AI in Human Sciences
https://polaris-aihs.com/journal/index.php/aihs/index

AI bringing old photos to life: User Attitude Differentiation and Driving Mechanisms in Image-to-Video Generation Technology in Diverse Scenarios

Li Linying [1], Li Huimin [2*]

[1,2]Faculty of Marxism, Yunnan College of Business Management, Kunming, China.

Article history:

Received May 22, 2025

Revised June 7, 2025

Accepted June 18, 2025

***Corresponding Author:**

Li Huimin

E-mail:

761616640@qq.com

Abstract: This study explores how users' attitudes toward AI-assisted image-to-video generation technology vary across different contexts, focusing on the proportion of these attitudes and the mechanisms behind their formation. Utilizing institutional contextualism theory, we analyze how the fit between context and content influences user attitudes in three scenarios: heroes, elders, and ancient paintings. Though employing text analysis to identify user attitudes and combining content analysis with word frequency evaluation, we assess the distribution of these attitudes and representative opinions across various scenarios. The findings indicate that most users maintain positive attitudes, with the 'heroes' scenario generating the highest level of positivity, likely due to feelings of national pride. In contrast, the 'elders' scenario evokes many negative emotions, reflecting the emotional complexity linked to both personal memories and virtual animations. Meanwhile, the 'ancient paintings' scenario is closely associated with cultural values, resulting in a neutral attitude. These findings underscore the importance of context fit in user acceptance and provide valuable insights for designing and disseminating AI-

generated content to enhance user well-being.

Keywords: image-to-video generation technology, diverse scenarios, user attitude, cultural cognition, collective memory

1. INTRODUCTION

The rapid development of artificial intelligence has brought profound changes to all areas of society, changing not only people's behavior patterns but also how we interact with technology (Mondal et al., 2023). We must capture users' changing needs, beliefs, preferences, and expectations (Marikyan et al., 2023), Understanding users' technology usage patterns (Pan & Nishant, 2023). Among the expanding applications of AI, 'AI bringing old photos to life,' the Image-to-Video Generation technology has garnered significant attention and sparked intense debates regarding user acceptance.

This technology, the Image-to-Video Generation, has not only pushed the boundaries of what we thought was possible but also continuously delivered surprises to humanity. From bringing historical figures to life to revitalizing cultural heritage, it has showcased the vast potential of AI in creative and meaningful ways (Fan et al., 2023). As AI-generated content becomes increasingly diversified (Wang et al., 2024), ranging from educational and historical recreations to artistic expressions and personalized experiences, users are presented with a wide array of content that caters to different interests and preferences (Liao & Cao, 2025; Pereira et al., 2022). However, users exhibit positive, neutral, and negative attitudes toward this technology. These varying attitudes reflect cognitive differences in how users perceive technology and the complex interplay between technological advancement and societal needs (Bao et al., 2022; Stolpe & Hallström, 2024). Understanding what makes people accept or reject AI is key to building healthy image-to-video technology. This understanding is crucial for managing how AI and humans interact. It will not only help us cope with the

challenges faced in the future human-AI symbiosis but also help us seize potential opportunities.

The extent to which users' attitudes toward AI-generated video content vary in application scenarios is unclear. What factors drive these different attitudes? Solving these two problems will help balance technological development with social needs. As AI technology penetrates all aspects of society (Storey et al., 2025), users show different attitudes toward AI-generated video content due to cultural, social, and psychological differences (Charfou & Naji, 2024). This attitude difference reflects the complex relationship between technological progress and user expectations (Park et al., 2022). A detailed analysis of attitudes' differences and underlying mechanisms can effectively identify the gap between technological innovation and user needs. This understanding promotes the responsible development and deployment of AI technology, ensuring that advancements align with social values and expectations. It also assists creators in designing AI-generated content that meets technical standards, reflects user values, and adheres to social norms, laying the groundwork for the sustainable development of technology.

2. LITERATURE REVIEW

Previous studies on user attitudes toward AI technologies, such as the Technology Acceptance Model (TAM), focus on the technology itself, explaining attitude differences through perceived ease of use and usefulness (Ma et al., 2025). While TAM effectively highlights why users accept or reject a technology, it fails to address how attitudes vary across diverse contexts, such as cultural, social, or emotional settings. Like user-centric perspectives, Charfou and Naji (2024) explore consumer attitudes toward AI-generated videos using the ABC model of attitudes, including Affective, Behavioral, and Cognitive components. This view emphasizes that cultural backgrounds, personal experiences, or emotional traits capture individual interpretations of contexts but struggle to explain sentiments like 'this is the meaning

of technology,' which reflect broader interactions between content and its environment. This study employs the theory of institutional contextualism to illuminate the complexities surrounding AI-generated video content (Lejano & Shankar, 2013). This theoretical framework examines the interaction between AI-generated content and specific contexts, highlighting the crucial role of the 'fit' between the content and its context in determining user acceptance. When the content closely matches users' preferences, they are more likely to accept it. This approach effectively considers users' attitudes and underlying reasons, creating a solid framework for understanding the complex relationship between AI-generated content and its social acceptance.

2.1. Theory of Institutional Contextualism

The theory of institutional contextualism offers valuable insights into understanding user attitudes toward AI-generated video content. It emphasizes the conceptualizing fit between technology and context, which is mediated through constructive and ecological mechanisms (Lejano & Shankar, 2013). Originating from institutional theory and organizational studies, institutional contextualism builds on the work of scholars on institutional isomorphism and institutional adaptation (Beckert, 2010; Sarta et al., 2021), integrating insights from contextual analyses in policy implementation. The concept of fit is essential, as it significantly influences how effectively AI-generated content aligns with users' expectations, cultural values, and social norms. Users are more inclined to accept and embrace the technology when there is a strong sense of fit—such as in scenarios that evoke national pride or cultural significance. Conversely, a low fit, which often occurs in contexts involving personal or sensitive content, can result in rejection due to concerns about privacy or emotional discomfort (Cho et al., 2020). Therefore, understanding and conceptualizing fit is crucial for shaping user attitudes and their acceptance of AI technologies.

Examining the use of AI technology and scenarios within the framework of Institutional Contextualism is highly significant. It provides a structured approach to

analyze the complex interplay between technology and its application contexts. By doing so, we can better comprehend the reasons behind varying user attitudes and identify the conditions under which AI-generated content is most likely to be accepted (Li et al., 2024). This understanding is essential for guiding the responsible development and implementation of AI technologies, ensuring they are designed in ways that are sensitive to user needs and societal values (Blut & Wang, 2020). It also contributes to the broader goal of fostering a more harmonious and beneficial integration of AI into society.

2.2. User Perspective

Users' acceptance of video content varies due to differences in their cognitive, emotional, and behavioral responses, which are deeply tied to their cultural backgrounds (Charfou & Naji, 2024). Cognitively, users evaluate content based on their knowledge, experiences, and expectations (Capatina et al., 2020), favoring videos that align with cultural heritage. Emotionally, content that evokes positive resonance, such as nostalgia or national pride, is more readily accepted (Danhe & Ma'rof), while sensitive topics like privacy may trigger rejection (Awan & Rao, 2023). Behaviorally, users' interactions are driven by cognitive and emotional alignment (Wu et al., 2024), with content-matching community norms more likely to prompt engagement. These differences shape content preferences regardless of whether the video is filmed by influencers, AI-enhanced, or fully AI-generated.

Cultural backgrounds shape user attitudes through values, social norms, and psychological resonance. Collectivist cultures may prefer content emphasizing community (Muthukrishna & Schaller, 2020), while individualist cultures lean toward videos highlighting personal expression (Liu & Herndon, 2022). Social norms influence acceptance through community pressures, with privacy-sensitive cultures often rejecting content involving personal imagery, regardless of its production method. Psychological resonance makes some cultures more receptive to nostalgic or traditional

themes, while unfamiliar cultural symbols may elicit indifference. These cultural factors underscore that content's alignment with context, rather than its technical origin, drives acceptance.

Users focus on content quality, emotional resonance, and cultural fit, not the method of creation (Song & Ho, 2024), whether by influencers, AI processing, or the whole AI generation. Institutional contextualism suggests that acceptance hinges on how well content aligns with context, facilitated by mechanisms like platform promotion and adaptation to local cultural practices (Covarrubias, 2024). For instance, a high-quality AI-generated video that accurately reflects cultural symbols may be preferred over a low-quality human-filmed one. Thus, designing context-sensitive content that aligns with users' cognitive, emotional, and cultural expectations is key to enhancing acceptance and minimizing resistance.

2.3. Content Perspective

From a content perspective, video content creation is shaped by the author's characteristics, with cultural inclinations and emotional expressions stemming from their subjective background (Riemer et al., 2014). The author's cultural context influences the themes, symbols, and narrative styles (Ladzekpo et al., 2024), such as traditional authors crafting videos reflecting historical or collective values (Birkner & Donk, 2020), while those from urban backgrounds may focus on individual emotions or innovative expressions. Emotional expressions convey the author's inner experiences, such as nostalgia or critique, directly impacting users' resonance with the content. Whether created through live filming, AI processing, or hand-drawn methods, the essence of the content is determined by the author's cultural perspective and intent, reflecting their unique interpretation of social and cultural contexts (Rice et al., 2020).

The source of creation is open and diverse, with authors' varied backgrounds injecting rich cultural and emotional elements into video content (Rice et al., 2020). Openness allows authors to freely choose their mode of expression through traditional

forms like hand drawing or modern techniques like AI generation, enabling content to reflect global cultural fusion or localized practices. Diversity arises from the heterogeneity of authors, encompassing different cultures, experiences, and values, which cater to a wide range of audience needs. For instance, one author might recreate cultural scenes through hand-drawn art (Mousavi, 2023), while another uses AI to produce modern-themed videos, offering users a broad spectrum of content choices.

Institutional contextualism suggests that the 'conceptualizing fit' between content and users' cultural backgrounds influences acceptance; for example, hand-drawn nostalgic videos may be more readily accepted due to emotional resonance, while AI-generated videos of deceased loved ones might face rejection due to privacy or emotional sensitivities. The creation method is not the sole determinant; its presentation affects users' emotional responses. Users can adopt an open attitude to embrace diverse cultures or reject content based on ethical or emotional grounds, ensuring a balance that encourages creators to prioritize context sensitivity to enhance acceptance.

2.4. Technical Perspective

From a technical perspective, Image-to-video generation technology represents only a minor component in the evolution of generative AI (Huang et al., 2023), and attitudes toward it should not be equated with views on AI. As a subset of generative AI, this technology facilitates the transformation of static images into dynamic videos, offering a convenient, creative tool (Yu et al., 2024). However, its scope and impact are far less significant than AI's broader applications in data analytics, natural language processing, or autonomous driving. User attitudes toward this technology vary by context (Breward et al., 2017), such as enthusiasm for entertainment content or resistance to privacy-sensitive applications, but these reflect opinions on specific use cases rather than AI in its entirety.

The strength of this technology lies in lowering the barriers to content creation, enabling creators without hand-drawing or filming skills to produce videos easily, and

showcasing the benefits of technological diversification. By converting static images into dynamic narratives through algorithms, Image-to-Video Generation Technology empowers users to craft nostalgic videos from old photos or create personalized clips. This accessibility fosters the democratization of content creation, allowing more individuals to engage in cultural and artistic expression (Park, 2024). However, the technology's success depends on its alignment with contextual factors, such as cultural relevance or emotional resonance, to gain user acceptance.

AI's applications extend beyond video creation from old photos, demonstrating its capacity to break conventional boundaries and enable diverse creative expressions (Davoodi, 2024). For instance, AI can animate static images of plants to depict growth, generate surreal artworks, simulate natural phenomena, or create virtual character animations. These examples highlight AI's flexibility in pushing creative limits, far surpassing the scope of Image-to-Video Generation Technology. Technical development must prioritize context sensitivity to ensure that AI's diverse applications meet users' cultural and emotional needs while advancing technological diversification (Pisoni et al., 2021).

2.5. Conceptualizing Fit

The theory of institutional contextualism provides a robust framework for understanding user attitude differentiation toward AI-generated video content by examining its core viewpoint—conceptualizing fit (Lejano & Shankar, 2013). When scenarios align more closely with users' values, cultural backgrounds, or personal experiences, users are more receptive to AI-generated content. This alignment interacts with diverse contexts—social norms, cultural values, and psychological predispositions—through constructive and ecological mechanisms, influencing acceptance. This dynamic explains the varied responses to AI-generated content across scenarios and underscores the importance of context-sensitive design to enhance acceptance.

The varying degrees of conceptualizing fit can directly shape users' attitudes toward Image-to-Video Generation Technology. Where content aligns with contextual expectations fosters positive attitudes. For instance, when AI-generated videos with nostalgic themes are used in memorial settings, they resonate with users' cultural values and emotional predispositions, leading to greater acceptance. Conversely, using AI to animate personal photographs in privacy-conscious contexts can trigger adverse reactions due to ethical sensitivities and privacy concerns (Choung et al., 2023), highlighting the importance of understanding the interplay between technology and context (Magliocca et al., 2024; Yadav & Pavlou, 2020).

Constructive mechanisms like norm diffusion influence this fit through platform promotion, and ecological mechanisms like content adaptation to local cultural practices (Sutton & Anderson, 2020). Cultural alignment plays a significant role, as content compatible with cultural heritage tends to gain favor (Kirk et al., 2024). Furthermore, psychological resonance is crucial, as emotional congruence between the content and users' expectations can enhance the appeal of AI-generated videos. By understanding how these factors contribute to the varying degrees of fit, we can gain insights into the mechanisms behind attitude differentiation. It underscores the need for context-sensitive content design to optimize acceptance and mitigate resistance, ensuring that AI technology is developed and applied in ways attuned to user needs and societal values.

3. METHOD

3.1. Research Design

This study examines user attitudes toward Image-to-Video Generation Technology across diverse scenarios and the underlying reasons. Using online ethnography, it collects and analyzes user comments from social media to identify positive, negative, and neutral attitudes and explores the core reasons behind them.

3.2 Instruments and Technologies

The data collection and analysis are fully implemented via Python programming, which offers rich libraries for efficient data crawling, structuring, word-frequency analysis, and word-cloud mapping (Navlani et al., 2021).

3.3 Data Collection

3.3.1 Participants

Data comes from four major social - media platforms: Bilibili, Xiaohongshu, Douyin, and Weibo. Their diverse user bases provide rich perspectives and data. Videos on these platforms about 'AI bringing old photos to life' generally have high views and rich comments, with over 1,000 comments per selected video.

3.3.2 Data Collection Process

Python-based data collection uses platform APIs or web crawler technology to obtain videos and comments about "AI makes old photos come alive." We started collecting data using specific search keywords (such as "AI makes old photos come alive") on the selected social media platforms. This enabled us to identify many videos related to the topic. After identifying these videos, we carefully screened them and selected videos with more than 1,000 comments. This threshold ensured that we focused on videos with a large amount of user engagement, which would provide a richer data set for analysis.

We used the web crawler's code to automate data collection with the selected relevant videos. The text data was organized into structured data, with each comment including detailed information such as the comment content, posting time, and number of likes. It is enabled to distinguish users' attitudes and emotions toward the technology accurately. Through manual annotation and natural language processing, we systematically cleaned 6,709 user reviews, removed non-substantive content, and obtained 4,136 valid reviews for further analysis.

3.4. Data Analysis

3.4.1. Content Analysis

This study highlights three main scenes from the retrieved video: heroes, elders, and ancient paintings. The users' attitudes toward these different scenes will be analyzed. The heroes' scenarios focus on historical figures who played a crucial role in China's

transformation, emphasizing individuals who sacrificed for national independence and liberation. The scenarios of elders showcase vintage photographs of ordinary people, capturing their life stages from childhood to old age while serving as memorials for the deceased. In this scenario, users may engage with their emotions sensitively. The scenarios of ancient paintings feature hand-painted portraits of historical figures who have profoundly influenced China's culture, art, and political landscape.

3.4.2. Text Analysis

In the Python environment, we first read the file using UTF-8 encoding to ensure compatibility with Chinese text, allowing the program to interpret and process Chinese characters accurately and without error. Next, we use the Jieba package to segment the Chinese text, which can be used to split Chinese sentences into individual words by identifying " 是(is)," "因为(because)," "非常(very)," "但(but)" etc. Next, punctuation marks and single-word words are removed to focus on more meaningful, content-rich, and sentiment-laden words. Finally, the remaining words are sorted in descending order of frequency and output in a list, providing valuable insights into the topics discussed.

4. RESULT

4.1. Scenarioal Analysis

The histogram illustrates the differentiation of user attitudes across three distinct contexts of diverse scenarios with AI-generated content—Heroes, Elders, and Ancient Paintings—based on an analysis of 4,136 valid comments. The histogram reveals an apparent variation in user attitudes across the three scenarios. Overall, most users hold positive attitudes toward Image-to-Video Generation technology. The heroes' scenario garners the most positive reactions, likely due to its connection with national pride. In contrast, the scenario of elders reveals the most significant negative sentiment, highlighting the emotional complexity tied to animating personal memories. Meanwhile, the ancient paintings' scenario, closely linked to cultural values, exhibits more balanced attitudes (Table 1, Figure 1).

Table 1. Proportion and frequency of user attitude

	Positive		**Neutral**		**Negative**		**Total**
	n	***f***	***n***	***f***	***n***	***f***	***n***
Ancient painting	780	30.71%	514	59.61%	246	9.69%	2540
Elders	20	67.52%	41	6.59%	161	25.88%	622
Heroes	97	92.09%	75	7.70%	2	0.21%	974

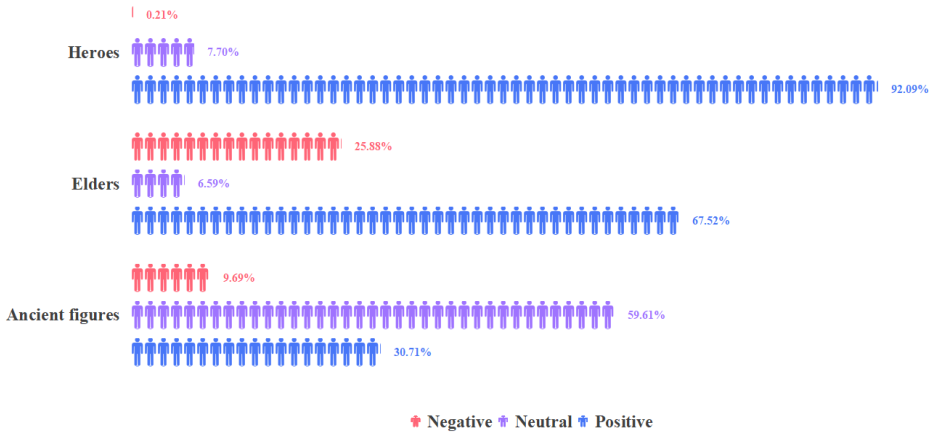

Figure 1. Histogram of user attitude differentiation in diverse scenarios

For the scenario of heroes, which features historical figures who sacrificed for China's independence, user attitudes are predominantly positive, with 92.09% of respondents expressing approval (positive). In comparison, only 0.21% are neutral, and 7.70% are negative. This suggests a strong societal reverence for these figures, enhancing technology acceptance in this context. In contrast, the elders' scenario, involving old photographs of ordinary individuals with sentimental value, shows a more polarized response, with 67.52% positive, 25.88% negative, and 6.59% neutral attitudes, indicating significant concerns likely tied to emotional or ethical sensitivities regarding the animation of personal memories. The scenario of ancient paintings, representing historical figures from China's cultural heritage, shows 59.61% neutral, 30.71% positive, and 9.69% negative attitudes, indicating moderate acceptance. The higher neutral responses may stem from concerns over preserving the authenticity of cultural artifacts when subjected to AI transformation.

4.2. Word Frequency Analysis

4.2.1. Positive Attitude: Scenario of Heroes

The word frequency analysis of user-generated content (UGC) from the scenario of heroes (Figure 2), where multimodal AI animates images of historical figures who sacrificed for China's independence, reveals a robustly positive user attitude deeply rooted in national pride and reverence. Terms such as 'heroes' (n = 557), 'motherland' (n = 176), and 'remenber' (n = 139) dominate the lexicon, reflecting a strong emotional connection to these revolutionary figures and their contributions, with 'prosperous' (n = 126) and 'immortal' (n = 66) underscoring a celebratory view of the technology's ability to honor their legacy. The frequent mention of 'revolutionary' (n = 65), 'ancestors' (n = 71), and 'sacrifice' (n = 31) highlights users' admiration for the historical sacrifices, amplified by emotive phrases like 'deep emotion' (n = 52) and 'tears' (n = 46), suggesting the AI-generated content evokes profound patriotism.

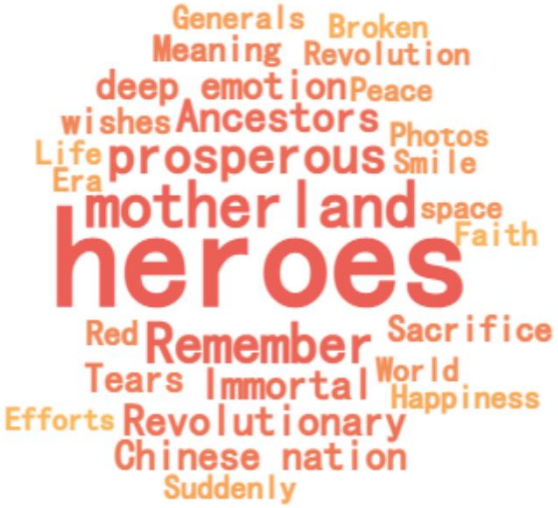

Figure 2. Word frequency cloud for the scenario of heroes

4.2.2. Neutral Attitude: Scenario of Ancient Paintings

The word frequency analysis of user-generated content (UGC) from the scenario of ancient paintings (Figure 3), where multimodal AI animates hand-drawn portraits of historical figures from China's cultural heritage, reveals a neutral user attitude

characterized by a balanced mix of curiosity and ambivalence. High-frequency terms such as ‘portrait’ (n = 246), ‘Du Fu’ (95), ‘Li Qingzhao’ (n = 62), and ‘Tang Bohu’ (n = 59) indicate a strong focus on the historical and artistic significance of these figures, with users engaging with specific cultural icons like poets and painters, suggesting an interest in the technology’s ability to revive ancient imagery. The presence of ‘official’ (n = 46), ‘history’ (n = 37), and ‘ancient’ (n = 34) reflects a context-aware discussion, while ‘modern’ (n = 49) and ‘modern people’ (n = 25) hint at comparisons between past and present representations, contributing to a neutral stance as users weigh the technology’s cultural relevance.

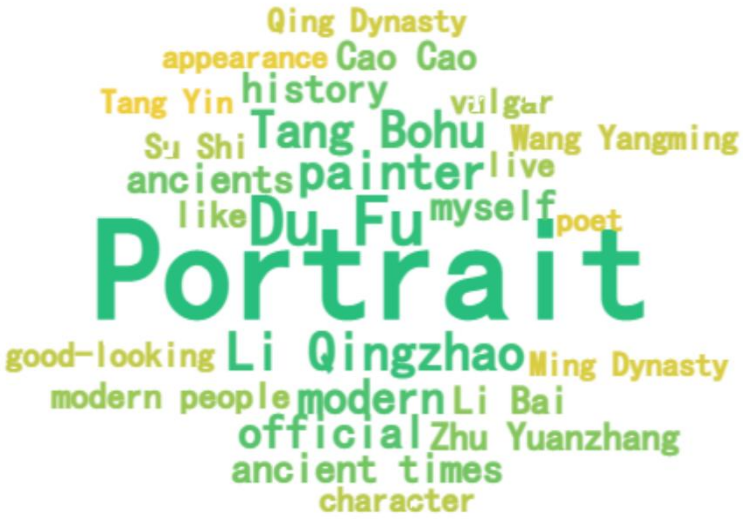

Figure 3. Word frequency cloud for the scenario of ancient paintings

4.2.3. Negative Attitude: Scenario of Elders

The word frequency analysis of user-generated content (UGC) from the scenario of elders (Figure 4), involving the animation of old photographs, reveals a strong negative user attitude driven by emotional and ethical concerns, as evidenced by terms like ‘tamper memory’ (n = 74) and ‘feeling unwell’ (n = 53). This suggests that users perceive the technology as disrupting cherished memories, with additional discontent reflected in ‘unlike’ (n = 20) and ‘unnatural’ (n = 13), pointing to dissatisfaction with the AI's authenticity and output quality, particularly in a context tied to deceased loved

ones or personal histories. This negative sentiment is further compounded by terms like 'horror' (n = 23) and 'algorithm' (n = 14), indicating an uncanny valley effect, while 'technology' (n = 14) and 'algorithm' (n = 4) highlight users' frustration with the mechanistic approach.

Figure 4. Word frequency cloud for the scenario of ancient paintings

5. DISCUSSION

5.1. Discussion of the Analysis Results

The Theory of Institutional Contextualism highlights how the interplay of content, context, and mechanisms shapes these varied responses (Lejano & Shankar, 2013). In the scenario of heroes, the high fit with cultural pride and minimal ethical concerns drives positive attitudes (Boer & Fischer, 2013). In the scenario of ancient paintings, the moderate fit with cultural heritage, tempered by concerns about authenticity, leads to neutral attitudes. In the scenario of elders, the low fit due to emotional and ethical sensitivities results in negative attitudes.

5.1.1. Scenario and Emotional Tendencies

The theory of institutional contextualism explains the variation in user attitudes toward AI-generated video content across different scenarios by focusing on the

conceptualizing fit between the content and its context, mediated through constructive mechanisms and ecological mechanisms. In the scenario of heroes, featuring AI-animated images of historical figures, the high conceptualizing fit with cultural values of national pride and reverence for revolutionary figures leads to the most positive user attitudes (Zhang et al., 2020). Platform promotion amplifies societal norms of patriotism, and the content adapts to a context where historical sacrifices are celebrated, evoking strong emotional responses (Wang, 2021).

In the context of ancient paintings, the use of AI to animate portraits of cultural figures has led to a moderate conceptual fit, resulting in generally neutral attitudes (Land, 2023). While the content resonates with cultural values associated with heritage, there are concerns about authenticity and the role of AI, which creates a sense of ambivalence. Discussions on various platforms emphasize the historical significance of these figures, but the technology used to animate cultural artifacts raises important questions about their authenticity, prompting a cautious response.

Furthermore, the concept of animating elderly individuals through AI-enhanced old photographs tends to provoke the most negative responses, primarily due to a poor alignment with users' emotional and ethical expectations (Ho et al., 2023). The act of animating personal memories raises important concerns regarding privacy, authenticity, and emotional sensitivity (Alabed et al., 2022). There is a conflict between social norms that emphasize respect for personal memories and the application of this technology. Consequently, the content often fails to adapt to a context where emotional sensitivity and authenticity are crucial, resulting in discomfort and rejection.

5.1.2. Image-to-video Generation Technology

The results show the potential advantages and risks of image-to-video technology in public domain network platforms. In the context of heroes, the technology effectively captures and amplifies national pride, showing positive reactions, and users are able to present content with strong social resonance (Song & Ho, 2024). In contrast, in the context of older adults, high-frequency words such as ‘tampered memory’ and ‘unnatural’ indicate that users question the authenticity and emotional accuracy of the output content (Land, 2023). Meanwhile, the neutral attitude towards ancient paintings indicates that, while the technology is appealing for cultural heritage, it falls short in

alleviating concerns regarding the authenticity of cultural relics. At the same time, the neutral attitude towards ancient paintings shows that the application of this technology to cultural heritage themes is attractive to users, but people may still have general concerns about the authenticity of cultural relics. Therefore, it is necessary to further optimize the subtle emotional and cultural fit between AI-generated content and users in different scenarios, which requires AI developers to fully consider the psychological characteristics of users.

5.1.3. Emotional Complexity

The scenario of heroes demonstrates how AI-generated content can evoke profound collective emotions when aligned with deeply rooted cultural narratives (Salmela, 2014). Lexical analysis reveals terms like ‘tears’ and ‘deep emotion’ signify affective mobilization through historical symbolism, where technology acts as a conduit for reinforcing national identity (Harrison & Johnson, 2009; Zhang, 2022). This suggests that algorithmic outputs leveraging shared heritage—particularly martyrdom narratives—can achieve high acceptance by fulfilling a socio-psychological need for commemorative cohesion (Figueiredo et al., 2017; Páez et al., 2015), though risks of emotional instrumentalization warrant caution.

Contrastingly, negative sentiment in the scenario of elders—marked by ‘feeling unwell’ and ‘horror’ —exposes the fragility of emotional boundaries when technology mediates private grief. The relationship between credibility and motivation attribution, user moral emotions, etc. is closely established (Fernandez et al., 2022). Such reactions align with theories of affective dissonance (Hillman et al., 2023), where posthumous digital animation disrupts mourning rituals, inducing uncanny valley effects (Kim et al., 2022). These findings underscore the necessity of participatory design frameworks that prioritize bereavement ethics, such as dynamic consent models (Lay et al., 2025; Prictor et al., 2020) to mitigate perceived violations of emotional sanctity.

The scenario of ancient paintings’ neutral stance reflects a tension between cultural curiosity and preservationist concerns. Frequent mentions of historical figures coupled with low negativity indicate conditional acceptance, contingent upon perceived fidelity to canonical representations. This implies that AI-mediated cultural revival must balance innovation with transparent historicity—perhaps through explainable AI

interfaces that disclose interpretive logic—to sustain user trust in algorithmic stewardship of collective memory (Varda et al., 2024).

5.1.4. Sociocultural Background

National Identity in Hero Representation. The scenario of heroes's 92.09% positive response rate demonstrates perfect alignment between technological outputs and culturally ingrained values of national reverence. As evidenced by high-frequency terms like 'motherland' and 'revolutionary', the AI-generated content successfully mirrors established narratives of historical commemoration prevalent in Chinese media and education systems. This congruence creates what institutional theorists call a "taken-for-granted" acceptance, where the technology becomes indistinguishable from traditional forms of patriotic expression.

The scenario of elders reveals significant cultural boundaries, with 25.88% negative responses highlighting discomfort around technological mediation of personal remembrance. Terms like 'feeling unwell' and 'tamper memory' indicate a perceived violation of fundamental norms governing private mourning practices. This tension emerges when technological capabilities conflict with deeply rooted cultural expectations regarding the appropriate treatment of personal and familial history.

scenario of ancient paintings responses (30.71% positive, 59.61% neutral) reflects a cautious equilibrium between cultural preservation and technological curiosity. The frequent mentions of historical figures like 'Du Fu' and 'Li Qingzhao' demonstrate engagement with cultural content, while the predominant neutral stance suggests reservations about altering traditional representations. This balanced response pattern indicates that user acceptance depends on technology's ability to maintain perceived authenticity while offering innovative engagement with cultural artifacts.

6. CONCLUSION AND IMPLICATION

Image-to-video generation technology will surely find a practical path that will benefit humanity. Considering the dual nature of AI's impact, AI is not just a tool but a transformative force that reshapes our interactions with media, culture, and each other. In this study, we mainly explored AI's ability to give life to static images, provide new

dimensions of expression and understanding, dynamically preserve cultural heritage, and provide emotional connections that were previously unattainable. However, this power is not without shadows. The ethical and emotional challenges it poses, especially in individual differences in sociocultural contexts, remind engineers and AI users that the advancement of artificial intelligence technology requires consideration of profound responsibilities. We urgently need to establish a dialogue mechanism that should not simply accept artificial intelligence at face value but should seek to understand, guide, and improve its integration into our lives. Only in this way can we use the potential of AI to enrich our world while mitigating its risks, ensuring that technology becomes a bridge, not an obstacle, to a more connected and empathetic future. Its risks are ensuring that technology becomes a bridge, not an obstacle, to a more connected and empathetic future.

To balance user attitudes toward AI - generated video content, context, and fit, we need to adjust users' understanding and set a stringent framework for AI users. Boosting users' basic AI knowledge and scientific literacy via popular science and education can reduce misunderstandings. Also, AI users should focus on the application process and context, using diversity and inclusion mechanisms to enhance content - context fit. The application of Image-to-Video Generation Technology should be context-specific. For the scenario of elders involving ordinary seniors, we can encourage users to explore the technology for personal use. This private engagement can offer a unique emotional experience and meet their emotional needs, but it's better not to share it in public online spaces. As for the heroes and ancient paintings' scenarios, which have higher user acceptance, we can promote the creation of such content during meaningful times or festivals. This can enrich users' spiritual lives. Policymakers need to create proper policy frameworks to align AI development with user interests. This means considering users' complex attitudes toward AI, and addressing their expectations and concerns in the policy - making process. Encouraging user participation in AI discussions and decisions can also build their confidence in technological development. Through education, transparency, and well - crafted policies, we can balance technology and user needs. This approach promotes healthy AI development and fosters harmony between technology and society.

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

REFERENCES

Alabed, A., Javornik, A., & Gregory-Smith, D. (2022). AI anthropomorphism and its effect on users' self-congruence and self–AI integration: A theoretical framework and research agenda. *Technological Forecasting and Social Change*, *182*, 121786.

Awan, J., & Rao, V. (2023). Privacy-aware rejection sampling. *Journal of machine learning research*, *24*(74), 1-32.

Bao, L., Krause, N. M., Calice, M. N., Scheufele, D. A., Wirz, C. D., Brossard, D., Newman, T. P., & Xenos, M. A. (2022). Whose AI? How different publics think about AI and its social impacts. *Computers in human behavior*, *130*, 107182.

Beckert, J. (2010). Institutional isomorphism revisited: Convergence and divergence in institutional change. *Sociological theory*, *28*(2), 150-166.

Birkner, T., & Donk, A. (2020). Collective memory and social media: Fostering a new historical consciousness in the digital age? *Memory studies*, *13*(4), 367-383.

Blut, M., & Wang, C. (2020). Technology readiness: a meta-analysis of conceptualizations of the construct and its impact on technology usage. *Journal of the Academy of Marketing Science*, *48*, 649-669.

Boer, D., & Fischer, R. (2013). How and when do personal values guide our attitudes and sociality? Explaining cross-cultural variability in attitude–value linkages. *Psychological bulletin*, *139*(5), 1113.

Breward, M., Hassanein, K., & Head, M. (2017). Understanding consumers' attitudes toward controversial information technologies: A contextualization approach. *Information Systems Research*, *28*(4), 760-774.

Capatina, A., Kachour, M., Lichy, J., Micu, A., Micu, A.-E., & Codignola, F. (2020). Matching the future capabilities of an artificial intelligence-based software for social media marketing with potential users' expectations. *Technological Forecasting and Social Change*, *151*, 119794.

Charfou, A. R., & Naji, J. (2024). The Era of Ai-Generated Video Production: Exploring Consumers' Attitudes. In.

Cho, H., Li, P., & Goh, Z. H. (2020). Privacy risks, emotions, and social media: A coping model of online privacy. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *27*(6), 1-28.

Choung, H., David, P., & Ross, A. (2023). Trust in AI and its role in the acceptance of AI technologies. *International Journal of Human–Computer Interaction*, *39*(9), 1727-1739.

Covarrubias, R. (2024). On being accepted: Interrogating how university cultural scripts shape personal and political facets of belonging. *Educational Psychology Review*, *36*(4), 1-26.

Danhe, Z., & Ma'rof, A. A. The Role of Emotional Response to Music, Media Influence, and Nationalistic Themes on Chinese National Identity.

Davoodi, A. (2024). EQUAL AI: A Framework for Enhancing Equity, Quality, Understanding and Accessibility in Liberal Arts through AI for Multilingual Learners. *Language, Technology, and Social Media*, *2*(2), 178-203.

Fan, F., Luo, C., Gao, W., & Zhan, J. (2023). Aigcbench: Comprehensive evaluation of image-to-video content generated by ai. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, *3*(4), 100152.

Fernandez, P., Hartmann, P., & Apaolaza, V. (2022). What drives CSR communication effectiveness on social media? A process-based theoretical framework and research agenda. *International Journal of Advertising*, *41*(3), 385-413.

Figueiredo, A., Martinovic, B., Rees, J., & Licata, L. (2017). Collective memories and present-day intergroup relations: Introduction to the special thematic section. *Journal of Social and Political Psychology*, *5*(2), 694-706.

Harrison, C. E., & Johnson, A. (2009). Introduction: Science and national identity. *Osiris*, *24*(1), 1-14.

Hillman, J. G., Fowlie, D. I., & MacDonald, T. K. (2023). Social verification theory: A new way to conceptualize validation, dissonance, and belonging. *Personality and Social Psychology Review*, *27*(3), 309-331.

Ho, M.-T., Mantello, P., & Ho, M.-T. (2023). An analytical framework for studying attitude towards emotional AI: The three-pronged approach. *MethodsX*, *10*, 102149.

Huang, Y., Lv, S., Tseng, K.-K., Tseng, P.-J., Xie, X., & Lin, R. F.-Y. (2023). Recent advances in artificial intelligence for video production system. *Enterprise Information Systems*, *17*(11), 2246188.

Kim, B., de Visser, E., & Phillips, E. (2022). Two uncanny valleys: Re-evaluating the uncanny valley across the full spectrum of real-world human-like robots. *Computers in human behavior*, *135*, 107340.

Kirk, H. R., Whitefield, A., Rottger, P., Bean, A. M., Margatina, K., Mosquera-Gomez, R., Ciro, J., Bartolo, M., Williams, A., & He, H. (2024). The PRISM alignment dataset: What participatory, representative and individualised human feedback reveals about the subjective and multicultural alignment of large language models. *Advances in Neural Information Processing Systems*, *37*, 105236-105344.

Ladzekpo, G. K., Amekor, C. K., & Davi, S. K. (2024). The Power of Storytelling and Narrative in Literature and Cultural Identity: A Review Article. *British Journal of Humanities and Social Sciences*, *1*(1), 20-31.

Land, J. (2023). Synthesizing collective memory and counter-memory in urban space. *Urban Geography*, *44*(5), 1011-1020.

Lay, W., Gasparini, L., Siero, W., & Hughes, E. K. (2025). A rapid review of the benefits and challenges of dynamic consent. *Research Ethics*, *21*(1), 180-202.

Lejano, R. P., & Shankar, S. (2013). The contextualist turn and schematics of institutional fit: Theory and a case study from Southern India. *Policy Sciences*, *46*(1), 83-102. https://doi.org/10.1007/s11077-012-9163-9

Li, J.-C., Lin, Y., & Yang, Y.-C. (2024). Extending the theory of planned behavior model to explain people's behavioral intentions to follow China's AI generated content law. *BMC psychology*, *12*(1), 367.

Liao, X., & Cao, P. (2025). Digital media entertainment technology based on artificial intelligence robot in art teaching simulation. *Entertainment Computing*, *52*, 100792.

Liu, N. H., & Herndon, J. L. (2022). A framework for culturally humble therapeutic responses using the deliberate practice multicultural orientation video prompts. *Practice Innovations*, *7*(3), 178.

Ma, J., Wang, P., Li, B., Wang, T., Pang, X. S., & Wang, D. (2025). Exploring user adoption of ChatGPT: A technology acceptance model perspective. *International Journal of Human–Computer Interaction*, *41*(2), 1431-1445.

Magliocca, P., Canestrino, R., Carayannis, E. G., & Gagliardi, A. R. (2024). Understanding human–technology interaction: evolving boundaries. *European Journal of Innovation Management*.

Marikyan, D., Papagiannidis, S., & Stewart, G. (2023). Technology acceptance research: Meta-analysis. *Journal of Information Science*, 01655515231191177.

Mondal, S., Das, S., & Vrana, V. G. (2023). How to bell the cat? A theoretical review of generative artificial intelligence towards digital disruption in all walks of life. *Technologies*, *11*(2), 44.

Mousavi, N. (2023). Cameras, Pencils, Traumas: Drawn Images in and as Documentary Practice. In *Truth Claims Across Media* (pp. 127-150). Springer International Publishing Cham.

Muthukrishna, M., & Schaller, M. (2020). Are collectivistic cultures more prone to rapid transformation? Computational models of cross-cultural differences, social network structure, dynamic social influence, and cultural change. *Personality and Social Psychology Review*, *24*(2), 103-120.

Navlani, A., Fandango, A., & Idris, I. (2021). *Python Data Analysis: Perform data collection, data processing, wrangling, visualization, and model building using Python*. Packt Publishing Ltd.

Páez, D., Rimé, B., Basabe, N., Wlodarczyk, A., & Zumeta, L. (2015). Psychosocial effects of perceived emotional synchrony in collective gatherings. *Journal of personality and social psychology*, *108*(5), 711.

Pan, S. L., & Nishant, R. (2023). Artificial intelligence for digital sustainability: An insight into domain-specific research and future directions. *International Journal of Information Management*, *72*, 102668.

Park, I., Kim, D., Moon, J., Kim, S., Kang, Y., & Bae, S. (2022). Searching for new technology acceptance model under social context: Analyzing the determinants

of acceptance of intelligent information technology in digital transformation and implications for the requisites of digital sustainability. *Sustainability*, *14*(1), 579.

Park, S. (2024). The work of art in the age of generative AI: aura, liberation, and democratization. *AI & society*, 1-10.

Pereira, S. W., Fishman, E. K., & Rowe, S. P. (2022). The future is now: how technology and entertainment are transforming education in the artificial intelligence era. *Journal of the American College of Radiology*, *19*(9), 1077-1078.

Pisoni, G., Díaz-Rodríguez, N., Gijlers, H., & Tonolli, L. (2021). Human-centered artificial intelligence for designing accessible cultural heritage. *Applied Sciences*, *11*(2), 870.

Prictor, M., Lewis, M. A., Newson, A. J., Haas, M., Baba, S., Kim, H., Kokado, M., Minari, J., Molnar-Gabor, F., & Yamamoto, B. (2020). Dynamic consent: an evaluation and reporting framework. *Journal of Empirical Research on Human Research Ethics*, *15*(3), 175-186.

Rice, C., LaMarre, A., Changfoot, N., & Douglas, P. (2020). Making spaces: Multimedia storytelling as reflexive, creative praxis. *Qualitative Research in Psychology*.

Riemer, H., Shavitt, S., Koo, M., & Markus, H. R. (2014). Preferences don't have to be personal: Expanding attitude theorizing with a cross-cultural perspective. *Psychological review*, *121*(4), 619.

Salmela, M. (2014). The functions of collective emotions in social groups. *Institutions, emotions, and group agents: Contributions to social ontology*, 159-176.

Sarta, A., Durand, R., & Vergne, J.-P. (2021). Organizational adaptation. *Journal of management*, *47*(1), 43-75.

Song, T., & Ho, Y.-C. (2024). Relational Resonance and Content Creation. *Management Science*.

Stolpe, K., & Hallström, J. (2024). Artificial intelligence literacy for technology education. *Computers and Education Open*, *6*, 100159.

Storey, V. C., Yue, W. T., Zhao, J. L., & Lukyanenko, R. (2025). Generative Artificial Intelligence: Evolving technology, growing societal impact, and

opportunities for information systems research. *Information Systems Frontiers*, 1-22.

Sutton, M. Q., & Anderson, E. N. (2020). *An introduction to cultural ecology*. Routledge.

Varda, C., Iordanou, K., Antoniou, J., Barrios, M. M. Z., Yalaz, E., Gurzawska, A., Szüdi, G., Bartar, P., & Häberlein, L. (2024). The Role of Stewards of Trust in Facilitating Trust in Science: A Multistakeholder View. *Journal of Academic Ethics*, 1-21.

Wang, Q. (2021). The cultural foundation of human memory. *Annual review of Psychology*, *72*(1), 151-179.

Wang, X., Hong, Y., & He, X. (2024). Exploring artificial intelligence generated content (AIGC) applications in the metaverse: Challenges, solutions, and future directions. *IET Blockchain*, *4*(4), 365-378.

Wu, C.-K., Lin, C.-A., & Hsu, Y.-M. (2024). Unveiling the intricate dynamics of user engagement in social media: the triad co-evolution through affordances and emotional attachment. *Communication Research and Practice*, *10*(2), 230-248.

Yadav, M. S., & Pavlou, P. A. (2020). Technology-enabled interactions in digital environments: A conceptual foundation for current and future research. *Journal of the Academy of Marketing Science*, *48*, 132-136.

Yu, T., Yang, W., Xu, J., & Pan, Y. (2024). Barriers to industry adoption of AI video generation tools: A study based on the perspectives of video production professionals in China. *Applied Sciences*, *14*(13), 5770.

Zhang, J. (2022). Hardening national boundaries in a globally-connected world: Technology, development and nationalism in China. *Journal of Contemporary Asia*, *52*(5), 783-804.

Zhang, X., Nekmat, E., & Chen, A. (2020). Crisis collective memory making on social media: A case study of three Chinese crises on Weibo. *Public Relations Review*, *46*(4), 101960.