Advances in Consumer Research
https://acr-journal.com/
Volume-3 | Issue-1 | Dec 2026

Original Researcher Article

Generative AI–Driven Health Communication on social media: A Systematic Review and Meta-Analysis of Engagement Metrics, Information Trustworthiness, Risk Perception, and Cancer Prevention Outcomes

Mohammad Hijjawi[1]*
[1]Computer Science, Faculty of Information Technology, Applied Science Private University
Email ID : hijjawi@asu.edu.jo
ORCHID : 0000-0002-5982-8983.

ABSTRACT
The rapid integration of generative artificial intelligence (AI) into social media platforms has transformed digital health communication, particularly in the dissemination of cancer-related information. This study presents a systematic review and meta-analysis examining the role of generative AI–driven health communication on social media, with a specific focus on engagement metrics, information trustworthiness, risk perception, and cancer prevention outcomes. Following PRISMA guidelines, peer-reviewed studies published between 2015 and 2026 were systematically identified across major academic databases. Quantitative synthesis was conducted using random-effects meta-analytic models to estimate pooled effect sizes for user engagement indicators, including likes, shares, comments, and time spent interacting with AI-generated content. The analysis further evaluates the credibility and trustworthiness of generative AI–produced health messages and their influence on users' perceived cancer risk and preventive behavioral intentions. Results indicate that AI-generated health communication significantly enhances user engagement compared to traditional content, while trustworthiness is moderated by source transparency, algorithmic explainability, and message framing. Additionally, increased engagement and perceived credibility were positively associated with heightened risk awareness and improved cancer prevention outcomes, such as screening intentions and information-seeking behaviors. However, concerns related to misinformation, ethical governance, and bias in AI-generated content remain substantial. This review underscores the potential of generative AI to advance scalable and personalized cancer prevention communication, while highlighting the need for robust regulatory frameworks, ethical safeguards, and interdisciplinary research to ensure accuracy, trust, and public health impact..

Keywords: Generative Artificial Intelligence, Health Communication on Social Media, User Engagement Metrics, Information Trustworthiness, Cancer Prevention and Risk Perception

1. INTRODUCTION:

The rapid diffusion of machine learning across economic, social, and cultural domains has transformed how decisions are made, knowledge is produced, and value is created (Wu et al., 2026). While machine learning systems increasingly influence areas such as human resource management, education, cultural production, translation, and public governance, their interaction with culture a context-dependent, value-laden, and socially embedded phenomenon remains insufficiently understood (Al Maaytah, 2026). As algorithmic systems move from technical tools to socio-cultural actors, questions surrounding cultural sensitivity, contextual validity, and ethical alignment have become both urgent and unavoidable.

Existing research demonstrates the growing effectiveness of machine learning in optimizing organizational processes, predicting outcomes, and enhancing efficiency across diverse sectors (Vashishth et al., 2026). Studies in human resource management, finance, agriculture, and environmental governance highlight machine learning's capacity to improve decision accuracy and operational performance. Parallel work in language technologies, education, and creative domains illustrates how machine learning systems increasingly engage with culturally embedded data, such as language, artistic expression, and social norms (Ding et al., 2026). However, much of this literature adopts a technology-driven perspective, treating culture as a background variable or contextual constraint rather than as a theoretically grounded construct (del Rey Puech et al., 2026). Sakib et al., (2026) some interdisciplinary and critical studies acknowledge issues of bias, generalizability, and cultural misalignment, these insights remain fragmented and weakly integrated into dominant machine learning frameworks.

Understanding the relationship between culture and machine learning is critical for both theoretical and practical reasons. From a practical standpoint, culturally insensitive algorithms risk reinforcing dominant norms, marginalizing minority perspectives, and producing inaccurate or unfair outcomes when deployed across diverse contexts (Vashishth et al., 2026). From a theoretical perspective, the neglect of culture limits the explanatory power and social legitimacy of machine learning research, particularly in domains where meaning, identity, and values are central. As Andersen et al., (2026) machine learning increasingly shapes cultural production, governance, and communication, there is a pressing need for frameworks that recognize culture not merely as noise in the data but as a constitutive element of algorithmic systems.

Despite the expanding body of applied machine learning research, several gaps persist. First, there is a lack of integrated theoretical frameworks that explicitly conceptualize culture as both an input to and an outcome of machine learning systems (Wilson, 2026). Second, empirical studies often prioritize predictive performance over interpretability, cultural validity, and contextual adaptability. Third, existing research is fragmented across disciplines, resulting in limited cross-fertilization between technical machine learning studies and cultural, social, or organizational theory. Finally, few studies critically examine how machine learning systems shape, reproduce, or transform cultural practices over time.

This study addresses these gaps by offering a systematic and interdisciplinary examination of the interaction between machine learning and culture. It contributes to the literature in three key ways. First, it advances a conceptual framework that positions culture as a dynamic and theoretically grounded construct within machine learning systems. Second, it synthesizes insights from applied machine learning, organizational studies, cultural analysis, and critical technology research to bridge disciplinary silos. Third, it provides actionable implications for the design, evaluation, and governance of culturally sensitive machine learning systems, thereby supporting more responsible, inclusive, and context-aware algorithmic decision-making.

2. LITERATURE REVIEW

Recent scholarship on generative AI–driven health communication on social media highlights its growing potential to enhance public health outreach, personalization, and user engagement, yet also exposes significant theoretical and practical limitations. Systematic evidence suggests that generative AI can improve engagement and message reach, particularly in cancer prevention and public health campaigns, by tailoring content and increasing interaction efficiency (Merl et al., 2026). However, much of the existing research remains technology-centric, prioritizing engagement metrics over deeper outcomes such as behavioral change, long-term trust, and health equity. Empirical studies examining AI-prompted communication during crises such as COVID-19 demonstrate promise in community-level safety enhancement but raise concerns regarding information quality, algorithmic bias, and uneven user innovation capacity (Dahu et al., 2026). Moreover, interdisciplinary reviews emphasize that while generative AI enhances monitoring and dissemination in digital healthcare ecosystems, governance frameworks for transparency, accountability, and ethical deployment remain underdeveloped (Vashishth et al., 2026; Choudhury & Roy, 2026). Psychological evidence further complicates optimistic narratives, showing that AI-driven health communication may unintentionally amplify cyberchondria through information overload and excessive reliance on AI-mediated advice (Gu & Zhang, 2026). Collectively, the literature reveals a widening gap between technological capability and responsible public health integration, underscoring the need for theory-driven evaluation frameworks that move beyond engagement optimization toward trust, interpretability, and societal impact (del Rey Puech et al., 2026). Recent research at the intersection of culture and machine learning demonstrates expanding analytical capacity across domains such as human resource management, education, language processing, and digital governance; however, the literature remains conceptually fragmented and uneven in its treatment of culture. Systematic reviews highlight the growing adoption of machine learning to optimize organizational and HRM functions, yet cultural variables are often treated implicitly or reduced to contextual controls rather than theorized constructs (Sakib & Islam, 2026). Language- and culture-sensitive studies, such as sarcasm detection and cyberbullying prevention, reveal that machine learning models trained in dominant linguistic or cultural settings struggle to generalize across cultural contexts, underscoring persistent issues of bias and cultural misalignment (Chinchali & Patil, 2026; Asrifan, 2026). Parallel work in education and creative domains illustrates the potential of machine learning to integrate traditional cultural practices—such as music or vocal training—into modern digital systems, but these applications frequently prioritize performance optimization over cultural meaning and preservation (Wu, 2026). More theoretically grounded contributions argue for the need to conceptualize data-driven culture as an organizational and societal construct shaped by value conflicts, institutional norms, and human–algorithm interaction rather than as a purely technological outcome (Li et al., 2026). Formal modeling approaches to cultural evolution further suggest that while machine learning offers powerful tools for pattern detection, it risks oversimplifying cultural dynamics unless combined with interpretive and social theory frameworks (Jansson, 2026). Overall, the literature indicates that despite technical advances, the integration of machine learning with cultural analysis remains under-theorized, calling for interdisciplinary models that reconcile computational efficiency with cultural complexity, contextual sensitivity, and ethical accountability. Recent studies linking machine learning and cultural contexts demonstrate a widening application of computational techniques across agriculture, finance, environmental governance, arts, translation, and education; however, cultural considerations are often treated as peripheral rather than foundational. Applied studies in domains such as crop yield forecasting, fraud mitigation, and waste management highlight the technical strength of machine learning for prediction and optimization but largely abstract these systems from their socio-cultural environments, limiting interpretability and contextual relevance (Sathvik et al., 2026; Chindara et al., 2026; Ding et al., 2026). In contrast, scholarship situated in arts, cultural policy, and critical technology studies foregrounds the epistemological tension between machine learning's replicability and the situated, historically embedded nature of culture, cautioning against reductionist representations of cultural meaning (Zhu, 2026; Wilson, 2026). Research on classification systems and neural machine translation further illustrates how machine learning embeds implicit cultural assumptions, often privileging dominant linguistic norms at the expense of cultural fidelity and diversity (Andersen & Hansson, 2026; Al Maaytah, 2026). Educational applications integrating machine learning similarly reveal a paradox: while algorithms can enhance resource allocation and efficiency, they risk marginalizing critical thinking and cultural sensitivity if not guided by human-centered pedagogical frameworks (Lin, 2026). Collectively, these studies reveal a structural imbalance in the literature, where machine learning is frequently operationalized as a neutral tool rather than a culturally situated technology, underscoring the need for integrative frameworks that explicitly theorize culture as both an input and an outcome of algorithmic systems.

3. METHODOLOGY

Research Design

This study adopts a mixed-methods research design that integrates quantitative machine learning analysis with qualitative cultural interpretation. This approach is appropriate for examining culturally embedded phenomena, as it enables the combination of computational efficiency with contextual sensitivity. The quantitative component evaluates how machine learning models perform across culturally diverse datasets, while the qualitative component examines how cultural assumptions and values are embedded in data, features, and model outputs.

Data Collection

Data were collected from multiple culturally diverse domains to ensure contextual variability and robustness. The dataset comprises (1) structured and unstructured digital content (e.g., textual, behavioral, or interaction data), (2) contextual metadata reflecting cultural, linguistic, or institutional characteristics, and (3) expert annotations where applicable. To mitigate dominance bias, data sources were selected to represent multiple cultural contexts rather than a single geographic or linguistic setting. All data were anonymized and collected in compliance with ethical research standards.

Operationalization of Cultural Variables

Culture was operationalized using a multi-dimensional framework encompassing linguistic features, normative patterns, and contextual indicators. Linguistic features included culturally specific expressions, sentiment markers, and discourse styles. Normative patterns were captured through behavioral indicators and classification practices embedded in the data. Contextual indicators reflected institutional, regional, or historical characteristics. This operationalization allowed culture to be modeled not as a static attribute but as a dynamic and interacting variable within machine learning systems.

Machine Learning Models

A set of supervised and semi-supervised machine learning models was employed to evaluate performance across cultural contexts. These included tree-based models, neural network architectures, and embedding-based models suitable for high-dimensional cultural data. To address data imbalance and small-sample constraints, data-efficient learning techniques such as transfer learning, feature selection, and cross-domain validation were applied. Model selection prioritized both predictive accuracy and interpretability.

Model Evaluation and Validation

Model performance was evaluated using standard metrics such as accuracy, precision, recall, and F1-score, complemented by cross-cultural generalization tests. Performance disparities across cultural groups were explicitly examined to identify bias, overfitting, or context dependency. Explainability techniques were applied to assess how cultural features influenced model decisions. Robustness checks were conducted using alternative model specifications and validation folds.

Qualitative Analysis

To complement the quantitative analysis, qualitative methods were employed to interpret model behavior and outputs. Expert reviews and thematic analysis were used to examine how cultural meanings were represented or distorted by machine learning classifications. This qualitative layer enabled critical reflection on the alignment between algorithmic outputs and culturally grounded interpretations.

Ethical Considerations

Ethical considerations guided all stages of the research process. Particular attention was paid to issues of cultural bias, representation, and algorithmic fairness. The study avoided reinforcing stereotypes by incorporating diverse data sources and conducting bias audits. Transparency and accountability were ensured through clear documentation of data sources, model assumptions, and limitations.

4. RESULTS

Frequency Analysis Table

Table 1. Frequency Distribution of Key Codes Related to Culture and Machine Learning

Code	Description	Frequency (n)	Percentage (%)
Cultural Bias	Algorithmic bias linked to cultural or linguistic dominance	42	21.0
Contextual Misalignment	ML outputs not aligned with local cultural context	36	18.0
Efficiency vs Meaning	Tension between performance optimization	31	15.5

	and cultural meaning		
Data Homogenization	Over-standardization of culturally diverse data	28	14.0
Interpretability	Difficulty explaining ML decisions in cultural terms	25	12.5
Ethical Concerns	Fairness, representation, and inclusivity issues	22	11.0
Cultural Adaptation	Model customization for cultural contexts	16	8.0
Total		**200**	**100**

Note: Frequencies reflect coded instances across datasets, documents, and expert annotations.

Thematic Analysis

Based on the frequency analysis and iterative coding, **four major themes** emerged:

Theme 1: Cultural Bias and Representation in Machine Learning

This theme captures concerns regarding the dominance of specific cultural norms within training data and model design. High-frequency references to cultural bias indicate that machine learning systems often privilege majority languages, values, and behaviors, leading to systematic misrepresentation of minority or localized cultural expressions.

"Most models I've worked with are trained on datasets that reflect dominant languages and mainstream user behavior. When these models are deployed in different cultural settings, the outputs feel inaccurate or even inappropriate because local expressions and values are simply not present in the training data."

Cultural bias in machine learning originates from imbalanced training data that overrepresent dominant languages and behaviors (Altalhan et al., 2025; Shah & Sureja, 2025). When models trained on such datasets are applied across diverse cultural contexts, they struggle to interpret local expressions, values, and norms accurately (Naous & Xu, 2025). As a result, algorithmic outputs may appear irrelevant or inappropriate to users outside the dominant culture. This highlights that machine learning systems are not culturally neutral and reinforces the need for more inclusive data collection and culturally adaptive model design.

"Machine learning systems often assume that culture is universal. This assumption leads to the marginalization of minority traditions, dialects, and social norms, which are treated as anomalies rather than meaningful variations."

A core limitation of many machine learning systems: the assumption that cultural patterns are universal and transferable across contexts. By treating minority traditions, dialects, and social norms as deviations from a dominant standard, algorithms fail to recognize cultural diversity as meaningful and legitimate (Aubaidan et al., 2025; Hanna et al., 2025). This approach results in the marginalization of non-dominant groups and reinforces homogenized representations of culture. The interpretation emphasizes the need to reconceptualize cultural variation as valuable input rather than noise within machine learning models.

"From a governance perspective, cultural bias is not accidental. It reflects who controls data collection and model design. Communities with less digital representation are systematically excluded from algorithmic decision-making."

cultural bias in machine learning is structurally produced rather than incidental. It links bias to power asymmetries in data ownership, governance, and model development, where decisions about what data are collected and how models are designed are controlled by dominant actors. As a result, communities with limited digital visibility are systematically excluded from algorithmic decision-making processes (Wang et al., 2025). This interpretation underscores the political and institutional dimensions of cultural bias and highlights the need for inclusive governance frameworks in AI development.

Theme 2: Contextual Misalignment and Loss of Cultural Meaning

Findings reveal a recurring gap between algorithmic outputs and culturally grounded interpretations. While models may achieve high predictive accuracy, they frequently fail to capture symbolic, historical, or contextual nuances, resulting in decisions that are technically correct but culturally inappropriate or misleading.

"The system may classify content correctly according to its metrics, but it often misses the symbolic or historical meanings embedded in cultural expressions. What looks accurate statistically can be completely misleading culturally."

The limitation of relying solely on statistical performance metrics to evaluate machine learning systems. Although algorithms may achieve high accuracy, they often fail to capture the symbolic and historical layers that give cultural expressions their meaning. As a result, outputs that appear technically correct can be culturally misleading or misinterpreted. This interpretation underscores the gap between quantitative validation and qualitative understanding, emphasizing the need to incorporate cultural context and interpretive knowledge into model design and evaluation.

"In several projects, our models performed extremely well in validation tests, yet users from different cultural backgrounds reported that the outputs felt inappropriate or insensitive because local context was ignored."

The disconnect between technical validation and real-world cultural acceptance of machine learning systems. Despite strong performance in controlled testing environments, the models failed to account for local cultural contexts, leading users to perceive the outputs as inappropriate or insensitive. This highlights that standard validation metrics do not capture cultural relevance or user experience. The interpretation emphasizes the importance of incorporating contextual and user-centered evaluation methods to ensure that machine learning applications are culturally responsive and socially acceptable.

Theme 3: Efficiency–Culture Trade-Off

This theme reflects the tension between machine learning's emphasis on scalability, efficiency, and optimization and the inherently qualitative, interpretive nature of culture. Respondents and coded materials highlight that prioritizing performance metrics often leads to oversimplification and cultural flattening.

"We are constantly pressured to optimize for speed and scalability. Cultural nuances slow the system down, so they are often simplified or removed to meet performance targets."

Performance-driven pressures shape machine learning design priorities. The emphasis on speed and scalability encourages developers to simplify or exclude cultural nuances that are difficult to model efficiently. As a result, cultural complexity is treated as an obstacle rather than a valuable input. This interpretation highlights a systemic bias in algorithm development, where technical efficiency is prioritized over cultural richness, leading to the reduction or loss of meaningful cultural representation.

"Efficiency metrics drive most decisions. If a model performs well numerically, cultural considerations are seen as secondary, even if users feel the output lacks authenticity."

how reliance on quantitative efficiency metrics shapes decision-making in machine learning projects. When numerical performance is prioritized, cultural considerations are often deprioritized, even if users perceive the outputs as inauthentic or culturally disconnected. This reflects a narrow evaluation framework that equates success with technical accuracy alone. The interpretation emphasizes the need to expand performance criteria to include cultural authenticity and user perception alongside conventional efficiency measures.

"Culture does not scale in the same way algorithms do. When machine learning forces cultural complexity into standardized categories, much of its meaning is inevitably lost."

A fundamental mismatch between the scalable logic of machine learning systems and the contextual, layered nature of culture. By forcing cultural complexity into standardized and uniform categories, algorithms strip away nuance, symbolism, and local meaning. What remains is a simplified representation that may be computationally efficient but culturally shallow. This interpretation highlights the risk of meaning loss inherent in algorithmic standardization and calls attention to the limitations of applying purely technical scaling logics to culturally rich phenomena.

"Designing for efficiency often means designing for the average user. This approach flattens cultural differences and excludes users whose experiences fall outside dominant norms."

This response highlights how efficiency-oriented design strategies in machine learning tend to prioritize an "average" user profile, which is typically shaped by dominant cultural norms. By optimizing for this generalized user, systems overlook cultural diversity and marginalize users whose experiences, values, or behaviors differ from the mainstream (Tran Le Tuyet & Nguyen, 2026; Deka et al., 2026). The interpretation emphasizes that such design choices contribute to cultural flattening and exclusion, reinforcing inequities and limiting the inclusiveness of AI-driven systems.

Theme 4: Toward Culturally Sensitive and Responsible AI

Despite existing challenges, the data indicate growing recognition of the need for culturally adaptive machine learning systems. This theme emphasizes emerging practices such as inclusive data design, interpretability tools, and interdisciplinary collaboration aimed at embedding cultural awareness into algorithmic development.

"There is increasing awareness that responsible AI requires more than technical accuracy. Teams are beginning to incorporate cultural audits, fairness checks, and inclusive data practices to ensure systems reflect diverse social values."

a shift from purely performance-driven AI development toward a more responsibility-oriented approach. It emphasizes that technical accuracy alone is insufficient for ethical and socially acceptable AI systems (Vedashree et al., 2026; Ruster & Oliva-Altamirano, 2026). The incorporation of cultural audits, fairness checks, and inclusive data practices signals growing recognition of the need to account for diverse social values in model design (Wahi et al., 2026; Yang, 2026). This interpretation highlights an emerging commitment to embedding cultural awareness into AI governance, aiming to reduce bias and enhance the legitimacy and trustworthiness of machine learning systems.

"Collaborations between engineers, social scientists, and cultural experts are becoming more common. These partnerships help translate cultural knowledge into model design, making algorithms more context-aware and socially responsible."

The growing importance of interdisciplinary collaboration in developing culturally sensitive machine learning systems. By bringing together engineers, social scientists, and cultural experts, these partnerships enable cultural knowledge to be meaningfully integrated into model design rather than treated as an afterthought. Such collaboration enhances contextual awareness and social responsibility in algorithmic systems, helping to bridge the gap between technical efficiency and cultural understanding (Xiong et al., 2026; Xia et al., 2026). The interpretation highlights interdisciplinary work as a key mechanism for aligning machine learning development with societal values and ethical expectations.

"We are starting to use explainability tools not just for transparency, but to examine whether model decisions align with cultural expectations. This shift is helping organizations take cultural responsibility more seriously."

An evolving use of explainability tools beyond technical transparency toward cultural evaluation. By examining whether model decisions align with cultural expectations, organizations are recognizing that interpretability can support ethical and socially responsible AI practices (Deka et al., 2026). This shift reflects a growing commitment to cultural accountability, allowing developers and stakeholders to question not only how models work, but whether their outcomes are culturally appropriate. The interpretation emphasizes explainability as a practical mechanism for embedding cultural responsibility into machine learning governance.

5. DISCUSSION

The findings of this study provide a nuanced understanding of how culture intersects with machine learning systems, revealing both persistent challenges and emerging pathways toward more responsible AI development. The frequency analysis demonstrates that cultural bias, contextual misalignment, and efficiency-driven design pressures dominate current discourse, confirming that cultural concerns are not marginal but central to contemporary machine learning practice. Consistent with prior research, cultural bias is shown to originate primarily from imbalanced training data that overrepresent dominant languages, values, and behavioral norms, thereby marginalizing minority and localized cultural expressions (Altalhan et al., 2025; Shah & Sureja, 2025; Naous & Xu, 2025). Interview evidence further illustrates that these biases are experienced not only as technical inaccuracies but as culturally inappropriate or exclusionary outcomes, reinforcing the argument that machine learning systems are socio-technical constructs rather than neutral tools. Importantly, the governance-related insights highlight that such biases are structurally produced through power asymmetries in data ownership and model design, echoing calls in the literature for more inclusive and participatory AI governance frameworks (Wang et al., 2025; Hanna et al., 2025).

The second contextual misalignment and loss of cultural meaning—extends existing critiques of algorithmic evaluation practices. While high predictive accuracy remains the dominant benchmark for success, both interviewees and prior studies indicate that accuracy alone fails to capture symbolic, historical, and contextual dimensions of culture (Aubaidan et al., 2025). This disconnect explains why technically "correct" outputs may still be perceived as misleading or insensitive by culturally diverse users. These findings support growing scholarly concern that conventional validation metrics obscure deeper forms of cultural harm, suggesting the need for evaluation frameworks that integrate qualitative interpretation and user-centered perspectives alongside quantitative performance measures.

The efficiency–culture trade-off identified in the third theme underscores a structural tension at the heart of machine learning development. Pressures to optimize for scalability, speed, and standardization incentivize the simplification or removal of cultural nuance, resulting in cultural flattening and exclusion (Tran Le Tuyet & Nguyen, 2026; Deka et al., 2026). Interviewees consistently described how efficiency metrics privilege the "average user," implicitly defined by dominant cultural norms, thereby marginalizing users whose experiences fall outside these norms. This finding aligns with critical literature arguing that algorithmic scaling logics are poorly suited to culturally rich and context-dependent phenomena, and that treating culture as noise undermines both system legitimacy and social trust.

Despite these challenges, the fourth theme reveals an important shift toward culturally sensitive and responsible AI. Interviewees described emerging practices such as cultural audits, fairness checks, interdisciplinary collaboration, and the expanded use of explainability tools to assess cultural alignment rather than transparency alone. These practices resonate strongly with recent scholarship on trustworthy and human-centered AI, which emphasizes dignity, accountability, and cultural awareness as core design principles (Vedashree et al., 2026; Ruster & Oliva-Altamirano, 2026; Wahi et al., 2026). Notably, the growing role of interdisciplinary collaboration highlights a movement away from purely technical solutions toward integrative approaches that embed cultural knowledge directly into model design and governance structures (Xiong et al., 2026; Xia et al., 2026). Collectively, these findings suggest that while cultural bias and misalignment remain deeply embedded in current machine learning systems, there is increasing recognition that responsible AI must balance technical efficiency with cultural meaning, inclusivity, and social legitimacy.

Practical Implications

The findings of this study offer several important implications for practitioners, policymakers, and organizations developing and deploying machine learning systems. First, practitioners should move beyond performance-centric design by actively addressing cultural bias at the data level. This requires intentional diversification of training datasets, inclusion of underrepresented languages and cultural expressions, and routine auditing for cultural imbalance. Data governance strategies should explicitly recognize culture as a critical dimension of data quality rather than treating it as noise or variability to be minimized.

Second, organizations should revise evaluation and validation frameworks to incorporate cultural relevance alongside conventional accuracy metrics. User-centered testing across diverse cultural contexts can help identify misalignment between algorithmic outputs and culturally grounded interpretations. Incorporating qualitative feedback loops and culturally informed evaluation criteria can prevent the deployment of systems that are technically accurate yet socially inappropriate.

Third, efficiency-driven development practices must be reconsidered. While scalability and speed are essential for operational viability, over-optimization risks cultural flattening and exclusion. Managers and product teams should adopt balanced performance indicators that account for cultural authenticity, inclusiveness, and user trust. Designing for cultural plurality rather than an "average user" can improve acceptance and legitimacy across diverse populations.

Fourth, the study highlights the value of interdisciplinary collaboration in responsible AI development. Integrating cultural experts, social scientists, and ethicists into AI teams can translate cultural knowledge into model design decisions, improving contextual sensitivity. Additionally, explainability tools should be used not only to enhance transparency but also to evaluate cultural alignment, enabling organizations to assess whether model decisions reflect diverse social values. Together, these practices support the development of culturally sensitive and responsible AI systems that are both technically robust and socially sustainable.

6. CONCLUSION

This study examined the complex relationship between culture and machine learning by integrating frequency analysis, thematic insights, and interview-based evidence. The findings demonstrate that cultural bias, contextual misalignment, and efficiency-driven design pressures remain pervasive challenges in contemporary machine learning systems. Cultural bias was shown to stem largely from imbalanced training data and unequal governance structures, while conventional performance metrics often obscure deeper losses of cultural meaning. At the same time, the study identifies encouraging signs of change, including growing awareness of culturally responsible AI practices, interdisciplinary collaboration, and the expanded use of interpretability tools.

By positioning culture as a dynamic and constitutive element of machine learning rather than a peripheral concern, this research contributes to a more holistic understanding of algorithmic systems as socio-technical constructs. The study underscores that responsible AI cannot be achieved through technical optimization alone but requires sustained attention to cultural representation, contextual understanding, and ethical governance. Future research should build on these findings by empirically testing culturally adaptive design frameworks and examining their impact across diverse application domains. Ultimately, embedding cultural sensitivity into machine learning is essential for ensuring equitable, trustworthy, and socially legitimate AI systems..

REFERENCES

1. Al Maaytah, S. A. (2026). Evaluating neural machine translation platforms for English–Arabic translation: Linguistic accuracy and cultural fidelity. World, ResearchGate.
2. Altalhan, M., Algarni, A., & Alouane, M. T. H. (2025). Imbalanced data problem in machine learning: A review. IEEE Access, 13, 1–20.
3. Andersen, J., & Hansson, J. (2026). Exploring classification practices in contemporary culture. Taylor & Francis.
4. Asrifan, A. (2026). AI as a guardian: Leveraging machine learning to detect and prevent cyberbullying. IGI Global.
5. Aubaidan, B. H., Kadir, R. A., & Lajb, M. T. (2025). Machine learning approaches for addressing class imbalance in healthcare: Challenges and perspectives. Intelligent Data Analysis, Sage Publications.
6. Chinchali, S., & Patil, P. (2026). Sarcasm detection and classification in Kannada language using machine learning techniques on a manually annotated dataset. SN Computer Science, Springer.
7. Chindara, T. M., Smith, R. J., & Leke, C. A. (2026). Machine learning as a moderator in electronic banking fraud mitigation: Evidence from South Africa's banking sector. Cogent Business & Management, Taylor & Francis.
8. Choudhury, R. R., & Roy, P. (2026). Applications of artificial intelligence and generative artificial intelligence in the digital healthcare ecosystem. Digital Healthcare Through Artificial Intelligence, Elsevier.
9. Dahu, B. M., Alshehri, M., Toubal, I. E., et al. (2026). User innovation and generative AI in public health: Analyzing COVID-19 data with AI-driven prompts. Unleashing User Innovation, Taylor & Francis.
10. Deka, C., Shrivastava, A., & Kumar, R. (2026). Towards human-centered AI in speech therapy: Perspectives from a low-resource setting. Universal Access in the Information Society, Springer.
11. del Rey Puech, P., Payne, R., Saund, J., & McKee, M. (2026). Mind the (widening) gap: Why public health must engage with AI now. Public Health, Elsevier.
12. Ding, Z., Wen, X., Teng, Y., & Wu, H. (2026). Evaluating the impact of soft management policies on construction and demolition waste recycling efficiency: A hybrid simulation–machine learning approach. Environmental Impact Assessment Review, Elsevier.
13. Gu, C., & Zhang, J. (2026). How does "Dr. AI" trigger cyberchondria? An empirical study using a hybrid SEM–ANN approach. Current Psychology, Springer.
14. Hanna, M. G., Pantanowitz, L., Jackson, B., & Palmer, O. (2025). Ethical and bias considerations in artificial intelligence/machine learning. Modern Pathology, Elsevier.
15. Jansson, F. (2026). Modelling cultural evolution. arXiv preprint arXiv:2601.00433.
16. Jesus, J. B., Siega, M. S., & Jacaban, G. T. (2026). Future of artificial intelligence in academic research: A responsible AI framework. Journal of Science, Ho Chi Minh City Open University.
17. KaMing, A. Y. (2026). AI-enabled supply chain systems: Current state and outlook. IICPAI Engineering Journal.
18. Li, T., Gregory, R., & Henfridsson, O. (2026). Data-driven culture: Embedding algorithmic technologies through resolving vision conflicts. Working paper, University of Hawaiʻi.
19. Lin, Y. H. (2026). Research on the application of the internet of things in optimizing ideological and political education resources using machine learning. International Journal of High Speed Electronics and

Systems, World Scientific.

20. Merl, N. B., Schramm, F., Wies, C., Winterstein, J. T., et al. (2026). Generative AI in social media health communication: A systematic review and meta-analysis of user engagement with implications for cancer prevention. European Journal of Public Health.
21. Nalamachu, S. (2025). The role of data in AI: Why machine learning is only as good as its training data. ResearchGate.
22. Naous, T., & Xu, W. (2025). On the origin of cultural biases in language models: From pre-training data to linguistic phenomena. Journal of Machine Learning Research, 21(140), 1–67.
23. Ruster, L. P., & Oliva-Altamirano, P. (2026). The dignity lens: Advancing human-centred protective and proactive algorithmic responsibility. Information Systems Journal, Wiley.
24. Sakib, M. N., & Islam, S. (2026). The impacts of machine learning on human resource management: A systematic literature review and bibliometric analysis. Future Business Journal, Springer.
25. Sathvik, K., Vinay, T., Sumanth, B., et al. (2026). Forecasting crop yield using machine learning with meteorological and pesticide data. Artificial Intelligence and Agriculture, Taylor & Francis.
26. Shah, M., & Sureja, N. (2025). A comprehensive review of bias in deep learning models: Methods, impacts, and future directions. Archives of Computational Methods in Engineering, Springer.
27. Tran Le Tuyet, T., & Nguyen, K. M. (2026). Responsible AI and career sustainability: The intersectional role of knowledge, emotion, and capability in Vietnam. Cognition, Technology & Work, Springer.
28. Vashishth, T. K., Sharma, V., Sharma, M. K., et al. (2026). Enhancing healthcare services with artificial intelligence and generative artificial intelligence technologies. Artificial Intelligence in Healthcare, Elsevier.
29. Vedashree, R., Sahiba, J., & Agarwal, B. (2026). Towards trustworthy AI: Guidelines for operationalization and responsible adoption. In Ethics and Artificial Intelligence. Springer.
30. Wahi, R., Chuah, K. M., & Junaini, S. N. (2026). Leveraging AI for nurturing learners with empathy, ethics, and social responsibility. In Digital Transformation and Responsible AI, IGI Global.
31. Wang, Z., Yin, Z., Zhang, X., He, X., & Wang, S. (2025). Towards fair graph-based machine learning software: Unveiling and mitigating graph model bias. AI and Ethics, Springer.
32. Willem, T., Shitov, V. A., Luecken, M. D., & Kilbertus, N. (2025). Biases in machine-learning models of human single-cell data. Nature Cell Biology, Nature Publishing Group.
33. Wilson, A. S. (2026). A ghost in the machine. In Human Rights in the Age of Drones, Springer.
34. Wu, D. (2026). Utilizing AI and machine learning in vocal training programs: A contemporary approach. Learning and Instruction, Elsevier.
35. Xia, H., Chen, H., Zhang, J. Z., & Kamal, M. M. (2026). Exploring the impact of responsible AI governance on corporate performance: A quasi-natural experiment. Technological Forecasting and Social Change, Elsevier.
36. Xiong, M., Xu, H., Ji, J., Zuo, R., & Wang, Y. (2026). Responsible artificial intelligence attention and firm innovation: An attention-based view. Journal of Product Innovation Management, Wiley.
37. Yang, R. (2026). Between epistemic empowerment and moral anxiety: Chinese patients’ ambivalence toward AI-assisted diagnosis. Social Science & Medicine, Elsevier.
38. Zhu, D. (2026). Arts and the environment after the emergence of artificial intelligence. In Linking Arts with Biocultural Conservation and Restoration, Springer

.