JURNAL STUDI KOMUNIKASI

Volume 9 **Ed 3, November 2025** **Page 793 - 810**

Artificial intelligence in social media: a systematic review of trends

I Kadek Aryasatya Danendra*),
Jonathan Jacob Paul Latupeirissa
Faculty of Social Sciences and Humanities, Universitas Pendidikan Nasional
39 Bedugul Street, Denpasar, Bali, Indonesia
Email: aryasatyadanendra120503@gmail.com, Phone: +62 361 723868

How to Cite This Article: Danendra, I.K.A. & Latupeirissa, J.J.P. (2025). Artificial intelligence in social media: a systematic review of trends. *Jurnal Studi Komunikasi*, 9(3). doi: 10.25139/jsk.v9i3.10287
Received: 19-05-2025, Revision: 05-10-2025, Acceptance: 16-10-2025, Published: 28-11-2025

Abstract This systematic literature review addresses critical gaps in understanding AI's role in social media communication strategies, specifically examining its impact on influencer marketing, crisis communication, and human-centred trust-building methodologies. Informed by conceptual frameworks of information ethics, user trust, and algorithmic influence, the study distinguishes itself by focusing on recent advancements in fairness-aware machine learning, AI-driven fact-checking, and evolving regulatory landscapes. Analysing 30 peer-reviewed Scopus articles from 2020 to 2025, this study employs a systematic literature review methodology guided by the PRISMA guidelines to ensure rigorous selection and analysis of relevant studies. Findings reveal AI's transformative potential in personalisation and efficiency, yet expose critical ethical challenges like AI-driven deepfakes, algorithmic polarisation, and privacy concerns. It is recommended to develop AI transparency mechanisms, accountability frameworks for public service algorithms, and targeted training for content detection by citizens, as well as provide practical guidance for responsible AI integration in digital communications.

Keywords: Artificial Intelligence; communication strategies; social media

INTRODUCTION

The emergence of social media has profoundly transformed the ways in which organisations engage and communicate with their audiences, allowing for more direct, immediate, and interactive exchanges between brands, institutions, or companies and individual users or communities they seek to reach. Social media platforms function as essential instruments for communication, facilitating real-time contact and feedback that improve public participation and institutional communication strategies (Albarrak & Sorour, 2024; Bokolo & Liu, 2024). Furthermore, the proliferation of misinformation on these platforms poses challenges, as it complicates the public's ability to discern credible information from false narratives, thereby impacting societal discourse (Aïmeur et al., 2023; Lampou & Antonopoulos, 2023; Momeni, 2024). The incorporation of artificial intelligence (AI) into social media has revolutionised content generation and distribution, facilitating individualised communication that can improve user engagement (Arkhipova & Janssen, 2024; Binlibdah, 2024). However, this also raises ethical concerns regarding privacy and the potential for manipulative practices in digital marketing (Pecheranskyi et al., 2024; Shoukat et al., 2024). As firms confront these difficulties, comprehending the ramifications of social media on communication tactics is crucial for cultivating trust and sustaining effective audience interactions (Boatman et al., 2024; Zuboff, 2022).

*) Corresponding Author

ISSN: 2549-7294 (Print), 2549-7626 (Online)

In recent years, the utilisation of artificial intelligence has significantly enhanced the efficacy of social media communication techniques. The incorporation of artificial intelligence (AI) into social media communication methods has markedly improved their efficacy and scope. AI-driven algorithms examine user behaviour and interests, allowing organisations to develop tailored marketing tactics that appeal to certain populations (Shoukat et al., 2024). Furthermore, AI techniques are progressively utilised in content generation and distribution, optimising procedures that were historically labour-intensive (Apablaza-Campos et al., 2024; Matsiola et al., 2024). This transition not only boosts productivity but also improves the quality of communication by guaranteeing that content is pertinent and timely (Kuzyk et al., 2023). Nonetheless, dependence on AI engenders apprehensions about privacy and data security, as users frequently need to disclose personal information for these technologies to operate efficiently (Lim & Shim, 2022). Thus, while AI expands the capabilities of social media communication through technological innovation, it necessitates a careful balance between advancing new solutions and ethical considerations.

AI-driven technologies and methodologies are progressively utilised by organisations ranging from brands, institutions, and companies to individual users or communities to enhance their social media communication strategies, significantly transforming how they engage with and reach their target audiences. AI can examine extensive datasets to discern trends and feelings, allowing firms to develop messages that align with their target populations (Binlibdah, 2024). Moreover, AI-driven chatbots and conversational agents facilitate real-time interactions, providing users with immediate responses and support (Singh et al., 2024). This not only enhances user experience but also allows organisations to maintain continuous engagement without the need for constant human oversight (Shoukat et al., 2024). Additionally, AI tools are employed for content creation, such as generating posts or analysing the effectiveness of different types of content, which streamlines the communication process (Sánchez-García et al., 2023). However, the integration of AI in social media also raises privacy concerns, as users often need to share personal data for these technologies to function optimally (Lim & Shim, 2022). Thus, while AI enhances communication strategies, it necessitates careful consideration of ethical implications.

Highlight the capacity of AI to transform social media communication through personalisation, content optimisation, and predictive analytics. Artificial intelligence (AI) has significant potential to revolutionise social media communication, particularly in areas like personalisation, content optimisation, and predictive analytics. In terms of content optimisation, AI tools assist in generating and curating content that aligns with audience preferences and trending topics. This capability not only improves the relevance of posts but also increases the likelihood of user interaction (Sánchez-García et al., 2023). Additionally, AI-driven analytics provide insights into content performance, enabling organisations to refine their strategies based on real-time feedback (Hua et al., 2024). Predictive analytics further enhance social media strategies by forecasting user behaviour and trends. By leveraging historical data, AI can predict future engagement patterns, allowing organisations to proactively adjust their content and outreach efforts (Mahr & Huh, 2022). This combination of personalisation, optimisation, and predictive capabilities positions AI as a transformative force in social media communication, fostering deeper connections between organisations and their target audiences.

As enterprises progressively integrate artificial intelligence (AI) into their Customer Relationship Management (CRM) strategies, they encounter several challenges alongside the benefits of enhanced efficiency and engagement. A significant issue is the inconsistency in consumer relations and perception that can arise from AI's application in personalised interactions. While AI can streamline communication and foster tailored experiences, it may also lead to misunderstandings or mistrust among consumers if not implemented transparently (Hua et al., 2024). Privacy concerns are a major barrier, as AI-driven CRM often involves collecting and analysing large volumes of personal data, raising ethical questions regarding data security and user consent, which may alienate consumers wary of their information being exploited (Lim & Shim, 2022). Additionally, reliance on AI can inadvertently perpetuate biases present in training data, resulting in distorted messaging that may not resonate equally across diverse audience segments (Shoukat et al., 2024). The rapid pace of AI technological development frequently outpaces existing legal and regulatory frameworks, creating a complex environment for compliance and ethical considerations (Kuzyk et al., 2023). To effectively leverage AI within their

CRM systems, organisations must address these challenges to build trust, ensure ethical practices, and sustain meaningful relationships with their audiences.

The role of AI in enhancing CRM strategies demonstrates its growing importance in generating and managing interactions with customers through social media channels. AI is utilised to automate the creation of relevant content, such as texts, images, and videos, designed to increase customer engagement and satisfaction. For example, research by Lim and Schmälzle (2023) shows that large language models (LLMs) can effectively generate health awareness messages with impactful content, allowing for tailored communication that meets specific audience needs and characteristics (Lim & Schmälzle, 2023). This not only accelerates content production but also improves the relevance and accuracy of the messages conveyed. Additionally, AI technologies facilitate visual content creation, such as images and videos, which can significantly enhance visual appeal and attract customer attention in a crowded digital landscape. Matsiola et al. (2024) highlight that AI-powered image processing techniques can improve the quality of visual materials used in social media, making them more engaging for viewers (Matsiola et al., 2024). Furthermore, AI-driven algorithms enable automated content curation by analysing user preferences and engagement metrics to recommend personalised content, thereby boosting interaction and customer satisfaction (Binlibdah, 2024). However, integrating AI into CRM content management also raises ethical concerns, particularly regarding the authenticity and transparency of AI-generated materials. Therefore, researchers and practitioners must address these challenges to ensure that AI's benefits in CRM strategy development are realised without compromising trust and openness (Aïmeur et al., 2023; Lampou & Antonopoulos, 2023).

Examine research on AI algorithms for audience segmentation, sentiment analysis, and behaviour prediction. Research on AI algorithms for audience segmentation, sentiment analysis, and behaviour prediction reveals their transformative potential in enhancing social media strategies. Audience segmentation is significantly improved through AI, which utilises machine learning algorithms to analyse user data and categorise audiences based on demographics, interests, and behaviours. This allows organisations to tailor their messaging effectively, ensuring that content resonates with specific segments (Kuzyk et al., 2023). For instance, AI can identify niche markets that traditional methods might overlook, thereby optimising marketing efforts (Lim & Shim, 2022). Sentiment analysis, powered by natural language processing (NLP), enables organisations to gauge public opinion and emotional responses to their content. Studies have shown that AI-driven sentiment analysis can accurately classify sentiments in social media posts, providing insights into audience reactions and preferences (Xie et al., 2023). This capability is crucial for brands to adjust their strategies in real-time, responding to negative sentiments or amplifying positive feedback (Afzal et al., 2023). Furthermore, AI algorithms excel in behaviour prediction by analysing historical data to forecast future user actions. This predictive capability allows businesses to anticipate trends and user needs, facilitating proactive engagement strategies (Bokolo & Liu, 2024). For example, AI can predict when users are most likely to engage with content, enabling brands to optimise posting times for maximum visibility (Binlibdah, 2024). However, the reliance on AI for these tasks raises ethical concerns regarding privacy and data security, necessitating a balanced approach to leveraging these technologies (Hua et al., 2024; Shoukat et al., 2024).

Review studies on AI-powered tools for social listening, brand reputation management, and performance measurement. Research on AI-powered tools for social listening, brand reputation management, and performance measurement highlights their significant role in modern marketing and communication strategies. Social listening technologies employ AI algorithms to scrutinise extensive data from social media platforms, allowing corporations to track public mood and discern emerging patterns pertinent to their brands or industries. For instance, a study by Boatman et al. (2024) emphasises the effectiveness of AI-driven social listening platforms in tracking misinformation, showcasing their potential in public health contexts. These tools can provide real-time insights into consumer opinions, allowing brands to respond proactively to public sentiment and manage their reputation effectively. In terms of brand reputation management, AI tools can automate the analysis of online mentions and sentiment, helping organisations to gauge their public image. By employing natural language processing (NLP), these tools can categorise sentiments expressed in user-generated content, enabling brands to identify both positive and negative feedback swiftly.

While the reference by Hua et al. (2024) discusses the interplay between AI and corporate social responsibility communication, it does not specifically address sentiment analysis in brand reputation management. The reference by Binlibdah (2024) highlights AI's role in strategic communication and personalised content, which can indirectly relate to brand reputation management but does not focus on sentiment analysis. AI enhances performance measurement by tracking key performance indicators (KPIs) across diverse social media platforms. Organisations can evaluate the efficacy of their social media initiatives and modify strategy by monitoring engagement metrics, reach, and conversion rates. The dependence on AI for these functions prompts issues over data privacy and the ethical ramifications of automated surveillance, requiring a meticulous approach to deployment. The references by Lim & Shim, 2022; Matsiola et al., (2024) discuss privacy concerns and ethical implications related to AI use, which are relevant to this discussion. However, the reference by Mahr & Huh (2022) does not directly address performance measurement in the context of AI. The reference by (Shoukat et al., 2024) also touches on privacy and intrusiveness concerns in AI-powered platforms, contributing to the ethical discussion.

Investigate the ethical dilemmas presented by AI in social media, encompassing algorithmic bias, misinformation, privacy issues, and the effects on human connection and social discourse. The ethical dilemmas presented by artificial intelligence (AI) in social media are complex, involving algorithmic bias, misinformation, privacy issues, and effects on human connection and social discourse. Algorithmic bias is a critical concern, since AI systems may unintentionally reinforce prevailing cultural biases embedded in training data. This can lead to skewed content recommendations and reinforce stereotypes, ultimately affecting how different demographic groups are represented and engaged on social media platforms (Kuzyk et al., 2023). Misinformation is another critical challenge, particularly with the rise of AI-generated content such as deepfakes. Research indicates that individuals often struggle to discern manipulated media from authentic content, which can distort public perception and influence political discourse (Aïmeur et al., 2023; Momeni, 2024).

The ethical implications of disseminating misinformation through AI tools necessitate a robust framework for accountability and transparency in content creation and sharing. Privacy concerns are paramount as well, given that AI systems often require extensive data collection to function effectively. The trade-off between personalised experiences and user privacy raises ethical questions about consent and data security (Lim & Shim, 2022; Shoukat et al., 2024). Users may be unaware of the extent to which their data is being utilised, leading to a potential erosion of trust in digital platforms. Furthermore, the impact of AI on human interaction and social discourse cannot be overlooked. The increasing reliance on AI-driven communication tools may diminish genuine human connections, altering the nature of social interactions (Swiatek et al., 2024). As AI systems mediate communication, there is a risk of reducing the richness of human discourse, leading to a more transactional and less empathetic form of interaction. In summary, whereas AI provides considerable benefits for social media interaction, it also poses ethical dilemmas that necessitate thorough examination and proactive oversight to guarantee that technology enhances rather than detracts from society ideals.

Identify nascent trends and prospective research avenues in AI and social media communication, including the application of AI in influencer marketing, crisis communication, and the advancement of more human-centric AI methodologies. Emerging developments in the utilisation of artificial intelligence (AI) in social media communication reveal substantial prospects and problems, especially in influencer marketing, crisis communication, and the advancement of more human-centric AI methodologies. In influencer marketing, AI tools are increasingly utilised to analyse audience engagement and optimise influencer partnerships. Research indicates that AI can identify the most effective influencers based on audience demographics and engagement metrics, enhancing the targeting of marketing campaigns (Kuzyk et al., 2023). Future studies could explore the ethical implications of AI-driven influencer selection processes and the potential for algorithmic bias in determining influencer effectiveness. Crisis communication is another area where AI is making strides.

AI-powered social listening tools can monitor social media platforms for emerging crises, allowing organisations to respond swiftly and effectively (Hua et al., 2024). The ability to analyse sentiment in real-time can help brands manage their reputation during crises. Future research should focus on refining these tools to improve accuracy and reduce false positives,

ensuring that organisations can respond appropriately to genuine threats. Moreover, the development of more human-centred AI approaches is crucial for fostering trust and enhancing user experience. As AI systems become more integrated into social media interactions, there is a growing need to ensure that these systems prioritise user privacy and ethical considerations (Shoukat et al., 2024). Research should explore the design of AI systems that are transparent and accountable, enabling consumers to comprehend the use of their data and the decision-making processes involved. In conclusion, although AI offers promising prospects for improving social media communication, it is crucial to confront the ethical dilemmas and guarantee that future advancements emphasise human values and society welfare.

While numerous studies detail AI's technical capabilities and general ethics, there is a critical void in synthesising this knowledge into an actionable, human-centred governance framework for high-stakes public communication, particularly in domains like crisis response and institutional reputation management. The opposing perspective suggests that while the integration of AI in social media presents certain challenges, it primarily offers significant benefits that can enhance digital communication. This study distinguishes itself by focusing on the latest advancements in AI technology and their specific applications within social media, providing a current and comprehensive understanding that differs from previous systematic literature reviews (SLRs). Specifically, recent developments in fairness-aware machine learning algorithms are highlighted, which aim to mitigate algorithmic bias and ensure that AI outputs are equitable across diverse user demographics (de-Lima-Santos & Jamil, 2024). Additionally, the research addresses AI-driven tools for combating misinformation, emphasising innovative fact-checking platforms that enhance the reliability of information shared on social media (Baig et al., 2024; Parra & Chatterjee, 2024).Moreover, the evolving regulatory landscape, including data protection regulations such as the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA), reflects an increasing emphasis on user privacy and transparency in AI applications (Sadiku et al., 2021). The unique contribution of this work lies in analysing how these technological advancements and regulatory frameworks collectively influence the future trajectory of responsible AI deployment in social media, advocating for a balanced perspective that recognises both the technological benefits and ethical safeguards necessary in this domain (Gorwa et al., 2020; Hakimi et al., 2024).

This study is essential as it aims to provide an exhaustive examination of the current state of AI-driven communication techniques on social media platforms. Specifically, the research seeks to address three main questions: (1) How can AI systems enhance influencer marketing by discovering and evaluating influencers, and what are the risks of algorithmic bias in these processes? (2) How can AI be applied effectively in crisis communication to monitor social media for potential issues and facilitate rapid organisational responses? (3) How can human-centred AI methodologies be developed to foster trust and improve user experience? While all three questions are interconnected, the primary research focus centres on understanding how AI can be leveraged responsibly across these domains to maximise benefits while mitigating ethical concerns. The findings will be beneficial to practitioners, policymakers, and researchers, guiding decision-making and aiding in the formulation of responsible AI implementation recommendations in social media. Overall, this research aims to elucidate AI's revolutionary role in shaping online communication habits, emphasising the importance of addressing ethical considerations for responsible and advantageous AI integration.

METHODOLOGY

This systematic literature review employed a rigorous multi-stage selection process to identify relevant research on the integration of AI in social media communication strategies. A comprehensive search was initially conducted on the Scopus database using the keywords 'AI' Or 'artificial intelligence' and 'social media' and 'communication' and 'strategies', with these terms targeted within the article title, abstract, and keywords to ensure direct relevance. The initial search yielded 649 documents, which were then refined by applying a series of filters: publication years limited to 2020-2025 (reducing to 591 documents), subject area restricted to 'social Sciences' (490 documents), and article types narrowed to 'article' and 'review' (216 documents). From these, abstracts were reviewed, and papers were included if they addressed AI's application or implications in social media communication strategies, featured empirical

data or robust theoretical discussions on AI's impact, explicitly discussed ethical dilemmas, benefits, or challenges, and were available in full-text English.

Papers were excluded if they lacked focus on social media or communication, omitted discussion of ethical/ societal/ practical implications, primarily focused on technical AI development without direct relevance to social media communication, or were preliminary studies/ opinion pieces. This screening resulted in 53 documents for full-text analysis, with 30 ultimately meeting all inclusion criteria and forming the basis of this review. Data from these 30 documents were then extracted, including research objectives, methodologies, main findings, identified benefits, challenges, ethical considerations, and theoretical implications, and synthesised using thematic synthesis to identify recurring themes and patterns. The studies referenced in this review represent diverse geographical origins, including Australia, Canada, Chile, China, France, Hungary, Indonesia, Iraq, Italy, Korea, Malaysia, Nepal, Netherlands, Pakistan, Saudi Arabia, South Africa, Spain, Ukraine, and United States, which, while not directly analysed for cultural differences in this specific review, strengthens the analysis by providing a broad perspective on global AI integration in social media.

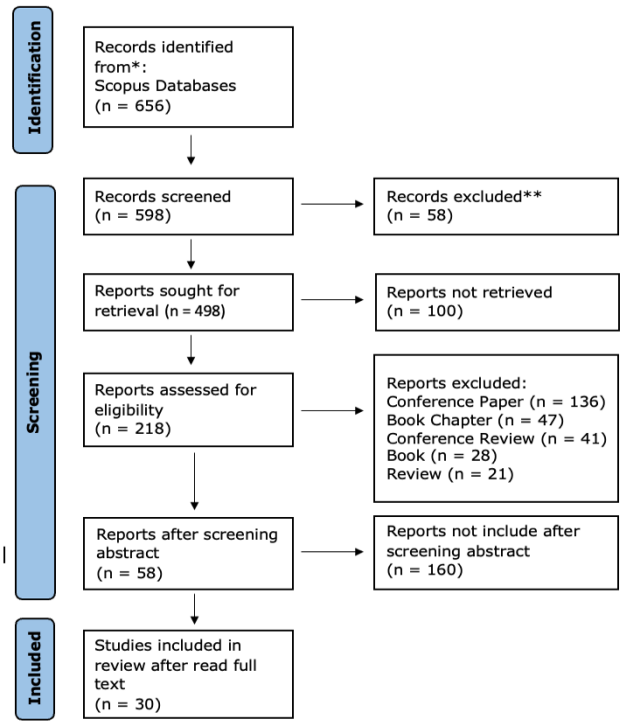

Figure 1. PRISMA flowchart of identification and selected studies.
Source: Processed by author (2025)

RESULTS AND DISCUSSION

The study emphasises the multifaceted implications of integrating artificial intelligence (AI) into social media communication strategies, highlighting both the advantages and the ethical challenges that arise from this integration. Key findings indicate that AI can significantly enhance operational efficiency, personalise user experiences, and provide data-driven insights that empower organisations to tailor their communication strategies more effectively. However, these benefits come with substantial ethical concerns, particularly regarding issues of privacy, bias, and manipulation of user data. The reviewed studies demonstrate that while AI-driven tools can facilitate improved engagement and interaction across diverse user demographics, they also risk reinforcing existing biases and compromising user privacy if not deployed responsibly. Moreover, the analysis underscores the importance of transparency and accountability in AI development, suggesting that organisations maintain a user-centric approach that prioritises ethical considerations in their deployment strategies. The diverse perspectives presented by the various studies, originating from countries like the United States, the United Kingdom, and India, further illustrate the global nature and the varying challenges associated with AI in social media contexts. These insights collectively argue for the necessity of developing robust ethical

guidelines and fostering ongoing dialogue among stakeholders, which is critical for harnessing AI's potential while mitigating its inherent risks, thus ensuring a more inclusive and equitable online environment for all users, as shown in the appendix.

Benefits of AI in social media communication strategies

The integration of Artificial Intelligence (AI) into social media communication tactics has produced substantial advantages, altering the manner in which corporations interact with consumers and oversee their brand narratives. AI technologies augment personalisation, enhance efficiency, and cultivate deeper customer interactions, ultimately resulting in more successful communication tactics. This synthesis will examine the diverse advantages of AI in social media communication, substantiated by several academic sources. A primary advantage of AI in social media communication is its capacity to provide individualised information dissemination. Binlibdah highlights that AI supports the creation of personalised media content, which significantly enhances consumer engagement and service efficiency in digital marketing strategies (Binlibdah, 2024). This personalisation is crucial in an era where consumers are inundated with information, as tailored content can capture attention more effectively than generic messaging. Furthermore, Hua et al. emphasise that AI-driven initiatives can enhance corporate social responsibility (CSR) communication by aligning messages with consumer values and preferences, thereby fostering trust and engagement (Hua et al., 2024). The ability to analyse consumer behaviour and preferences through AI allows organisations to craft messages that resonate on a personal level, which is essential for building lasting relationships with audiences.

AI enhances the efficiency of communication strategies by automating various processes. For instance, Hau and Schmälzle demonstrate that AI can generate health messages that are comparable in quality to those created by humans, thereby streamlining the content creation process (Lim & Schmälzle, 2023). This capability not only saves time but also allows organisations to respond to emerging trends and consumer needs more rapidly. Additionally, Mahr and Huh discuss how AI technologies can improve service communication, suggesting that similar methodologies can be applied to enhance communication strategies in social media contexts (Mahr & Huh, 2022). Through the automation of regular processes, firms can optimise resource allocation, concentrating on strategic initiatives that necessitate human creativity and intelligence. Artificial intelligence is crucial in overseeing and evaluating customer engagements on social media sites. Shoukat et al. examine the impact of psychological factors, including worries about privacy and intrusiveness, on user intentions towards AI-enabled social commerce systems (Shoukat et al., 2024). Understanding these factors is vital for organisations aiming to optimise their communication strategies, as it allows them to address consumer concerns proactively. Furthermore, the work of Bokolo and Liu highlights the application of AI in social media forensics, which can help organisations monitor and respond to misinformation, cyberbullying, and other harmful behaviours that can undermine brand reputation (Bokolo & Liu, 2024). By leveraging AI tools for social listening and sentiment analysis, organisations can gain valuable insights into public perception and adjust their communication strategies accordingly.

The ethical implications of AI in social media communication cannot be overlooked. Lim and Shim raise important questions regarding privacy concerns associated with AI agents, emphasising that while AI offers convenience, it also necessitates a careful approach to data handling and user privacy. Organisations must navigate these ethical challenges to maintain consumer trust and ensure that their communication strategies are not only effective but also responsible. The integration of AI in CSR communication, as discussed by Hua et al., underscores the importance of aligning AI-driven initiatives with ethical standards and consumer expectations (Hua et al., 2024). By prioritising transparency and ethical considerations, organisations can enhance their credibility and foster stronger connections with their audiences. In addition to enhancing communication strategies, AI also facilitates the development of new communication practices among users. Arkhipova and Janssen's comparative analysis reveal how AI recommendations influence social media practices among Generation Z users, highlighting the contextual significance of AI in shaping communication behaviours (Arkhipova & Janssen, 2024). This generational shift underscores the need for organisations to adapt their strategies to align with evolving consumer expectations and preferences. By embracing AI technologies,

organisations can not only enhance their communication effectiveness but also remain relevant in a rapidly changing digital landscape.

The potential of AI to generate content that resonates with diverse audiences is significant. Apablaza-Campos et al. discuss the growing interest in AI for journalistic content production, indicating that AI can assist in creating tailored narratives that appeal to specific demographic segments (Apablaza-Campos et al., 2024). This capability is particularly valuable in social media communication, where diverse audiences require nuanced messaging that reflects their unique perspectives and experiences. By harnessing AI's content generation capabilities, organisations can craft compelling narratives that engage and inform their audiences effectively. The role of AI in enhancing user engagement through interactive communication is another critical benefit. Technologies such as chatbots and virtual assistants, powered by AI, enable organisations to provide real-time responses to consumer inquiries, thereby improving customer service and satisfaction (Mahr & Huh, 2022). This immediacy in communication fosters a sense of connection and responsiveness, which is essential for building brand loyalty in the digital age. Additionally, the findings of Hu and Rui suggest that algorithmic beliefs can influence interpersonal communication outcomes on dating apps, indicating that AI's impact extends beyond traditional marketing to shape user experiences across various platforms (Hu & Rui, 2023).

Figure 2. AI's Role in social media Strategies
Source: Processed by author (2025)

As corporations increasingly depend on AI to enhance their social media communication strategy, the necessity for continuous research and adaptation becomes essential. The evolving landscape of social media and consumer behaviour requires ongoing assessment of AI's efficacy in communication settings. Subsequent research needs to concentrate on the enduring impacts of AI-driven communication methods on customer trust and participation, alongside the ethical ramifications of AI's influence on public discourse (Hua et al., 2024). By staying aware of these trends, firms may guarantee that their communication strategies are both effective and connected with ethical norms and consumer expectations. The incorporation of AI into social media communication tactics provides several advantages, such as increased personalisation, heightened efficiency, and the capacity to manage intricate consumer relationships. Organisations must face the ethical ramifications of AI utilisation and remain adaptive to the changing digital environment. By leveraging AI technologies responsibly, organisations can enhance their communication effectiveness and foster deeper connections with their audiences. From personalised content delivery and automated task management to predictive analytics and enhanced safety, AI has become an integral component of effective social media communication, as demonstrated by the following examples in Figure 2.

Enhanced efficiency and automation

Artificial Intelligence streamlines and automates various tasks in social media communication, including content creation, scheduling, and community management. AI-powered tools can generate engaging captions for posts, schedule content for optimal times when the target audience is most active, and even moderate online communities by detecting and addressing inappropriate or harmful content. This automation of routine tasks allows marketing and communication teams to focus their efforts on more strategic initiatives, such as developing creative campaigns, analysing performance data, and making informed decisions to enhance their social media presence. Furthermore, AI can play a significant role in automating social

media interactions, such as responding to customer inquiries, providing personalised recommendations, and maintaining a consistent brand voice across multiple platforms. Organisations can utilise AI-driven chatbots and virtual assistants to provide prompt and effective customer assistance, thereby liberating human resources for more intricate duties and enhancing the overall user experience on social media.

Hyper-personalisation and targeted reach
Artificial Intelligence enables highly personalised content recommendations, precisely targeted advertising, and tailored customer interactions. By analysing substantial user data, AI can produce content that deeply aligns with individual preferences and habits, enhance advertising campaigns for optimal effectiveness, and offer personalised assistance via intelligent chatbots. The enhanced ability to customise communication experiences has considerable ramifications, as AI emerges as a potent intermediary in social media interactions. AI-driven customisation enables firms to establish deeper, more significant connections with their target audience, enhancing brand loyalty and engagement. The growing dependence on AI presents ethical issues related to privacy, transparency, and potential misuse, which must be meticulously addressed to guarantee the appropriate and ethical application of these transformative technologies in social media communication methods.

Improved content quality and performance; Enhanced brand reputation and crisis management
AI optimises content for better engagement, analyses sentiment to refine messaging, and identifies trending topics to ensure relevance. AI-powered tools can suggest improvements to existing content, gauge audience reactions, and uncover emerging trends, enabling organisations to create more effective and timely content. These AI-driven capabilities explore the various aspects of social media, including content creation, content optimisation, audience analysis, and trend identification. By leveraging AI, organisations can enhance their social media communication strategies, creating more engaging and relevant content that resonates with their target audience.

AI facilitates reputation monitoring, automated responses to customer inquiries, and sentiment analysis during crises. By tracking brand mentions and analysing public sentiment, AI can alert organisations to potential reputational issues and enable proactive crisis management. Specifically, AI-powered tools can continuously monitor online conversations and social media platforms for any mentions or discussions related to the brand. This allows organisations to quickly identify emerging concerns or negative perceptions that could impact their reputation. Furthermore, AI can perform real-time sentiment analysis on these brand mentions, providing valuable insights into the emotional tone and sentiment of the public discourse. Armed with these insights, organisations can respond proactively to address customer complaints, clarify misinformation, and take corrective actions to mitigate potential crises. This capability is crucial in an era where news and information spread rapidly online, and a single crisis can quickly escalate and cause significant damage to a brand's reputation if not addressed swiftly and effectively. By leveraging AI for reputation monitoring and crisis management, organisations can stay ahead of potential issues, maintain transparency, and demonstrate their responsiveness to customer concerns, ultimately strengthening their brand image and consumer trust. discusses the impact of AI on communication, including its potential for reputation management.

Data-driven insights and decision-making
Artificial intelligence offers significant data-driven insights regarding audience behaviour, advertising efficacy, and competitor assessment. This data enables firms to make informed decisions regarding their social media strategies, optimise campaigns, and secure a competitive advantage. The study by Pauwels and Joshi examines the application of AI in the analysis of social media data and its ramifications for data-driven decision-making. AI algorithms can analyse extensive volumes of user-generated material, consumer interactions, and platform analytics to reveal patterns, discern new trends, and produce actionable insights. These insights enable firms to enhance their social media marketing strategy, optimise resource allocation, and surpass their competition. Utilising AI-driven data analysis, organisations can maintain a competitive edge, predict client preferences, and provide tailored experiences that align with their audience.

Challenges and ethical considerations in AI-driven social media communication

The integration of artificial intelligence (AI) into social media communication has ushered in a myriad of challenges and ethical considerations that warrant thorough examination. As AI technologies evolve, they increasingly shape how information is disseminated and consumed, leading to significant implications for user trust, privacy, and the integrity of communication. This discourse synthesises various scholarly perspectives on the challenges and ethical dilemmas posed by AI-driven social media communication. One of the primary challenges associated with AI in social media is the issue of misinformation and disinformation. The proliferation of AI-generated content, particularly deepfakes and misleading narratives, has raised concerns about the authenticity of information shared online. Research indicates that individuals often struggle to discern deepfake videos from genuine content, which can significantly influence public opinion and perceptions (Momeni, 2024). The ethical implications of this phenomenon are profound, as the manipulation of information can undermine democratic processes and erode trust in media institutions (Aïmeur et al., 2023). Furthermore, AI technologies are often employed to create and disseminate fake news, complicating efforts to combat misinformation and posing a significant challenge to public discourse (Aïmeur et al., 2023).

The rapid spread of false information, particularly during crises such as the COVID-19 pandemic, has underscored the necessity for effective misinformation detection mechanisms. Research indicates that large language models (LLMs) like ChatGPT can be trained to identify misinformation, but their efficacy is contingent upon the quality of the prompts used during interaction (Haupt et al., 2024). This highlights a significant challenge: the need for robust systems that can adapt to the dynamic nature of misinformation while maintaining accuracy and reliability. Furthermore, the study by Boatman et al. emphasises the potential of AI-driven social listening platforms for public health surveillance, yet it also acknowledges the inherent challenges related to data quality and representativeness (Boatman et al., 2024). This duality illustrates the complexity of leveraging AI for social good while navigating the pitfalls of misinformation. In addition to misinformation, privacy concerns are paramount in the context of AI-driven social media communication. Users frequently express apprehension regarding the intrusiveness of AI technologies, particularly in relation to data collection and surveillance practices (Shoukat et al., 2024).

The theory of planned behaviour suggests that psychological factors, such as perceived privacy risks, significantly influence user intentions toward AI-enabled applications (Shoukat et al., 2024). This raises ethical questions about the extent to which companies should leverage user data for targeted marketing and communication, particularly when such practices may infringe upon individual privacy rights (Lim & Shim, 2022). The balance between utilising AI for enhanced user experiences and respecting user privacy remains a contentious issue that requires careful consideration. The ethical implications of AI in social media extend to the potential for job displacement within the communication sector. As AI technologies become more sophisticated, there is a growing concern that human roles in public relations and communication may be rendered obsolete (Swiatek et al., 2024). This shift not only poses economic challenges but also raises ethical questions about the value of human interaction in communication processes. The potential loss of jobs in favour of AI-driven solutions could exacerbate existing social inequalities, necessitating a re-evaluation of the ethical frameworks guiding the deployment of AI in communication (Swiatek et al., 2024).

The role of AI in shaping user interactions and communication dynamics is another critical area of concern. AI-driven platforms often foster a sense of intimacy and social presence, which can paradoxically lead to increased privacy concerns among users (Lim & Shim, 2022). The interplay between user motivations for engaging with AI agents and their subsequent privacy anxieties highlights the complex relationship between technology and human behaviour. As users navigate these interactions, the ethical implications of AI's influence on social relationships and communication norms must be scrutinised (Lim & Shim, 2022). Furthermore, the ethical considerations surrounding AI in social media are compounded by the rapid pace of technological advancement. The emergence of generative AI has transformed the landscape of digital communication, prompting discussions about the ethical responsibilities of developers and organisations in deploying these technologies (Pecheranskyi et al., 2024). As AI continues to evolve, the potential for misuse or unintended consequences increases, necessitating a proactive approach to ethical governance in AI applications (Pecheranskyi et al.,

2024). This includes establishing frameworks that prioritise transparency, accountability, and user empowerment in AI-driven communication.

The implications of AI-driven communication extend beyond individual interactions to encompass broader societal impacts. The concept of surveillance capitalism, as articulated by Zuboff, underscores the need for critical examination of how AI technologies are utilised to monitor and influence user behaviour (Zuboff, 2022). The commodification of personal data raises ethical questions about consent, autonomy, and the potential for manipulation in digital spaces. As AI becomes increasingly integrated into social media platforms, the ethical ramifications of these practices must be addressed to safeguard user rights and promote equitable communication practices (Zuboff, 2022). The challenge of maintaining user trust in AI systems is another critical aspect of this discourse. The study by Кузик et al. discusses how discrepancies in AI performance can undermine user trust, particularly in contexts such as voice marketing (Kuzyk et al., 2023). Trust is a fundamental component of effective communication, and any erosion of this trust can have far-reaching implications for user engagement and brand loyalty. As AI technologies become more integrated into social media platforms, ensuring transparency and accountability in their operations will be essential to maintaining user confidence.

Furthermore, the role of AI in shaping public discourse cannot be overlooked. The research by Carral and Elías highlights how AI tools can be utilised to analyse toxicity in online conversations, revealing that real users often generate more toxicity (Carral & Elías, 2024). This insight underscores the need for AI systems that can not only detect harmful content but also promote healthier discourse. The ethical challenge here involves developing AI technologies that can effectively moderate conversations without infringing on free speech or stifling legitimate expression. The ethical landscape of AI-driven social media communication is further complicated by the need for interdisciplinary collaboration. As AI technologies intersect with various fields, including communication, ethics, and public policy, a holistic approach is essential to address the multifaceted challenges posed by these advancements (Kim, 2024). The implications of AI on social media communication extend to issues of accessibility and inclusivity. The integration of AI technologies must consider the diverse needs of users, particularly those from marginalised communities. The findings of Baskota and Poudel suggest that AI can enhance digital literacy among students, enabling them to navigate online spaces more effectively (Baskota & Poudel, 2024). However, this potential benefit must be weighed against the risk of exacerbating existing inequalities in access to technology and information. Ensuring that AI-driven social media tools are accessible to all users is a critical ethical consideration that must be addressed.

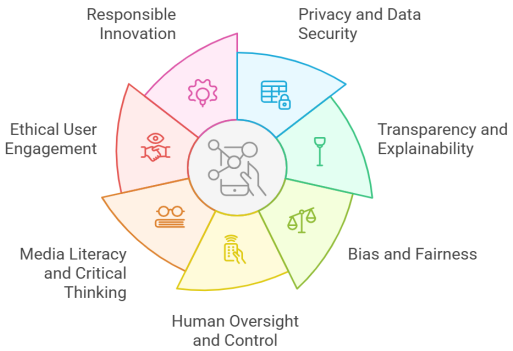

Figure 3. Navigating AI in social media
Source: Processed by author (2025)

Engaging diverse stakeholders in discussions about the ethical implications of AI can facilitate the development of comprehensive frameworks that prioritise user welfare and societal benefit. In summary, the challenges and ethical considerations surrounding AI-driven social media communication are complex and multifaceted. From the proliferation of misinformation to privacy concerns and the potential for job displacement, the implications of AI technologies necessitate rigorous examination and proactive governance. As the landscape of digital communication continues to evolve, it is imperative that stakeholders prioritise ethical

considerations to ensure that AI serves as a tool for enhancing communication rather than undermining it. These concerns span various aspects, ranging from data privacy and algorithmic bias to the potential for manipulation and the erosion of human connection, as outlined in Figure 3 and Table 1.

Table 1. Literature of AI integration dimensions.

No.	Dimension	Article
1	Privacy and data security	(Lim & Shim, 2022; Pecheranskyi et al., 2024; Shoukat et al., 2024; Zuboff, 2022)
2	Transparency and explainability	(Boatman et al., 2024; Kim, 2024; Swiatek et al., 2024)
3	Bias and fairness	(Aïmeur et al., 2023; Haupt et al., 2024; Momeni, 2024)
4	Human oversight and control	(Carral & Elías, 2024; Kim, 2024; Lim & Shim, 2022; Shoukat et al., 2024)
5	Media literacy and critical thinking	(Momeni, 2024; Pecheranskyi et al., 2024; Swiatek et al., 2024; Zuboff, 2022)
6	Ethical user engagement	(Kuzyk et al., 2023; Lim & Shim, 2022; Pecheranskyi et al., 2024)
7	Responsible innovation	(Baskota & Poudel, 2024; Haupt et al., 2024; Pecheranskyi et al., 2024)

Source: Processed by author (2025)

Privacy and data security; Transparency and explainability

The deployment of artificial intelligence (AI) fundamentally requires comprehensive data governance frameworks to address emerging privacy concerns and ensure ethical data handling practices. The necessity for organisations to not only comply with existing regulations but also anticipate future challenges is underscored by the implications of AI on individual privacy rights and societal integrity. These frameworks ought to navigate the inherent tension between innovation and prudent data stewardship, striking a balance that fosters responsible AI usage while maximising its potential benefits for society. A proactive approach, establishing policies and practices that prioritise data protection, is essential in mitigating risks associated with AI applications, particularly as technologies evolve rapidly. By implementing robust governance mechanisms, organisations can ensure a sustainable path forward in AI development, aligning technological advancements with ethical considerations that uphold the principles of privacy and social responsibility.

A deep understanding of AI's decision-making mechanisms is pivotal for fostering user trust and ensuring their effective engagement with AI technologies. The advancement of explainable AI (XAI) methodologies plays a crucial role in enhancing transparency and allowing users to grasp the reasoning behind algorithmic outputs. By providing clear insights into how decisions are made, organisations can improve accountability, giving users more confidence in the systems they interact with. Such understanding not only empowers users but also enables them to evaluate the fairness and reliability of the AI tools they use. Emphasising transparency and interpretability in AI systems can lead to more substantiated user experiences, consequently optimising the decision-making processes inherent to these technologies. Therefore, the development and implementation of XAI methodologies are integral to promoting informed user interactions and fostering a more trusting relationship between humans and AI systems.

Bias and fairness

Mitigating algorithmic bias is a critical imperative in the development and deployment of AI systems. It necessitates a multifaceted approach, emphasising the creation of effective tools for bias detection to identify and rectify prejudices inherent in algorithms. Employing fairness-oriented algorithms ensures equitable treatment across diverse demographic groups, fostering inclusivity and representation in AI outputs. Moreover, utilising diverse datasets that encompass various human experiences is crucial to prevent the perpetuation of harmful stereotypes and the marginalisation of underrepresented communities. Establishing explicit ethical guidelines and frameworks is essential for steering the responsible advancement of AI technologies, ensuring compliance with principles of equity and non-discrimination. By adopting these proactive strategies, organisations can work towards developing AI systems that prioritise fairness and reflect the diverse needs of their stakeholders, ultimately contributing to a more equitable society.

Human oversight and control
Preserving human oversight in AI systems is imperative to ensure that technology enhances rather than replaces human judgment. This requires clearly defined responsibilities and accountability structures, alongside robust human monitoring and intervention systems. A critical aspect of this oversight involves maintaining human agency, particularly in domains such as social media, where ethical considerations and moral reasoning are essential in guiding AI applications. The integration of AI must prioritise the values and ethics that underpin human interactions, safeguarding autonomy while harnessing technology's potential. Moreover, explicit ethical frameworks will offer directives for the responsible use of AI, ensuring alignment with principles of equity and inclusiveness. By emphasising significant human engagement and supervision, organisations can leverage AI capabilities while preserving the integrity of human decision-making and fostering a collaborative future between humans and technology.

Media literacy and critical thinking
Empowering users to critically evaluate information is essential in today's digital landscape, where misinformation is prevalent. This can be achieved through comprehensive media literacy education, equipping individuals with the skills necessary to discern credible information sources. Additionally, the development of advanced technological tools for identifying and flagging misinformation is crucial; these AI-driven solutions can aid users in making informed choices and navigating the vast amount of content available. By fostering a culture of critical thinking and intellectual curiosity, organisations can encourage users to approach social media content thoughtfully, recognising both the benefits and drawbacks of these platforms. It is important to engage users in this way to highlight how media literacy can improve social dynamics and individual well-being. Through these initiatives, organisations can cultivate an informed user base that actively contributes to a healthier online environment, promoting constructive discourse and improving societal outcomes.

Ethical user engagement
Empowering users to critically evaluate information is essential in today's digital landscape, where misinformation is prevalent. This can be achieved through comprehensive media literacy education, equipping individuals with the skills necessary to discern credible information sources. Additionally, the development of advanced technological tools for identifying and flagging misinformation is crucial; these AI-driven solutions can aid users in making informed choices and navigating the vast amount of content available. By fostering a culture of critical thinking and intellectual curiosity, organisations can encourage users to approach social media content thoughtfully, recognising both the benefits and drawbacks of these platforms. Notably, the work of Momeni emphasises the importance of engaging users in this manner, highlighting how media literacy can improve social dynamics and individual well-being. Through these initiatives, organisations can cultivate an informed user base that actively contributes to a healthier online environment, promoting constructive discourse and improving societal outcomes.

Responsible innovation
Ensuring ethical considerations are central to the development and deployment of AI technologies is vital for safeguarding user well-being and societal benefit. This requires a commitment to responsible data practices, where user privacy and autonomy are upheld, and AI innovations align with fundamental human values. The immense potential of AI to drive positive societal changes can only be harnessed through a proactive approach that thoughtfully addresses ethical challenges such as privacy, bias, transparency, and the preservation of meaningful human interactions. For example, incorporating frameworks that emphasise ethical practices into AI development can help mitigate risks associated with misinformation and manipulation, ensuring these technologies serve the public good. Furthermore, the dialogue surrounding AI's ethical dimensions must be inclusive, engaging diverse perspectives to cultivate solutions that promote human flourishing. Ultimately, embedding ethical frameworks in AI development can unlock opportunities for technology to enhance lives while minimising unintended consequences.

Toward responsible and beneficial integration of AI in social media

The integration of artificial intelligence (AI) into social media platforms has emerged as a critical area of research, particularly as it pertains to enhancing user engagement, improving communication strategies, and addressing social issues. This integration can lead to both beneficial outcomes and potential risks, necessitating a responsible approach to its implementation. The following synthesis explores the various dimensions of AI integration in social media, drawing on recent studies that highlight its applications, challenges, and implications for society. One of the primary benefits of AI in social media is its ability to enhance communication mechanisms. For instance, Luo and Yang propose a social media communication framework that leverages deep learning and AI-generated content (AIGC) to facilitate psychological crisis interventions during disasters (Luo & Yang, 2024). This framework not only aids in real-time communication but also emphasises the importance of timely and accurate information dissemination, which is crucial during emergencies. Similarly, Singh et al. highlight the use of conversational agents (CAs) in social media to manage dialogues effectively, particularly in interventions for nicotine cravings (Singh et al., 2024). This demonstrates how AI can be harnessed to provide personalised support to users, thereby improving their overall experience on social media platforms.

The application of AI in social media extends to the analysis of user interactions and sentiments. Khan et al. discuss the significance of politeness in online conversations, especially given the anonymity that social media provides (Khan et al., 2023). The use of advanced natural language processing (NLP) techniques allows for the detection of politeness, which can enhance user interactions and reduce conflicts. Furthermore, Xie et al. explore sentiment analysis during the COVID-19 pandemic, emphasising the need for public health agencies to understand public sentiments and concerns through social media discussions (Xie et al., 2023). Their research underscores the potential of AI-driven 'infoveillance' systems to provide granular insights into public discourse, which can inform more effective communication strategies. However, the integration of AI in social media is not without its challenges. The rapid advancements in AI technologies, including the Internet of Things (IoT) and AIoT (Artificial Intelligence of Things), present both opportunities and hurdles for social media applications (Hou et al., 2023). As these technologies evolve, they necessitate robust frameworks to ensure interoperability and data security. Additionally, the ethical implications of AI use in social media must be carefully considered. The potential for misinformation and manipulation of public opinion, as seen in the context of populist political narratives during the COVID-19 pandemic, raises concerns about the responsible use of AI (Souza, 2024). This highlights the need for regulatory frameworks that govern AI applications in social media to mitigate risks associated with misinformation and user exploitation.

The educational sector also benefits from AI integration in social media, as evidenced by the development of AI-powered online learning platforms. Wahjusaputri et al. discuss how AI can enhance vocational education through organised online learning management systems (Wahjusaputri et al., 2024). This integration not only improves educational outcomes but also fosters a collaborative learning environment that can extend to social media platforms. Similarly, Farhan et al. demonstrate that AI techniques significantly improve the quality of educational services, suggesting that social media can serve as an effective tool for educational engagement and knowledge sharing (Farhan et al., 2024). Furthermore, the role of AI in supporting emergent farmers through knowledge management frameworks illustrates the versatility of AI applications (Buitendag & Hattingh, 2024). The integration of AI in agricultural practices can be mirrored in social media strategies that aim to empower communities and facilitate knowledge exchange. This cross-domain applicability of AI emphasises the need for interdisciplinary approaches to harness its full potential in social media contexts.

The responsible and beneficial integration of AI in social media is a multifaceted endeavour that requires careful consideration of its applications, challenges, and societal implications. While AI offers significant opportunities for enhancing communication, understanding user sentiments, and improving educational outcomes, it also poses ethical risks that must be addressed through appropriate regulatory measures. As the landscape of social media continues to evolve, ongoing research and dialogue will be essential to ensure that AI is utilised in ways that promote positive social outcomes while safeguarding user rights and privacy. Organisations must prioritise user privacy, combat misinformation, ensure transparency,

maintain human oversight, address algorithmic bias, and balance personalisation with user autonomy. These are several key considerations:

Prioritising user privacy and data security
Prioritising user privacy and data security beyond just compliance with rules. A proactive approach to data governance is essential, wherein firms foresee potential privacy problems and establish ethical data management procedures that emphasise user control and openness. This underscores the intrinsic societal challenge in AI development, accentuating the conflict between innovation and appropriate data utilisation. Organisations ought to investigate sophisticated methodologies such as differential privacy and federated learning to reduce data exposure while facilitating AI progress. Moreover, granting customers detailed control over their data via user-friendly privacy dashboards and permission protocols is essential. This method cultivates trust and enables consumers to make informed choices regarding their personal information. By emphasising user privacy and data security, enterprises may establish sustainable and reliable AI-driven social media ecosystems that honour individual rights and encourage responsible innovation.

Ensuring transparency and explainability
Transparency encompasses not just the disclosure of algorithms but also the elucidation of the decision-making processes of AI systems for users. This necessitates the development of explainable AI methodologies that offer transparent insights into the reasoning behind algorithmic decisions. Highlighting the significance of transparency and explainability in AI systems is essential, since it enables consumers to understand the rationale behind the content or recommendations they get. Propose the integration of XAI elements within social media platforms to facilitate users' comprehension of the reasoning and logic underlying the content presented to them. This method cultivates trust and enables users to make more informed decisions, since they can more effectively understand the mechanisms underlying content distribution. Enhancing the transparency of AI systems enables users to attain a more profound comprehension of the technology and make more judicious decisions regarding the information they engage with.

Mitigating bias and promoting fairness
Mitigating algorithmic bias is a continuous and complex endeavour. It necessitates not only varied and representative datasets but also the creation of fairness-conscious algorithms that proactively reduce bias throughout both training and deployment phases. It is essential to underscore the substantial issues presented by algorithmic bias in AI systems and stress the pressing necessity for thorough ethical norms to regulate their development and application. Methods such as adversarial debiasing and fairness constraints may be investigated to promote more equitable results. Furthermore, routine audits and effect evaluations are essential to detect and rectify any nascent biases that may develop, even in systems that first seem impartial. This proactive strategy for bias mitigation is crucial to maintain the concepts of justice, inclusivity, and non-discrimination in AI-driven social media platforms. Persistent endeavours to mitigate algorithmic bias through various approaches, including as data curation, model architecture, and ongoing evaluation, are essential for developing reliable and accountable AI systems that cater to the requirements of all users.

Maintaining human oversight and control
Human discernment is crucial, particularly in content filtering and the management of sensitive circumstances. This includes creating clear lines of responsibility between humans and AI, developing methods for human involvement, and ensuring that AI augments, not replaces, human decision-making. Artificial intelligence ought to be developed to augment and enhance human abilities, rather than entirely supplanting them. Conversations regarding the significance of human values, ethics, and social impact must be prioritised in AI research. Consider the implementation of hybrid models in which AI algorithms identify potentially problematic information for human assessment and final decision-making, so ensuring a balance between automation and human oversight. This enables AI to aid with content moderation by highlighting potential flaws, while preserving human oversight and discernment, especially in

complex or sensitive circumstances. A human-centric strategy that emphasises openness, responsibility, and the safeguarding of human agency is essential for the proper incorporation of AI in social media communication.

Fostering media literacy and critical thinking

Empowering users to navigate the complex digital landscape requires a multifaceted approach that goes beyond just media literacy education. It involves developing comprehensive tools and resources to help users identify misinformation, evaluate the credibility of information sources, and understand the potential biases inherent in algorithms. The work by Dwivedi et al.(INSERT YEAR) explores the complex nature of social media, highlighting both the positive and negative aspects, and emphasising the need for critical engagement. To encourage more discerning content consumption, consider integrating targeted critical thinking prompts and fact-checking tools directly within social media platforms. These features can empower users to approach online information with a more analytical mindset, evaluating the accuracy and reliability of the content they encounter. By equipping users with the knowledge and tools to navigate the digital landscape, we can foster a more informed and discerning online community that is better equipped to identify and resist the spread of misinformation.

Enhancing user engagement ethically

Personalisation should empower users, not manipulate them. This means providing individuals with greater control over their online experiences, respecting their autonomy, and avoiding practices that exploit vulnerabilities or promote addictive behaviours. The cited work on AI in social media discusses ethical considerations related to AI's influence on social media interactions. The focus should be on personalisation that enhances user agency, allowing users to customise their feeds, control recommendations, and manage their online identities. This empowerment enables users to tailor their experiences according to their preferences, fostering a sense of autonomy and control over their digital lives. By prioritising user-centric design and ethical principles, the integration of AI can create personalised social media environments that empower and engage users, rather than manipulate or exploit them.

Driving innovation responsibly

Ethical considerations must be paramount in the development and deployment of AI. This necessitates a multi-stakeholder strategy, engaging researchers, developers, politicians, and users in determining the future of AI in social media. This strategy should examine the prospects and obstacles of AI for societal benefit, while also instituting ethical review committees for AI systems to formulate comprehensive industry-wide ethical standards. Moreover, it is essential to promote open discourse regarding the societal ramifications of AI in social media, incorporating varied viewpoints to guarantee the responsible and ethical implementation of AI technology. This collaborative initiative is crucial to tackle the intricate ethical dilemmas associated with AI utilisation in social media and to guarantee that its application advantages all societal members (See Figure 4).

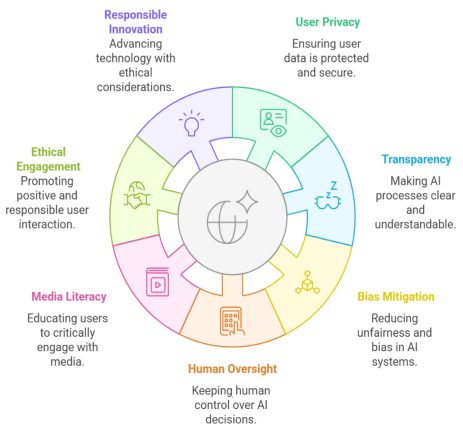

Figure 4. Responsible AI integration in social media
Source: Processed by author (2025)

CONCLUSION

The conclusion of this study shows that the main finding of the Systematic Literature Review (SLR) analysis is that the development and application of AI in social media communication is currently characterised more by challenges related to ethical aspects, transparency, and data management, rather than technical technology constraints. Based on the clustering of scientific journals, it was found that the main influential dimensions are (1) data privacy and security aspects that require a more proactive and ethical approach to data management; (2) transparency and explainability that are important for building user trust through AI that can be explained and understood; and (3) the development of ethical and regulatory frameworks that are able to overcome algorithm bias and manipulation. These results confirm that the main challenge is no longer the innovation of AI technology itself, but rather the existence of governance and oversight deficits in the public sphere.

For future research, it is recommended to explore the development of adaptive ethical standards and regulations that keep pace with technological advancements. Empirical studies should also examine the effectiveness of implementing these frameworks across diverse social and media communication contexts. Additionally, further investigation into the role of media literacy as a strategy to enhance public awareness of AI's risks and benefits is essential. Research into effective multidisciplinary collaboration among researchers, regulators, and industry stakeholders will also be valuable. This future research aims to provide comprehensive insights into building a safe, transparent, and ethical AI ecosystem, addressing the governance challenges that currently hinder its responsible deployment.

REFERENCES

Afzal, S., Ghani, S., Hittawe, M. M., Rashid, S. F., Knio, O. M., Hadwiger, M., & Hoteit, I. (2023). Visualization and Visual Analytics Approaches for Image and Video Datasets: A Survey. ACM Transactions on Interactive Intelligent Systems, 13(1). https://doi.org/10.1145/3576935

Aïmeur, E., Amri, S., & Brassard, G. (2023). Fake news, disinformation and misinformation in social media: a review. Social Network Analysis and Mining, 13(1). https://doi.org/10.1007/s13278-023-01028-5

Albarrak, K. M., & Sorour, S. E. (2024). Web-Enhanced Vision Transformers and Deep Learning for Accurate Event-Centric Management Categorization in Education Institutions. Systems, 12(11). https://doi.org/10.3390/systems12110475

Apablaza-Campos, A., Wilches Tinjacá, J. A., & Salaverría, R. (2024). Generative Artificial Intelligence for Journalistic Content in Ibero-America: Perceptions, Challenges and Regional Projections. BiD, 52. https://doi.org/10.1344/bid2024.52.06

Arkhipova, D., & Janssen, M. (2024). AI recommendations' impact on individual and social practices of Generation Z on social media: a comparative analysis between Estonia, Italy, and the Netherlands. Semiotica. https://doi.org/10.1515/sem-2023-0089

Baig, K., Altaf, A., & Azam, M. (2024). Impact of AI on Communication Relationship and Social Dynamics: A qualitative Approach. Bulletin of Business and Economics (BBE), 13(2), 282–289. https://doi.org/10.61506/01.00283

Baskota, P., & Poudel, T. (2024). Artificial intelligence and computer-mediated communication: the text analysis and undergrad's class observation. Discover Education, 3(1). https://doi.org/10.1007/s44217-024-00187-7

Binlibdah, S. (2024). Investigating the Role of Artificial Intelligence to Measure Consumer Efficiency: The Use of Strategic Communication and Personalized Media Content. Journalism and Media, 5(3), 1142–1161. https://doi.org/10.3390/journalmedia5030073

Boatman, D., Starkey, A., Acciavatti, L., Jarrett, Z., Allen, A., & Kennedy-Rea, S. (2024). Using Social Listening for Digital Public Health Surveillance of Human Papillomavirus Vaccine Misinformation Online: Exploratory Study. JMIR Infodemiology, 4. https://doi.org/10.2196/54000

Bokolo, B. G., & Liu, Q. (2024). Artificial Intelligence in Social Media Forensics: A Comprehensive Survey and Analysis. Electronics (Switzerland), 13(9). https://doi.org/10.3390/electronics13091671

Buitendag, A., & Hattingh, F. (2024). A smart agricultural knowledge management framework to support emergent farmers in developmental settings. Interdisciplinary Journal of Information, Knowledge, and Management, 19. https://doi.org/10.28945/5315

Carral, U., & Elías, C. (2024). Application of AI tools as methodology for the analysis of toxicity in social media: A case study of Spanish politics on Twitter. Revista Latina de Comunicacion Social, 2024(82). https://doi.org/10.4185/rlcs-2024-2205

de-Lima-Santos, M.-F., & Jamil, S. (2024). Bridging the AI Divide: Human and Responsible AI in News and Media Industries. Emerging Media, 2(3), 335–346. https://doi.org/10.1177/27523543241291229

Farhan, N. D., Sadiq, B. H., Zwayyer, M. H., & Arnout, B. A. (2024). The impact of using artificial intelligence techniques in improving the quality of educational services/case study at the University of Baghdad. Frontiers in Education, 9. https://doi.org/10.3389/feduc.2024.1474370

Gorwa, R., Binns, R., & Katzenbach, C. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. Big Data & Society, 7(1), 205395171989794. https://doi.org/10.1177/2053951719897945

Hakimi, M., Sazish, B., Rastagari, M. A., & Shahidzay, K. (2024). Artificial Intelligence for Social Media Safety and Security: A Systematic Literature Review. Studies in Media, Journalism and Communications, 1(1), 10–21. https://doi.org/10.32996/smjc.2023.1.1.2x

Haupt, M. R., Yang, L., Purnat, T., & Mackey, T. (2024). Evaluating the Influence of Role-Playing Prompts on ChatGPT's Misinformation Detection Accuracy: Quantitative Study. JMIR Infodemiology, 4.
Hou, K. M., Diao, X., Shi, H., Ding, H., Zhou, H., & de Vaulx, C. (2023). Trends and Challenges in AIoT/IIoT/IoT Implementation. Sensors, 23(11). https://doi.org/10.3390/s23115074
Hu, J., & Rui, J. R. (2023). The Relationship Between Preference for Online Social Interaction and Affective Well-Being via Compulsive Dating App Use: The Moderating Role of Algorithmic Beliefs. Cyberpsychology, 17(4). https://doi.org/10.5817/CP2023-4-2
Hua, X., Hasan, N. A. M., De Costa, F., & Qiao, W. (2024). Opportunities or Challenges? The Interplay between Artificial Intelligence and Corporate Social Responsibility Communication. Business Systems Research, 15(1), 131–157. https://doi.org/10.2478/bsrj-2024-0007
Khan, S., Fazil, M., Imoize, A. L., Alabduallah, B. I., Albahlal, B. M., Alajlan, S. A., Almjally, A., & Siddiqui, T. (2023). Transformer Architecture-Based Transfer Learning for Politeness Prediction in Conversation. Sustainability (Switzerland), 15(14). https://doi.org/10.3390/su151410828
Kim, D. K. D. (2024). An investigation of public trust in autonomous humanoid AI robot doctors: a preparation for our future healthcare system. Frontiers in Communication, 9. https://doi.org/10.3389/fcomm.2024.1420312
Kuzyk, O., Kabanova, O., Chyrva, H., Vlasenko, D., & Komarnytska, H. (2023). Trends and perspectives on the impact of digital technologies on the efficacy of marketing communication. Financial and Credit Activity: Problems of Theory and Practice, 6(53), 471–486. https://doi.org/10.55643/fcaptp.6.53.2023.4259
Lampou, E., & Antonopoulos, N. (2023). Ranked by Truth Metrics: A New Communication Method Approach, on Crowd-Sourced Fact-Checking Platforms for Journalistic and Social Media Content. Studies in Media and Communication, 11(6), 231–243. https://doi.org/10.11114/smc.v11i6.6166
Lim, S., & Schmälzle, R. (2023). Artificial intelligence for health message generation: an empirical study using a large language model (LLM) and prompt engineering. Frontiers in Communication, 8. https://doi.org/10.3389/fcomm.2023.1129082
Lim, S., & Shim, H. (2022). No Secrets Between the Two of Us: Privacy Concerns Over Using AI Agents. Cyberpsychology, 16(4). https://doi.org/10.5817/CP2022-4-3
Luo, J., & Yang, A. (2024). Human Adaption to Climate Change: Marine Disaster Risk Reduction in the Era of Intelligence. Sustainability (Switzerland), 16(22). https://doi.org/10.3390/su16229647
Mahr, D., & Huh, J. (2022). Technologies in service communication: looking forward. Journal of Service Management, 33(4–5), 648–656. https://doi.org/10.1108/JOSM-03-2022-0075
Matsiola, M., Lappas, G., & Yannacopoulou, A. (2024). Generative AI in Education: Assessing Usability, Ethical Implications, and Communication Effectiveness. Societies, 14(12). https://doi.org/10.3390/soc14120267
Momeni, M. (2024). Artificial Intelligence and Political Deepfakes: Shaping Citizen Perceptions Through Misinformation. Journal of Creative Communications. https://doi.org/10.1177/09732586241277335
Parra, J. L., & Chatterjee, S. (2024). Social Media and Artificial Intelligence: Critical Conversations and Where Do We Go from Here? Education Sciences, 14(1), 68. https://doi.org/10.3390/educsci14010068
Pecheranskyi, I., Oliinyk, O., Medvedieva, A., Danyliuk, V., & Hubernator, O. (2024). Perspectives of Generative AI in the Context of Digital Transformation of Society, Audio-Visual Media and Mass Communication: Instrumentalism, Ethics and Freedom. Indian Journal of Information Sources and Services, 14(4), 48–53. https://doi.org/10.51983/ijiss-2024.14.4.08
Sadiku, M. N. O., Ashaolu, T. J., Ajayi-Majebi, A., & Musa, S. M. (2021). Artificial Intelligence in Social Media. International Journal Of Scientific Advances, 2(1). https://doi.org/10.51542/ijscia.v2i1.4
Sánchez-García, P., Merayo-álvarez, N., Calvo-Barbero, C., & Diez-Gracia, A. (2023). Spanish technological development of artificial intelligence applied to journalism: companies and tools for documentation, production and distribution of information. Profesional de La Informacion, 32(2). https://doi.org/10.3145/EPI.2023.MAR.08
Shoukat, M. H., Elgammal, I., Selem, K. M., & Shehata, A. E. (2024). Fostering social media user intentions: AI-enabled privacy and intrusiveness concerns. Spanish Journal of Marketing - ESIC. https://doi.org/10.1108/SJME-07-2023-0205
Singh, T., Truong, M., Roberts, K., & Myneni, S. (2024). Sequencing conversational turns in peer interactions: An integrated approach for evidence-based conversational agent for just-in-time nicotine cravings intervention. Digital Health, 10. https://doi.org/10.1177/20552076241228430
Souza, C. S. E. (2024). The Evidence of Populism in the Narratives of the President of Brazil during the Covid-19 Pandemic. KOME, 12(1), 111–140. https://doi.org/10.17646/KOME.of.11
Swiatek, L., Galloway, C., Vujnovic, M., & Kruckeberg, D. (2024). Humanoid artificial intelligence, media conferences and natural responses to journalists' questions: The end of (human-to-human) public relations? Public Relations Inquiry, 13(1), 113–121. https://doi.org/10.1177/2046147X231221828
Wahjusaputri, S., Nastiti, T. I., Bunyamin, & Sukmawati, W. (2024). Development of artificial intelligence-based teaching factory in vocational high schools in Central Java Province. Journal of Education and Learning, 18(4), 1234–1245. https://doi.org/10.11591/edulearn.v18i4.21422
Xie, T., Ge, Y., Xu, Q., & Chen, S. (2023). Public Awareness and Sentiment Analysis of COVID-Related Discussions Using BERT-Based Infoveillance. AI (Switzerland), 4(1), 333–347. https://doi.org/10.3390/ai4010016
Zuboff, S. (2022). Surveillance Capitalism or Democracy? The Death Match of Institutional Orders and the Politics of Knowledge in Our Information Civilization. Organization Theory, 3(3).

APPENDIX

Appendix can be accessed via QR: