Frontiers in Business Innovations and Management (Open Access)

Volume 03, Issue 06, June 2026,
Publish Date: 30-06-2026
DOI: - https://doi.org/10.64917/fbim/Volume03Issue06-01 PageNo.01-10

The Evolution of The Theoretical Foundations of Crisis Communication from The Pre-Internet Era to The Age of Generative Artificial Intelligence

Ali Hajizada Nizami
Sawol Communications Inc., Washington, USA Hajizade Group, Baku, Azerbaijan

RECEIVED - 27-04-2026, **RECEIVED REVISED VERSION** - 19-05-2026, **ACCEPTED**- 07-06-2026, **PUBLISHED**- 30-06-2026

Abstract

This article examines the evolution of crisis communication theory from pre-internet models of institutional response to conditions shaped by generative artificial intelligence. The study addresses a growing mismatch between inherited crisis communication principles and a communication environment in which reputational threats can be produced, amplified, personalized, and interpreted through AI systems. The objective is to explain how crisis communication theory has changed across three stages: the pre-internet period, digital adaptation, and the age of generative AI. The methodology relies on comparative source analysis, conceptual synthesis, typologization, and analytical generalization of ten recent scholarly publications in crisis communication, public relations, misinformation studies, and human-machine communication. The results identify a shift from event-centered response logic to dynamic, adversarial, and machine-mediated crisis conditions. The article proposes a revision path based on anticipatory monitoring, human-machine response governance, authenticity infrastructure, and adaptive trust repair. Its practical value lies in a decision model for organizations facing AI-generated misinformation, synthetic evidence, and algorithmically accelerated reputation crises.

Keywords: *crisis communication, generative artificial intelligence, public relations, misinformation, reputation, trust*

Introduction

Crisis communication theory grew out of a media order in which organizations usually dealt with identifiable events, visible stakeholders, official channels, and response cycles that could be planned with some confidence. A product failure, public accusation, accident, scandal, or institutional controversy required explanation, correction, apology, compensation, or reassurance. The crisis manager's task was to interpret the event, evaluate responsibility, choose a response, and protect trust before reputational damage became durable.

The internet weakened that order but did not fully dismantle it. Digital platforms increased speed, diversified publics, multiplied unofficial sources, and made crisis interpretation more participatory. Organizational routines still kept much of the inherited logic: monitor public reaction, assess responsibility, choose a response strategy, deliver messages, and evaluate reputational effects. Digital communication changed channels, tempo, and stakeholder visibility. It left much of the theoretical core in place.

Generative artificial intelligence creates a sharper break. Synthetic text, audio, images, video, automated accounts, chatbot-mediated interaction, and AI-assisted disinformation campaigns alter the conditions under which crises begin. A reputational threat no longer requires a factual incident. A crisis can grow from fabricated evidence, manipulated archives, cloned voices, synthetic screenshots, automated narrative flooding, or ambiguous use of AI by the organization or individuals associated with it. This produces a velocity gap between crisis generation and institutional response. Classical theory was built around interpretation after an event. Generative AI can manufacture the appearance of the event.

The aim of this article is to develop an analytical interpretation of how crisis communication theory has evolved from pre-internet response models to the conditions of generative artificial intelligence.

The study pursues three objectives. The first is to reconstruct the shift from pre-internet crisis communication logic to digital crisis communication without reducing the transition to a change of media channels. The second is to identify why generative AI places inherited crisis response principles under strain. The third is to formulate a revised conceptual model for crisis communication in conditions of algorithmically generated, synthetic, and machine-mediated crises.

The novelty of the article lies in treating the current transformation of crisis communication as a theoretical rupture in the relation between crisis evidence, attribution, response speed, and trust repair. The working hypothesis states that the theoretical core of crisis communication remains useful only after revision from a response-centered model into an anticipatory and authenticity-centered model designed for crises that AI systems can generate, amplify, and personalize.

Materials and Methods

The material base consists of ten scholarly publications issued between 2021 and 2025 and selected from journals and platforms covering crisis communication, public relations, misinformation research, risk communication, information management, and media studies. The search strategy combined the terms “crisis communication,” “artificial intelligence,” “generative AI,” “AI-driven disinformation,” “crisis misinformation,” “corrective communication,” “ChatGPT crisis communication,” “public relations misinformation,” “dynamic crisis theory,” and “AI-triggered organizational threats.” The search covered ScienceDirect, Taylor & Francis, Emerald, Wiley Online Library, SAGE Journals, Frontiers, and publisher-level journal pages. Inclusion criteria required direct relevance to crisis communication theory, AI-mediated communication, crisis misinformation, organizational readiness, or generative AI in communication industries. Excluded works dealt mainly with technical deepfake detection, general cybersecurity, electoral disinformation without organizational communication relevance, PR measurement, and industry advice without scholarly grounding. A wider pool of more than thirty texts was screened. Ten sources remained after removing works outside the thematic boundary, items with weak journal credibility, studies focused only on measurement metrics, and publications centered narrowly on deepfake trust restoration. The final corpus covers five groups of questions: the modernization of crisis communication theory in dynamic crisis situations (Jong, 2025), AI applications in crisis communication (Cheng et al., 2025), AI-mediated crisis interaction and chatbot response (Xiao & Yu, 2025), organizational preparation for AI-driven disinformation (Karinshak & Jin, 2023), crisis misinformation detection and management (Mehta et al., 2021), corrective and prebunking strategies (Lu & Jin, 2024), crisis readiness against misinformation (Liu & Zhang, 2025), reputational bandwagon effects in misinformation diffusion (Kim & Lim, 2026), practitioner views of AI-triggered threats (Zhao et al., 2025), and the broader consequences of generative AI for public relations and media industries (Guzman & Lewis, 2024).

The study applies comparative analysis to trace changes between historical crisis response logic, digital adaptation, and AI-era conditions. Source analysis separates academic models from practice-oriented claims. Conceptual synthesis connects crisis responsibility, misinformation, synthetic media, machine mediation, and organizational readiness. Typologization supports the classification of crisis communication theory by technological stage. Analytical generalization supports the revised model for generative AI conditions.

Results

The first stage in the evolution of crisis communication theory formed around a fairly stable relation between crisis event, organizational responsibility, public interpretation, and institutional response. Recent studies that revisit classical theories describe their value through structured responsibility assessment, response matching, stakeholder interpretation, and reputational repair (Jong, 2025). This logic developed in a media environment where organizations could treat communication as a sequence. An incident occurred, stakeholders interpreted it, media actors framed it, and the organization selected a response. Theoretical strength came from the ability to turn reputational turbulence into a decision problem.

The same strength now reveals a limit. A model built around response selection works best when the object of

response remains stable enough for classification. In pre-internet and early mass-media settings, crisis managers did not have to deal with the endless multiplication of synthetic crisis objects. Organizations faced hostile reporting, blame attribution, rumors, activist pressure, and institutional distrust, but they could still orient themselves toward a recognizable event. Recent work on dynamic crisis situations reports that crisis responsibility rarely stays fixed and can shift as new information appears, stakeholders reinterpret evidence, and accountability spreads across several actors (Jong, 2025). This matters for the current argument because inherited theory had already started to outgrow static categories before generative AI. AI increases that instability and makes it harder to decide what exactly requires a response.

The second stage developed through the internet and social media. Digital platforms did not erase classical crisis communication theory. They stretched it. Organizations kept using familiar ideas such as responsibility attribution, corrective communication, apology, denial, rebuilding, bolstering, transparency, and stakeholder engagement. Speed and visibility changed. Stakeholders began interpreting crises in public, through platforms, comment threads, reposts, memes, and informal expert communities. Research on crisis misinformation from the perspective of public relations professionals reports that misinformation varies by formation, severity, source, and format, while verification processes shape the timing and quality of correction (Mehta et al., 2021). The value of this finding lies in its process view of misinformation. False content does not circulate as one isolated message. Public relations professionals have to verify, classify, and decide whether the intervention will reduce harm or amplify the falsehood.

Digital adaptation added monitoring and correction to the classical response apparatus, but it left the center of theory largely intact. Organizational practice still leaned toward reaction. The organization detected misinformation or disinformation, vetted it, and chose correction, silence, legal action, or stakeholder engagement. That logic assumes that the false claim has a manageable form. It assumes that public relations professionals can identify the object of correction and the channel where correction should appear. Social media had already weakened this assumption. Generative AI weakens it further because misleading or intentionally deceptive claims can be reformulated, localized, translated, personalized, visually supported, and redistributed through semi-automated networks.

Recent studies on corrective communication show that crisis misinformation response cannot rely on refutation as a single instrument. Experimental work on correction placement, refutation detail, and corrective narrative type reports that prebunking in narrative form can work when misleading or intentionally deceptive content circulates as a narrative (Lu & Jin, 2024). This contributes to the second objective of the article because it shows movement away from bare factual rebuttal. Correction has to compete with the form, timing, and emotional structure of misleading or intentionally deceptive content. Crisis communication theory, in turn, has to account for narrative competition. Audiences do not treat truth as self-explanatory once an organization releases an official statement.

Research on the bandwagon effect deepens the same problem. Crisis misinformation with visible popularity cues can damage organizational reputation because perceived social endorsement affects how audiences evaluate the disputed message (Kim & Lim, 2026). The issue extends beyond social media metrics. A crisis message gains force through signs of collective uptake. Likes, shares, repeated claims, influencer involvement, or bot-amplified visibility can create the appearance of consensus before audiences evaluate evidence. This finding explains why inherited crisis response principles often fall behind platform dynamics. Traditional response theory asks what the organization should say. Digital crisis theory has to ask whether the organization can enter the interpretive space before crowd signals begin to look like reputational evidence.

Generative AI marks the third stage because it changes the production conditions of the crisis itself. Studies of AI-driven disinformation treat AI as a force that can support the creation, targeting, and scaling of deceptive content, which requires organizational preparation before a specific attack occurs (Karinshak & Jin, 2023). This work carries direct relevance for the present article because it moves crisis communication from message response to readiness architecture. A company, public institution, or public figure no longer prepares only for accidents, scandals, service failures, or public criticism. It prepares for fabricated crisis triggers that imitate the evidentiary qualities of real events.

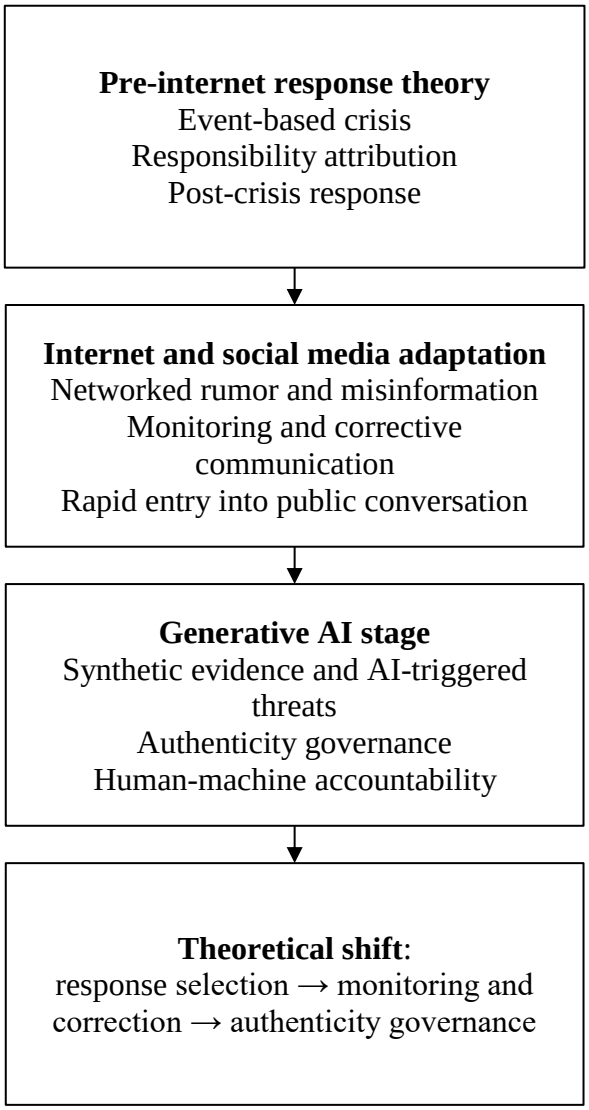
The same shift appears in practitioner-oriented research on AI-triggered organizational threats. Survey-based findings on communication practitioners indicate that AI-related threats can arise from external malicious use and from internal organizational use of AI that publics interpret as ethically questionable (Zhao et al., 2025). This distinction changes the theoretical map. AI is a response tool, an external source of misinformation and disinformation, and part of the organization's own risk surface. A chatbot response, an automated statement, an undisclosed synthetic spokesperson, an AI-generated image, or an algorithmic personalization decision can become the crisis trigger. Pre-internet theory separated the organization from the media environment. Generative AI blurs that separation because the organization's communication infrastructure can become a contested object.

Research on AI-mediated crisis communication further complicates the response problem. Experimental findings on ChatGPT-mediated crisis communication report that users may value AI crisis chatbots in scenarios requiring immediate and instructive information, while human agents retain value where empathy, responsibility, and relational judgment are needed (Xiao & Yu, 2025). This result rejects a simple replacement thesis. AI can accelerate response, but speed does not settle trust. Crisis communication theory, built around spokesperson credibility, now has to include nonhuman touchpoints. The question becomes who or what speaks, under which disclosure rules, with what emotional adequacy, and through which escalation path to human responsibility.

The broader communication studies on generative AI support the same conclusion. Researchers argue that generative AI changes media industries because practitioners in journalism, advertising, and public relations use the same technologies for content production, audience engagement, and operational support. These uses bring shared concerns about ethics, transparency, bias, accuracy, and the human-machine relationship (Guzman & Lewis, 2024). For crisis communication, the old division between source, channel, and audience becomes harder to maintain. AI can operate as a content producer, a monitoring tool, an audience interface, a rumor generator, and a corrective assistant. Crisis communication theory, therefore, needs a broader unit of analysis: the crisis communication system, not only the crisis message.

The literature on AI in crisis communication as a field confirms that this theoretical expansion has begun but remains unfinished. A systematic review of AI applications in crisis communication reports growing research interest in how AI supports information processing, response, decision-making, and risk detection, while researchers call for stronger theoretical foundations, ethical governance, and empirical testing (Cheng et al., 2025). The review diagnoses uneven development. Practitioners move toward AI-supported crisis functions at speed, while theory remains split across technical, behavioral, ethical, and managerial lines. Figure 1 presents the condensed trajectory of crisis communication theory across three technological stages.

Figure 1. Condensed evolution of crisis communication theory (compiled by the author based on Jong, 2025; Cheng et al., 2025; Karinshak & Jin, 2023; Guzman & Lewis, 2024)



The figure clarifies why the generative AI stage should not be treated as a faster version of social media crisis communication. Social media accelerated public interpretation. Generative AI alters the evidentiary basis of interpretation. A false claim can now arrive with synthetic visual proof, cloned voice, fabricated leaked memo, automated comment consensus, or AI-generated search-like summaries. The object of crisis becomes unstable before response begins. Under these conditions, crisis managers cannot begin only with the question "Which response strategy fits this crisis type?" They first need to secure a credible reality layer before any response strategy will be judged as believable.

A comparison of several sources reveals the same movement from response to infrastructure. Mehta et al. (2021) show that crisis misinformation management depends on detecting, vetting, deciding, responding, and evaluating. Lu and Jin (2024) show that the correction must match the disinformation form and timing. Karinshak and Jin (2023) move further by framing AI-driven disinformation as a preparedness problem. Cheng et al. (2025) add that AI in crisis communication requires integration of theory, ethics, and governance. Taken together, these works support a layered conclusion for the first and second objectives. The internet forced crisis communication to add speed and monitoring. Generative AI forces it to add authenticity systems, AI governance, and pre-crisis legitimacy reserves.

A second comparison concerns the human-machine boundary. Xiao and Yu (2025) indicate that AI-mediated responses may satisfy stakeholders when immediacy and instructing information dominate. Zhao et al. (2025) show that AI-triggered threats create ethical and organizational dilemmas for practitioners. Guzman and Lewis (2024) argue that generative AI changes the shared professional environment of public relations, journalism, and advertising. The combined implication is that AI cannot be understood only as a channel or tool. In a crisis, it may act as a responder, accelerant, evidence fabricator, decision support layer, or reputational liability. Theoretical revision

has to place accountability over the whole communication chain, from data and prompts to approval, disclosure, monitoring, correction, and post-crisis learning.

A third comparison concerns trust. Kim & Lim (2026) focus on bandwagon cues and reputational effects. Lu and Jin (2024) focus on correction timing and narrative detail. Liu and Zhang (2025) examine corporate readiness through coordinated actors such as the corporation, employees, and social media influencers in combating crisis misinformation. These works point to a shared problem: trust repair depends on more than an official statement. It depends on whether the organization can coordinate several credible sources before a misleading interpretation gains social reinforcement. In the AI stage, such reinforcement may be artificial, semi-artificial, or human-machine mixed. Audiences may not know whether visible consensus reflects public judgment or coordinated simulation. Crisis communication theory has to treat consensus signals themselves as objects requiring verification.

The resulting finding is that crisis communication has moved through three theoretical orders. The first order was response-centered and event-centered. The second was platform-adaptive and correction-centered. The third has to be authenticity-centered and system-centered. The difference changes crisis work at the level of sequence. In pre-internet theory, the organization defended its reputation after a crisis event. In digital theory, it fought interpretation in fast-moving networks. In the AI era, it has to defend the conditions under which publics decide whether an event, source, image, voice, transcript, chatbot, or correction deserves trust.

Discussion

The reviewed research points to a shift in crisis communication theory away from response-centered models. Response remains necessary, but synthetic and machine-mediated crises make it too narrow as a starting point. A revised model has to begin before the incident, operate during the first minutes of narrative formation, and continue after correction because disproving a false claim does not always repair damaged trust.

The implementation model proposed here consists of four linked functions. The first is anticipatory sensing. Organizations need monitoring systems that track weak signals, suspicious narrative formation, synthetic media indicators, and unusual amplification patterns. This function differs from ordinary social listening. It requires authentication workflows, source mapping, escalation rules, and documented thresholds for intervention. The purpose is to reduce the interval between first appearance and institutional awareness.

The second function is authenticity verification. A crisis communication unit has to maintain a verified archive of official statements, executive voice samples under protected control, media asset provenance, domain authentication, crisis landing pages, and rapid evidence validation channels. The practical aim is to create a trusted reference layer before fabricated messages spread. Without such a layer, every correction starts from a weaker position because audiences compare one claim with another claim.

The third function is human-machine response governance. AI tools may draft, classify, translate, summarize, or assist in triage, yet the organization cannot delegate responsibility for crisis meaning to the system. Organizations need rules specifying which AI outputs require human approval, when chatbot response is allowed, when a human spokesperson must intervene, and how the organization discloses AI involvement. The fastest response is not always the safest response. In moral, safety-related, and high-blame crises, audiences need signs of accountable human judgment.

The fourth function is trust repair after correction. Refutation settles only part of the problem. A synthetic accusation, manipulated image, or AI-generated rumor can leave doubt after technical disproval. Trust repair requires repeated explanation, stakeholder dialogue, employee alignment, influencer or expert support where appropriate, and public education about verification procedures. The organization has to show how it knows what it claims to know. Table 1 compares the revision logic across three periods. Its purpose is to show which assumptions can be retained and which require replacement under generative AI conditions.

Table 1. Theoretical revision of crisis communication principles across three technological periods (compiled by the author based on Jong, 2025; Mehta et al., 2021; Karinshak and Jin, 2023; Cheng et al., 2025)

Classical principle	Pre-internet meaning	Internet adaptation	Revision for generative AI
Responsibility assessment	Determine whether the organization caused, suffered, or failed to prevent the crisis	Account for visible public blame and platform-based interpretation	Assess factual responsibility, perceived responsibility, synthetic attribution, and AI-system responsibility separately
Timely response	Issue a statement within the expected media cycle	Respond before rumors dominate social platforms	Establish pre-authorized micro-response protocols for the first minutes of synthetic crisis formation
Message consistency	Maintain a unified official line	Coordinate statements across websites, social media, and spokespersons	Maintain a verified reality layer with provenance, authentication, and version control
Corrective communication	Refute inaccurate claims after verification	Use platform-specific corrections and narrative rebuttal	Combine correction with prebunking, source authentication, and synthetic evidence explanation
Stakeholder trust	Restore confidence through apology, explanation, or compensation	Engage publics through dialogue and transparency	Protect trust through human accountability, AI disclosure rules, and durable verification infrastructure
Evaluation	Assess reputational recovery after the crisis	Track reach, engagement, sentiment, and correction spread	Track velocity gap, authenticity doubts, AI disclosure effects, amplification anomalies, and residual distrust

The table indicates that older principles retain value but need more precise operational meaning. Responsibility, timeliness, consistency, correction, trust, and evaluation still matter. Each term now refers to a different task set. A crisis team that treats an AI-generated accusation as an ordinary rumor will lose time. A team that treats every synthetic incident as a technical problem will miss the moral and emotional layer. The revised approach has to join technical verification with communicative legitimacy.

A practical decision logic follows from this revision. When a crisis signal appears, the first question should concern the kind of reality claim being made. If the disputed object is a synthetic video, cloned voice, fabricated document, or AI-generated screenshot, the response sequence begins with authentication. If the issue arises from the organization's own AI use, the sequence begins with responsibility clarification and disclosure. If the crisis spreads through apparent public consensus, the sequence begins with amplification, diagnosis, and stakeholder mapping. If the threat involves safety, financial risk, or public health, human spokesperson escalation has to override automated interaction.

This decision logic reduces the risk of two common errors. Reactive literalism appears when the organization answers the surface claim but ignores the mechanism that gives it force. Technological overcorrection appears when

the organization treats detection as communication. Detection establishes what happened. It does not restore trust by itself. Table 2 translates the revised theory into an applied governance model for organizations facing AI-era crises. It compares the decision layer, operational task, failure risk, and monitoring signal.

Table 2. Implementation model for crisis communication governance in the age of generative AI (compiled by the author based on Xiao and Yu, 2025; Zhao et al., 2025; Lu and Jin, 2024; Liu and Zhang, 2025; Kim & Lim, 2026; Guzman and Lewis, 2024)

Governance layer	Operational task	Decision rule	Failure risk	Monitoring signal
Signal detection	Identify emerging claims, synthetic media, abnormal amplification, and stakeholder concern	Escalate when a claim combines reputational harm with rapid spread or evidentiary imitation	Late recognition creates narrative lock-in	Time from first appearance to internal alert
Authenticity verification	Check media provenance, document origin, source legitimacy, and technical manipulation indicators	Publish only verified claims, but acknowledge uncertainty when delay would deepen harm	Silence is interpreted as evasion	Share of disputed materials with verified status
AI response use	Deploy AI for drafting, translation, classification, and triage under human supervision	Permit AI support for speed, reserve moral judgment and apology for accountable humans	Automated tone appears evasive or inhuman	Ratio of AI-assisted outputs reviewed by humans
Disclosure and accountability	Explain whether AI was used in organizational communication or implicated in the crisis	Disclose when AI involvement affects trust, evidence, authorship, or stakeholder rights	Non-disclosure becomes a secondary crisis	Disclosure clarity and stakeholder follow-up questions
Corrective strategy	Match correction to misinformation format, timing, and narrative strength	Use prebunking where risk is predictable, use detailed correction where false belief has formed	Dry rebuttal loses to vivid synthetic narrative	Correction uptake and residual belief indicators
Trust repair	Coordinate employees, experts, credible partners, and direct stakeholder channels	Move from denial to sustained explanation when doubt persists	The crisis formally ends while distrust remains	Recurrent questioning after correction

The table shows why AI-era crisis communication cannot stay inside the communications department. Legal teams, cybersecurity staff, data governance specialists, executive leadership, HR, customer support, and platform relations all become part of the crisis architecture. Public relations retains a central coordinating function, but it cannot verify synthetic evidence, govern AI disclosure, and control automated response tools without cross-functional authority.

Implementation should follow a staged sequence. First, organizations should create an AI crisis inventory. It records where AI is used in external communication, internal decision support, customer interaction, content production, monitoring, and campaign automation. Second, each AI use should receive a crisis risk category: low reputational sensitivity, moderate public sensitivity, or high trust sensitivity. Third, the organization should prepare pre-approved response templates that do not sound mechanical, including uncertainty statements, authentication updates, and escalation notices. Fourth, a verified source hub should exist before any crisis occurs. It should contain official statements, media verification notes, executive authentication channels, and updates in chronological order. Fifth, crisis drills should simulate synthetic accusations, cloned audio, AI-generated employee leaks, chatbot errors, and coordinated disinformation. Sixth, the post-crisis review should assess the velocity gap, decision delay, correction effectiveness, disclosure clarity, and trust residue.

Metrics deserve narrower use here than in PR measurement studies. The focus is not general campaign effectiveness. Crisis communication in the AI age needs operational indicators tied to threat containment and trust preservation. Useful indicators include time to detection, time to first holding statement, time to verification, percentage of claims classified by authenticity status, correction penetration across channels, recurrence of the same false claim, stakeholder confusion after disclosure, human escalation rate from AI-mediated touchpoints, and unresolved trust questions after the formal end of the crisis. These indicators serve governance, not promotional reporting.

The theoretical revision proposed in this article has a boundary. It does not replace established crisis response theories. It relocates them inside a wider system. Responsibility attribution still guides message choice. Corrective communication still matters. Stakeholder trust remains the central outcome. The difference lies in sequence and infrastructure. Inherited theory begins too late when a crisis is synthetic, automated, or AI-mediated. A contemporary model has to begin with the protection of evidentiary conditions.

A second boundary concerns deepfakes. They are relevant, yet they should not dominate the entire article. Deepfakes represent one visible form of synthetic crisis evidence. Generative AI creates wider problems: text fabrication, automated consensus, conversational hallucination, falsified screenshots, fake expert commentary, cloned documents, and AI-enabled impersonation. Centering the article only on deepfakes would narrow the theoretical problem. The deeper issue is cognitive and institutional: publics have to judge reality in an environment where the signs traditionally used to verify it can be simulated.

The proposed model gives practitioners a realistic path. It does not ask organizations to predict every attack. It asks them to reduce response latency, define AI accountability, build authentication channels, and sustain trust repair after correction. The communication unit becomes less like a message factory and more like a steward of organizational credibility under conditions of synthetic uncertainty.

Conclusion

Crisis communication theory has moved from event-centered response models to platform-adaptive correction models and now faces a new stage shaped by generative artificial intelligence. The earlier theoretical base remains valuable for responsibility assessment, stakeholder interpretation, and reputational repair, yet it no longer covers the full formation of contemporary crises. Inherited principles were designed for situations in which the event, source, and evidence could be identified with relative stability. Generative AI weakens that premise by enabling fabricated evidence, automated amplification, synthetic spokespersons, and machine-mediated organizational responses.

The transition from the internet stage to the AI stage changes the central problem of crisis communication. The main difficulty is no longer speed alone. The deeper problem lies in the loss of a stable evidentiary ground on which the public judges whether a crisis is real, fabricated,

exaggerated, or ethically linked to organizational AI use. Correction, apology, denial, and transparency remain necessary, but they work only when supported by authentication, governance, and credible human accountability.

The hypothesis of the article is confirmed. The theoretical core of crisis communication remains useful only after revision from a response-centered model into an anticipatory and authenticity-centered model. The literature reviewed in the article supports three conclusions: crisis responsibility has become dynamic, misinformation response has become narrative and platform-dependent, and AI has become both a response tool and a source of reputational threat. The practical implication is a need for crisis communication systems that combine early sensing, authenticity verification, human-machine response rules, disclosure standards, coordinated correction, and long-term trust repair.

References

1. Cheng, Y., Li, J., Lee, J., & Liu, Y. (2025). How is artificial intelligence shaping crisis communication? A systematic review and future research agenda. *Journal of Risk Research, 1*–21. https://doi.org/10.1080/13669877.2025.2579317

2. Guzman, A. L., & Lewis, S. C. (2024). What generative AI means for the media industries, and why it matters to study the collective consequences for advertising, journalism, and public relations. *Emerging Media, 2*(3), 347–355. https://doi.org/10.1177/27523543241289239

3. Jong, W. (2025). Beyond the snapshot: Rethinking crisis communication theories in dynamic crisis situations. *Public Relations Review, 51*(3), Article 102586. https://doi.org/10.1016/j.pubrev.2025.102586

4. Karinshak, E., & Jin, Y. (2023). AI-driven disinformation: A framework for organizational preparation and response. *Journal of Communication Management, 27*(4), 539–562. https://doi.org/10.1108/JCOM-09-2022-0113

5. Kim, Y., & Lim, H. (Dana). (2026). Alleviating the bandwagon effect of crisis misinformation on social media: Understanding social media users' bandwagon perceptions and the credibility of crisis misinformation to protect organizational reputation. *Communication Studies, 77*(1), 22–50. https://doi.org/10.1080/10510974.2025.2485368

6. Liu, J., & Zhang, L. (2025). Strengthening corporations' crisis readiness: The roles of corrective communication, employee backup and social media influencer in debunking crisis misinformation. *Journal of Communication Management*. Advance online publication. https://doi.org/10.1108/JCOM-12-2024-0283

7. Lu, X., & Jin, Y. (2024). Optimizing organizational corrective communication: The effects of correction placement timing, refutation detail level, and corrective narrative type on combating crisis misinformation narratives. *Public Relations Review, 50*(5), Article 102514. https://doi.org/10.1016/j.pubrev.2024.102514

8. Mehta, A. M., Liu, B. F., Tyquin, E., & Tam, L. (2021). A process view of crisis misinformation: How public relations professionals detect, manage, and evaluate crisis misinformation. *Public Relations Review, 47*(2), Article 102040. https://doi.org/10.1016/j.pubrev.2021.102040

9. Xiao, Y., & Yu, S. (2025). Can ChatGPT replace humans in crisis communication? The effects of AI-mediated crisis communication on stakeholder satisfaction and responsibility attribution. *International Journal of Information Management, 80*, Article 102835. https://doi.org/10.1016/j.ijinfomgt.2024.102835

10. Zhao, W., Rachwalski, A., Berndt-Goke, M., & Jin, Y. (2025). An examination of management of AI-triggered organisational threats from communication practitioners' perspective. *Journal of Contingencies and Crisis Management, 33*(1), Article e70031. https://doi.org/10.1111/1468-5973.70031