Content in the Esthetic Reproduction and Sustainable Innovative Design of Ming-Style Furniture. *Sustainability*, *16*(12), Article 5173. <https://doi.org/10.3390/su16125173>
57. Weston, S. J., Shryock, I., Light, R., & Fisher, P. A. (2023). Selecting the number and labels of topics in topic modeling: A tutorial. *Advances in Methods and Practices in Psychological Science*, *6*(2), 25152459231160105. <https://doi.org/10.1177/25152459231160105>
58. Zhai, Y., Song, X., Chen, Y., & Lu, W. (2022). A study of mobile medical app user satisfaction incorporating theme analysis and review sentiment tendencies. *International Journal of Environmental Research and Public Health*, *19*(12), 7466. <https://doi.org/10.3390/ijerph19127466>
59. Zhan, J. T., Ai, Q. Y., Liu, Y. Q., Pan, Y. W., Yao, T., Mao, J. X., Ma, S. P., & Mei, T. (2024). Prompt Refinement with Image Pivot for Text-to-Image Generation. *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 941–954. <https://doi.org/10.18653/v1/2024.acl-long.53>

60. Zhang, L., Liao, X., Yang, Z., Gao, B., Wang, C., Yang, Q., & Li, D. (2024). Partiality and Misconception: Investigating Cultural Representativeness in Text-to-Image Models. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 1–25. <https://doi.org/10.1145/3613904.3642877>
61. Zhong, S., Huang, Z., Wen, W., Qin, J., & Lin, L. (2023). Sur-adapter: Enhancing text-to-image pre-trained diffusion models with large language models. *Proceedings of the 31st ACM International Conference on Multimedia*, 567–578. <https://doi.org/10.1145/3581783.3611863>
62. Zhu, Q., & Rahman, R. (2025). Authenticity in souvenir design integrating cultural features of Dunhuang's mural heritage: a qualitative inquiry. *Humanities and Social Sciences Communications*, 12(1), 1–11. <https://doi.org/10.1057/s41599-025-04710-5>

## Appendix A

### Basic Information

- 1.Name (Optional):
- 2.Educational Background:
- 3.Years of Design Study/Work Experience:
- 4.Major/Field of Study:

### Subjective Impressions and Experience Differences

- 5.In this cultural souvenir design task, what was your initial impression when using a structured prompt (Prompt) for the first time?
- 6.Do you feel that the Prompt guidance was helpful to your design process or did it impose certain limitations?

### Performance of Specific Structural Elements

- 7.Among the structural elements you used in the Prompt (e.g., landmark symbols, narrative embellishments, color expression, etc.), which were the most inspiring or helpful for your cultural souvenir design?
- 8.Did you find any structural elements unclear or hard to apply?

### Technical Challenges and Barriers

- 9.What practical operational or comprehension difficulties did you encounter while using a structured Prompt for AI-generated cultural souvenir design?
- 10.In your opinion, what aspects of this Prompt structure need the most improvement for cultural souvenir design: the content structure, language formulation, or Topic Alignment?

### Impact on Creative Process

- 11.Did this AI cultural souvenir design process change your original design direction or thinking approach? Could you share a representative example?

### Industry Application Barriers

- 12.In an industry like cultural souvenir design, which places a strong emphasis on cultural expression, what do you think is the greatest obstacle to the widespread use of structured Prompts?

### Creativity and Originality

- 13.Do you feel that using a structured Prompt enhanced your creativity in cultural souvenir design, or did it sometimes lead your design to become formulaic?
- 14.Could you share a design detail that sparked unexpected original inspiration during your use of the Prompt? How does this detail relate to the cultural tourism theme?

### Future Development Trends

- 15.How do you view the future application prospects of structured Prompts in cultural souvenir or broader cultural and creative design fields?
- 16.How would you like future Prompt systems to better serve cultural tourism-related design tasks?

### Open-Ended Suggestions and Comments

- 17.What advice would you give to other designers who are trying to use Prompts for cultural souvenir design?
- 18.Is there any additional experience, thought, or suggestion regarding this Prompt-assisted AI cultural design task that you would like us to know?

Table B Inter-rater reliability (Cronbach's  $\alpha$ ) of expert ratings across topics and dimensions

Topic	Dimension	Cronbach's $\alpha$
T0	Souvenir Types	0.918
	Landmarks and Cultural Elements	0.932
	Color Expression	0.956
	Visual and Design Style	0.951
	Narrative and Embellishment	0.922
	Overall Topic-Level Expression	0.926
T1	Souvenir Types	0.939
	Cultural Heritage and Identity	0.953
	Narrative and Embellishment	0.932
	Visual and Design Style	0.914
	Overall Topic-Level Expression	0.907
T2	Souvenir Types and Architectural Elements	0.874
	Local Cultural and Artistic Motifs	0.942
	Narrative and Embellishment	0.950
	Modern Visual Foundations	0.943
	Overall Topic-Level Expression	0.901
T3	Primary Iconography and Souvenir Types	0.930
	Narrative and Embellishment	0.927
	Hyperreality Factor	0.855
	Visual and Design Style	0.941
	Overall Topic-Level Expression	0.793
T4	Souvenir Types	0.967
	Global Landmarks and Culture Symbols	0.951
	Cross-Cultural Expression	0.940
	Narrative and Embellishment	0.926
	Environment and Background	0.973
	Overall Topic-Level Expression	0.931

Table C Some Results of the Qualitative Content Analysis

Meaning Units	Condensed Meaning Units	Codes	Categories
P01: The prompt has already defined my design direction, providing clear guidance.	Clear structure helps define design direction.	Structural Clarity	Prompt Structure Controllability
P04: Display products should highlight materials and background, while clothing focuses on color and pattern.	Structure enhances generation precision.		
P10: Too rigid a structure sometimes limits my creative freedom	Excessive structure constrains design freedom	Design Freedom	
P16: When designing freely, I mostly rely on intuition and experience, which may lack systematic direction.	Too much freedom may lead to loss of focus.		
P01 : The prompt structure must match semantics, otherwise the generated image just “looks similar”.	Structure must align with semantics.	Semantic Accuracy	
P06 : The more specific the structure, the more accurate the AI output.	Concrete description outperforms abstract terms.		
P10 : Emphasizing product type and visual style helps AI recognize more accurately.	Product type and visual style are easily identified.	Output Predictability	
P22: Especially for material and craft details, the AI’s understanding exceeded my expectations.	High controllability for Certain structures.		
P05 : When prompts include “function + texture + cultural elements”, the AI output is closest to my expectation.	Proper element combination improves output.		
P09 : Overly complex prompts like Topic 4 with too many elements weaken control.	Excessive complexity reduces controllability.	Design Quality	
P11 : The AI result was even better than I expected, making me consider using it as inspiration.	AI output provides valuable references.		
P12 : Certain structure items indeed help generate high-quality images.	Clear structure improves image quality.		
P03 : Topic 4 prompts included too many landmarks, leading to high information load and low controllability.	Too many prompts cause unstable output.	Perceived Usefulness	
P02 : At first, I was not used to it, but later found it helpful for controlling output.	Perceived usefulness of the tool.		
P06: Topic 4 prompts included too many landmarks, leading to high information load and low controllability.	Good thematic expression effect.	Efficiency Experience	
P22 : AI can quickly test different materials and crafts with low time cost and good results.	AI greatly improves design iteration and decisions.		
P19 : Structured prompts act like a design map, reducing detours.	AI helps designers quickly verify direction.		
P18 : It helps designers generate drafts fast, sparking more ideas and improving efficiency.	AI generates fast and saves effort.		Trust and Acceptance of the Tool

P06 : Niche cultural symbols are easily replaced by generic Chinese elements, causing design homogeneity.	Niche symbols tend to be generalized.		
P03: The AI generated a Ming Dynasty princess, but I intended a Qing Dynasty one.	AI output does not match intended symbols.	Cultural Symbol Recognition	
P21 : For “roof ridge beasts”, the AI understood it as a roof animal but missed detailed features and meaning.	AI struggles with detailed cultural semantics.		
P07 : In Topic 4, the AI just listed three landmarks without blending them—no sense of integration.	Cross-cultural prompts tend to stack elements.		
P13: A clear, detailed cross-cultural description improved the result.	Detailed cross-cultural prompts improve accuracy.	Cross-Cultural Coordination	Cultural Translation
P08 : Doing cross-cultural design taught me that “fusion” means finding commonality, not just stacking.	Fusion requires semantic alignment, not mere stacking.		
P01 : Current AI struggles with symbolic or narrative cultural nuances, producing only “lookalike” visuals.	Hard to convey complex cultural stories.		
P17 : Historical scenes often remain superficial visuals lacking cultural depth.	Output lacks cultural context depth.	Depth of Cultural Narratives	
P14 : AI mostly imitates images and cannot truly participate in cultural creation.	Stuck at imitation, low cultural participation.		
P01: I used to sketch first, now I write the prompt first and let AI generate; the process is reversed.	Workflow order fundamentally changed.	Design Workflow	
P11 : The design process now feels like constructing a blueprint through language.	Language-driven design workflow.		
P01 : The prompt reshaped my expression strategy.	Prompt restructured design strategy.		
P15 : I used to focus on my own ideas, now I follow prompt instructions more.	Shift from self-expression to task guidance.	Design Method	Design Thinking
P21 : I used to do concept-driven art; now I start from product usability.	Shift from abstract concepts to functional details.		
P10: The prompt gave me new ideas, like combining ceramics with new Chinese style—never tried before.	Prompt inspired new creative paths.	Creative Stimulation	
P01: I just wanted to make a palace lipstick and powder box, but the AI generated a surprising inner design.	AI output sparked unexpected inspiration.		

P15 : I believe structured prompts have great potential for cultural souvenir and creative design.	Broad application potential.		
P20 : Now I see souvenirs as interactive, tech-enabled, even digital experiences.	Souvenirs can be experiential products.	Practical Potential	
P05 : Structured prompts are best for quick prototyping and multi-style parallel design.	Suitable for industry workflows.		Market Demand
P16 : Adding interaction allows real-time feedback and adjustment.	Adds real-time feedback mechanism.		
P07 : Providing culture fusion examples would help designers and AI align better.	Include culture fusion method examples.	Tool Optimization	
P17: There should be a complete traditional cultural knowledge base.	Need a complete traditional cultural database.		
P11 : Overall, the prompt guidance is more helpful than restrictive.	Guidance meets expectations.		
P15 : The prompt changed my original design direction.	Guided the design plan.	Topic Guidance	Design Topic Control
P17: A clear definition of “souvenir type” avoided deviation from design focus.	Avoided design deviation.		