How AIGC Influences Audience Engagement in Chinese Museum Videos on Douyin?

Anonymous authors

No Institute Given

Abstract. Artificial intelligence-generated content (AIGC) is increasingly adopted by museums on short-video platforms, yet its actual impact on audience engagement remains unclear. This extended abstract presents a large-scale mixed-methods study analyzing 72,445 videos from 270 Chinese museum accounts on Douyin. While raw metrics suggest AI videos outperform non-AI videos, our regression analysis reveals this is largely a nominal effect driven by confounders like follower base and content orientation. However, specific AI applications - namely, artifact anthropomorphism and mixed-type AI - independently drive deeper cognitive engagement (comments, shares) and shift audience discourse toward specific cultural artifacts rather than general logistics. We propose an artifact-centric interaction perspective for AI-mediated cultural heritage communication.

Keywords: AI-generated content · Short videos · Douyin · TikTok · Audience engagement · Cultural heritage

1 Introduction

The rapid digital transformation of cultural heritage has pushed museums to transition from object-centered physical institutions to audience-oriented digital communication platforms [2]. Among emerging formats, short-video platforms like Douyin emphasize immediacy, entertainment, and algorithm-driven dissemination. Concurrently, generative artificial intelligence (AIGC) has reshaped media production, enabling museums to create digital humans, synthetic voices, and animated heritage artifacts at scale.

Despite the growing adoption of AIGC in cultural communication, its actual impact on audience engagement remains empirically ambiguous. Existing literature often treats AI as a binary "gimmick" and engagement as a monolithic metric, failing to capture the multidimensional nature of user interaction (e.g., shallow likes vs. deep cognitive discourse) [3].

This study aims to bridge this gap by investigating how specific AI-generated elements influence the structure and depth of audience engagement. We address the following core question: *Does AI inherently boost engagement in cultural heritage videos, or do specific AI modalities drive distinct forms of audience interaction?*

2 Methodology

To systematically evaluate the impact of AIGC, we employed a mixed-methods approach utilizing a large-scale dataset of museum short videos.

Data Collection: We collected 72,445 videos from 270 official Chinese museum accounts on the Douyin platform. Of these, 2,055 videos contained identifiable AI-generated elements, while the remaining 70,390 served as a non-AI comparison group.

AI Element Typology: To move beyond binary analyses, we manually coded AI videos into a six-category typology: AI-generated Voice (47.9%), Mixed Type (23.8%), Artifact Anthropomorphism (13.1%), AI Images/Visual Effects (11.7%), Digital Human/Virtual Avatar (1.8%), and AI Animation (1.7%).

Analytical Strategy: We differentiated between shallow engagement (likes) and deep engagement (comments, shares, favorites, and engagement rate). We conducted descriptive and inferential statistical analysis, as well as qualitative topic modeling to obtain an in-depth understanding of audience engagement.

3 Key Findings

3.1 The "Nominal Effect" Paradox

At the bivariate level, AI-generated videos significantly outperformed non-AI videos across likes, favorites, shares, and overall engagement rate (ER). However, our hierarchical regression model revealed that this raw advantage is heavily confounded. When controlling for account size, publication timing, and content orientation (entertainment vs. knowledge), the independent coefficient for the presence of AI became non-significant. AI videos generated higher raw metrics primarily because they tended to be more entertainment-focused and were often posted by smaller accounts (where ER is mathematically inflated). Therefore, the mere presence of AI is not an inherent engagement booster.

3.2 The Power of Embodied Agents

When comparing the performance of specific AI types, stark differences emerged. "Artifact Anthropomorphism" (endowing cultural relics with human-like emotions, voices, or behaviors) and "Mixed Type" AI (combining multiple modalities like visual reconstruction + synthetic voice) consistently and significantly outperformed isolated tools like plain AI voiceovers or basic AI animations. Notably, Mixed Type AI was the only category to significantly boost comments—a high-effort, deep cognitive behavior. This aligns with anthropomorphism theory [**?**], demonstrating that human-like presentation of artifacts enhances social presence and narrative immersion.

3.3 The Engagement Shift

The LDA topic modeling of comment sections revealed that AI does not just inflate engagement volume; it fundamentally alters its focus. Comments on non-AI videos were more likely to feature broad historical knowledge exchange, practical trip-planning discussions (e.g., ticket reservations), and general positive affective reactions. Conversely, comments on AI videos were heavily concentrated on specific museum collections, the provenance of individual artifacts, and reactions to the media format itself. AI acts as a focal lens, directing public discourse away from institutional logistics and toward artifact-centered controversies and narratives.

4 Discussion and Implications

This study contributes an artifact-centric interaction perspective to the social computing and digital heritage discourse. We demonstrate that AI should not be viewed as a universal engagement booster. Instead, it is a highly targeted storytelling tool that reshapes the nature of audience participation. Our study shows strategic implications for museums. We suggest that cultural institutions should move beyond the assumption that any AIGC will yield higher traffic. Relying on simple AI voiceovers or disjointed animations is largely ineffective. To foster meaningful, deep public engagement, museums could invest in artifact anthropomorphism and multimodal storytelling. Furthermore, practitioners should tailor their content format to their communicative goals: use traditional video formats to support general historical education and museum visitation, but deploy AIGC when the goal is to spotlight specific cultural relics and stimulate issue-oriented public discussion.

References

1. Epley, N., Waytz, A., Cacioppo, J.T.: On seeing human: a three-factor theory of anthropomorphism. Psychological Review, **114**(4), 864 (2007)
2. Hellwig, A.F.: Exuberant and fantastical: How flythroughs are changing the way museums construct their identities and build relationships with their publics. Multimodality & Society, **4**(3), 354-371 (2024)
3. Shi, Y., Gao, H.: The Triumph of Substance: Decoding the "Functional Infotainment" Model for Sex Education on Douyin. Behavioral Sciences, **15**(9), 1226 (2025)