Journal of Digital Marketing and Communication

Review
Volume 6(2), 2026, 318−333
https://doi.org/10.53623/jdmc.v6i2.1476

The Role of Artificial Intelligence in Sustainable Marketing: A Systematic Literature Review Between Products and Social Media

Ratu Ratna Juwita

Bachelor of Communication Science Program, Faculty of Law, Social and Political Sciences, Indonesia Open University, Jakarta, Indonesia

Correspondence Author: juwitaratu@gmail.com

SUBMITTED: 15 October 2026; REVISED: 26 October 2026; ACCEPTED: 3 November 2026

ABSTRACT: The increasing adoption of Artificial Intelligence (AI) in digital marketing has transformed conventional message automation into interactive communication between consumers and intelligent agents. However, the prevalence of greenwashing on social media and limited transparency in AI-driven communication have intensified consumer skepticism toward sustainability claims. This study aims to examine the role of AI as a communication agent in sustainable marketing and develop a conceptual framework for strengthening consumer trust through Meta-Authenticity. A qualitative research approach was employed using a Systematic Literature Review (SLR) guided by the PRISMA 2020 framework. The review examined 21 peer-reviewed journal articles, five theoretical textbooks, supplemented by four methodological references. The selected literature was analyzed using thematic analysis and systematic coding to identify recurring patterns, conceptual relationships, and emerging challenges. The findings indicate that AI-driven communication can enhance the perceived credibility of sustainable products by providing consistent, evidence-based, and technically objective information. However, excessive algorithmic personalization may contribute to sustainable filter bubbles, selective information exposure, and algorithmic greenwashing, potentially undermining consumer trust. This study concludes that effective AI-driven sustainable marketing requires more than technological efficiency and personalized communication. Consumer trust depends on algorithmic transparency, data accountability, and ethical consistency in human–machine interactions. By integrating Human–Machine Communication (HMC) theory and Platform Political Economy, the proposed Meta-Authenticity framework offers a conceptual foundation for developing transparent, accountable, and trustworthy AI-mediated sustainability communication.

KEYWORDS: Artificial Intelligence (AI); Sustainable marketing; human-machine communication; meta-authenticity; algorithmic greenwashing

1. Introduction

Sustainable marketing has become an important communication strategy for promoting environmentally responsible products, increasing consumer awareness, and encouraging

sustainable consumption. Effective sustainability communication is expected to establish a credible connection between environmentally friendly products and consumers by providing reliable information about environmental benefits and corporate sustainability practices. The rapid development of digital technologies and social media platforms has significantly transformed how sustainability-related information is produced, distributed, and consumed. In particular, Artificial Intelligence (AI) has increasingly been integrated into digital marketing through recommendation systems, automated content generation, personalized advertising, and predictive targeting [1−3]. These developments have transformed AI from a passive technological tool into an active communication agent capable of shaping sustainability narratives and influencing consumer perceptions. Consequently, understanding the communicative role of AI has become increasingly important for evaluating the credibility and effectiveness of sustainable marketing in contemporary digital environments.

Despite these technological advancements, sustainability communication faces significant challenges associated with greenwashing, misleading environmental claims, and declining consumer trust. The widespread dissemination of sustainability-related content on social media has created a paradox in which greater access to environmental information does not necessarily lead to greater confidence in sustainable products. Instead, excessive promotional messages, unverified environmental claims, and complex algorithmic recommendation mechanisms may intensify consumer skepticism toward green advertising [1, 2]. Furthermore, AI-driven personalization can influence which sustainability messages consumers encounter, potentially limiting exposure to diverse perspectives and reinforcing selective environmental narratives. These challenges raise concerns about whether AI-mediated communication genuinely improves public understanding of sustainability or primarily serves commercial objectives. Accordingly, the credibility of sustainable marketing increasingly depends not only on the environmental characteristics of products but also on the transparency, reliability, and ethical integrity of the communication processes through which sustainability claims are presented [1, 4].

Previous studies have investigated various aspects of AI-driven marketing and sustainability communication, including comparisons between AI-generated and human influencers [5], emerging trends in green marketing [6], and the contribution of marketing strategies to brand equity. Other research has examined algorithmic communication, consumer perceptions, and ethical challenges associated with automated digital marketing practices [7, 8]. These studies provide valuable insights into the technological capabilities of AI and its potential influence on consumer behavior. However, the existing literature remains fragmented across marketing, communication, technological, and ethical perspectives. Most importantly, the relationships among AI communicative agency, platform commercial interests, sustainable product attributes, and consumer perceptions of authenticity have received limited integrated theoretical attention. This fragmentation highlights the need for a comprehensive synthesis that considers AI not only as a marketing technology but also as a communicative actor operating within economically driven digital platforms [1−3].

Human–Machine Communication (HMC) provides a theoretical foundation for understanding how AI participates in the construction and interpretation of communication. According to Guzman et al. [3], AI can be conceptualized as a communicative actor rather than merely a passive medium for transmitting information. Within this perspective, AI-mediated communication can be examined through three interrelated dimensions: action, meaning-

making, and relating. Action refers to the capacity of AI systems to autonomously perform tasks, generate recommendations, and respond to user interactions. Meaning-making describes the ability of AI to construct, frame, and communicate narratives that influence audience interpretations. Relating concerns the development of personalized interactions and perceived social relationships between users and intelligent systems [3]. In sustainable marketing, these dimensions explain how AI can communicate environmental information, construct sustainability narratives, and establish personalized connections with consumers. Consequently, the credibility of sustainability messages may depend not only on corporate reputation and product attributes but also on consumers' perceptions of the intelligence, objectivity, and communicative agency of AI systems [3, 5].

Although HMC explains the communicative capabilities of AI, it does not fully account for the economic structures and institutional interests that influence algorithmic communication. Therefore, this study incorporates Platform Political Economy, particularly the perspective developed by Fuchs [2], to examine the relationship between digital communication and commercial platform structures. From this perspective, social media platforms operate within the broader dynamics of digital capitalism, in which user interactions, behavioral information, and personal data are increasingly commodified to support economic accumulation. These commercial structures influence how algorithms prioritize, distribute, and personalize digital content. While AI-driven personalization can improve the relevance and accessibility of sustainability information, platform optimization mechanisms may prioritize engagement, advertising performance, and sales conversions over comprehensive environmental education [2, 3]. This tension may contribute to the emergence of sustainable filter bubbles, characterized by repeated exposure to selectively personalized environmental content, and algorithmic greenwashing, in which AI-mediated communication amplifies exaggerated, misleading, or insufficiently substantiated sustainability claims [1, 2, 6, 7]. Integrating HMC with Platform Political Economy therefore provides a more comprehensive explanation of how AI communicative agency interacts with platform commercial logic to shape sustainability narratives and influence consumer perceptions.

Within this theoretical context, Meta-Authenticity provides a conceptual perspective for understanding how consumers evaluate the credibility of AI-mediated sustainability communication. Drawing on Kim [1], Meta-Authenticity extends traditional interpretations of authenticity beyond the physical characteristics or perceived naturalness of products to encompass the trustworthiness of technologically mediated communication processes. In digital environments, consumers may recognize that sustainability messages are generated, personalized, or distributed by AI while continuing to perceive such messages as credible when the underlying communication processes demonstrate transparency, accountability, and ethical consistency [1, 4]. In this study, Meta-Authenticity is examined through three interrelated dimensions. First, algorithmic transparency refers to the clarity with which AI systems disclose how information is collected, processed, and used to generate sustainability-related messages [1, 8, 9]. Second, data accountability concerns the availability of verifiable evidence, scientific justification, and reliable information supporting environmental claims [1, 10, 11]. Third, ethical consistency refers to the alignment between AI-generated sustainability narratives and the actual environmental practices, commitments, and performance of the organizations communicating those messages [1, 7, 12, 13]. Together, these dimensions provide a conceptual

basis for evaluating the authenticity of AI-mediated environmental communication and its potential implications for consumer trust.

Despite the growing body of research on AI-driven marketing, sustainable communication, and digital consumer behavior, a significant theoretical gap remains in understanding how AI communicative agency and platform commercial structures jointly influence perceived authenticity and consumer trust. Existing studies have generally examined these dimensions separately, providing limited systematic integration of HMC, Platform Political Economy, and Meta-Authenticity within sustainable marketing [1−3, 14, 15]. The novelty of this study lies in integrating these theoretical perspectives into a unified conceptual framework that explains the relationships among AI agency, platform commercial logic, sustainability narrative construction, and consumer trust. Specifically, the proposed framework conceptualizes AI communicative agency and platform commercial logic as interacting influences on sustainability narratives, including the potential formation of sustainable filter bubbles and algorithmic greenwashing. These communication processes are subsequently evaluated through Meta-Authenticity, in which algorithmic transparency, data accountability, and ethical consistency provide the basis for understanding perceived credibility and consumer trust. As illustrated in Figure 1, the proposed framework integrates the communicative functions of AI with platform commercial mechanisms to explain how sustainability narratives and Meta-Authenticity may influence consumer trust. Unlike approaches that primarily emphasize technological efficiency or marketing performance, this framework incorporates the structural and ethical dimensions of AI-mediated communication, offering a more comprehensive perspective on the opportunities and risks associated with sustainable digital marketing [1−3, 10, 13, 16].

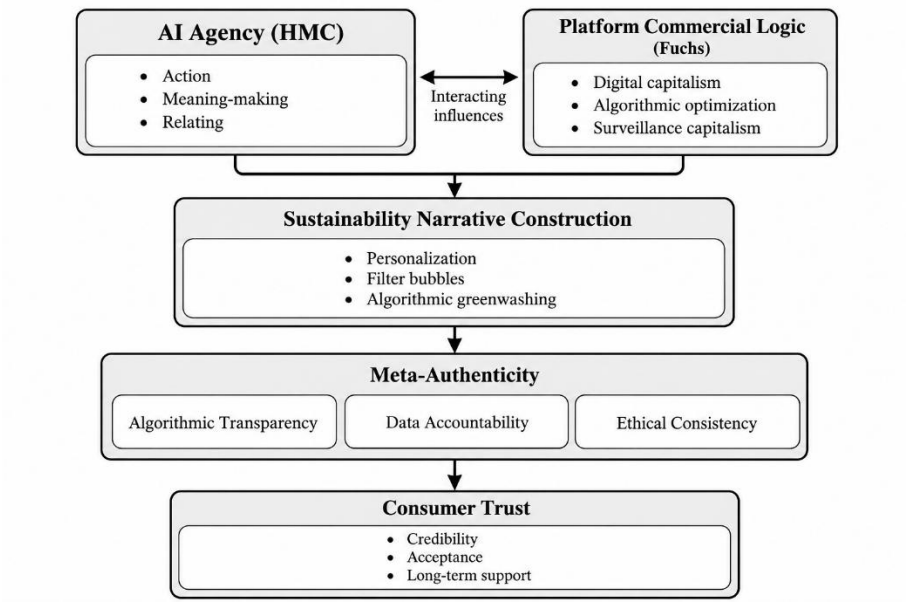


Figure 1. Proposed Conceptual Framework Integrating AI Agency, Platform Commercial Logic, Meta-Authenticity, And Consumer Trust.

Therefore, this study aims to systematically review and synthesize the existing literature on the role of AI as a communication agent in sustainable marketing and to develop an integrative Meta-Authenticity framework for understanding consumer trust. Specifically, the study seeks to (1) examine the communicative functions of AI in presenting and promoting

sustainable products; (2) analyze how platform algorithms and commercial structures influence the construction and distribution of sustainability narratives; (3) investigate the relationships among brand equity, perceived authenticity, and consumer trust; (4) identify ethical challenges associated with algorithmic personalization, sustainable filter bubbles, and greenwashing; and (5) integrate HMC and Platform Political Economy into a conceptual framework explaining Meta-Authenticity in human–machine interactions [1-3, 9]. Through a Systematic Literature Review (SLR) guided by PRISMA 2020, this study seeks to consolidate fragmented theoretical perspectives and contribute to the development of more transparent, accountable, and ethically responsible AI-mediated sustainable marketing practices. The proposed framework also provides a theoretical foundation for future empirical investigations into the relationships among AI communication, platform governance, perceived authenticity, and consumer trust.

2. Materials and Methods

2.1. Research Design and Literature Search Strategy.

This study employed a qualitative research approach using a Systematic Literature Review (SLR) to examine the role of Artificial Intelligence (AI) as a communication agent in sustainable marketing and its implications for consumer trust. The review was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) framework to ensure transparency, methodological consistency, and reproducibility in the identification, screening, and selection of relevant literature [17]. The SLR approach was selected to synthesize fragmented theoretical and empirical findings concerning AI-mediated communication, platform commercial logic, sustainability narratives, and Meta-Authenticity.
The literature search was conducted using Google Scholar and Scopus, covering peer-reviewed journal articles relevant to AI-driven marketing, sustainability communication, digital platforms, and consumer behavior. Boolean operators were applied to combine the principal research concepts and identify relevant publications. The following search string was used:
("AI" OR "artificial intelligence") AND ("sustainable marketing" OR "green marketing" OR "sustainability marketing") AND ("social media" OR "digital platform" OR "social network") AND ("product" OR "brand" OR "consumer")

The selection criteria emphasized peer-reviewed journal articles published in English or Indonesian that directly addressed the communicative, marketing, or ethical dimensions of AI in digital sustainability contexts. Publications focusing exclusively on technical algorithm development without relevance to marketing communication, consumer trust, or sustainability were excluded. Grey literature, duplicate records, and publications with insufficient bibliographic information were also excluded. The final analytical corpus comprised 21 peer-reviewed journal articles, supplemented by five academic textbooks providing theoretical foundations for Human–Machine Communication, Platform Political Economy, and authenticity in digital communication. Thus, 26 publications contributed to the substantive literature synthesis. Four additional methodological references were used to guide the systematic review procedures, thematic analysis, and coding reliability assessment [18−21]. The methodological references were treated separately from the substantive analytical corpus.

2.2. Literature Screening and Selection.

The literature identification and selection process followed the PRISMA 2020 framework, comprising identification, screening, eligibility assessment, and inclusion. As illustrated in Figure 2, the initial search identified 2,450 records, including 2,015 from Google Scholar and 435 from Scopus. After removing 355 duplicate records, 2,095 publications remained for title and abstract screening. At this stage, 1,585 records were excluded because they did not meet the predefined eligibility criteria, including language restrictions, publication type, and topical relevance. The remaining 510 publications underwent full-text eligibility assessment. Of these, 489 were excluded because their primary focus was unrelated to the research objectives, their full texts were inaccessible, or their bibliographic information was insufficient for reliable evaluation. Consequently, 21 journal articles met the inclusion criteria and were retained for the primary thematic synthesis. Additionally, five theoretical textbooks and four methodological references were included separately to support conceptual interpretation, and framework development, resulting in a total of 30 references included in the study.

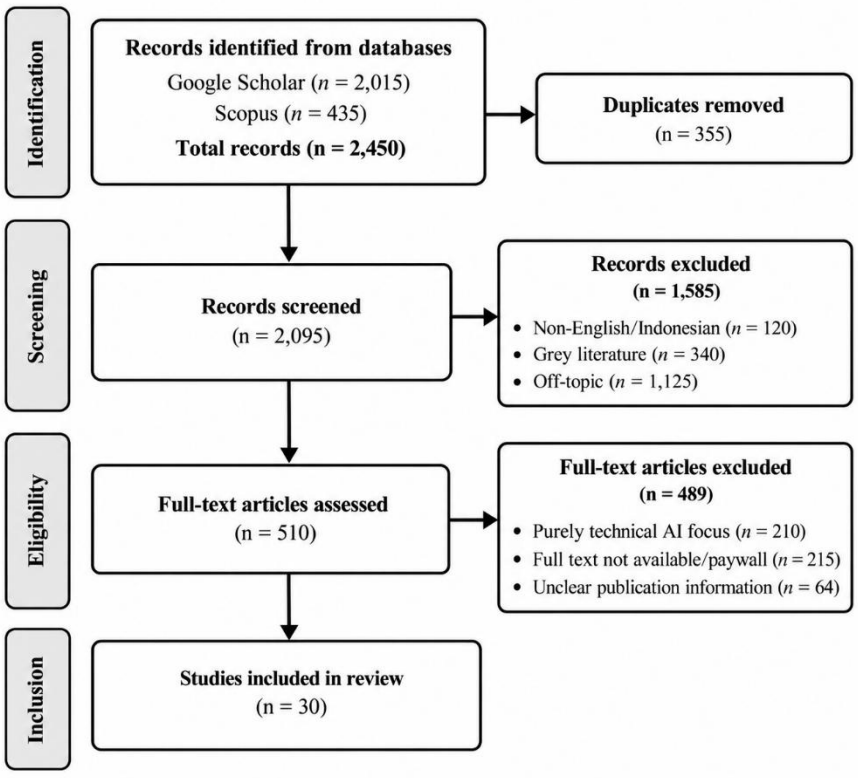


Figure 2. PRISMA 2020 flow diagram of literature selection and supplementary references.

2.3. Thematic Analysis and Coding Reliability.

Qualitative data analysis was conducted using thematic analysis based on the six-phase framework proposed by Braun and Clarke [18]. The process involved familiarization with the selected literature, generation of initial codes, identification of potential themes, review of themes, definition and refinement of themes, and interpretation of the findings. The analysis focused on the 21 journal articles and five theoretical textbooks included in the substantive analytical corpus. During the initial coding phase, relevant statements, theoretical arguments, and empirical findings were systematically extracted from the selected publications, producing

approximately 225 descriptive codes. Following the data reduction and categorization principles described by Miles, Huberman, and Saldaña [21], overlapping and conceptually similar codes were consolidated into 68 thematic codes. These codes were subsequently organized into broader analytical categories corresponding to the five research questions, including AI communicative agency, platform commercial logic, sustainability narrative construction, consumer trust, and ethical challenges. The resulting thematic structure provided the foundation for developing the proposed Meta-Authenticity framework, particularly its dimensions of algorithmic transparency, data accountability, and ethical consistency.

Because the analysis was conducted by a single researcher, coding stability was assessed using an intra-coder reliability procedure based on the code–recode method. This approach evaluates the consistency of coding decisions made by the same researcher at different points in time rather than agreement between independent coders. The procedure was informed by established qualitative content analysis and reliability assessment principles [20, 21]. A random sample of 50 coded text excerpts was selected from the analytical corpus for the stability assessment. The researcher initially coded these excerpts during the first assessment (T1) and independently recoded the same material after a 14-day interval (T2) to reduce potential recall effects. The comparison identified 44 consistent coding decisions and six discrepancies between the two assessments. Coding stability was evaluated using the Holsti agreement coefficient, which measures the proportion of matching coding decisions between two assessments relative to the total number of coding decisions. The resulting coefficient was 0.88, indicating 88% agreement between the initial coding and subsequent recoding. This result suggests a relatively high level of coding stability across the two assessment periods. Nevertheless, the coefficient represents intra-coder consistency rather than complete objectivity or the absence of interpretive bias [20, 21]. Accordingly, the findings were interpreted through systematic comparison of the selected literature, refinement of thematic categories, and consideration of the theoretical relationships underlying the proposed conceptual framework [17−19].

3. Results and Discussion

3.1. Thematic Synthesis of AI Communication Benefits and Ethical Risks.

The thematic analysis of 21 peer-reviewed journal articles, supported by five theoretical textbooks, identified several interconnected themes concerning the opportunities and ethical challenges associated with Artificial Intelligence (AI) in sustainable marketing communication. The findings indicate that AI offers substantial opportunities to improve information accessibility, communication efficiency, message personalization, and the presentation of evidence-based environmental claims. However, these advantages are accompanied by ethical concerns related to algorithmic opacity, selective information exposure, misleading sustainability narratives, and the commercialization of environmental information. Table 2 summarizes the principal communication benefits and corresponding ethical risks identified through the literature synthesis. The communication benefits of AI are primarily associated with its capacity to process complex information and translate technical environmental data into accessible consumer-oriented messages. This capability is particularly relevant for products whose sustainability performance can be supported by measurable indicators, including energy efficiency, carbon emissions, resource consumption, and life cycle

assessment results. AI systems can organize these data into personalized explanations, potentially improving consumers' understanding of product characteristics and environmental performance [5, 22−24]. Gerlich [5] further highlights the relevance of AI communication in contexts involving technically oriented products, where consumers may place greater emphasis on objective information than on emotional persuasion. Nevertheless, the perceived objectivity of AI should not be equated with factual accuracy, because the reliability of AI-generated communication remains dependent on the quality of its underlying data, verification procedures, and system design.

Table 2. Thematic Synthesis of AI Communication Benefits and Ethical Risks in Sustainable Marketing.

Thematic dimension	Communication benefits	Potential ethical risks	Supporting references
Information accuracy	Efficient processing and presentation of environmental data	Dissemination of inaccurate or unverified sustainability claims	[3, 5, 22−25]
Content personalization	Delivery of relevant sustainability messages to targeted consumers	Sustainable filter bubbles and selective information exposure	[2, 11, 14, 23, 26, 27]
Automated communication	Consistent and scalable dissemination of sustainability information	Algorithmic greenwashing and reduced human oversight	[1, 10, 12, 13, 23]
Consumer interaction	Personalized responses and improved information accessibility	Excessive data collection and privacy concerns	[3, 9, 28]
Brand communication	Greater visibility of sustainable products and environmental initiatives	Manipulative messaging and superficial environmental positioning	[7, 8, 11]
Platform optimization	Efficient content distribution and audience engagement	Prioritization of commercial performance over environmental accuracy	[2, 7, 13]

In contrast, the synthesis identifies significant ethical concerns when AI-driven communication operates without adequate transparency and verification. Generative AI can produce persuasive sustainability narratives even when the underlying environmental claims are incomplete, exaggerated, or unsupported by empirical evidence [1, 10, 11]. This creates conditions under which algorithmic greenwashing may emerge, particularly when automated content production is guided primarily by engagement optimization and commercial performance. From the perspective of Platform Political Economy, such practices reflect the tension between the economic objectives of digital platforms and the informational requirements of credible sustainability communication [2, 10, 13]. Consequently, the effectiveness of AI in sustainable marketing cannot be evaluated solely through communication reach, engagement rates, or conversion performance.

Another important finding concerns the potential development of sustainable filter bubbles through algorithmic personalization. Recommendation systems may repeatedly expose environmentally conscious consumers to similar sustainability narratives while limiting their engagement with alternative viewpoints or competing environmental evidence [2, 14]. Although personalized communication can increase message relevance and strengthen consumer interest, excessive targeting may reduce informational diversity and reinforce pre-existing preferences. These findings suggest that sustainable marketing strategies should balance personalization with transparency, information diversity, and independent verification [2, 26, 29]. Overall, the thematic synthesis indicates that the benefits of AI-mediated

sustainability communication depend on the extent to which computational efficiency is supported by ethical governance and credible environmental evidence [1, 9, 22].

3.2. Development of the Meta-Authenticity Conceptual Framework.

The integration of Human–Machine Communication (HMC), Platform Political Economy, and authenticity-related perspectives provides the theoretical foundation for the proposed Meta-Authenticity framework. HMC explains how AI participates in communication through automated actions, narrative construction, and personalized interactions, whereas Platform Political Economy examines how commercial incentives and algorithmic governance influence the distribution and prioritization of sustainability messages [2, 3]. Meta-Authenticity connects these perspectives by providing a conceptual basis for evaluating whether AI-mediated sustainability communication is perceived as transparent, accountable, and ethically consistent [1, 9].

The framework begins with sustainability-related information, including product environmental attributes, ecological certifications, and life cycle assessment data. These inputs are processed and communicated through AI systems that perform the HMC functions of action, meaning-making, and relating. However, the resulting messages are also shaped by the commercial structures of digital platforms, including recommendation algorithms, engagement optimization, and data-driven advertising. This interaction influences how sustainability narratives are constructed, personalized, and presented to consumers. Consequently, communication outcomes depend not only on the technical capabilities of AI but also on the credibility of environmental information and the institutional conditions governing its distribution [2, 3, 10].

Within the proposed framework, Meta-Authenticity is represented by three interrelated dimensions: algorithmic transparency, data accountability, and ethical consistency. Algorithmic transparency concerns the disclosure and explainability of AI-mediated communication processes. Data accountability emphasizes the availability of verifiable evidence supporting sustainability claims, while ethical consistency concerns the alignment between communicated environmental commitments and actual organizational practices. Table 3 summarizes these dimensions and their proposed implications for consumer trust.

Table 3. Dimensions of Meta-Authenticity and Their Implications for Consumer Trust.

Dimension	Conceptual definition	Key evaluation indicators	Expected implications
Algorithmic transparency	Openness regarding AI-mediated communication processes	AI disclosure, explainability, data-use transparency	Greater perceived communication credibility
Data accountability	Verification of environmental information and claims	Scientific evidence, certification, traceable environmental data	Reduced uncertainty about sustainability claims
Ethical consistency	Alignment between sustainability messages and organizational practices	Environmental performance, consistency of claims, responsible communication	Stronger perceived authenticity and trust

As summarized in Table 3, these dimensions provide complementary mechanisms for evaluating the credibility of AI-mediated sustainability communication. Algorithmic transparency enables consumers to understand whether and how AI influences the information they receive, while data accountability establishes the evidential basis for environmental claims. Ethical consistency extends this assessment beyond communication practices by considering whether organizational actions support the sustainability narratives being

promoted. Together, these dimensions suggest that consumer trust may be strengthened when AI-generated information is transparent, verifiable, and consistent with actual environmental performance [1, 7, 9, 10, 11, 13, 15]. The proposed framework also recognizes the possibility of contrasting communication outcomes. When transparency, accountability, and ethical consistency are maintained, AI-mediated sustainability communication may enhance perceived credibility, consumer acceptance, and green brand equity. Conversely, limited transparency, inadequate verification, and inconsistencies between environmental claims and corporate practices may increase skepticism and perceptions of greenwashing. These relationships should be interpreted as theoretical propositions derived from the literature synthesis rather than empirically established causal effects. Further quantitative and qualitative investigations are therefore necessary to validate the proposed framework across different product categories, digital platforms, and consumer populations.

3.3. Discussion of the Five Research Questions.

3.3.1. AI as a Primary Communication Agent.

The first research question examines how AI functions as a communication agent in sustainable marketing. The literature synthesis indicates that AI has progressed beyond conventional marketing automation to participate actively in the generation, framing, and delivery of sustainability narratives. Within the HMC perspective, this transformation can be understood through action, meaning-making, and relating, which collectively describe the operational, interpretive, and relational functions of intelligent communication systems [3, 5]. AI enables organizations to translate complex environmental information into accessible communication formats, particularly when sustainability claims involve measurable technical characteristics. For example, information concerning energy efficiency, carbon emissions, and resource conservation can be processed into comparative product descriptions or personalized consumer recommendations. Previous findings suggest that AI-generated communication may be perceived as informative and credible when it presents technical evidence in a consistent and understandable manner [5, 23]. However, perceptions of technological objectivity should not be interpreted as evidence that AI systems are inherently neutral or free from commercial influence. The findings also indicate that the effectiveness of AI communication depends on product characteristics and consumer expectations. For technically oriented products, objective environmental indicators may play a central role in establishing credibility. In contrast, products associated with identity, lifestyle, and emotional expression may require communication that incorporates social meaning, cultural context, and interpersonal authenticity. Therefore, AI should be understood as a potentially influential communication agent whose effectiveness varies according to the informational and relational requirements of sustainable marketing contexts.

3.3.2. Interactions Among Sustainable Products, AI Algorithms, and Digital Platforms.

The second research question concerns the interaction among sustainable product attributes, AI algorithms, and digital platform structures. The findings suggest that these components operate as an interconnected communication system rather than independent determinants of consumer behavior. Product attributes provide the substantive environmental information, AI

algorithms influence how this information is interpreted and personalized, and digital platforms determine the conditions under which messages are distributed and encountered [2, 14]. From a Platform Political Economy perspective, algorithmic recommendation systems are influenced by commercial incentives associated with advertising revenue, engagement, and user retention [2]. These mechanisms may increase the visibility of sustainability content that generates strong audience responses while reducing exposure to less engaging but scientifically relevant environmental information. Such conditions create potential tensions between marketing effectiveness and the communication of comprehensive sustainability evidence. The sustainable filter bubble represents one possible consequence of this interaction. Repeated exposure to similar environmental narratives may reinforce existing consumer preferences while restricting opportunities to encounter alternative perspectives or critically evaluate competing sustainability claims. Nevertheless, the extent to which this phenomenon occurs is likely to depend on platform design, recommendation mechanisms, and user behavior. Accordingly, sustainable marketing communication should consider not only the accuracy of individual messages but also the broader informational environment created by algorithmic content distribution [2, 14, 26].

3.3.3. Contributions to Brand Equity and Consumer Behavior.

The third research question investigates how AI-mediated sustainability communication contributes to brand equity and consumer behavior. The thematic synthesis identifies four potential pathways through which AI may influence consumer perceptions and relationships with sustainable brands: brand awareness, stakeholder engagement, innovation, and environmental information transparency . These pathways illustrate how AI communication capabilities may contribute to brand development while highlighting the importance of credible sustainability information. Table 4 presents the proposed relationships and their theoretical implications.

Table 4. Potential Pathways Linking AI Communication to Green Brand Equity and Consumer Behavior.

Pathway	AI communication mechanism	Potential contribution	References
Brand awareness	Automated targeting and content recommendations	Increased visibility and recognition of sustainable products	[14]
Stakeholder engagement	Conversational agents and personalized interaction	Improved responsiveness and consumer engagement	[8, 28]
Innovation	Analysis of consumer feedback and crowdsourced information	Identification of sustainability-related product preferences	[7, 9]
Environmental transparency	Communication of verifiable ecological information	Improved perceived credibility and green brand trust	[9, 11, 15]

As indicated in Table 4, AI may support green brand equity by improving the accessibility, relevance, and consistency of sustainability-related information. Automated targeting can increase brand exposure, while conversational agents facilitate interactions that may strengthen consumer engagement. Furthermore, AI-assisted analysis of consumer feedback can provide insights into environmental preferences and emerging sustainability expectations, potentially informing product development and communication strategies [6, 7, 10, 14]. sHowever, increased brand visibility and engagement do not necessarily translate into stronger consumer trust or sustainable purchasing behavior. The credibility of environmental claims remains essential, particularly when consumers encounter conflicting sustainability information.

Acatrinei et al. [9] provide relevant insights into the relationship between consumer experiences with AI-mediated interactions and subsequent behavioral responses. Nevertheless, the direction and strength of these relationships may vary across consumer groups, product categories, and platform environments. The findings therefore suggest that AI contributes most meaningfully to sustainable brand development when communication efficiency is accompanied by transparent and verifiable environmental information.

3.3.4. Ethical Implications and Distortions of Meta-Authenticity.

The fourth research question examines the ethical challenges associated with AI-mediated sustainability communication. The thematic synthesis identifies algorithmic greenwashing, limited transparency, selective information exposure, and inconsistencies between environmental claims and organizational practices as major concerns. These challenges are particularly relevant because AI systems can generate large volumes of persuasive communication without independently establishing the scientific validity of the information being presented [9, 21, 27−30]. Algorithmic greenwashing can be conceptualized as the generation, amplification, or selective distribution of misleading sustainability claims through AI-driven communication processes. Unlike conventional greenwashing, which primarily concerns organizational misrepresentation of environmental performance, algorithmic greenwashing emphasizes the role of automated content generation, recommendation mechanisms, and personalized persuasion in potentially reinforcing misleading environmental narratives. This distinction highlights the need to evaluate both the accuracy of sustainability claims and the technological processes through which they are communicated [1, 9, 24, 28].

The Meta-Authenticity perspective provides a basis for addressing these concerns by emphasizing algorithmic transparency, data accountability, and ethical consistency. Transparency allows consumers to recognize the role of AI in communication, accountability establishes the evidential basis of environmental information, and ethical consistency connects sustainability narratives with actual organizational behavior. When these conditions are absent, AI-mediated communication may intensify skepticism and weaken perceived authenticity. Conversely, their implementation may support more credible and responsible sustainability communication [1, 9, 27]. These findings suggest that responsible AI adoption in sustainable marketing requires organizational governance mechanisms extending beyond technological performance. Such mechanisms may include independent verification of environmental claims, disclosure of AI-generated promotional content, regular assessment of algorithmic outputs, and consistency checks between marketing messages and documented environmental performance. These measures represent practical implications of the conceptual synthesis and require further evaluation to establish their effectiveness in different organizational and regulatory contexts.

3.3.5. Integration of HMC and Platform Political Economy.

The fifth research question addresses how HMC and Platform Political Economy can be integrated to explain the formation of Meta-Authenticity and consumer trust. The synthesis indicates that the two perspectives offer complementary explanations of AI-mediated sustainability communication. HMC emphasizes the communicative capabilities of AI, including autonomous action, narrative construction, and personalized interaction, whereas Platform Political Economy examines the commercial structures that influence how these

capabilities are deployed [2, 3]. The integration of these perspectives suggests that AI-generated sustainability messages emerge through the interaction between technological agency and platform commercial incentives. Although AI can improve the accessibility and personalization of environmental information, the commercial objectives of digital platforms may influence which messages receive visibility and how sustainability narratives are framed. Consumer evaluations of these messages may therefore depend on both the perceived credibility of AI communication and the transparency of the broader digital environment. Meta-Authenticity provides a conceptual connection between these processes and consumer trust by emphasizing the conditions under which technologically mediated sustainability communication may be perceived as credible. Rather than assuming a direct relationship between AI adoption and consumer trust, the proposed framework suggests that transparency, accountability, and ethical consistency influence how consumers interpret AI-generated environmental narratives. This perspective extends conventional approaches to sustainable marketing by incorporating communicative agency, platform governance, and ethical evaluation within a single analytical structure [1−3]. Overall, the findings contribute to the theoretical understanding of AI-mediated sustainable marketing by demonstrating the importance of examining technological capabilities alongside commercial incentives and ethical considerations. The proposed Meta-Authenticity framework provides an integrated basis for future research investigating how AI communication, platform structures, and environmental claim verification interact to influence consumer trust. However, because the framework was developed through qualitative literature synthesis, its proposed relationships remain theoretical and should be empirically tested before broader causal conclusions can be established.

4. Conclusions

This study systematically examined the role of Artificial Intelligence (AI) as a communication agent in sustainable marketing and its implications for consumer trust. Based on a qualitative synthesis of 21 journal articles supported by theoretical literature, the findings indicate that AI enhances sustainability communication through automated content generation, personalized interactions, and the presentation of evidence-based environmental information. However, these benefits are accompanied by ethical challenges, including algorithmic greenwashing, sustainable filter bubbles, limited transparency, and the prioritization of commercial interests over environmental credibility. By integrating Human–Machine Communication (HMC) and Platform Political Economy, this study proposes the Meta-Authenticity framework, emphasizing algorithmic transparency, data accountability, and ethical consistency as essential dimensions of trustworthy AI-mediated communication. The framework suggests that consumer trust depends not only on technological efficiency and message personalization but also on the credibility of environmental claims and their alignment with actual organizational practices. These findings contribute to sustainable marketing theory by connecting AI communicative agency, platform governance, and consumer perceptions of authenticity within an integrated conceptual framework. Nevertheless, the study is limited by its qualitative literature-based approach. Future research should empirically validate the proposed framework across different industries, digital platforms, and consumer groups to establish its practical applicability.

Author Contributions

Ratu Ratna Juwita was solely responsible for the study's conceptualization, methodology, literature search, data collection, analysis, interpretation, manuscript preparation, review, and editing.

Acknowledgments

The author independently conducted this systematic literature review, including protocol development, PRISMA-guided literature screening, data extraction, thematic analysis, and manuscript preparation, without external assistance.

Conflicts of Interest

The author declares no conflicts of interest.

Competing Interests

The author declares no financial, personal, or professional competing interests that could have influenced the research or its findings.

References

[1] Kim, D.O. (2025). Meta-authenticity and fake but real virtual influencers: A framework for artificial sociality analysis and ethics. *New Media & Society*, *27*, 5568–5591. https://doi.org/10.1177/14614448251338272.

[2] Fuchs, C. (2021). Social Media: A Critical Introduction, 3rd ed.; SAGE Publications: London, UK.

[3] Guzman, A.L.; McEwen, R.; Jones, S. (Eds.) (2023). The SAGE Handbook of Human-Machine Communication; SAGE Publications: London, UK.

[4] Evans, W.D.; Bardus, M.; French, J. (2024). A vision of the future: Harnessing artificial intelligence for strategic social marketing. *Businesses*, *4*, 196–210. https://doi.org/10.3390/businesses4020013.

[5] Gerlich, M. (2025). The shifting influence: Comparing AI tools and human influencers in consumer decision-making. *AI*, *6*, 11. https://doi.org/10.3390/ai6010011.

[6] Alabi, R.O.; Dampare Addo, L.A.; Ijebor, L.; Anaege, A.E. (2025). The intersection of Green AI, digital advertising, and corporate sustainability: A systematic review. *American Journal of Economics and Business Innovation*, *4*, 234–244. https://doi.org/10.54536/ajebi.v4i2.5373.

[7] Mohammad-Ebrahimzadeh, S.; Khajeheian, D.; Arbatani, T.R.; Labafi, S.; Sepasgozar, S. (2025). Sustainable knowledge-based enterprise products using AI-powered social media for enhancing brand equity: A scientometric review. *Sustainability*, *17*, 8427. https://doi.org/10.3390/su17188427.

[8] Dash, G.; Sharma, C.; Sharma, S. (2023). Sustainable marketing and the role of social media: An experimental study using Natural Language Processing (NLP). *Sustainability*, *15*, 5443. https://doi.org/10.3390/su15065443.

[9] Acatrinei, C.; Apostol, I.G.; Barbu, L.N.; Popa, R.-G.C.; Orzan, M.-C. (2025). Artificial intelligence in digital marketing: Enhancing consumer engagement and supporting sustainable behavior through social and mobile networks. *Sustainability*, *17*, 6638. https://doi.org/10.3390/su17146638.

[10] Alsaffarini, E.; Awwad, B.S. (2026). The integration between green marketing and artificial intelligence to achieve corporate sustainability. *Sustainability*, *18*, 3597. https://doi.org/10.3390/su18073597.
[11] Nicolau, A.-M.; Petcu, P. (2026). Beyond green labels: Leveraging blockchain, IoT, and AI for enhanced traceability and verification of green marketing claims in transnational agri-food supply chains. *Sustainability*, *18*, 782. https://doi.org/10.3390/su18020782.
[12] Farid, F.; Putra, S.M.; Bachri, S.; Sutomo, M. (2025). Digital marketing and sustainable marketing to create competitive advantage in SMEs: The moderating effect of value co-creation. *Jurnal Manajemen Bisnis*, *16*, 249–280. https://doi.org/10.18196/mb.v16i1.23033.
[13] Israfilzade, K. (2026). AI-generated vs. human-created sustainable advertising: Effects of source disclosure and environmental claim strength on perceived greenwashing and green trust among Generation Z consumers. *Sustainability*, *18*, 7270. https://doi.org/10.3390/su18147270.
[14] Garg, V.; Bohara, S.; Srivastav, A. (2025). AI-driven sustainability marketing transforming consumers' perception toward eco-friendly brands. *Discover Sustainability*, *6*, 984. https://doi.org/10.1007/s43621-025-01934-y.
[15] Gündüzyeli, B. (2024). Artificial intelligence in digital marketing within the framework of sustainable management. *Sustainability*, *16*, 10511. https://doi.org/10.3390/su162310511.
[16] Jakubczak, J.K.; Chmielewska-Muciek, D.; Iwanicka, K. (2026). Evaluating multimodal AI for greenwashing detection: A comparative analysis of ChatGPT, Claude, and Gemini in ESG reports. *Sustainability*, *18*, 236. https://doi.org/10.3390/su18010236.
[17] Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; Chou, R.; Glanville, J.; Grimshaw, J.M.; Hróbjartsson, A.; Lalu, M.M.; Li, T.; Loder, E.W.; Mayo-Wilson, E.; McDonald, S.; Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, n71. https://doi.org/10.1136/bmj.n71.
[18] Braun, V.; Clarke, V. (2021). Thematic Analysis: A Practical Guide; SAGE Publications: London, UK.
[19] Sugiyono. (2022). Metode Penelitian Kuantitatif, Kualitatif, dan R&D, 27th ed.; Alfabeta: Bandung, Indonesia.
[20] Krippendorff, K. (2019). Content Analysis: An Introduction to Its Methodology, 4th ed.; SAGE Publications: Thousand Oaks, CA, USA. https://doi.org/10.4135/9781071878781.
[21] Miles, M.B.; Huberman, A.M.; Saldaña, J. (2014). Qualitative Data Analysis: A Methods Sourcebook, 3rd ed.; SAGE Publications: Thousand Oaks, CA, USA.
[22] Mei, Y.; Geng, L.; Cao, X.; Xie, Y. (2025). Artificial intelligence in green marketing: A systematic literature review. *Sustainability*, *17*, 10382. https://doi.org/10.3390/su172210382.
[23] Alsaffarini, E.; Awwad, B.S. (2026). Artificial intelligence in sustainable marketing: How AI personalization impacts consumer purchase decisions. *Sustainability*, *18*, 1123. https://doi.org/10.3390/su18021123.
[24] Obi-Okoye, S.O.; Nofal, R. (2026). AI-enabled sustainable digital marketing systems: A Stimulus–Organism–Response framework of customer engagement and institutional sustainability capabilities. *Sustainability*, *18*, 8925. https://doi.org/10.3390/su18178925.
[25] Redda, E.H. (2026). AI-driven digital marketing and responsible consumption: The mediating role of marketing intelligence in advancing SDG 12. *Sustainability*, *18*, 3912. https://doi.org/10.3390/su18083912.
[26] Tuten, T.L. (2024). Social Media Marketing, 5th ed.; SAGE Publications: London, UK.
[27] Mousa, M.M.; Rashed, A.S.; Akaileh, M.; Zamil, A.M.; Ahmed, H.A.M.; Abdelghani, A.A.A. (2026). Artificial intelligence marketing technologies and consumer purchasing decisions: The moderating role of virtual customer experience and implications for sustainable consumption in

telecommunications service environments. *Sustainability*, *18*, 2674. https://doi.org/10.3390/su18062674.

[28] Walters, H.D.; Hammond, R.M. (Eds.) (2025). AI in Marketing: Applications, Insights, and Analysis; Routledge: London, UK.

[29] Chaffey, D.; Ellis-Chadwick, F. (2022). Digital Marketing, 8th ed.; Pearson Education: Harlow, UK.

[30] Salhab, H.; Zoubi, M.; Khrais, L.T.; Estaitia, H.; Harb, L.; Al Huniti, A.; Morshed, A. (2025). AI-driven sustainable marketing in Gulf Cooperation Council retail: Advancing SDGs through smart channels. *Administrative Sciences*, *15*, 20. https://doi.org/10.3390/admsci15010020.

© 2026 by the authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).