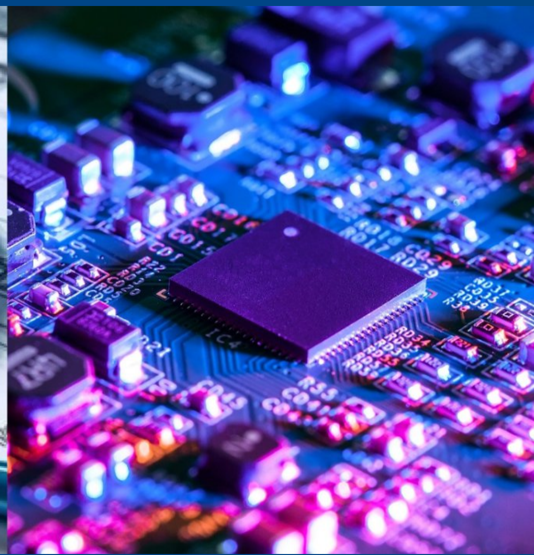
ISSN: 2582-7219

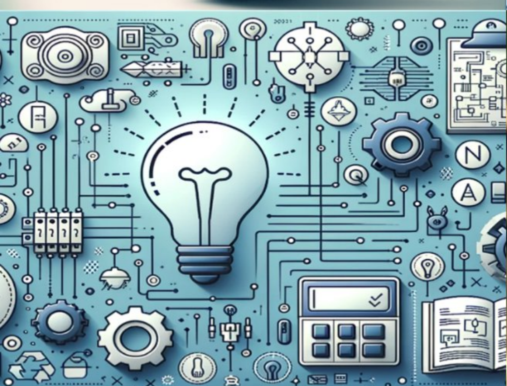




International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 9, Issue 4, April 2026

www.ijmrset.com ijmrset@gmail.com +91 6381 907 438 +91 63819 07438

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904218

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Public Perception of AI-Generated Video Tools

An Analysis of YouTube Videos and User Comments

Prabin PB Raj, Dr. Avinash Rana,

MBA, Business Analytics, CMS Business School, JAIN (Deemed-to-be University), Bengaluru, Karnataka, India

Associate Professor, Business Analytics, CMS Business School, JAIN (Deemed-to-be University), Bengaluru, Karnataka, India

ABSTRACT: This study investigates how YouTube audiences perceive, react to, and evaluate AI-generated video tools through systematic analysis of large-scale, organic digital discourse. Using a three-phase mixed-methods analytical framework, the research analyses 84,905 YouTube comments from 50 videos across 44 channels, complemented by video transcript analysis and engagement metric examination.

Sentiment analysis via VADER revealed a predominantly positive distribution (47.9%), with negative comments at 31.5% and neutral at 20.6%. Latent Dirichlet Allocation (LDA) topic modelling identified five coherent thematic clusters: (1) AI authenticity and content detection, (2) creative labour and displacement, (3) AI video aesthetics and experimentation, (4) internet culture and platform dynamics, and (5) human-AI interaction and systemic implications. Transformer-based emotion detection identified surprise (0.107) as the dominant emotion, followed closely by happiness (0.101) and anger (0.096). Video transcripts were classified 88.5% positive, markedly diverging from comment-level ambivalence. Findings confirm that public sentiment toward AI-generated video is inherently ambivalent, structured around simultaneous fascination with creative democratisation and apprehension about authenticity erosion, deepfake proliferation, and human creative displacement.

KEYWORDS: AI-generated video, sentiment analysis, YouTube, VADER, LDA topic modelling, deepfakes, public perception, NLP

I. INTRODUCTION

The AI video generation industry has emerged as one of the most rapidly evolving frontiers of generative artificial intelligence. The global AI video generator software market, valued at approximately USD 1.23 billion in 2025, is projected to reach USD 21.6 billion by 2034, expanding at a compound annual growth rate of 46.0% (Grand View Research, 2025). Key actors — including OpenAI (Sora), Runway, Pika Labs, and Kling AI — continuously compete to establish new benchmarks in text-to-video synthesis and video extension.

This transformation is not merely technological but deeply social. Research from Runway's Turing Reel study (2026) found that 90% of participants could not reliably distinguish AI-generated footage from authentic video, raising urgent questions about media authenticity, public trust, creative authorship, and platform governance. The same tools empowering independent creators also enable deepfake proliferation and large-scale misinformation.

YouTube occupies a uniquely central position in this landscape. As the world's largest video-sharing platform, its comment sections constitute one of the richest available repositories of organic, unmediated public discourse on AI-generated video — capturing genuine viewer responses spanning technical enthusiasm, ethical alarm, creative anxiety, and cultural commentary in real time and at scale. Despite rapid proliferation of AI video tools, a systematic understanding of how audiences perceive these technologies through organic digital discourse remains limited, a gap this study addresses directly.

II. LITERATURE REVIEW

2.1 Sentiment Analysis and NLP of Digital Discourse

Hutto and Gilbert (2014) introduced VADER (Valence Aware Dictionary and Sentiment Reasoner), a lexicon and rule-based tool specifically designed for social media text, demonstrating strong performance on the short, informal,

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

context-dependent language characteristic of YouTube comments. Devlin et al. (2019) introduced BERT, establishing new benchmarks across NLP tasks; this study employs a fine-tuned DistilRoBERTa derivative for emotion detection. Blei et al. (2003) introduced Latent Dirichlet Allocation (LDA), the probabilistic generative model applied in the topic modelling phase, confirmed by Egger and Yu (2022) as the most interpretable method for short-text social media corpora.

2.2 Public Perception of AI Video Content

Bal et al. (2024) applied opinion mining to 66 YouTube videos discussing generative AI, finding that video content predominantly conveyed optimistic sentiment while comment sections exhibited greater ambivalence — directly informing this study's decision to analyse comments and transcripts as distinct layers. A landmark MDPI Systems study (Recoding Reality, 2025) applied mixed-methods LDA to 11,418 YouTube comments, identifying a three-tiered framework: Socio-Technical Systems, Aesthetics, and Ethics. Xu et al. (2025), analysing Reddit discussions, identified a near-even sentiment split (47% positive, 36.8% negative) closely mirrored by the present study's comment distribution (47.9% positive, 31.5% negative), providing cross-platform validation for structurally ambivalent public sentiment.

2.3 Theoretical Framework

The study integrates four theoretical perspectives. The Technology Acceptance Model (Davis, 1989) and its extension UTAUT (Venkatesh et al., 2003) explain adoption motivations. The Stimulus-Organism-Response (SOR) paradigm frames how AI video content triggers cognitive and emotional responses shaping behavioural intentions. Human-AI co-creation frameworks (Boden, 2004; Elgammal et al., 2017) address creative agency tensions. Diffusion of Innovations Theory (Rogers, 2003) situates attitudinal diffusion within the social ecology of platform-mediated discourse.

III. METHODOLOGY

3.1 Research Design

This study employs a pragmatic mixed-methods sequential explanatory design integrating quantitative computational analysis with qualitative interpretive analysis across three data layers. The three-phase framework encompasses: (1) YouTube comment analysis, (2) video transcript analysis, and (3) engagement metric analysis.

3.2 Data Collection

YouTube comments were extracted using ytdt.digitalmethods.net and the YouTube Data API v3. Video selection employed systematic keyword searches including 'AI video generation', 'AI-generated video', 'AI slop', 'deepfake tools', and 'Sora review'. Inclusion criteria required minimum 5-minute duration, at least 50 comments, and primary substantive focus on AI video content published between 2023–2026. The resulting dataset of 84,905 comments spans 50 videos, 44 channels, and 26 classification categories. Video transcripts were extracted via auto-generated captions; engagement metadata (views, likes, comment count) was retrieved concurrently.

3.3 Analytical Pipeline

All comments underwent standardised preprocessing: URL removal, non-alphabetic character stripping, lowercasing, NLTK tokenisation, stopword removal (English corpus augmented with domain-specific terms), and WordNet lemmatisation. Four analytical techniques were applied: (1) VADER sentiment classification (compound > 0.05 = Positive; < -0.05 = Negative; otherwise Neutral); (2) LDA topic modelling via scikit-learn ($84{,}905 \times 48{,}972$ document-term matrix, k=5 topics); (3) N-gram frequency analysis (unigrams, bigrams, trigrams); and (4) transformer-based emotion detection using j-hartmann/emotion-english-distilroberta-base on 84,902 comments in batches of 500.

3.4 Research Hypotheses

Hypothesis	Statement
H1	YouTube viewers express predominantly mixed/ambivalent sentiment toward AI video tools
H2	Sentiment varies significantly across content categories; tutorials generate more positive sentiment than news
H3	Video transcripts exhibit more positive framing than comment sections

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Hypothesis	Statement
H4	Engagement metrics positively correlate with negative sentiment intensity

IV. RESULTS

4.1 Sentiment Distribution (Comment-Level)

VADER classification of 84,905 comments produced a positively skewed but clearly polarised distribution. Nearly one in three comments was negative, confirming substantial critical discourse co-existing with the positive majority.

Sentiment	Count	Percentage
Positive	40,703	47.9%
Negative	26,750	31.5%
Neutral	17,452	20.6%

Disaggregating by the 26 video classification categories reveals meaningful variation. Tutorial/Review content reached 100% positive; Tutorial recorded 71.0%; Tutorial/Case Study achieved 66.2%. Conversely, Educational Commentary registered 48.9% negative (the highest negative proportion), followed by News Report (41.1%), Investigative Documentary (41.4%), and Talk/Presentation (43.0%). This confirms that audiences engaging with instructional content are more favourable, while content foregrounding ethical and regulatory concerns activates more critical responses.

4.2 LDA Topic Modelling

A five-topic LDA model was selected as offering the most coherent thematic structure from the 84,905 × 48,972 document-term matrix. The five topics and their interpretive labels are summarised below.

Topic	Label	Key Terms
1	AI Authenticity & Content Detection	ai, real, slop, generated, art, tell, cant, never
2	Creative Labour & Economic Displacement	ai, content, job, people, work, artist, tool, keep
3	AI Video Aesthetics & Experimentation	ai, image, stuff, would, sora, probably, short
4	Internet Culture & Platform Dynamics	ai, internet, company, tech, year, new, wont
5	Human-AI Interaction & Systemic Implications	human, ai, world, brain, humanity, machine, data

Topic 1 (AI Authenticity) is the dominant cluster, reflecting persistent detection anxiety — the concern that AI content is becoming indistinguishable from authentic human output. The prominence of 'ai slop' (3,491 bigram occurrences) signals that critical vocabulary has become an established discourse marker. Topic 2 aggregates discourse around economic implications for creative workers, with 'hate ai' appearing 833 times. Topic 5 engages the most philosophically oriented discourse, with vocabulary centred on 'humanity', 'machine', and 'system' reflecting existential commentary on AI's civilisational implications.

4.3 Emotion Detection

Transformer-based emotion detection on 84,902 comments produced the following average scores across five categories. All five emotions score within a 0.02 range, indicating a genuinely mixed emotional landscape with no single dominant emotion.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Emotion	Avg. Score	Interpretation
Surprise	0.107	Dominant; reflects cognitive disruption by rapidly evolving technology
Happy	0.101	Near-equal with surprise; enthusiastic responses to creative capabilities
Angry	0.096	Near-equal with happy; hostile responses to deepfakes and displacement
Sad	0.087	Existential and threat-oriented responses; consistent with Topic 5
Fear	0.087	Present but not numerically dominant emotional register

4.4 Transcript-Level Analysis

VADER classification of 50 video transcripts (58,015 words) revealed a strongly positive corpus: 88.5% classified Positive, 5.8% Negative, 5.8% Neutral — a striking contrast to comment-level ambivalence. However, transformer emotion detection on transcripts revealed Fear (avg. 0.200) and Surprise (0.198) as dominant, with Happy at only 0.159. This divergence indicates creators employ energetic, engaging language driving positive VADER valence, while the underlying subject matter simultaneously generates apprehension. Aspect-based analysis further showed that the aspect 'people' carried the heaviest Fear loading (67 of 136 occurrences), while platform/creator aspects ('channel', 'content') attracted more Happy scores — suggesting creators are optimistic about their own prospects but anxious about broader societal consequences.

4.5 Engagement Metrics

The 50-video dataset showed highly right-skewed view distributions (mean 1,129,312; median 260,374; SD 2,136,055), consistent with power-law dynamics on digital platforms. Interview/Talk content dominated both reach and discussion (avg. 5,466,777 views; 20,096 comments), followed by News/Commentary and Investigative Documentary formats — categories that also exhibited elevated negative sentiment proportions. Tutorial and Tech Demo videos, despite prevalence in the dataset, ranked among the lowest for engagement. Pearson correlations between video duration and both views and comments were weak ($|r| < 0.3$), confirming that content type and emotional valence are far more influential than video length.

V. HYPOTHESIS OUTCOMES

Hypothesis	Outcome	Evidence
H1: Ambivalent sentiment	Supported	47.9% positive / 31.5% negative; five-emotion range within 0.02; 'ai slop' (3,491) and 'hate ai' (833) as prominent bigrams
H2: Sentiment varies by category	Supported	Tutorial/Review: 100% positive vs Educational Commentary: 48.9% negative — self-selection effect confirmed
H3: Transcripts more positive than comments	Supported	88.5% of transcripts positive vs 47.9% of comments; creator-audience discourse asymmetry documented
H4: Engagement correlates with negative sentiment	Partially Supported	High-engagement formats (interviews, investigative) align with higher negative sentiment; direct correlation not computed

VI. DISCUSSION

6.1 Theoretical Implications

The co-existence of high positive sentiment with persistent authenticity anxiety suggests that TAM's conventional binary between acceptance and rejection is insufficient for AI video discourse. Future extensions of TAM should incorporate perceived authenticity risk and creative displacement anxiety as independent constructs. The divergence between VADER-classified positive sentiment and transformer-detected fear at the transcript level confirms that the SOR framework is particularly apt for AI video discourse, where the same stimulus simultaneously generates wonder and apprehension. The LDA Topic 2 creative displacement cluster — organised around 'job', 'artist', and 'work' —

confirms that audiences frame AI as a competitive creative agent rather than mere tool, consistent with Boden's account of transformational creativity but revealing that public reception of AI as collaborator remains deeply contested.

6.2 Managerial Implications

For AI tool developers, authenticity anxiety and deepfake concerns dominate public discourse, constituting a clear mandate for transparency investment. Tools such as Sora, Runway, and Pika Labs should integrate watermarking, provenance metadata, and disclosure frameworks as standard features. The emergence of 'ai slop' as an established negative discourse marker indicates that perceived quality degradation is a direct reputational risk. For content creators, strategies that acknowledge authenticity concerns directly and engage critically with ethical implications are likely to generate higher trust and deeper engagement than purely promotional approaches. For platform operators, recommendation algorithms that privilege emotionally provocative content may be inadvertently amplifying fear responses; governance frameworks should consider balancing engagement maximisation with epistemic quality objectives. For media literacy policymakers, the presence of authenticity anxiety across all 26 classification categories — not confined to news content — underscores the need for broad-based public education programmes on AI content detection and evaluation.

VII. LIMITATIONS AND FUTURE RESEARCH

Three primary limitations constrain the generalisability of findings. First, the study is restricted to English-language YouTube content, introducing acknowledged linguistic and cultural sampling bias; public discourse on AI-generated video in Chinese, Spanish, and Arabic communities may exhibit substantially different patterns given divergent regulatory environments. Second, VADER's performance on domain-specific nuances — irony, sarcasm, and colloquialisms such as 'ai slop' — may introduce classification noise, and the 512-token truncation in transformer models constrains analysis of longer comments. Third, the cross-sectional design (2023–2026) cannot track how individual users' sentiment evolves across the full diffusion cycle.

Future research should pursue: (1) longitudinal panel designs tracking sentiment evolution across the AI video diffusion cycle; (2) cross-linguistic extensions to non-English YouTube communities for comparative cultural analysis; (3) integration of NLP-based discourse analysis with experimental methods to strengthen causal inference regarding AI video exposure and attitude formation; and (4) examination of how platform recommendation algorithms specifically shape thematic cluster amplification in AI video discourse.

VIII. CONCLUSION

This study provides a comprehensive, evidence-based understanding of public perception toward AI-generated video tools through the largest-scale systematic YouTube comment analysis in this domain to date. The three-phase analytical framework — combining VADER sentiment classification, LDA topic modelling, transformer emotion detection, and engagement metric analysis — reveals that public discourse is structured around simultaneous enthusiasm for creative democratisation and apprehension about authenticity erosion, deepfake proliferation, and human creative displacement.

The pronounced divergence between creator framing (88.5% positive transcripts) and audience response (47.9% positive, 31.5% negative comments), alongside the discovery of Fear as the dominant emotion in transcript-level analysis despite positive surface valence, constitutes the study's most novel theoretical contribution. These findings advance understanding of technology acceptance in the generative AI era and carry practical implications for AI developers, content platforms, media professionals, and policy stakeholders. As AI video tools continue to mature and public discourse evolves, the analytical framework established here provides a replicable methodology for ongoing longitudinal monitoring of platform-native public perception.

REFERENCES

1. Bal, A. S., Weidner, K., Hanna, R., & Mills, A. J. (2024). Sentiment and opinion mining of YouTube videos on generative AI in language education. Computers & Education: Artificial Intelligence, 6, 100201.
2. Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. Journal of Machine Learning Research, 3, 993–1022.
3. Boden, M. A. (2004). The creative mind: Myths and mechanisms (2nd ed.). Routledge.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

4. Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. California Law Review, 107(6), 1753–1820.
5. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340.
6. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. NAACL-HLT, 4171–4186.
7. Egger, R., & Yu, J. (2022). A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify Twitter posts. Frontiers in Sociology, 7, 886498.
8. Elgammal, A., Liu, B., Elhoseiny, M., & Mazzone, M. (2017). CAN: Creative adversarial networks. ICCC Proceedings, 96–103.
9. Grand View Research. (2025). AI Video Generator Software Market Size, Share & Trends Analysis Report. Grand View Research.
10. Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. ICWSM, 8(1), 216–225.
11. Kaya, T. (2025). Public reactions to AI-generated video content on YouTube. Journal of Media and Communication Studies, 17(1), 45–62.
12. Liu, Y., Ott, M., Goyal, N., et al. (2019). RoBERTa: A robustly optimized BERT pretraining approach. arXiv:1907.11692.
13. Rogers, E. M. (2003). Diffusion of innovations (5th ed.). Free Press.
14. Runway. (2026). Turing Reel Study: Human Perception of AI-Generated Video. Runway Research.
15. Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., & Ortega-Garcia, J. (2020). Deepfakes and beyond: A survey. Information Fusion, 64, 131–148.
16. Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS Quarterly, 27(3), 425–478.
17. Xu, Z., Li, Y., & Wang, S. (2025). Public discourse on AI-generated video: A topic modelling analysis of Reddit communities. Social Media + Society, 11(1).

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com