# From Myth to Interface: An AI-Augmented Interactive Visual System for Exploring Artifact Interactions in *Journey to the West*


Zeyu Chen
Zeyu.Chen21@student.xjtlu.edu.cn
Xi'an Jiaotong-Liverpool University
Suzhou, Jiangsu, China

Anqi Xie
Xi'an Jiaotong-Liverpool University
Suzhou, Jiangsu, China
anqixie2024@163.com

Lingyun Yu
Xi'an Jiaotong-Liverpool University
Suzhou, Jiangsu, China
lingyun.yu@xjtlu.edu.cn

Yu Liu*
Xi'an Jiaotong-Liverpool University
Suzhou, Jiangsu, China
yu.liu02@xjtlu.edu,cn

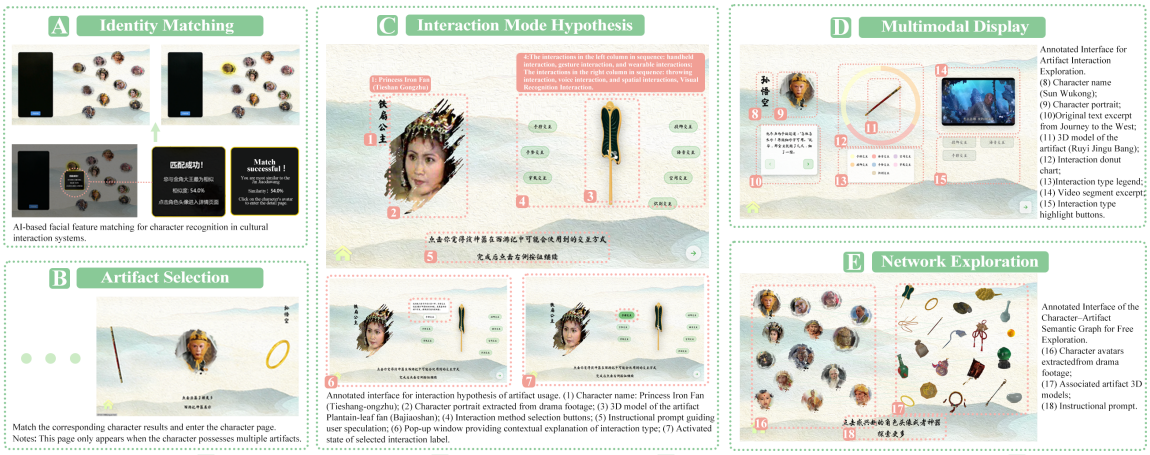




**Figure 1: System Interface. A Identity matching, B Artifact selection, C Interaction mode hypothesis, D Multimodal display, and E Network exploration.**


## Abstract

We developed an AI-assisted interactive visual system that allow users to explore *Journey to the West*'s magical artifacts. Although prior studies have explored its narrative and symbolism, the embedded technological metaphors and speculative interaction concepts in the novel remain underexplored. Leveraging ChatGPT-4 for deep semantic analysis, we extracted mappings between principal characters and their signature artifacts and embedded interaction patterns reflecting mythical affordances. These affordances guided the creation of a progressive, narrative-driven interactive visual system built on large displays. In a user study with 20 participants, the system achieved high usability scores and strong preference with facial-matching and multimodal artifact exploration feature. Our work provides a framework for reinterpreting and showcasing mythological narratives. The Appendix is available at our GitHub repository.



## CCS Concepts

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

## Keywords

Visualization, Journey to the West, Superpower, Cultural Heritage


## 1 Introduction

As one of the Four Great Classical Novels of Chinese literature, *Journey to the West* is not only a monumental work in vernacular fiction but also a narrative tapestry interwoven with mythological symbolism, religious allegory, and imaginative fantasy. Written in the 16th century during the Ming dynasty, the novel recounts the pilgrimage of Tang Sanzang and his disciples—Sun Wukong (Monkey King), Zhu Bajie, and Sha Wujing—as they travel westward in search of sacred scriptures. Across its 100 chapters, the narrative unfolds in an episodic structure that blends Buddhist, Taoist, and Confucian worldviews while vividly portraying a universe filled with gods, demons, monsters, and magical artifacts. The work has endured for over four centuries not only as literary art but as a



*Conference acronym 'XX, Woodstock, NY*



ACM ISBN 978-1-4503-XXXX-X/2018/06
https://doi.org/XXXXXXX.XXXXXXX

cultural system continually reinvented across opera, television, animation, and now digital media.

Recent decades have seen an array of digital reinterpretations of *Journey to the West*, including interactive visualizations, immersive installations, and transmedia exhibitions. At Peking University, for example, students have built data-driven comparisons between the historical monk Xuanzang's route and the fictional 81 trials, while others statistically analyzed demon battles to reveal hidden narrative patterns [17]. Artists and curators have transformed episodes like "Sun Wukong Borrowing the Iron Fan" into spatial installations, bringing mythic figures into four-dimensional performance environments [1]. Theme parks have embedded these stories into interactive rides, allowing visitors to embody the Monkey King in virtual battle [13, 18]. These creative efforts have succeeded in reenergizing the text, but most focus on its storyline, visual aesthetics, or cultural symbolism.

Our design motivation emerged from a different line of inquiry. Willett et al.'s work on *"superpowers as inspiration for visualization"* [20] demonstrated how fictional abilities from science fiction can inspire novel interaction metaphors. While such research emphasizes futuristic imaginaries, we were struck by a related but underexplored question: *could the classical mythological novels that have shaped our cultural imagination for centuries also embed forward-looking interaction paradigms?*

To investigate this possibility, we focus on *Journey to the West*—a canonical text in Chinese mythology and literature. We systematically extracted artifacts described in the novel, coded their associated modes of interaction, and presented them in an interactive visual system on large displays. A user study confirmed that this system not only supports engaging exploration, but also enriches participants' interpretive experience of the novel.

This work makes three contributions:

(1) a dataset of artifacts from *Journey to the West* and the interaction metaphors they embody;
(2) an interactive visualization system that enables the public to explore these artifacts and their associated interaction styles;
(3) empirical insights on how mythological narratives can be re-appropriated through visualization and AI technologies to foster playful, immersive, and culturally situated forms of engagement.

By shifting attention from science fiction to traditional mythology, our work broadens the design space for interaction metaphors. It highlights how classical cultural texts can be reinterpreted through contemporary visualization and AI techniques, opening new opportunities for digital humanities and interactive system design.

## 2 Related Work

### 2.1 *Journey to the West*

Recent study on *Journey to the West* (JttW) has expanded beyond traditional literary analysis, embracing interdisciplinary approaches. For example, Chan [4] examines cross-media adaptations of JttW, including literature, stage performances, films, and video games, highlighting its global influence and commercial potential. Fu [7] reviews the diverse interpretations from Ming and Qing commentaries to modern perspectives. Cai [12] shows the textual evolution of the story spans from oral traditions to digital media, reflecting its adaptability across different platforms.

*Journey to the West* has also inspired rich media reinterpretations. A prominent case is *Black Myth: Wukong*, a high-fidelity video game adaptation praised for its deep mythological grounding and visual innovation. Zhang and Zhou [23] interpret this adaptation through the lens of cultural empowerment, highlighting how such reimaginings engage young audiences with classical narratives via interactive technologies.

However, despite this growing body of work, existing digital adaptations and research tend to focus either on retelling the narrative or analyzing character networks and thematic arcs. Few efforts have examined the internal mechanics of interaction within the story—specifically how magical artifacts in JttW operate in terms of activation methods, user intent, and interface metaphor. Yet many of these artifacts appear ro anticipate contemporary HCI modes. The Gold-Banded Cudgel, for instance, lengthens at the single spoken word "Grow!", mirroring today's voice-activated assistants.

### 2.2 Digital Technologies for Cultural Heritage

The different digital technologies have transformed how we interpret and present cultural heritage, with classical literature forming a key subdomain. Techniques such as natural language processing, narrative graph extraction, and multimodal visualization now enable systematic analyses of canonical texts and artifacts. For example, Hoyt et al. [9] introduced ScripThreads, an interactive visual analytic system mapping character involvement and scene structure in film scripts to clarify narrative complexity; DeBuse et al. [5] developed automated plot-extraction that uses entity density patterns to segment scenes and quantify narrative flow.

In the domain of character analysis, Hoque et al. [8] proposed Portrayal, leveraging deep language models to automatically extract linguistic indicators and visualize character development without manual annotation. In Chinese classical literature, Liu et al.[15] applied computational methods to map character–monster encounters in *Journey to the West*, revealing power dynamics and narrative structures. Multimodal models such as CLIP have further bridged text and visual media by aligning literary descriptions with related imagery, enriching interpretations of artworks and historical documents [19].

Beyond text analytics, researchers now integrate AI and immersive technologies to deliver culturally grounded experiences. Zhang et al. [22] developed a depth-sensor and gesture-based Chinese shadow puppetry system, allowing users to perform traditional folktales in a modern interactive setting. Huang et al. [10] employed AI-generated content to visualize dream sequences from classical narratives, while Feng et al. [6] demonstrated augmented reality applications in Hainan's cultural creative industries to disseminate mythological elements in immersive formats.

Concurrently, the notion of "digital mythology" has highlighted how online cultural practices reshape collective mythic structures[2], and Linked Open Data initiatives have enhanced semantic accessibility of cultural archives [16]. Gamified adaptations, such as mobile games based on Turkic and Mongolian epics [24], and visual design

projects inspired by the *Shan Hai Jing* [21] illustrate the breadth of interactive, media-rich approaches to heritage dissemination.

Despite these advances, two critical gaps remain. First, no prior work has systematically analyzed *Journey to the West*'s activation methods and interaction metaphors from an HCI perspective. Second, there is a lack of framework to transform traditional literary classics into interactive visual systems. To address these gaps, our research systematically deconstructs the activation methods and operational logic of *Journey to the West*'s artifacts as interactive metaphors and builds an AI-augmented, multimodal visual system.

## 3 Design and Implementation

### 3.1 Design Considerations

Inspired by Willett et al. 's work [20], our project seeks to showcase the interaction methods involved by these magical tools and weapons, wielded by gods, demons, and spirits in *Journey to the West*. Our system is guided by the following design considerations (**DC**):

**DC1** **Narrative progression.** Rather than presenting information all at once, we emphasize a gradual, layered unfolding of content that mirrors the episodic and journey-like structure of the original narrative. This approach allows users to engage with the mythical world through a sense of discovery, fostering deeper cognitive and emotional resonance with the embedded interaction metaphors.

**DC2** **Intuitive interaction.** Given that *Journey to the West* is a widely recognized and culturally significant mythological narrative with a broad popular readership, it is essential that the corresponding interaction system maintain a high degree of accessibility and immediacy. We prioritize intuitive, easy-to-understand modes of engagement that minimize learning barriers, making the experience enjoyable and inclusive for diverse user groups.

**DC3** **Multimodal representation.** *Journey to the West* is not only a literary work but also a cultural phenomenon that has evolved through multiple media over time. Therefore, we consider it essential to support diverse forms of representation—textual, visual, and auditory—to do justice to the richness of its transmission and reinterpretation. A multimodal approach encourages users to interpret the narrative through multiple sensory and symbolic channels, amplifying the cultural depth of their experience.

### 3.2 System Implementation

Figure 2 illustrates the overall architecture of our interactive system. The development process presents three distinct but interrelated stages. These stages are: (1) **Text and video analysis**, (2) **Visual design**, and (3) **Narrative interaction design**. The system is developed using a lightweight architecture based on HTML, JavaScript, and Tailwind CSS, enabling modular interaction logic and cross-platform accessibility. Interactive 3D models are rendered using the `<model-viewer>` web component.

*3.2.1 Text and Video Analysis.* We employed a four-step pipeline—conducted by two researchers—to extract and encode interaction metaphors from JttW using both text and video sources:

(1) Text-Based Extraction of Character–Artifact Pairs.
- Segmentation: The novel was divided into chapter-aligned paragraphs to preserve narrative context and manage tokenizer limits.
- Prompt-Driven Extraction: We used GPT-4 with the template:
  `"Identify the artifact used by [character name] in this passage. Extract the description of how the artifact is activated."`
- Independent Validation: Both researchers performed extraction independently. Their combined outputs then underwent manual review and verification—discrepancies were discussed and resolved collaboratively, ensuring each character–artifact mapping was accurately finalized.

(2) Video Clip Collection for Artifact Usage.
- Scene Identification: For each confirmed artifact, we located and clipped its activation sequence from a mainstream TV adaptation[1], preserving visual context of its use (e.g., voice cues, gestures, spatial actions).

(3) Dual Coding of Interaction Methods.
- Taxonomy Definition: Each researcher first extracted initial interaction categories from established HCI frameworks [11, 20], then refined them through open coding of the text and video sources from JttW. To capture actions prevalent in traditional Chinese mythology, we added "Throwing" as a culturally specific modality. This yielded the final 7 interaction taxonomy : voice commands, spatial manipulation, wearable triggers, gestures, throwing actions, handheld controls, and visual recognition.
- Coding Procedure: Two researchers independently reviewed both text excerpts and video clips for every artifact and assigned one or more interaction methods from our taxonomy. Discrepancies were discussed and reconciled weekly in our research team; iterative rounds of joint review ensured unanimous consensus.

(4) Ternary Mapping Table Construction.
- Integration: We merged the validated character–artifact pairs with their jointly confirmed interaction methods into a comprehensive **Character–Artifact–Interaction** table.

In total, we extracted 85 artifacts associated with 65 mythological figures, 20 of which demonstrated interaction features analogous to premodern HCI modes (see Appendix B.1 for the full mapping table). The typology of interaction modes identified is as follows:

**Voice Interaction** (N=12): These artifacts respond to auditory commands, such as incantations, keywords, or spoken instructions. The user's voice acts as the trigger for functional transformations. The Gold Banded Cudgel (Ruyi Jingu Bang) responds to Sun Wukong's voice commands to shrink or expand. Similarly, activating spells for seals or talismans are also voice-triggered.

**Spatial Interaction** (N=7): Artifacts here are activated based on their position, orientation, or spatial relationship to the user or environment. The function may be proximity-based, directionally dependent, or triggered by movement within a specific space. For

[1]Specifically, the 2011 adaptation of *Journey to the West* directed by Zhang Jianya. This version was selected for its higher fidelity to the original text compared to other adaptations, as it preserves detailed scenes and character logic essential for interaction analysis.

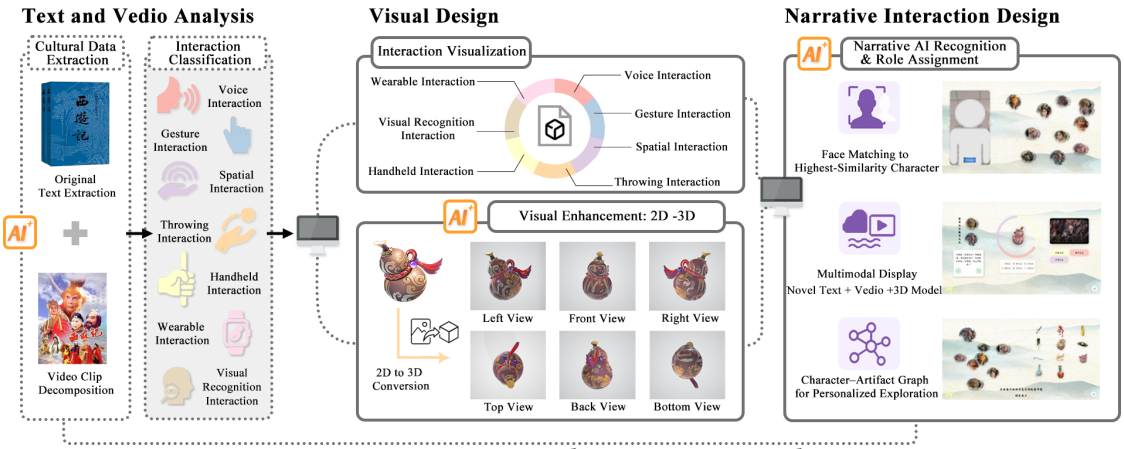



**Figure 2: Interactive Visual System Design Pipeline**

instance, both Red gourd (Zijin Hong Hulu) and Yin-yang vital principles jar (Yinyang Erqi Ping) need to be aimed at the enemy to interact.

**Wearable Interaction** (N=7): This category is triggered by being worn on the body or ingested, including items that act passively once attached or consumed, like Wind settling Pill (Dinfeng Dan) – ingested to remain still in strong wind and Invisibility Cloak – worn to become invisible.

**Gesture Interaction** (N=6): This type of artifact is driven by bodily gestures or posture. The motion itself serves as the command. Fanning with the Plantain leaf Fan (Bajiao Shan) to extinguish fire, shaking a Golden Bells (Zijin Lin) to summon spirits, or waving to summon the Somersault Cloud (Jindou Yun) are all representative gesture-based activations.

**Throwing Interaction** (N=6): Artifacts in this category are activated by releasing them through throwing motions, including both upward throws (e.g., summoning to the sky) and downward tosses (e.g., grounding or sealing). For example, Diamond Jade (Jingang Zhuo) throw attack and Human seed bag (Renzhong Dai) throw to suck in the enemy.

**Handheld Interaction** (N=4): Artifact is controlled through direct physical contact, typically by grasping or wielding it during use. This includes all close-range weapons manipulated through body strength or manual coordination. The Demon-quelling Pole (Jiuchi Dinpa) used by Zhu Bajie is a typical handheld combat weapon.

**Visual Recognition Interaction** (N=1): The artifact is activated by visually detecting or recognizing a target, often revealing hidden identities or truths, like Demon revealing Mirror (Zhaoyao Jin) – revealing demon form when facing the target.

*3.2.2 Visual Design.* After establishing the core dataset of characters and artifacts, the second phase focused on visual representation and user-perceivable structuring of the data. Our design approach was guided by **DC2** considering the wide range of users engaging with the system.

We employed a globally consistent color-coding scheme to differentiate each interaction type. Building on this foundation, we then rendered equal-segment donut charts as shown in Figure 1 (12) for every artifact—each sector corresponding to one interaction mode (as shown in Figure 1 (13)) and tinted according to its assigned hue. Also, we aligned the matching color tags along the video scrub bar (as shown in Figure 1 (15)): as the video playback reaches a segment associated with a given interaction, the corresponding tag on the scrub bar illuminate simultaneously. This synchronized highlighting allows users to instantly correlate on-screen moments with their respective interaction categories.

To enhance perceptual engagement, we performed 2D-to-3D transformations of selected magical artifacts (as shown in Figure 3). These reconstructions were based on keyframes or stills extracted from *Journey to the West* video adaptations. For this process, we used Tripo AI, a diffusion-based generative modeling engine capable of synthesizing 3D structures from 2D images through depth estimation and texture reconstruction. The generated models are optimized for multi-angle interactive exploration via WebGL, enabling users to view and manipulate artifacts in a spatial context. To ensure visual fidelity and usability, all models were manually reviewed for texture accuracy, structural clarity, and overall quality. In terms of interaction design, the system enables users freely to interact with the model using touch gestures, including rotation, lifting, two-finger zooming in and out, and clicking, among others (as illustrated in Figure 3 C). These operations enable users to observe the appearance details of the treasure from multiple angles, enhancing the immersion in understanding and participation.

*3.2.3 Narrative Interaction Design.* With the textual and visual foundations in place, the final phase focused on designing a narrative-centric interaction model, grounded in the previously outlined design considerations (**DC1**–**DC3**).

This system uses the Face++ API to match the facial features of users with characters from *Journey to the West*. After capturing real-time images through a camera, the system uses the detection interface to confirm the facial region, encodes it in base64 format, and then compares it with character images one by one using the compare interface for similarity analysis. Face++ returns a confidence score for each pair of images, and the system selects the character with the highest similarity as the matching result. The matching process is based on the similarity calculation of high-dimensional feature vectors (such as cosine similarity), achieving a personalized mapping entry for mythological characters based on visual features. This process allows for a low-barrier yet engaging entry into the system, anchoring the user within a mythological context


**Figure 3: AI-Augmented Visualization of Artifacts: 2D-to-3D Model Generation and Multi-Angle Representation. (A) Generate 3D models from screenshots of TV shows, using Yin Yang Dual Energy Bottle (Yinyang Erqi Ping) as an example. (B) Other 3D models. (C) Five gesture-based interaction modes.**

that feels both personal and meaningful. Once assigned a character, users gain access to a tailored interaction environment featuring all the magical artifacts associated with that figure (as shown in Figure 1 B). Each artifact can presented through a multimodal display, combining narrative excerpts from the novel, video scenes, and an interactive 3D model. This design enables users to explore the artifact's imagined functionality through multiple interpretive channels, supporting both **DC1** and **DC3**. Interaction is facilitated through a large-scale touchscreen interface, allowing users to navigate content fluidly, including zooming in on text, watching related video clips, or rotating 3D objects (as shown in Figure 1 D). Finally, users may enter a network visualization mode to explore the broader constellation of characters and artifacts, supporting open-ended discovery and reinforcing systemic understanding of the mythological ecosystem (as shown in Figure 1 E).

## 3.3 System Usage Flow and Interface Structure

The system usage workflow is designed to support a culturally interpretable and exploratory digital experience, structured through five sequential stages (see Figure 1): A identity matching, B artifact selection, C interaction mode hypothesis, D multimodal display, and E network exploration. These stages are interconnected through a nonlinear navigation flow, enabling both guided and self-directed pathways for users to explore traditional cultural content through modern interaction paradigms.

*3.3.1 Identify Matching Interface.* The interaction journey begins with a facial recognition module that personalizes the user experience by connecting the user with a character from *Journey to the West*. Upon camera access, the system employs the Face++ API to extract facial feature vectors and matches them with character portraits in the database. This identification result serves as both a narrative anchor and the entry point into the system.

As shown in Figure 1 A, this module comprises 3 stages: (1) **Image Capture Interface**. Users are guided to align their face within the frame for feature extraction; (2) **Matching Feedback Screen**. The most similar character is highlighted visually (e.g., with a yellow border) among all available avatars, providing clear and immediate feedback; (3) **Result Confirmation Page**. A pop-up confirms successful matching. By clicking the selected character's avatar, users enter the corresponding interaction environment, thereby initiating a culturally contextualized and emotionally resonant journey.

*3.3.2 Interaction Mode Hypothesis Interface.* After matching a character, the user is directed to the divine artifact interface (see Figure 1) C. When an artifact is chosen, the system does not immediately reveal its function. Instead, it presents a transitional "Interaction Mode Hypothesis" page. This interface uses a two-step "hypothesis-then-reveal" sequence, which applies the predict-observe-explain learning model to improve user retention. First, users are prompted to form a hypothesis by considering the question: "what modern human-computer interaction functions does this divine artifact have?" To answer, users can select one or more interaction modes, such as handheld, throwing, voice, and spatial. To assist their choice, each mode is a cue button that, when hovered over or clicked, displays a pop-up with a concise definition. A selected mode is highlighted to indicate it is part of the user's hypothesis (as seen in Figure 1 ⑦). This entire phase emphasizes active thinking rather than correctness. Only after the user submits their hypothesis does the system proceed to the next interface, where it reveals the artifact's actual interaction method, thereby closing the cognitive feedback loop.

*3.3.3 Interactive Artifact Exploration Interface.* Following the reasoning phase, users enter a rich, multimodal display interface (as illustrated in Figure 1 D) that contextualizes the artifact through linked text, media, and visualization. This interface integrates three content modalities:

- Narrative texts from the original novel,
- Interactive 3D artifact model,
- Clips from film/TV adaptations,
- A donut chart showing recognized interaction modes.

*3.3.4 Characters-Artifacts Network.* In the final stage, users can switch from guided learning to open-ended discovery - a semantic network linking all characters and artifacts in the system. Users can click on the character's avatar to learn which artifacts the character possesses, and then further explore the artifacts they are interested in. Users can also click on the artifact model to directly explore the artifact themselves about which they want to know more.

As shown in Figure 1 E, this interface includes: (1) A panel of character avatars (left), (2) A panel of artifact icons (right). Node size is dynamically scaled based on the number of artifacts associated with each character, and users may click on any node to explore its connections freely. This network supports bidirectional navigation—users may start from a character to view all related artifacts or from an artifact to trace its narrative ownership. Additionally, a navigation bar at the bottom allows users to return to the identity matching interface at any point. This structure enables a nonlinear cultural exploration.

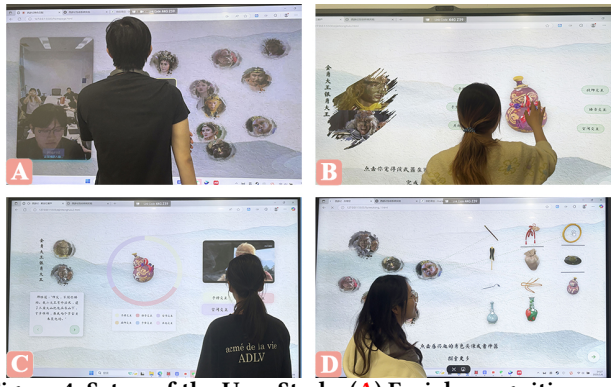



**Figure 4: Setup of the User Study. (A) Facial recognition and character matching; (B) Interaction mode hypothesis; (C) Multimodal display; (D) Network exploration.**

## 4 User Study

The purpose of this user study was to evaluate whether our interactive system—designed around premodern cultural artifacts from *Journey to the West*—could offer users a fresh and accessible way of interpreting classical literature through intuitive interaction and visual storytelling. We conducted the user study in which participants interacted with the full system deployed on a 65-inch touchscreen equipped with a camera, allowing access to facial recognition, 3D model manipulation, multimodal media exploration, and dynamic interaction reasoning (as shown in Figure 4) Our study was approved by school ethical committee.

### 4.1 Procedure

The user study was conducted in three stages and the testing time lasted from 10-20 minutes:

**Stage 1: Pre-experience Questionnaire and Introduction.** Participants first completed a short demographic and background questionnaire, which collected basic information such as age, gender, prior familiarity with *Journey to the West*, exposure to interactive systems.Afterward, a brief introduction to the project background and system objectives was provided, along with a walkthrough of the five key modules. This ensured that participants understood the purpose of the study and had sufficient context to engage with the system.

**Stage 2: System Interaction Session.** Participants then independently interacted with the system, without operational guidance. The process included: first undergoing facial recognition to obtain their mythic persona, then hypothesizing the Interaction mode of one artifact, next exploring its multimodal display (text, video, 3D model, donut chart), and finally roaming the character–artifact semantic map. Two observers recorded the duration of each user's experience, as well as behavioral patterns and notable hesitations or errors.

**Stage 3: Post-experience Feedback.** Upon completing the interaction session, participants filled out a questionnaire based on a 5-point Likert scale. The questionnaire design drew inspiration from standardized tools such as the SUS (System Usability Scale) [3] and UEQ (User Experience Questionnaire) [14], but was adapted to the goals of this study. Key focus areas covered ease of use and clarity of system modules, enjoyment and engagement with interactive features, cultural understanding and interpretation, and feedback on individual components such as 3D models, visual graphs, and dynamic interaction highlights. Finally, each participant took part in a short semi-structured interview, during which they were invited to share their overall impressions of the system, comment on its strengths and weaknesses, and suggest improvements.

**Table 1: Feedbacks of the overall system experience.**

| Questionnaire | | Mean | Score Distribution |
|---|---|---|---|
| **Q18** | Easy to use | 3.9 | 1 1 16 2 |
| **Q19** | Interestingness | 4.6 | 1 6 13 |
| **Q20** | Enhance immersion | 4.3 | 3 8 9 |
| **Q21** | Intention to continue using | 4.55 | 2 5 13 |
| **Q22** | Recommendation intention | 4.65 | 2 3 15 |
| **Q23** | Overall satisfaction | 4.6 | 8 12 |
| **Q24** | Knowledge enhancement | 4.45 | 1 1 6 12 |

1 (Strongly Disagree) 2 3 4 5 (Strongly Agree)

### 4.2 Participants

This study recruited 20 participants from local university (denoted as **P1–P20**) through social media who were interested in JttW, including 13 females and 7 males aged 18–30. They were from diverse fields including engineering, science, business, and humanities. Notably, **P13**, from a non-Chinese cultural background, had basic knowledge of Chinese and a keen interest in Chinese culture. All had prior interactive installation experience in museums, amusement parks, or exhibitions. Most of them (N = 16) reported moderate to high familiarity with the literary content of *Journey to the West*, while 4 participants indicated little prior knowledge.

### 4.3 Results and Findings

*4.3.1 Quantitative Results.* According to Table Table 1 (**Q18–Q24**), users gave generally positive feedback on the system as a whole. The system was rated as highly interesting (**Q19**, M=**4.6**) and immersive (**Q20**, M=**4.3**), with strong intention to continue using (**Q21**, M=**4.55**) and to recommend it to others (**Q22**, M=**4.65**). Overall satisfaction was also high (**Q23**, M=**4.6**). Notably, most participants agreed that the system helped them enhance their understanding of both *Journey to the West* and interactive design metaphors (**Q24**, M=**4.45**). However, the "ease of use" score was slightly lower (**Q18**, M=**3.9**), with a few low ratings (**1** or **2**), suggesting that some participants encountered usability difficulties. These responses indicated a need to improve onboarding instructions and simplify interface navigation.

From the results of **Q13-Q17** in Table 2, users showed high recognition of the design of each key module of the system. Among them, the multimodal display interface (**Q15**), which integrates original text excerpts, video clips, interaction diagrams, and 3D artifact models into a single interface, received the most enthusiastic user feedback. A majority of participants (N = 13) gave it a full score of **5**, strongly affirming that the AI-enhanced multimodal content significantly deepened their understanding and interest in the mythological artifacts. Other modules—such as facial recognition

**Table 2: Feedbacks on the designed features.**

| Questionnaire | | Mean | Score Distribution |
|---|---|---|---|
| **Q13** | Facial recognition and identify matching. | 4.4 | 1 10 9 |
| **Q14** | Interaction mode hypothesis. | 4.4 | 3 6 11 |
| **Q15** | Multimodal display (text, video, 3D model, donut chart). | 4.6 | 1 6 13 |
| **Q16** | Network exploration (character–artifact map). | 4.4 | 1 8 11 |
| 1 (Very Dissatisfied) 2 3 4 5 (Very Satisfied) | | | |
| **Q17 Preferences on design features** | | **Percent** | **Distribution** |
| A | Facial recognition and identify matching. | 75% | 15 |
| C | Interaction mode hypothesis. | 60% | 12 |
| D | Multimodal display (text, video, 3D model, donut chart). | 60% | 12 |
| E | Network exploration (character–artifact map). | 40% | 8 |

and character matching (**Q13**), interaction hypothesis (**Q14**), and free exploration (**Q16**)—also received high ratings (each with an average score of **4.4**), reflecting the overall ssatisfactory of design modules.

The results of the multiple-choice question (**Q17**) further confirmed users' preference for AI-enhanced modules: facial recognition and character matching was selected by 75% of participants as the most engaging feature, while multimodal display was selected by 60%, tied with interaction hypothesis. These results suggest that modules incorporating AI-driven personalization and rich media experiences were most appealing to users.

*4.3.2 Qualitative Results.* We conducted brief post-experience interviews with 20 participants, lasting 5 minutes on average. The AI-enhanced elements drew the most user interest. 75% participants stated that the 3D artifact interactions improved their exploratory interest. **P2** remarked, "I could freely explore the artifacts' details through touching and rotating." **P7** emphasized, "The interactive 3D models left a deeper impression on me than the plain text in novels or the 2D images in TV shows." The AI facial recognition also intrigued several users. **P1** found the system matched her blonde hair to "Golden Horn King," while **P3** noted the costume color matching to "Zhu Bajie." Notably, interviews revealed a recurring keyword: "cultural export." **P1** and **P3** cited the game *Black Myth: Wukong*, which has sparked overseas interest in the original JttW. They believe this system could serve as a vital auxiliary tool for foreign users to deeply explore culture of JttW through cross-cultural language adaptation. The strongest evidence came from international participant **P13**, who voluntarily engaged with the system for more than 30 minutes and expressed a clear intention to recommend it to others. She pointed out that the system's visual interaction and semantic association design effectively reduced cultural barriers for users, especially those from non-Chinese backgrounds, in understanding classical Chinese culture.

# 5 Discussion and Reflection

## 5.1 Reflection

In our interactive installation, AI serves as both a functional tool and a narrative medium, enhancing efficiency while deepening user engagement. Below is a critical examination of AI's application across different stages of the installation and its underlying objectives.

First, in the multimodal text-based key information extraction phase, AI primarily functions as an efficiency optimizer. Unlike rule-based or manual approaches, AI-driven natural language processing rapidly analyzes complex texts —*Journey to the West* source material to identify and extract core entities—characters, settings, and artifacts—with minimal human intervention. This automation ensures data richness and accuracy, forming a reliable foundation for subsequent interactive content. Second, the facial feature-based character matching system at the interaction entry point leverages AI for playful personalization. By mapping users' facial attributes to mythological characters, the system creates an immediate sense of identification, fostering emotional investment in the narrative. Lastly, AI enables the digitization of artifacts and supports 3D interaction by transforming static 2D film stills into manipulable 3D models. In this way, participants can actively engage with the artifact, enhancing their experience.

## 5.2 Limitation and Future Work

While this study demonstrates the potential of AI-augmented interaction and visualization in enhancing engagement with classical cultural narratives, several limitations remain, reflecting both technical constraints and areas for future exploration. First, the 2D-to-3D reconstruction of artifacts was conducted using existing commercial platforms (e.g., Tripo AI). While these tools enabled rapid generation of interactive models, the resulting geometric and textural accuracy was limited, particularly for complex or abstract artifact forms. More refined reconstruction could be achieved in future iterations by incorporating manual 3D modeling, neural radiance field (NeRF) techniques, or fine-tuned training on customized datasets. Second, the character matching process was based primarily on facial keypoint feature extraction using the Face++ API. Although effective for basic similarity mapping, the system does not currently support richer, more expressive forms of matching. Future improvements could include overlaying mythical character features (e.g., facial attributes of demons or gods) onto users' faces to generate immersive avatar-based mappings, which could enhance engagement and personalization. Third, the deployment environment was limited to a large-format touchscreen display. While this format was effective for public interaction and observation, it does not fully capture the potential embodied or immersive qualities of many artifacts. In future work, the system could be adapted for VR/AR environments to allow users to physically enact or inhabit the interaction metaphors—e.g., waving an actual fan, performing a throwing gesture, or experiencing spatial interaction through movement—thereby creating a more visceral and embodied experience. Finally, the user study was conducted in a controlled setting with a specific cultural background and limited participant diversity. Broader testing across different demographic and cultural contexts would be valuable to assess the system's accessibility and

interpretability for users unfamiliar with *Journey to the West* or Chinese mythological themes. Despite these limitations, the study offers a promising direction for bridging traditional cultural content and modern interactive paradigms, while also pointing to a rich set of technical and experiential opportunities for future development.

## 6 Conclusion

In this work, we developed an interactive system that reimagines *Journey to the West*—a canonical mythological novel and one of the Four Great Classical Novels of Chinese literature—through the lens of modern interaction design and multimodal visualization. The system integrates facial recognition, interaction reasoning, 3D artifact manipulation, audiovisual linkage, and semantic network exploration to construct an accessible and exploratory pathway for users to reinterpret the imaginative interaction logic embedded in traditional mythological narratives. Through a structured user study, we assessed the overall experience and examined the effectiveness and usability of key design modules—including character matching via facial recognition, analogical interaction reasoning, 3D artifact exploration, and multimodal interaction feedback. Results indicated that users not only found the system intuitive and enjoyable, but also reported that these individual components played a meaningful role in helping them understand the artifact's interactive logic and cultural significance.

This study demonstrates how interaction design can serve as a vehicle for cultural reinterpretation—not only enhancing accessibility to classical novels, but also transforming narrative engagement into a participatory and interpretive experience. By leveraging interaction as both a metaphor and a method, this work contributes to new forms of cultural interface design that bridge the imaginative frameworks of traditional literature with contemporary digital practice. Future work will focus on enhancing the technical fidelity of the system (e.g., high-precision 3D reconstructions, richer facial-avatar mappings), and exploring immersive extensions such as AR/VR environments that allow users to experience artifact interactions in more embodied and context-sensitive ways.

## References


[1] ALIEN Art Centre. 2019. Journey to the West – Flaming Mountains. https://www.alien.com.tw/u/en/pages/Journey_to_the_West_Flaming_Mountains. Accessed: 2025-05-15.

[2] Denis Artamonov. 2021. Digital Mythology: A New Direction In The Study Of Social Myths. In *Proceedings of the International Scientific Conference «Perishable and Eternal: Mythologies and Social Technologies of Digital Civilization-2021»*. Saratov State University. doi:10.15405/epsbs.2021.12.03.1

[3] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.

[4] Kenneth Chan. 2021. Transforming Tripitaka: Toward a (Buddhist) Planetary Ethics in Stephen Chow's Adaptation of Journey to the West. In *Sino-Enchantment: The Fantastic in Contemporary Chinese Cinemas*. Edinburgh University Press, Edinburgh, 222–244. doi:10.3366/edinburgh/9781474460842.003.0012

[5] Michael A. DeBuse and Sean Warnick. 2024. Plot Extraction and the Visualization of Narrative Flow. *Natural Language Engineering* 30, 3 (2024), 480–524. doi:10.1017/S1351324923000232

[6] Liying Feng, Edmund Ng Giap Weng, Liyao Ma, and Maozheng Fu. 2021. Research on the Application of Augmented Reality Technology in the Transformation and Development of Cultural and Creative Industries. *EAI Endorsed Transactions on Creative Technologies* 8, 28 (October 2021). doi:10.4108/eai.12-10-2021.171249

[7] Binbin Fu. 2024. A Review of Research on Journey to the West: A Multi-Dimensional Analysis from Ming and Qing Commentaries to Modern Interpretations. *Asia-Pacific Journal of Humanities and Social Sciences* 4, 3 (2024), 114–119. doi:10.53789/j.1653-0465.2024.0403.014

[8] Md Naimul Hoque, Bhavya Ghai, Kari Kraus, and Niklas Elmqvist. 2023. Portrayal: Leveraging NLP and Visualization for Analyzing Fictional Characters. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference (DIS '23)*. ACM, 74–94. doi:10.1145/3563657.3596000

[9] Eric Hoyt, Kevin Ponto, and Carrie Roy. 2014. Visualizing and Analyzing the Hollywood Screenplay with ScripThreads. *Digital Humanities Quarterly* 8, 4 (2014). http://www.digitalhumanities.org/dhq/vol/8/4/000190/000190.html

[10] Jiayang Huang, Yiran Chen, and David Yip. 2023. Dreaming Phantom in Immersive Experience: AIGC For Artistic Practice. In *Proceedings of the 16th International Symposium on Visual Information Communication and Interaction*. 1–3.

[11] Robert J.K. Jacob, Audrey Girouard, Leanne M. Hirshfield, Michael S. Horn, Orit Shaer, Erin Treacy Solovey, and Jamie Zigelbaum. 2008. Reality-based interaction: a framework for post-WIMP interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. Association for Computing Machinery, New York, NY, USA, 201–210. doi:10.1145/1357054.1357089

[12] Cai Jun. 2022. From brain text to electronic text: On the Textual Evolution of The Story of Westward Journey. *FilmTelevision and Theatre Review* (2022). https://api.semanticscholar.org/CorpusID:247263109

[13] Lagotronics Projects. 2021. Journey to the West – Wuhan Movie Park, China. Retrieved 2025-05-15 from https://www.lagotronicsprojects.com/en/projects/journey-to-the-west

[14] Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and usability engineering group*. Springer, 63–76.

[15] Chao-Lin Liu, Guan-Tao Jin, Hongsu Wang, Qing-Feng Liu, Wen-Huei Cheng, Wei-Yun Chiu, Richard Tzong-Han Tsai, and Yu-Chun Wang. 2015. Textual Analysis for Studying Chinese Historical Documents and Literary Novels. In *Proceedings of the ASE BigData & SocialInformatics 2015* (Kaohsiung, Taiwan) *(ASE BDSI '15)*. Association for Computing Machinery, New York, NY, USA, Article 30, 10 pages. doi:10.1145/2818869.2818912

[16] Valentina Pasqual and Francesca Tomasi. 2022. Linked Open Data per la valorizzazione di collezioni culturali: il dataset mythLOD. *AIB Studi* 62, 1 (2022), 149–168. doi:10.2426/aibstudi-13301

[17] Peking University. 2023. Peking University Students Use Journey to the West as a New Gateway to Chinese Traditional Culture. Retrieved 2025-05-15 from https://newsen.pku.edu.cn/news_events/news/campus/13043.html

[18] SDVoE Alliance. 2021. Christie Solutions and SDVoE Deliver Immersive Visuals in Journey to the West Theme Park. Retrieved 2025-05-15 from https://sdvoe.org/technology/case-studies/christie-solutions-and-sdvoe-deliver-immersive-visuals-in-journey-to-the-west-theme-park/

[19] Carmen Silva and Lídia Oliveira. 2024. Artificial Intelligence at the Interface between Cultural Heritage and Photography: A Systematic Literature Review. *Heritage* 7, 7 (2024), 3799–3820. doi:10.3390/heritage7070180

[20] Wesley Willett, Bon Adriel Aseniero, Sheelagh Carpendale, Pierre Dragicevic, Yvonne Jansen, Lora Oehlberg, and Petra Isenberg. 2022. Perception! Immersion! Empowerment! Superpowers as Inspiration for Visualization. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (Jan. 2022), 22–32. doi:10.1109/TVCG.2021.3114844

[21] Wenxuan Yang and Wenhua Sun. 2023. Research on Visual Design based on Chinese Mythological Elements: – Taking the Shan Hai Jing as an Example. *Journal of Innovation and Development* 3, 3 (Jul. 2023), 22–28. doi:10.54097/jid.v3i3.9787

[22] Hui Zhang, Yuhao Song, Zhuo Chen, Ji Cai, and Ke Lu. 2012. Chinese Shadow Puppetry with an Interactive Interface Using the Kinect Sensor. In *Proc. ECCV 2012 Workshops (Lecture Notes in Computer Science, Vol. 7583)*. Springer, Berlin, 352–361. doi:10.1007/978-3-642-33863-2_35

[23] Jijiao Zhang and Zixiang Zhou. 2025. Thick Power: the adaptations of Journey to the West and the interaction between the Chinese youth and Black Myth: Wukong from the perspective of Cultural Empowerment Theory. *International Journal of Anthropology and Ethnology* 9, 6 (2025). doi:10.1186/s41257-025-00129-7

[24] Bahar Özden Yıldırım, Bahadır Uçan, Çağatay Bilsel, and Feyza Yücebaş. 2024. Mobile Game Design Based on Turkic and Mongol Myths. *AURUM Journal of Engineering Systems and Architecture* 8, 1 (2024), 11–22. doi:10.53600/ajesa.1401252

# Appendix

## A Data Availability Statement

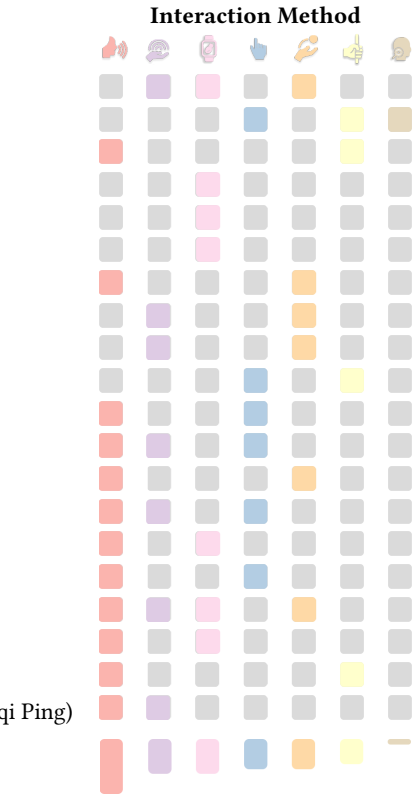
This project is intended solely for non-commercial academic research. Some visual materials (2D images and video clips) were sourced from publicly available television adaptations of *Journey to the West* and used only within the AI-based interaction reconstruction module. All media were functionally transformed for visualization purposes and fall within the scope of fair use as defined by Article 24 of the Copyright Law of the People's Republic of China.

## B Artifact and Interaction

### B.1 Interactive Classification

We reviewed the original work and TV series of *Journey to the West*, and summarized the interactive classification of 20 artifacts, which demonstrated premodern HCI modes. The boxes with different colors represent Voice Interaction, Spatial Interaction, Wearable Interaction, Gesture Interaction, Throwing Interaction, Handheld Interaction, Visual Recognition Interaction, and Non-Interaction.

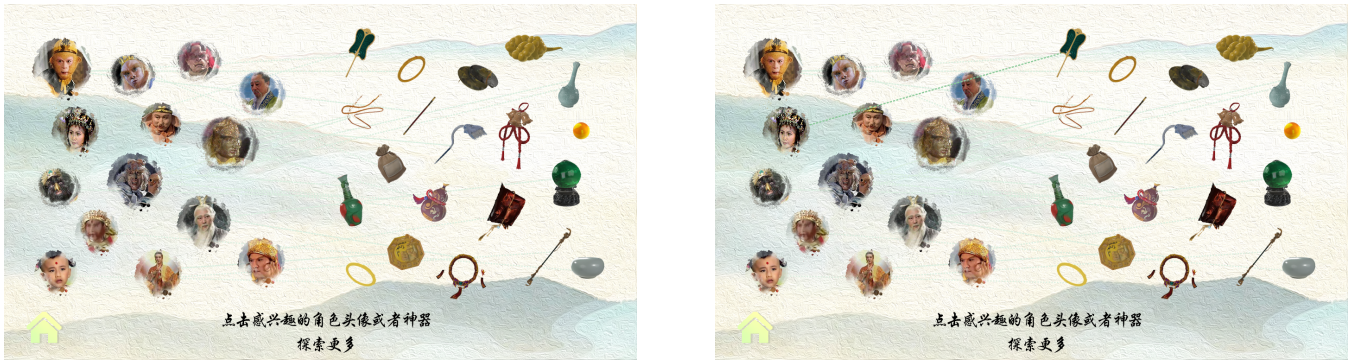
| Character | Artifact |
|---|---|
| Lute Heavenly King | **A1** Anti fire Cover (Bihuozhao) |
| Bodhisattva Manjusri | **A2** Demon revealing Mirror (Zhaoyao Jin) |
| Zhu Bajie | **A3** Demon-quelling Pole (Jiuchi Dinpa) |
| Dragon King of the Well | **A4** Face preserving Pearl (Dinyan Zhu) |
| Lao Jun | **A5** Diamond Jade (Jingang Zhuo) |
| | **A6** Wind settling Pill (Dinfeng Dan) |
| Lingji Bodhisattva | **A7** Flying Dragon Staff (Feilong Baozhang) |
| Yellow Eyebrow Ancestor | **A8** Gold cymbals (Jin Nao) |
| | **A9** Human seed bag (Renzhong Dai) |
| Sai Tai Sui | **A10** Golden Bells (Zijin Lin) |
| Princess Iron Fan | **A11** Plantain leaf Fan (Bajiao Shan) |
| Golden Horn King & Silver Horn King | **A12** Red gourd (Zijin Hong Hulu) |
| | **A13** Rope (Huangjin Sheng) |
| | **A14** Jade Vase (Yangzhiyu Jinpin) |
| Red Boy | **A15** Gold Band (Jin Gu) |
| Sun Wukong | **A16** Somersault Cloud (Jindou Yun) |
| | **A17** Gold Banded Cudgel (Ruyi Jingu Bang) |
| | **A18** Tight Band (Jingu Er) |
| King of Wuji Kingdom | **A19** White Jade Bowl (Baiyu Yu'er) |
| The Three Great Kings of Shituoling | **A20** Yin-yang Vital Principles Jar (Yinyang Erqi Ping) |




### B.2 3D Artifacts Models


**Figure 1: The 3D models of the 20 selected artifacts**

### B.3 Supplementary Interface Visuals




(a) Before activation (b) After activation: Plantain Leaf Fan (Bajiao Shan)

**Figure 2: Free exploration on the character-artifact interaction network. Users can explore the system freely before and after activation.**

## C Evaluation Instruments

### C.1 Questionnaire Overview

A concise list of post-experience survey questions covering multiple aspects of the system.

| Section | No. | Question |
|---|---|---|
| **Demographics** | **Q1** | Age Group |
| | **Q2** | Gender |
| | **Q3** | Educational Background |
| | **Q4** | Previous knowledge about JttW |
| | **Q5** | Previous knowledge about HCI methods |
| **Interactive Device Experience** | **Q6** | What interactive devices have you used? |
| | **Q7** | Which interaction types do you prefer? |
| | **Q8** | What features do you want in future devices? |
| | **Q9** | Are you willing to try new devices? |
| **Interest in JttW** | **Q10** | What are you most interested in? (Interest areas) |
| | **Q11** | How would you like to explore JttW? (Learning formats) |
| | **Q12** | Want to explore JttW via interaction? (Willingness) |
| **Project Module Feedback** | **Q13** | Satisfaction with facial recognition and character matching |
| | **Q14** | Satisfaction with interaction mode hypothesis classification |
| | **Q15** | Satisfaction with multimodal display (original text, video, 3D model, donut chart). |
| | **Q16** | Satisfaction with network exploration (through the character–artifact semantic map) |
| | **Q17** | The most engaging part of the experience |
| **Overall Experience** | **Q18** | Easy to use |
| | **Q19** | Interestingness |
| | **Q20** | Enhance immersion |
| | **Q21** | Intention to continue using |
| | **Q22** | Recommendation intention |
| | **Q23** | Overall satisfaction |
| | **Q24** | Knowledge enhancement |

## C.2 Interview Outline

A semi-structured interview guide conducted after the interactive experience, focusing on users' first impressions, sense of immersion, multimodal engagement, and interaction reasoning processes.

| Section | Question |
|---|---|
| **First Impressions** | What was your overall first impression of the system?<br>Which feature caught your attention first, and why?<br>What novelty did the AI face-matching add to your experience? |
| **Interaction Reasoning** | How did you guess the artifact's interaction method?<br>Did the "guess-then-reveal" flow enhance engagement or recall?<br>Which artifact stood out most, and for what reason?<br>Did you want to explore more artifacts/roles? What motivated you?<br>Did you feel immersed? Which part contributed most?<br>Is this format more memorable than reading or watching? |
| **3D/Multimodal Content** | Was the artifact info (text/image/diagram) clear and helpful?<br>How was the 3D interaction (rotate, zoom in, zoom out, lift, click)?<br>How did 3D models help compared to 2D or text? |
| **Cultural Understanding** | For non-Chinese users: Did it help cultural understanding?<br>Would you recommend this system to others? Why?<br>Would you want similar systems for other cultures or myths?<br>Summarize your overall impression and imagined future use. |
| **Usability Feedback** | Did you face any usability issues (e.g., unclear buttons)?<br>Was the interface layout clear? Any overload or gaps?<br>What features would you like to add in future versions?<br>How do you evaluate the system's visual and color design? |

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009