Archipel: Journal of Indonesian Interdisciplinary Studies
E-ISSN : 3124-6621
Vol.1 No. 11 July 2026
https://ojsap.com/index.php/archipel/index
DOI : https://doi.org/10.65739/archipel.v1i11.70

# Generative AI and Strategic Communication: Examining Its Impact on Organizational Communication Effectiveness and Public Trust

**La Iba**
Universitas Haluoleo, Indonesia

Email: iba@uho.av.id

Submit : May 18,, 2026
Revised : June 20, 2026,
Accepted: July 02, 2026
Published : July 27, 2026

**ABSTRACT**

*Generative artificial intelligence (GenAI) has rapidly become embedded in the strategic communication function, reshaping how organizations produce messages, engage stakeholders, and manage crises. Yet evidence on whether these capabilities strengthen organizational communication effectiveness or undermine public trust remains fragmented across disciplines. This article presents a systematic narrative literature review of twenty-five peer-reviewed studies published largely between 2022 and 2026, synthesizing findings from public relations, information systems, organizational behavior, and communication research. Using thematic synthesis, the review identifies four interrelated clusters: generative AI-enabled message production and personalization, algorithmic transparency and disclosure, trust calibration through tone and competence signaling, and organizational or institutional adaptation. Results show that efficiency and personalization gains are frequently offset by a transparency-trust paradox, whereby disclosing AI authorship can simultaneously legitimize and erode credibility depending on stakeholder AI literacy, message context, and institutional responsibility signaling. The novelty of this study lies in proposing an integrative Trust-Transparency-Effectiveness framework connecting micro-level message design, meso-level organizational practice, and macro-level public trust outcomes. The framework offers strategic communicators, corporate leaders, and policymakers a structured lens for deploying generative AI responsibly while safeguarding organizational legitimacy, stakeholder relationships, and public confidence in an increasingly algorithmically mediated communication environment*

***Keywords:*** *generative artificial intelligence; strategic communication; organizational effectiveness; algorithmic transparency; public trust*

## INTRODUCTION

The emergence of generative artificial intelligence (GenAI) systems, from large language models to multimodal content generators, marks one of the most consequential technological shifts to affect organizational communication since the arrival of social media. Since late 2022, tools capable of producing human-like text, images, and structured messages on demand have moved from experimental novelty to routine infrastructure inside marketing departments, public relations agencies, government press offices, and corporate communication teams (Storey et al., 2025). What began as a productivity aid for drafting press releases or social posts has evolved into a strategic capability that touches message ideation, audience segmentation, tone calibration, translation, crisis response drafting, and even real-time conversational engagement through AI-powered chatbots (Mondal et al., 2023; Zhou et al., 2024). This diffusion has occurred with unusual speed: surveys of communication practitioners show adoption rates climbing sharply within a single year, alongside a parallel rise in ethical unease about authenticity, disclosure, and accountability (Henke, 2025; Ishag et al., 2026).

Strategic communication, understood as the purposeful use of communication by an organization to fulfil its mission, has traditionally rested on assumptions of human

authorship, deliberate message design, and accountable spokespersons. GenAI unsettles each of these assumptions simultaneously. On one hand, organizations report meaningful gains in communication effectiveness when GenAI is used well: faster content turnaround, more consistent brand voice across channels, expanded capacity for audience personalization, and the ability to scale stakeholder engagement without proportionally scaling headcount (Maldonado-Canca et al., 2024; Vyas & Vishwakarma, 2026). Optimized AI social chatbots, for instance, have been shown to improve relational outcomes when profile design and message framing are carefully managed (Zhou et al., 2024), and broader organizational research suggests that GenAI is reshaping not only external messaging but also internal work practices, decision routines, and organizational culture more generally (Murire, 2024). On the other hand, an expanding body of evidence documents a countervailing risk: the erosion of public trust when audiences perceive, or discover, that communication was generated or mediated by an algorithm rather than a human being (Schilke & Reimann, 2025; Piller, 2024).

This tension between effectiveness and trust is not merely anecdotal; it is theoretically grounded in longstanding debates about transparency and credibility in communication. Trust in a message source is conventionally understood to depend on perceived competence, benevolence, and integrity, and GenAI complicates each of these dimensions simultaneously. Competence perceptions can rise because AI systems produce polished, error-free, rapidly iterated content, yet integrity perceptions can simultaneously fall if audiences interpret AI involvement as a form of concealment or inauthenticity (Zerilli et al., 2022; Park & Yoon, 2024). Recent experimental work on algorithm transparency shows that disclosure is not a simple switch that automatically increases or decreases trust; rather, its effect depends on how transparency is framed, whether it functions as a genuine "pipeline" toward accountability or merely a superficial "prism" that draws attention to the technology without building confidence (Park & Yoon, 2025). Similarly, research on AI-generated apologies finds that warmth-toned and competence-toned framing produce divergent emotional and trust outcomes depending on context and perceived sincerity (Lim, Hong, & Schneider, 2025).

At the organizational level, these dynamics interact with institutional responsibility. Lim, Lee, Shin, Kim, and Zhang (2025) demonstrate that stakeholder trust in generative AI systems is mediated by perceptions of algorithmic and institutional responsibility, meaning that trust in the technology is inseparable from trust in the organization deploying it. This insight is echoed in public sector research, where Lovari and De Rosa (2025) document that European government communicators face acute legitimacy challenges when adopting GenAI, given the heightened public expectation of accountability from state institutions and the risk that automated outreach may be read as evasive rather than efficient. Zhan, Zheng, Dong, and Thorson (2025) extend this reasoning to government crisis response, finding that AI-generated, care-based messages can increase citizen trust, but only when audiences possess sufficient AI knowledge to interpret the message appropriately, underscoring that trust outcomes hinge as much on audience literacy as on message design. These findings collectively suggest that organizational communication effectiveness and public trust are not independent outcomes of GenAI adoption; they are entangled through the mediating role of transparency, disclosure practice, audience literacy, and institutional reputation.

Despite this growing body of work, the literature remains conspicuously fragmented. Studies on GenAI and communication effectiveness are concentrated in marketing, brand management, and information systems journals, with a focus on productivity, personalization, and consumer response (Vyas & Vishwakarma, 2026; Mahmoud et al., 2025). Studies on GenAI and trust, by contrast, are scattered across

public relations, organizational behavior, human-computer interaction, and public administration outlets, each employing different theoretical vocabularies, from algorithm aversion to transparency signaling to institutional responsibility (Schilke & Reimann, 2025; Zerilli et al., 2022; Gabelaia & Gedmine, 2026). Few studies attempt to connect these two outcome domains within a single explanatory structure that spans multiple levels of analysis. Most existing reviews either summarize GenAI's general societal impact without a communication-specific lens (Storey et al., 2025) or examine a narrow application domain, such as crisis communication (Piller, 2024) or public health messaging (MacKay et al., 2025), without generalizing to the broader strategic communication function. Moreover, the role of communication professionals themselves as ethical gatekeepers, translators, and trainers navigating this technology has only recently begun to receive systematic attention (O'Neil & Vasquez, 2026).

This fragmentation carries practical consequences. Communication executives who consult the marketing-oriented literature encounter optimistic evidence about efficiency and personalization but little guidance on trust risk; those who consult the trust-oriented literature encounter cautionary evidence about disclosure and credibility but little guidance on how to remain competitive and efficient. Practitioners are therefore left to reconcile two partially contradictory literatures on their own, often under time pressure and without a shared conceptual vocabulary. A further complication is that the underlying evidence spans strikingly different organizational contexts, corporate brands, government agencies, healthcare providers, and universities, each with different baseline levels of public trust and different regulatory exposure (Ishag et al., 2026; Gabelaia & Gedmine, 2026). Without a framework that explicitly accounts for these contextual differences while still identifying cross-cutting mechanisms, comparative learning across sectors remains difficult, and organizations risk either overestimating the safety of GenAI adoption or overestimating its dangers.

This article addresses that gap by conducting an integrative, cross-disciplinary literature review that deliberately bridges the effectiveness and trust literatures. Its novelty is threefold. First, rather than treating communication effectiveness and public trust as separate research streams, this review synthesizes them within a single analytical frame, revealing recurring mechanisms, such as the transparency-trust paradox, that operate across both outcome domains. Second, the review introduces a multi-level Trust-Transparency-Effectiveness (TTE) framework that organizes findings according to three nested levels of analysis: the micro level of individual AI-generated messages and their design features (tone, disclosure, personalization), the meso level of organizational practices and governance (institutional responsibility, communicator roles, training), and the macro level of societal and public trust outcomes (citizen trust in institutions, media credibility, algorithm aversion). No prior review, to the authors' knowledge, has organized the GenAI-strategic communication literature according to this explicit multi-level structure. Third, by drawing on twenty-five recent, cross-sectoral studies spanning corporate communication, public relations, government communication, journalism, education, and healthcare, the review offers a more comprehensive and comparative account than single-domain studies typically provide.

The objectives of this study are therefore twofold: (1) to map and synthesize current empirical and conceptual evidence on how generative AI influences organizational communication effectiveness and public trust, and (2) to propose an integrative framework that can guide both future research and practical decision-making about responsible GenAI deployment in strategic communication. The remainder of this article proceeds as follows. The next section describes the systematic narrative review method used to identify, screen, and synthesize the twenty-five sources analyzed. The

subsequent section presents results organized around four thematic clusters and discusses their implications through the proposed TTE framework, supported by a comparative summary table. The final section concludes with theoretical and practical implications, acknowledges limitations, and outlines directions for future research.

## METHODS

This study employs a systematic narrative literature review, a method well suited to synthesizing conceptually diverse but topically related evidence across disciplines when the goal is theory-building rather than statistical meta-analysis (Storey et al., 2025). The review followed a structured four-stage process broadly informed by PRISMA-style reporting logic: identification, screening, eligibility assessment, and inclusion, as depicted in Figure 1.

In the identification stage, records were retrieved from academic aggregation and citation-mapping platforms covering major indexed databases (including Scopus-, Web of Science-, and Crossref-indexed journals), using combinations of the keywords "generative artificial intelligence," "strategic communication," "organizational communication effectiveness," "algorithmic transparency," and "public trust." The search was restricted to peer-reviewed journal articles published between 2021 and 2026 to ensure currency, given the rapid evolution of generative AI capabilities since the public release of large language model applications in late 2022.

In the screening stage, titles and abstracts were reviewed to remove records unrelated to organizational, institutional, or public-facing communication contexts, such as purely technical machine-learning papers with no communication or trust dimension. In the eligibility stage, full texts were assessed against three inclusion criteria: (a) the study addressed generative AI in relation to communication, messaging, or stakeholder engagement; (b) the study reported empirical findings, a validated instrument, or a substantive conceptual contribution relevant to organizational effectiveness or public trust; and (c) the study was published in a peer-reviewed outlet with an active, resolvable digital object identifier (DOI). Studies were excluded if they addressed generative AI purely from a technical, engineering, or algorithmic-performance standpoint without any communication, trust, or organizational implication.

Twenty-five sources met all inclusion criteria and form the basis of this review. Twenty were drawn from a curated bibliographic corpus assembled around the review's core research questions, and five additional sources were subsequently identified through supplementary targeted searches to strengthen coverage of organizational adoption, human resource communication, higher-education institutional response, and journalistic trust, all published within the preceding five years. The final corpus spans public relations and corporate communication (e.g., Ishag et al., 2026; Gabelaia & Gedmine, 2026), information systems (Storey et al., 2025), organizational behavior (Schilke & Reimann, 2025), public administration (Lovari & De Rosa, 2025; Zhan et al., 2025), higher education (Henke, 2025; Gering et al., 2025), human resource management (Prasad & De, 2024), journalism (Haq, 2025), and collective-intelligence and organization design research (Woolley, 2026), providing a cross-sectoral evidence base.

Data extraction followed a standardized matrix capturing author(s), year, discipline, methodological approach, unit of analysis, and key findings relevant to communication effectiveness, transparency, or trust. Thematic synthesis, following an inductive-deductive hybrid coding approach, was then used to group findings into recurring themes. Two researchers' coding logic (represented analytically in a single-author review through iterative re-coding) was cross-checked against the extraction matrix to reduce interpretive drift, and disagreements in theme labeling were resolved

by returning to the original source text. Four thematic clusters emerged inductively from this process and were subsequently mapped deductively onto the three levels, micro, meso, and macro, of the proposed Trust-Transparency-Effectiveness framework, which structures the presentation of results in the following section.

As a literature-based synthesis rather than a primary empirical study, this review does not involve human participants, surveys, or experimental manipulation, and therefore required no separate data-collection instrument beyond the extraction matrix described above. Its principal methodological limitation, namely reliance on published, English-language, DOI-indexed sources, is addressed in the concluding section.

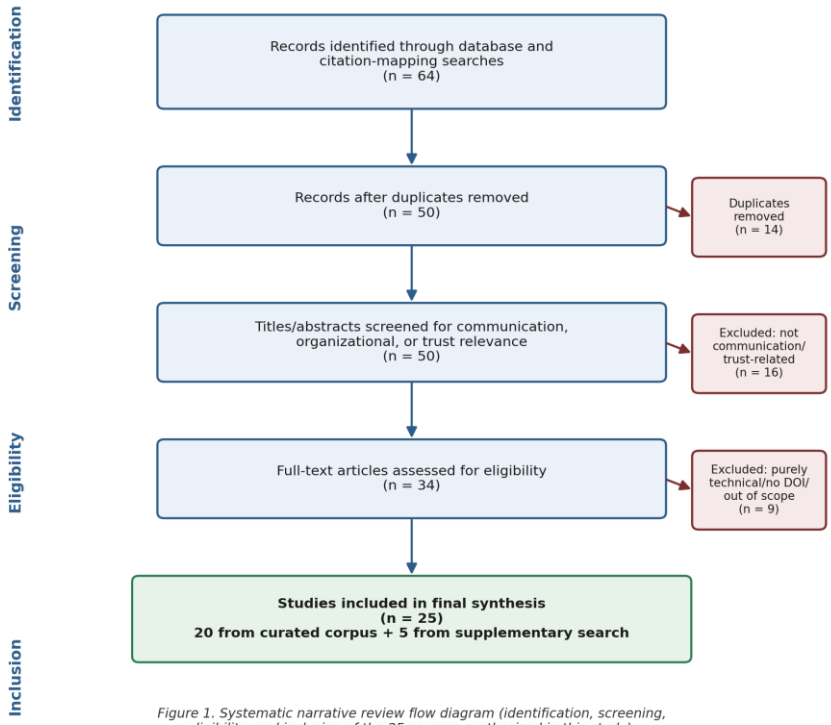


*Figure 1. Systematic narrative review flow diagram (identification, screening, eligibility, and inclusion of the 25 sources synthesized in this study).*

## RESULTS AND DISCUSSION

Thematic synthesis of the twenty-five reviewed sources produced four recurring clusters of findings, which this section organizes according to the three levels of the proposed Trust-Transparency-Effectiveness (TTE) framework: micro-level message design, meso-level organizational practice, and macro-level public trust outcomes. Table 1 summarizes the clusters, representative studies, and key findings before the discussion elaborates each in turn.

*Table 1. Thematic synthesis of the reviewed literature mapped to the Trust-Transparency-Effectiveness (TTE) framework.*

| Level (TTE Framework) | Thematic Cluster | Representative Studies | Key Findings |
|---|---|---|---|
| **Micro (message)** | GenAI-enabled message production, personalization, and disclosure framing | Storey et al. (2025); Zhou et al. (2024); Lim, Hong & Schneider (2025); Piller (2024) | GenAI speeds drafting and personalization, but relational and trust outcomes depend heavily on chatbot profile design, message framing, and perceived sincerity of disclosure. |

| | | | |
|---|---|---|---|
| **Meso (organization)** | Organizational governance and pace of adoption | Gering et al. (2025); Henke (2025); Lovari & De Rosa (2025); Murire (2024) | Adoption of GenAI in communication units is outpacing governance maturity across universities, government agencies, and firms, creating uneven risk exposure. |
| **Meso (organization)** | Evolving role of the communication professional | O'Neil & Vasquez (2026); Prasad & De (2024) | Communicators increasingly act as ethical guardians, relationship champions, and trainers; internal trust in GenAI mediates employee engagement and commitment. |
| **Meso (organization)** | Institutional responsibility as a trust precondition | Lim, Lee, Shin, Kim & Zhang (2025); Gabelaia & Gedmine (2026) | Trust in GenAI-generated corporate messaging depends on prior institutional credibility and communicated data/AI governance responsibility, not message content alone. |
| **Macro (public trust)** | Transparency-trust paradox (pipeline vs. prism) | Schilke & Reimann (2025); Park & Yoon (2025, 2024); Zerilli et al. (2022) | AI disclosure can raise or erode trust depending on whether transparency is substantive (accountability) or cosmetic (attention-drawing) signaling. |
| **Macro (public trust)** | Audience AI literacy and sector heterogeneity | Zhan et al. (2025); Mahmoud et al. (2025); Haq (2025); Woolley (2026) | Trust effects of AI-generated government, journalistic, and organizational messaging vary with citizens' AI knowledge and differ meaningfully across sectors. |

***Micro level: message production, personalization, and disclosure design***

At the micro level, the reviewed literature converges on the finding that GenAI substantially expands what a single communicator can produce, personalize, and iterate within a given time window. Storey, Yue, Zhao, and Lukyanenko (2025) frame this as a structural shift in information systems capability, in which generative models move from narrow task automation toward general-purpose content generation across text, image, and structured formats. Applied to communication practice, this translates into faster drafting cycles, more consistent tone across channels, and expanded capacity for micro-targeted messaging (Vyas & Vishwakarma, 2026; Mondal et al., 2023). Zhou, Tsai, and Men (2024) provide some of the most granular evidence at this level, showing experimentally that AI chatbot profile design and message framing jointly shape relational outcomes such as perceived warmth, satisfaction, and continued engagement intention; poorly designed disclosure or an overly mechanical profile undermines the very relational benefits the chatbot was deployed to achieve.

Crucially, message-level disclosure design does not operate in isolation from perceived sincerity. Lim, Hong, and Schneider (2025) find that AI-generated apologies framed with warmth cues generate different emotional and forgiveness trajectories than those framed with competence cues, and that perceived sincerity mediates both pathways. This suggests that organizations cannot treat "AI disclosure" as a single, uniform design choice; rather, the tone, framing, and context surrounding disclosure interact to determine whether transparency reads as authentic accountability or as a hollow procedural gesture. Piller (2024) extends this logic to crisis communication specifically, arguing that GenAI-generated crisis messages risk what he terms an

"inhuman rhetoric," in which technically fluent language fails to carry the moral seriousness that stakeholders expect during a crisis, unless deliberately counterbalanced by human oversight and emotionally calibrated framing.

***Meso level: organizational adaptation, governance, and the communicator's role***

At the meso level, the review identifies a consistent pattern of organizational adaptation occurring faster than governance structures can absorb. Gering, Feher, Harmat, and Tamassy (2025) document that even leading global universities exhibit wide variation in how systematically they have developed GenAI governance strategies, ranging from comprehensive institution-wide frameworks to fragmented, faculty-level improvisation. Henke (2025) corroborates this from a communication-practice angle, finding that university communication departments moved from cautious experimentation in 2023 to substantially broader integration by 2024, with efficiency and channel adaptability rising sharply, yet with concerns about factual accuracy and privacy persisting largely unresolved. This pattern, rapid adoption outpacing governance maturity, recurs across sectors: Lovari and De Rosa (2025) find a similar dynamic in European public sector communication, where practical experimentation with GenAI has outstripped the development of clear ethical and legal guardrails.

A second meso-level pattern concerns the evolving role of the communication professional. O'Neil and Vasquez (2026) argue that senior communicators are increasingly expected to act simultaneously as ethical guardians who vet AI-generated content for accuracy and appropriateness, as relationship champions who preserve the human touch in stakeholder engagement, as trainers who build organizational AI literacy, and as strategic advisors who help leadership calibrate the pace of adoption. This multidimensional role expansion helps explain why organizational communication effectiveness cannot be reduced to output volume or speed; effectiveness increasingly depends on the quality of human oversight embedded around GenAI outputs. Prasad and De (2024) offer complementary evidence from human resource management, showing that trust partially mediates the relationship between employees' perception of generative AI tools and their organizational commitment, engagement, and performance, indicating that the trust dynamics documented in external-facing communication research also operate internally, shaping how employees relate to AI-mediated organizational communication.

A third meso-level pattern relates to institutional responsibility as a precondition for effective GenAI-enabled stakeholder engagement. Lim, Lee, Shin, Kim, and Zhang (2025) show that perceived stakeholder engagement in corporate data responsibility communication predicts trust in generative AI systems only when mediated by perceptions of algorithmic and institutional responsibility; organizations that communicate about their data and AI governance practices, rather than simply deploying the technology silently, achieve materially better trust outcomes. Gabelaia and Gedmine (2026) reinforce this point, finding that AI-generated corporate messaging can function as a credibility signal, but only within organizations that have already established a baseline of institutional trustworthiness; the same message content produces divergent credibility effects depending on the sending organization's prior reputation.

***Macro level: transparency, algorithm aversion, and public trust***

At the macro level, the clearest and most consequential finding across the reviewed literature is what this review terms the transparency-trust paradox: disclosing AI involvement in communication does not straightforwardly increase or decrease public trust, but instead produces divergent effects depending on framing, audience literacy, and institutional context. Schilke and Reimann (2025) provide the most direct evidence of

this paradox, demonstrating across multiple organizational and interpersonal contexts that AI disclosure can erode trust even when the disclosed content is materially unchanged, because disclosure itself signals reduced human effort or investment. Park and Yoon (2025, 2024) complicate this picture further, showing that transparency signaling functions as a genuine trust-building "pipeline" only when it is substantive and verifiable; when transparency functions merely as a superficial "prism," drawing attention to the technology without offering accountability, it can heighten rather than reduce negative attitudes. Zerilli, Bhatt, and Weller's (2022) earlier theoretical account anticipates precisely this nuance, arguing that transparency modulates trust conditionally rather than unconditionally, depending on what is made transparent and to whom.

Government and public-sector contexts illustrate the macro-level stakes most clearly. Zhan, Zheng, Dong, and Thorson (2025) find that AI-generated, care-based crisis messages from government sources can increase citizen trust, but this effect is conditional on citizens' existing AI knowledge; audiences with low AI literacy respond less favorably, suggesting a public-education dimension to responsible GenAI communication that has received little attention in organizational practice. Mahmoud, Kumar, and Spyropoulou (2025), using large-scale public discourse data, similarly find that public beliefs about generative AI are heterogeneous and unstable, shaped by media narratives, personal experience, and sector-specific trust baselines rather than by a single, generalizable public attitude. This heterogeneity helps explain why single-sector studies produce seemingly contradictory conclusions about whether GenAI helps or harms public trust: both conclusions are locally valid, but neither generalizes cleanly across institutional contexts.

Journalism offers a particularly instructive macro-level case because it sits at the intersection of organizational communication effectiveness and public trust in information systems more broadly. Haq (2025) finds that generative AI's impact on journalistic credibility is similarly double-edged: efficiency and consistency gains in news production coexist with heightened audience skepticism when AI involvement is suspected or disclosed, echoing the transparency-trust paradox identified at the organizational level. Woolley (2026) situates these tensions within a broader argument about collective intelligence, contending that GenAI's organizational value depends on whether it is deployed as an individual production technology, which risks fragmenting workflows and isolating human judgment, or as a coordination technology that augments collective reasoning, memory, and attention; the latter deployment mode is more consistent with sustaining both effectiveness and trust simultaneously.

### *Integration through the Trust-Transparency-Effectiveness framework*

Read together, these four clusters support the central claim of the proposed TTE framework: communication effectiveness and public trust are not parallel, independent outcomes of GenAI adoption but are causally linked through transparency practice at each level of analysis. At the micro level, disclosure design and tonal framing determine whether individual AI-generated messages are read as authentic or evasive. At the meso level, organizational governance maturity and the evolving role of communicators determine whether transparency is substantive (a "pipeline") or cosmetic (a "prism"). At the macro level, institutional reputation and audience AI literacy determine whether transparency signals accountability or incompetence. Effectiveness gains achieved by bypassing transparency, for instance, by concealing AI authorship to preserve a "human" brand voice, appear to generate short-term efficiency at the cost of long-term trust erosion once AI involvement is eventually discovered, a risk documented directly by Schilke and Reimann (2025) and indirectly corroborated by Haq's (2025) journalism

findings. Conversely, transparency pursued without corresponding investment in message quality, institutional accountability, or audience education risks the "prism" effect identified by Park and Yoon (2025), in which disclosure alone fails to build trust. The strategic implication is that organizations must pursue transparency and effectiveness jointly rather than trading one for the other, a conclusion that few single-domain studies in the reviewed corpus state explicitly but that emerges clearly once the effectiveness and trust literatures are read together.

Several practical implications follow for strategic communication practice. First, disclosure policies should be designed at the message level with explicit attention to tone and framing, rather than treated as a binary legal or ethical checkbox; the evidence from Lim, Hong, and Schneider (2025) and Piller (2024) indicates that how disclosure is framed matters as much as whether it occurs. Second, organizations should invest in meso-level governance capacity, including clear internal policies on AI use, human review checkpoints, and communicator training, before scaling GenAI deployment, since Gering et al. (2025) and Henke (2025) both show that governance maturity, not technological sophistication, differentiates organizations that adapt successfully from those that experience reputational friction. Third, institutional trust-building should be treated as a prerequisite for, not a byproduct of, GenAI-enabled stakeholder engagement; Gabelaia and Gedmine (2026) and Lim, Lee, Shin, Kim, and Zhang (2025) both suggest that the same AI-generated content performs very differently depending on the sending organization's prior credibility, meaning that GenAI adoption should be sequenced after, not instead of, foundational trust-building work. Fourth, public-facing organizations, particularly government agencies and public-interest institutions, should consider audience AI literacy as a variable they can actively influence rather than a fixed constraint, given Zhan et al.'s (2025) finding that citizen trust in AI-generated crisis messages depends partly on citizens' own understanding of the technology. Taken together, these implications position the TTE framework not merely as a descriptive summary of the literature but as a practical diagnostic tool: communicators can use it to ask, at each level of their organization's GenAI deployment, whether transparency is functioning as genuine accountability or merely as a symbolic gesture, and to adjust practice accordingly before effectiveness gains are undermined by preventable trust erosion.

## CONCLUSIONS

This review synthesized twenty-five recent, cross-sectoral studies to examine how generative artificial intelligence reshapes organizational communication effectiveness and public trust. The evidence shows that GenAI reliably enhances communication capacity, speed, personalization, and scale, but that these effectiveness gains are neither automatically nor unconditionally accompanied by stronger public trust. Instead, the reviewed literature converges on a transparency-trust paradox operating at every level of communication practice: disclosure and algorithmic transparency can either build or erode trust depending on framing, sincerity, institutional credibility, and audience AI literacy. The study's principal contribution is the integrative Trust-Transparency-Effectiveness framework, which links micro-level message design, meso-level organizational governance, and macro-level public trust outcomes within a single explanatory structure, addressing a gap left by prior reviews that treat communication effectiveness and public trust as separate research streams.

Theoretically, this review extends transparency and trust scholarship by demonstrating that its core mechanisms, the pipeline-versus-prism distinction, warmth-competence framing, and institutional responsibility mediation, recur consistently across

corporate, governmental, educational, and journalistic contexts, suggesting a generalizable pattern rather than sector-specific anomalies. Practically, the review offers strategic communicators a diagnostic framework for evaluating whether GenAI deployment is likely to sustain or undermine stakeholder trust, and it underscores that governance maturity and institutional credibility, not technological capability alone, determine long-term success.

This study is not without limitations. As a narrative literature review rather than a systematic meta-analysis, it does not statistically aggregate effect sizes and remains subject to the interpretive judgment inherent in thematic synthesis, even with a structured extraction matrix. The corpus, while cross-sectoral, is limited to English-language, DOI-indexed journal articles, which may underrepresent practitioner reports, non-English scholarship, and rapidly evolving industry practice not yet captured in peer-reviewed outlets. Additionally, because generative AI capabilities and organizational practices are evolving rapidly, some findings, particularly those concerning specific tools or platforms, may have a limited shelf life.

Future research should pursue longitudinal designs that track how organizational trust trajectories evolve as GenAI disclosure norms mature and as public AI literacy increases, potentially altering the strength of the transparency-trust paradox over time. Comparative cross-national studies could further test whether the TTE framework's mechanisms hold across differing regulatory and cultural trust baselines, particularly outside the North American and European contexts that dominate the current evidence base. Finally, future work should empirically test the TTE framework itself, using structural modeling or experimental designs, to validate the proposed pathways linking message-level transparency, organizational governance, and public trust outcomes, thereby moving the field from descriptive synthesis toward predictive and prescriptive theory for the responsible strategic use of generative AI in organizational communication.

## REFERENCES


Gabelaia, I., & Gedmine, M. (2026). The impact of AI-generated corporate messaging as a credibility signal in algorithmic communication. International Journal of Management, Knowledge and Learning. https://doi.org/10.53615/2232-5697.15.271-284

Gering, Z., Feher, K., Harmat, V., & Tamassy, R. (2025). Strategic organisational responses to generative AI-driven digital transformation in leading higher education institutions. International Journal of Organizational Analysis, 33(12), 132-152. https://doi.org/10.1108/IJOA-09-2024-4850

Haq, A. (2025). The impact of generative AI on journalistic credibility and trust. Global Social Sciences Review, 10(3), 260-270. https://doi.org/10.31703/gssr.2025(X-III).22

Henke, J. (2025). The new normal: The increasing adoption of generative AI in university communication. Journal of Science Communication, 24(2), A07. https://doi.org/10.22323/2.24020207

Ishag, K. I. A., Setoutah, S., Mahmoud, K., Al-Obaidi, E., & Mohamed, E. (2026). The impact of generative artificial intelligence on public relations strategies: A shift in practitioner roles and the future of corporate communication. Studies in Media and Communication, 14(3). https://doi.org/10.11114/smc.v14i3.8719

Lim, J. S., Hong, N., & Schneider, E. (2025). How warm- versus competent-toned AI apologies affect trust and forgiveness through emotions and perceived sincerity.

Computers in Human Behavior, 172, 108761. https://doi.org/10.1016/j.chb.2025.108761

Lim, J. S., Lee, C., Shin, D., Kim, J., & Zhang, J. (2025). Perceived stakeholder engagement in corporate data responsibility (CDR) communication and its relationship with trust in generative AI systems: The mediating role of algorithmic and institutional responsibility. Journal of Public Relations Research, 37(6), 447-469. https://doi.org/10.1080/1062726X.2025.2501552

Lovari, A., & De Rosa, F. (2025). Exploring the challenges of generative AI on public sector communication in Europe. Media and Communication. https://doi.org/10.17645/mac.9644

MacKay, M., Kukan, A., & McWhirter, J. (2025). The double-edged algorithm: A rapid review exploring the trustworthy and responsible use of generative AI in public health. Digital Health, 11. https://doi.org/10.1177/20552076251393302

Mahmoud, A., Kumar, V., & Spyropoulou, S. (2025). Identifying the public's beliefs about generative artificial intelligence: A big data approach. IEEE Transactions on Engineering Management, 72, 827-841. https://doi.org/10.1109/TEM.2025.3534088

Maldonado-Canca, L.-A., Casado-Molina, A.-M., Cabrera-Sánchez, J.-P., & Bermúdez-González, G. (2024). Beyond the post: An SLR of enterprise artificial intelligence in social media. Social Network Analysis and Mining, 14. https://doi.org/10.1007/s13278-024-01382-y

Mondal, S., Das, S., & Vrana, V. (2023). How to bell the cat? A theoretical review of generative artificial intelligence towards digital disruption in all walks of life. Technologies, 11(2), 44. https://doi.org/10.3390/technologies11020044

Murire, O. (2024). Artificial intelligence and its role in shaping organizational work practices and culture. Administrative Sciences, 14(12), 316. https://doi.org/10.3390/admsci14120316

O'Neil, J., & Vasquez, R. (2026). The role and value of senior communicators in the AI era: Ethical guardians, relationship champions, trainers, and strategic advisors. Corporate Communications: An International Journal. https://doi.org/10.1108/CCIJ-11-2025-0374

Park, K., & Yoon, H. Y. (2024). Beyond the code: The impact of AI algorithm transparency signaling on user trust and relational satisfaction. Public Relations Review. https://doi.org/10.1016/j.pubrev.2024.102507

Park, K., & Yoon, H. Y. (2025). AI algorithm transparency, pipelines for trust not prisms: Mitigating general negative attitudes and enhancing trust toward AI. Humanities and Social Sciences Communications, 12. https://doi.org/10.1057/s41599-025-05116-z

Piller, E. (2024). Inhuman rhetoric: Generative AI and crisis communication. Journal of Business and Technical Communication, 39(1), 42-50. https://doi.org/10.1177/10506519241280594

Prasad, K. D. V., & De, T. (2024). Generative AI as a catalyst for HRM practices: Mediating effects of trust. Humanities and Social Sciences Communications, 11, 1362. https://doi.org/10.1057/s41599-024-03842-4

Schilke, O., & Reimann, M. (2025). The transparency dilemma: How AI disclosure erodes trust. Organizational Behavior and Human Decision Processes. https://doi.org/10.1016/j.obhdp.2025.104405

Storey, V., Yue, W., Zhao, J., & Lukyanenko, R. (2025). Generative artificial intelligence: Evolving technology, growing societal impact, and opportunities for information

systems research. Information Systems Frontiers, 27, 2081-2102. https://doi.org/10.1007/s10796-025-10581-7
Vyas, Y., & Vishwakarma, P. (2026). Artificial intelligence in brand communication: A comprehensive literature review and research outlook. International Journal of Consumer Studies. https://doi.org/10.1111/ijcs.70170
Woolley, A. W. (2026). Generative AI and collaboration: Opportunities for cultivating collective intelligence. Journal of Organization Design, 15, 45-50. https://doi.org/10.1007/s41469-025-00199-z
Zerilli, J., Bhatt, U., & Weller, A. (2022). How transparency modulates trust in artificial intelligence. Patterns, 3. https://doi.org/10.1016/j.patter.2022.100455
Zhan, E., Zheng, Q., Dong, C., & Thorson, E. (2025). Does AI-generated care-based message increase trust in government? The pivotal role of AI knowledge in government crisis response. International Journal of Strategic Communication, 19(3), 176-199. https://doi.org/10.1080/1553118X.2025.2471527
Zhou, A., Tsai, W.-H. S., & Men, L. (2024). Optimizing AI social chatbots for relational outcomes: The effects of profile design, communication strategies, and message framing. International Journal of Business Communication, 63, 648-673. https://doi.org/10.1177/23294884241229223
Verga Syaharani Sukma, Lia Nuraini & Muhammad Fajar Hidayat. (2025). Perlindungan Konsumen terhadap Makanan Impor Tanpa Label Bahasa Indonesia yang Dijual Melalui E-Commerce. Aliansi: Jurnal Hukum, Pendidikan, dan Sosial Humaniora.
W., & Dwitami, A. D. (2025). AKIBAT HUKUM PEREDARAN MAKANAN DAN MINUMAN OLAHAN ILEGAL DALAM PERSPEKTIF PERLINDUNGAN KONSUMEN. Justitia et Pax. https://doi.org/10.24002/jep.v41i1.8385
Wibowo, D. E., & Madusari, B. (2018). Legal Protection for Consumers on Unlabelled Processed Food from Seaweed in Brebes Regency. 06010. https://doi.org/10.1051/shsconf/20185406010
Yuanitasari, D., & Suwandono, A. (2024). Upaya Peningkatan Pemahaman Pelaku Usaha Mikro, Kecil, Dan Menengah (UMKM) Terkait Pencantuman Label Halal Pada Produk Pangan. Proficio, 5(2), 174-182.