



Responsible AI-Enabled Corporate Communication for Managing Digital Misinformation and Stakeholder Trust in UK Organisations

A Critical Literature Review and Philosophical Analysis

Assignment Title: Report

Course: Doctor of Business Administration (DBA)

Module: LCDB7007 – Research Principles and Philosophy

Lecturer: Dr Ebad Banissi

Student ID: 2536867

Abstract

Purpose: This report examines how Responsible AI can help UK organisations manage digital misinformation through corporate communication while maintaining stakeholder trust.

Design/Methodology/Approach: The report uses a structured critical literature review to bring together research on strategic communication, misinformation and crisis response, stakeholder trust and legitimacy, and Responsible AI governance. It also includes a philosophical and methodological review of Buhmann et al.'s (2026) qualitative study of AI adoption in corporate communication departments.

Findings: The review shows that generative AI can support monitoring, evidence searches and the preparation of responses, but it cannot protect reputation or trust by itself. Its value depends on sound governance, human judgement, careful verification, transparent communication, and clear decision-making responsibility. The evidence is still fragmented. Much of it comes from experiments, practitioner accounts or settings outside the UK, and few studies examine how communication, data, governance, and leadership teams work together during fast-moving misinformation incidents.

Originality/Value: The report proposes an initial framework that treats Responsible AI as a strategic communication capability. It connects governance and ethical principles with AI-supported monitoring and verification, accountable approval of responses, transparent stakeholder engagement and learning after an incident. Together, these elements can support trust, reputation, and organisational resilience. Future DBA research is proposed using a critical-realist, qualitative multiple-case study across UK sectors to explore how and when this capability works.

Keywords: Responsible AI; corporate communication; digital misinformation; stakeholder trust; AI governance; UK organisations.

Table of Contents

Abstract.....	2
1. Introduction.....	4
1.1 Background.....	4
1.2 Business Problem	4
1.3 Research Gap.....	5
1.4 Aim.....	5
Research Question.....	5
Objectives	5
2. Method for Conducting the Literature Review	6
3. Critical Literature Review	8
3.1 Responsible AI in Corporate Communication	8
3.2 Digital Misinformation and Organisational Reputation.....	9
3.3 Stakeholder Trust and Organisational Legitimacy	12
3.4 Responsible AI Governance in Communication	13
3.5 Research Gaps	15
4. Philosophical and Methodological Analysis of One Selected Study.....	17
4.1 Overview of the Selected Study	17
4.2 Ontology	17
4.3 Epistemology.....	18
4.4 Axiology and Researcher Positionality.....	18
4.5 Reliability, Validity, Trustworthiness and Bias.....	19
4.6 Ethical Considerations	20
4.7 Critical Appraisal and Alternative Philosophical Approach	20
5. Discussion and Contribution.....	22
5.1 Integrated Discussion.....	22
5.2 Philosophy and Research Practice	22
5.3 Practical and Theoretical Contribution	23
6. Conclusion	25
References	26

1. Introduction

1.1 Background

Corporate communication is part of organisational strategy. It helps organisations explain decisions, protect reputation and maintain relationships with employees, customers, investors, regulators and communities. The digital environment is harder to manage. Text, images, audio and video can be created and shared quickly through social media, platform amplification and generative artificial intelligence (AI). Broda and Strömbäck (2024) show that misinformation, disinformation and fake news differ by falsity, intention and format. For organisations, a harmful claim may be deliberate deception, an honest mistake or manipulated synthetic content. People may then share it without knowing that it is inaccurate.

The UK shows this uncertainty. Four in ten UK adults surveyed by Ofcom (2024b) had seen content they thought was misinformation or a deepfake in the previous four weeks. A separate Ofcom study found that 43 per cent of people aged 16 and over had seen at least one deepfake in the previous six months, but only one in ten felt confident that they could identify one (Ofcom, 2024a). These findings do not show that most misinformation concerns companies. They do show that UK organisations operate in an environment where false claims, impersonation and manipulated media can challenge official information.

AI is also changing internal corporate communication. Yue et al. (2024b) interviewed 20 public-relations leaders with extensive AI experience. They used AI for research, data analysis, content creation and other communication functions, but they raised concerns about accuracy, ethics and professional responsibility. Bowen (2024) identifies two roles for communication professionals: using AI responsibly and helping to oversee organisational AI systems that may later require public communication. Responsible AI is therefore a communication and governance issue as well as a technical issue.

1.2 Business Problem

The business problem has two sides. AI can create misinformation, fake social media campaigns, brand impersonation, deepfakes and fabricated claims against UK organisations. Organisations can also use AI to monitor stakeholders, analyse data and create corporate communication. This creates two risks: a delayed response to a false story and a rapid response based on AI-generated content that has not been checked properly. Both can damage credibility. Stakeholder trust depends on speed, accuracy, transparency, accountability and consistency with stakeholder interests.

This is an issue of organisational capability. Early detection, verification and preparation are more important with generative AI, yet responses to misinformation often remain reactive and crisis driven. Responsible AI principles such as fairness, transparency and accountability are also easier to state than to apply. Papagiannidis, Mikalef and Conboy (2025) find that Responsible AI research is fragmented and that organisations need structural, relational and procedural practices. Corporate communication therefore needs a stronger link between AI governance and misinformation management.

1.3 Research Gap

Research on misinformation, AI in public relations, stakeholder trust and Responsible AI governance remains fragmented. Misinformation research often examines production, dissemination and correction. Communication research is starting to examine professional use of AI, while Responsible AI research focuses on governance, accountability and ethical risk. Few studies combine these areas in one organisational communication framework, especially in the UK. The main gap is the weak link between technological capability, communication strategy, stakeholder trust and governance. This report addresses that gap through a critical literature review and philosophical analysis.

1.4 Aim

This report critically examines how Responsible AI can improve corporate communication strategies for combating digital misinformation without compromising stakeholder trust.

Research Question

What role can Responsible AI in corporate communication play in enhancing organisational capability to manage digital misinformation and maintain stakeholder trust in UK organisations?

Objectives

1. Critically analyse research on AI-supported corporate communication.
2. Discuss ways of combating digital misinformation.
3. Analyse the relationship between Responsible AI and stakeholder trust.
4. Critically assess the philosophical underpinnings of research.
5. Develop a conceptual model of Responsible AI in corporate communication.

2. Method for Conducting the Literature Review

I used a structured literature search to find recent research on Responsible AI, corporate communication, digital misinformation, and stakeholder trust. I searched Scopus, Web of Science, Emerald Insight, ScienceDirect, Business Source Complete and Google Scholar. Most sources were published between 2020 and 2026. I kept older sources when they provided foundational work on Stakeholder Theory, organisational legitimacy or crisis communication. Search terms covered Responsible AI, AI governance, corporate and strategic communication, misinformation, deepfakes, stakeholder trust, corporate reputation and the United Kingdom. I focused on peer-reviewed business, management, public relations, organisational governance and stakeholder research. I included relevant international studies but prioritised UK evidence. I excluded technical AI research without an organisational communication or governance focus. The method followed management-review guidance on clear boundaries and defensible source selection (Kraus et al., 2022; Sauer and Seuring, 2023).

I screened titles and abstracts before reading the full text. For each selected source, I recorded the publication year, context, theoretical framework, method, findings, limitations, and relevance in a review matrix. I organised the evidence into four linked themes: Responsible AI in corporate communication; misinformation and organisational reputation; stakeholder trust and legitimacy; and Responsible AI governance. I used thematic synthesis and critical comparison rather than a study-by-study summary. I compared agreements, disagreements, conceptual and empirical evidence, methodological strengths and weaknesses, UK relevance, and ontological, epistemological, and axiological assumptions. This analysis supports the argument and conceptual framework presented later (Paul and Criado, 2020).

This was a critical review, not a systematic review or meta-analysis. It covered English-language content from selected databases, and rapid changes in generative AI may make some findings outdated. I did not use a statistical quality score. I used judgement to assess relevance, research quality, and transferability, compared evidence across methods and settings, and explained the basis for source selection and evaluation.

Table 1. Literature-review synthesis

Theme	Key evidence and synthesis	Implication for the report
Responsible AI in corporate communication	AI can support research, listening, content development and internal communication. Its strategic value depends on organisational capability, learning, professional judgement, disclosure, and authenticity (Yue et al., 2024b; Uysal and Deng, 2025; Buhmann et al., 2026).	Move beyond output efficiency: develop communication capability, role clarity, and accountable human oversight. Evidence is primarily practitioner-centred and non-UK.
Misinformation, correction, and reputation	Detection, verification, response selection and evaluation are interdependent. Corrective timing, message design, and prebunking influence outcomes (Mehta et al., 2021; Lu and Jin, 2024; Boman and Schneider, 2021).	Pre-establish monitoring and verification protocols to balance rapid response with evidential certainty. Experimental and non-UK evidence limits direct transferability.
Stakeholder trust and legitimacy	Trust is shaped by experienced human agency, transparent processes, procedural fairness and congruence between claims and organisational practice (Oh and Ki, 2024; Park and Yoon, 2024; Martin and Waldman, 2023).	Transparency alone is insufficient: stakeholders need credible accountability and meaningful routes to question or challenge decisions.
Responsible AI governance	Structural, relational and procedural mechanisms translate Responsible AI principles into organisational practice. Governance spans corporate, IT, data and communication functions (Mäntymäki et al., 2022; Papagiannidis et al., 2025; Hadley et al., 2025).	Specify decision rights, data controls, audit processes, escalation routes, and post-publication correction. Implementation evidence is still sector specific.
Integrative UK research gap	The literatures rarely examine AI adoption, misinformation, governance, and stakeholder outcomes together in a UK organisational setting (Broda and Strömbäck, 2024; Horneber, 2025; Park and Nan, 2026).	A critical-realist, cross-sector multiple-case study should test the proposed mechanisms and include employees, leaders, and external stakeholders.

Source: Author's synthesis of the reviewed literature.

3. Critical Literature Review

3.1 Responsible AI in Corporate Communication

The role of AI in corporate communication starts with strategic communication. Zerfass et al. (2018) define strategic communication as purposeful communication about issues linked to organisational objectives. Cornelissen (2023) describes corporate communication as a management activity that manages internal and external stakeholder relationships. AI should therefore be assessed by whether it improves organisational listening, professional judgement, relationship management and communication strategy, not only efficiency.

AI use is growing, but good integration is still unclear. Interviews with 20 public-relations leaders found that AI supported research, data analysis, content development and workflow, while raising concerns about accuracy, ethics and job loss (Yue et al., 2024b). Uysal and Deng (2025) show that organisational dynamics and capability shape its value. Laskin and D'Agostino's (2024) Delphi study in investor relations identifies efficiency, insight and investor communication as benefits, but misinformation, bias and disclosure as risks. Adoption is moving faster than shared professional standards. These expert methods give useful access to professional judgement across communication settings. However, much of the evidence relies on practitioner perceptions and self-reported benefits rather than observed behaviour or long-term organisational outcomes.

Yue et al. (2024a) report that AI can improve internal information flow, listening, efficiency and employee experience. Their participants also raised concerns about AI knowledge, job security, bias and authenticity. AI may help teams process more information, but automated messages may weaken credibility if employees find them impersonal or unclear about accountability. The social setting of implementation therefore matters as much as technical output quality. The professional sample gives organisational insight, but it does not directly test employee trust or long-term effects.

Research on organisation-public relationships shows mixed effects. Oh and Ki (2024) found stronger control mutuality when a human, rather than an AI agent, ran an organisation's social media account. A human conversational voice also improved the relationship. In contrast, Jiang, Cheng and Wang (2025) found that socially oriented AI chatbots could improve perceived listening, shared meaning and supportive intentions among 606 Chinese consumers in corporate social-responsibility discussions. These findings are conditional rather than contradictory. AI can support dialogue when it feels responsive and socially meaningful, but automation may reduce relationship quality. Cultural and experimental limits also restrict generalisation to UK organisations and high-risk misinformation incidents.

Implementation research supports a social and technical approach rather than a technology-first approach. In a Danish municipal communication department, Christensen and Andersson (2025) found that a year-long learning process, employee involvement and slow implementation supported more sustainable use of generative AI. This also supports Uysal and Deng's (2025) finding that organisational dynamics and evolving capabilities shape adoption. Communication

capability develops through learning, clear roles and participation. Giving staff access to generative AI does not create strategic competence. The case provides useful depth, but its single public-sector setting limits direct transferability to UK corporations.

Responsible AI also affects professional effectiveness and ethical responsibility. Bowen's (2024) longitudinal mixed-method study gives communication professionals two roles: communicating ethically about AI and advising on AI systems that may later need public explanation or defence. Germinder and Capizzo (2025), based on 21 practitioner interviews, add standards for access, process, truth and disclosure. Dong and van den Berg (2025) identify mis- and disinformation, authorship, accountability and disclosure as common ethical concerns. They also question whether current professional codes provide enough guidance for generative AI. Together, the studies show a gap between broad ethical duties and specific decisions about tools, evidence and approval.

Overall, Responsible AI should be treated as a strategic communication capability, not only as a content tool. It can support analysis, listening and stakeholder interaction, but it also needs organisational learning, human judgement, authenticity, disclosure and ethical accountability. Most evidence comes from expert interviews and short-term audience research outside the UK. It remains unclear how AI-supported communication performs when speed, fact-checking, reputation and stakeholder trust are all at risk.

3.2 Digital Misinformation and Organisational Reputation

Digital misinformation threatens organisational reputation because many sources, not only official communication, shape stakeholder views. Wardle and Derakhshan's (2017) Information Disorder framework separates misinformation, disinformation and malinformation by falsity, intention and harm. This helps distinguish an honest mistake from a deliberate reputation attack. The distinction matters because monitoring, correction and crisis response may differ by intent and harm. Synthetic media makes this task harder. Corsi, Marino and Wong (2024) found a sharp rise in AI-generated media on X (formerly Twitter) after better image tools became available, but most items were not malicious. Organisations should not treat all AI-generated content as disinformation. They should assess intent, reach, credibility and possible harm.

Corporate Reputation Theory adds further complexity. Bigus, Hua and Raithel's (2024) review of 173 accounting and management studies shows that corporate reputation includes different stakeholder views of competence, integrity, emotional appeal and prominence. Misinformation can therefore harm different parts of reputation for different groups. Hollebeek et al. (2024) also describe consumers as active agents who may share, question or correct brand misinformation. The focus should extend from the false message to stakeholder relationships and the organisation's response.

Situational Crisis Communication Theory (SCCT) links crisis perceptions to response choices. Coombs (2022) argues that organisations should assess attributed responsibility and reputational threat before choosing a defensive, diminishing or rebuilding response. Misinformation challenges SCCT because an organisation may suffer without causing the event. Jahng's (2021)

experiment found that fake news designed to harm a brand was seen as a more serious crisis and produced stronger expectations for an accommodating response. Stakeholders may therefore judge a crisis as serious when they believe the organisation has acted wrongly, even if it is not responsible.

Digital-crisis research broadly supports SCCT but shows its limits. Zhang et al. (2024) analysed 770 corporate crisis posts and 931,668 public comments on Sina Weibo. Responses that broadly matched SCCT were linked to more positive public sentiment and supportive behaviour. Humanised language and executive involvement also mattered. The large real-world dataset is a strength, although an observational association cannot prove that the response alone caused the outcome. The Chinese platform and culture also limit direct use in the UK. SCCT assumes that the organisation can identify the crisis type quickly, which is difficult while misinformation evidence is still being checked.

Corrective communication research shows that message quality matters after an organisation decides to respond. Jin et al. (2020) found that detailed factual explanations and employee support improved message quality, organisational reputation and perceived crisis responsibility. Lu and Jin (2024) show that correction timing, rebuttal detail and narrative form also matter. Mehta et al. (2021) found that verification can delay a response, correction is not always the best choice, and evaluation is often informal. These studies show a conflict between speed and certainty. A fast response may limit a false claim but create more harm if the organisation later has to correct it.

Proactive approaches aim to reduce harm before misinformation spreads. Boman and Schneider (2021) found that prebunking reduced reputational damage after an astroturf attack. Boman (2021) found that clear detail and language that supports audience choice can work against public relations disinformation. This shifts practice from reactive crisis communication towards preparation and inoculation. However, both United States studies used controlled experiments and measured immediate attitudes. They do not show whether protection lasts or works during a real organisational crisis. Prebunking is also difficult when the future attack is unknown, and too many warnings may increase suspicion.

Amazeen's (2024) Misinformation Recognition and Response Model is shown in Figure 1. It links personal and message factors to problem identification, motivation, misinformation awareness, coping, and cognitive, emotional and behavioural responses.

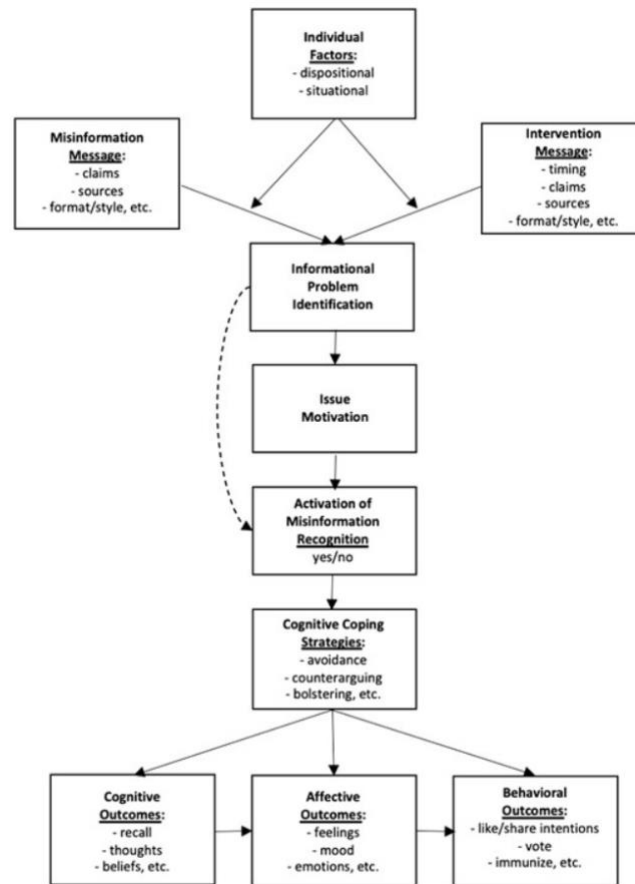


Figure 1. Conceptual process of the Misinformation Recognition and Response Model.

Source: Amazeen (2024, Figure 1). Reproduced under CC BY 4.0.

Reactive correction and proactive inoculation are both useful, but neither is enough. Organisations need continuous monitoring, an agreed verification process, clear response criteria, proactive communication about predictable risks and correction when false stories appear. SCCT supports response selection, the Information Disorder framework explains the threat, and reputation research shows how stakeholders may be affected in different ways. An integrated model is still needed to explain how AI-supported detection and monitoring can aid these decisions without weakening evidence quality, ethical accountability or stakeholder trust in UK organisations.

3.3 Stakeholder Trust and Organisational Legitimacy

Stakeholder Theory shows why AI-powered communication should be examined through relationships, not only technology. Freeman (1984) argues that organisations create and maintain

value through interactions with stakeholder groups. Mahajan et al. (2023) connect stakeholder management with strategy, sustainability and organisational performance. Customers, employees, investors, regulators and communities may have different expectations about accuracy, privacy, transparency and human involvement. They may also hold different levels of power and face different risks. A practice that works for one group may not work for another, so managers must explain whose interests shape the decision.

Recent studies show both positive and negative effects on organisation-public relationships. Men et al. (2023) surveyed 778 Facebook users who engaged with AI-powered corporate chatbots. A social and conversational approach improved views of corporate character and relationship quality. However, Oh and Ki (2024) found stronger control mutuality when a human managed a social media account, although a conversational human voice improved the relationship. Trust therefore depends on agency, responsiveness and control, not simply the presence of AI. These mediated interactions may not apply directly to serious misinformation incidents.

Transparency and explainability may protect trust, but their effects depend on context. Park and Yoon (2024) found that algorithm transparency signalling increased trust in an AI system and its parent organisation. Their later study also found that transparency could reduce negative attitudes and support trust (Park and Yoon, 2025). However, Martin and Waldman (2023) found that notices and impact statements did not automatically make algorithmic decisions legitimate. Appeal processes mattered more when decisions caused harm, and questionable data use led to strong legitimacy penalties. The contrast shows that low-risk disclosure and harmful decisions create different stakeholder needs. Transparency is most credible when supported by procedural protections, not disclosure alone.

Legitimacy Theory helps explain these procedures. Suchman (1995) separates pragmatic legitimacy, based on stakeholder self-interest; moral legitimacy, based on approval of what is right; and cognitive legitimacy, based on understanding and acceptance of organisational action. AI-powered corporate communication can affect all three. Stakeholders may accept AI when it improves service, but not when it seems deceptive, unfair or unclear. Hillo, Vento and Erkkilä (2025) studied 842 public administrators and 3,245 citizens in Finland and found that algorithmic transparency and human discretion improved perceived legitimacy. Wang et al. (2026), using survey experiments in China and the United States, found that disclosing AI use could reduce legitimacy and that explanations were not always effective. Transparency is therefore shaped by expectations and institutional setting; it is not a universal route to trust.

Institutional Theory provides an organisational explanation. DiMaggio and Powell (1983) argue that organisations respond to coercive, normative and imitative pressures in their institutional fields. These may include regulation, professional standards, competitors and stakeholder expectations for AI. Horneber (2025) uses a neo-institutional lens to show that Responsible AI policies can be separated from practice or fail to meet their goals. This policy-practice separation explains why formal adoption may be symbolic rather than operational. It is harmful in corporate communication because an organisation may promise transparency and accountability while employees use AI in unclear or inconsistent ways.

Trust is also shaped by competing ways of thinking. In a nested case study of a European airline, Giannitsas, Sun and Baptista (2026) identify tension between instrumental-analytic logic, which values automation, efficiency and standardisation, and contextual-normative logic, which values human judgement and ethical responsiveness. Staff managed this tension through trust and distrust in AI, careful evaluation and active safeguards. The case gives detailed organisational evidence, but one airline cannot represent all sectors. Corporate communication still faces the same conflict between automation, authenticity, contextual judgement and accountability.

Overall, stakeholder trust and organisational legitimacy do not come from technological transparency alone. They develop through relationships, fair procedures, credible human responsibility and consistency between statements and practice. Stakeholder Theory identifies relevant interests. Legitimacy Theory explains how behaviour is judged, while Institutional Theory shows how regulation and professional norms shape practice. Much evidence comes from experiments, public-sector settings and countries outside the UK. Evidence on AI-supported communication during UK misinformation incidents and its long-term effects across stakeholder groups remains limited.

3.4 Responsible AI Governance in Communication

Responsible AI governance concerns the gap between moral intent and organisational practice. Berente et al. (2021) identify three AI characteristics that require management: autonomy, learning and inscrutability. These features make outcomes harder to predict and explain, and responsibility harder to assign to a person. Governance is therefore a management issue as well as a technical issue. Jöhnk, Weißert and Wyrski (2021) show that AI implementation depends on organisational readiness, including resources, knowledge, strategic alignment and change management. An ethical AI policy works in corporate communication only when staff know when AI may be used, what evidence must be checked, who makes the final decision and how issues are escalated.

Mäntymäki et al. (2022) define organisational AI governance as rules, practices, processes and technical tools that align AI with strategy, law and ethical principles. They place AI governance within corporate and IT governance and show its overlap with data governance. This overlap matters because communication tools depend on data quality, access rights, lawful processing and technical controls. AI-supported monitoring and message generation also need clear corporate accountability. However, the model shows where governance sits more clearly than it shows which mechanisms produce responsible outcomes.

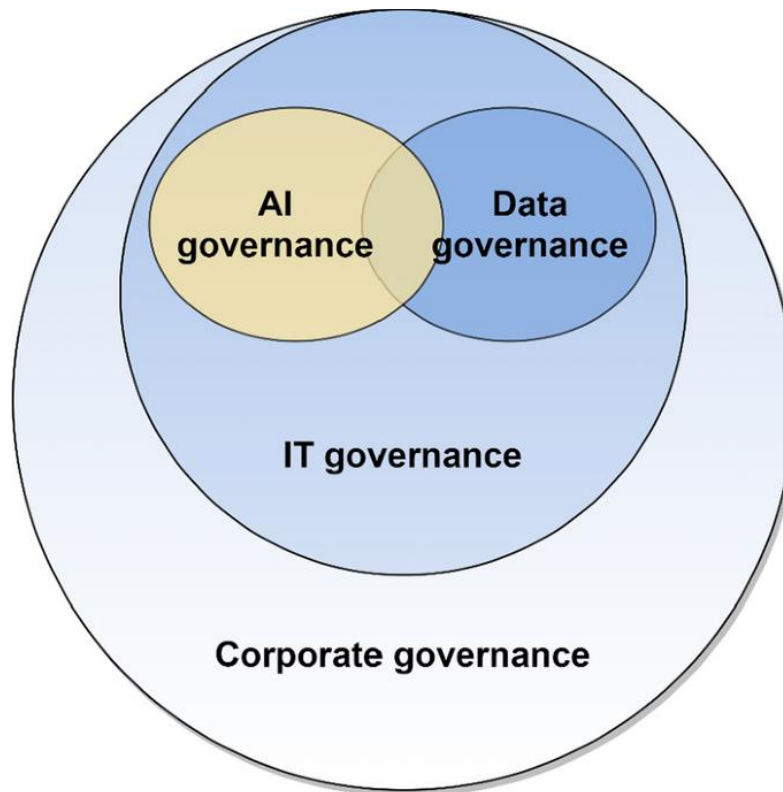


Figure 2. Artificial intelligence governance within an organisation's governance structure.

Source: Mäntymäki et al. (2022, Figure 1). Reproduced under CC BY 4.0.

Papagiannidis, Mikalef and Conboy (2025) provide a more practical approach by separating structural, relational and procedural mechanisms. Structural mechanisms allocate authority, relational mechanisms combine expertise and stakeholder engagement, and procedural mechanisms turn Responsible AI principles into repeatable action. Akbarighatar (2025) adds a capability perspective based on understandable models, action on bias, responsiveness, harm prevention and support for the common good. One framework focuses on governance structure and the other on organisational capabilities. Both require formal decision-making authority and professional skills.

Organisational evidence shows that formal structures alone do not ensure good implementation. Hadley, Blatecky and Comfort (2025) interviewed 17 technical contributors and found that

algorithm review boards worked best when linked to existing governance and supported by senior leaders. Financial and operational pressures still caused problems. Mökander and Floridi's (2023) longitudinal AstraZeneca case reaches a similar conclusion: ethics-based auditing needed shared standards, a clear scope, internal communication and change management, not only technical assessment. These studies provide rare evidence of governance in practice, but their sectors and participants are specific. One model may therefore not suit every communication function.

External frameworks add further expectations. The NIST AI Risk Management Framework organises AI risk activity into govern, map, measure and manage, and treats oversight as continuous (NIST, 2023). The UK principles-based approach sets cross-sector expectations for safety, transparency and explainability; fairness, accountability and governance; and contestability and redress (Department for Science, Innovation and Technology, 2023). The Information Commissioner's Office (2023) also requires lawfulness, fairness, transparency, accuracy, security, accountability and individual rights when personal data is used. NIST provides a risk-management process, UK policy provides regulatory principles, and data-protection guidance sets duties for personal information.

Corporate communication therefore needs more than one governance model. Board and senior-management responsibility should link with cross-functional AI oversight, communication approval processes, data controls, documented verification, staff skills and routes for stakeholder challenge. Governance should cover the AI lifecycle, from purchasing tools and using training data to checking and correcting published content. It should also assign owners for monitoring, authorisation, escalation and post-incident learning. The literature supports Responsible AI governance as part of strategic communication capability, not as a later compliance step. It remains unclear how these arrangements work during fast-moving misinformation incidents in UK organisations, where speed, legal duties and stakeholder trust compete.

3.5 Research Gaps

The four themes show a fast-growing field divided by theory and evidence. Broda and Strömbäck's (2024) systematic review finds that misinformation research is split across disciplines and does not always build consistently on earlier work. Park and Nan's (2026) scoping review also finds wide differences in methods, uses and outcome measures. Communication studies often focus on adoption or message effects, misinformation research on detection and correction, and Responsible AI research on governance principles and implementation. Few studies connect all three in one organisational explanation.

A second gap lies between frameworks and practice. Papagiannidis, Mikalef and Conboy (2025) find that organisations still lack clear ways to implement Responsible AI governance. Horneber (2025) also shows that policies may remain separate from daily work or fail to meet their goals. Communication research has the same weakness. Yue et al. (2024b) provide detailed practitioner

views on AI in public relations, but most newer evidence comes from interviews, surveys or short experiments rather than observation of real misinformation incidents.

A third gap concerns trust over time. Studies often measure immediate attitudes, perceived transparency, relationship satisfaction or organisational reputation. These measures do not show whether trust is maintained, repaired or slowly lost. Giannitsas, Sun and Baptista (2026) show that trust and distrust in AI are shaped by competing demands for automation, efficiency, contextual judgement and ethical responsiveness. Long-term research should examine how repeated AI-assisted communication decisions affect organisational reliability and legitimacy.

A fourth gap is the lack of UK-specific evidence. Few reviewed peer-reviewed empirical studies focus on UK organisations. Much communication research comes from the United States, China, Australia or continental Europe, while UK sources mainly cover regulation and policy. It is therefore unclear whether the findings fit the UK's media environment, data-protection rules, professional communication norms and media-sector expectations.

Finally, research rarely examines the cross-functional coordination needed for Responsible AI-supported communication. Studies often separate communication professionals, AI governance teams, data functions and senior leaders. Misinformation management requires these groups to coordinate monitoring, evidence checks, response authority, stakeholder engagement and evaluation. The main gap is a well-tested framework that integrates these elements as a strategic communication capability. This gap supports the Responsible AI-enabled corporate communication framework proposed later.

4. Philosophical and Methodological Analysis of One Selected Study

4.1 Overview of the Selected Study

The selected study is Buhmann et al. (2026), “‘We’re right in the foothills’: the adoption of artificial intelligence in corporate communication departments’. It examines AI in corporate communication rather than as a secondary use of a wider technology. The authors use sociotechnical systems theory and qualitative interviews with 30 chief communication officers (CCOs) from the 300 largest companies in Europe. The study offers senior decision-maker evidence from complex organisations and fits this report's focus on Responsible AI as a strategic communication capability.

The study examines AI across communication, organisation, processes and technical systems. Its findings challenge claims that AI is already transforming communication rapidly. The authors found little structural change in communication departments. Training and staff development changed most, while advanced uses remained experimental or limited to low-risk tasks. AI may improve daily efficiency, but it also raises problems in governance, technical integration and the role of people in creative and strategic work. The study therefore links AI adoption to organisational capability and governance rather than treating adoption as progress by itself.

4.2 Ontology

Buhmann et al. (2026) do not state their research philosophy, so their ontological stance must be inferred carefully from their theory and methods. Ontology concerns assumptions about what is real in the social world and which structures are real (Saunders, Lewis and Thornhill, 2023). The sociotechnical approach treats communication departments as systems in which technology, tasks, structures and people interact. These are organisational conditions that can support or restrict AI implementation, not only personal opinions. The classification offered here is therefore an informed interpretation, not a philosophical claim made by the authors.

This position fits a moderate realist, especially critical-realist, view of organisations. Critical realism separates what is observed from the underlying structure or mechanism that produces it (Fletcher, 2017). The study analyses executive accounts through technology, organisational structures and professional work. Its finding that AI often remains in ‘sandboxes’ or low-risk tasks assumes that deeper structures, including governance processes, roles, technical infrastructure and professional norms, shape what organisations can do. The interview accounts provide observations, while the sociotechnical analysis points towards mechanisms that may explain those observations.

This approach avoids technological determinism. AI is not presented as a tool that will automatically transform communication departments; it works only in certain contexts. However, the study relies mainly on CCOs, so implementation is seen from a senior-management perspective. A stronger critical-realist study would compare these accounts with organisational documents, AI policies, workflow records and direct observation. This would help separate

executive beliefs from mechanisms that operate in practice. It would also support triangulation and show whether formal policy matches daily work.

4.3 Epistemology

The study is mainly interpretive in epistemology, but it is also theory-driven. Knowledge comes from CCOs' experiences and explanations rather than controlled variables or statistical prediction. This suits AI use in corporate communication because it involves meaning, judgement, organisational politics and changing professional practice. Technical performance measures alone cannot explain these issues. Interpretivism accepts that roles, experience and institutional context can lead people to interpret the same situation differently (Saunders, Lewis and Thornhill, 2023). Knowledge is treated as situated and shaped through the interview rather than as a neutral measurement of one fixed reality.

The study is not fully inductive because sociotechnical systems theory guides the analysis. Its focus on technology, tasks, structures and processes shapes what counts as relevant evidence. This gives the interviews a clear analytical focus and supports comparison across cases. It may also limit attention to evidence outside the framework. If AI adoption is examined mainly through system components, stakeholder legitimacy, organisational power, professional identity and ethical disagreement may receive too little attention. The theory is therefore both a strength and a boundary on interpretation.

A positivist design would answer different questions. A survey could measure AI adoption and its drivers in a larger sample, but it would explain less about how senior communicators understand governance and AI's limits. Qualitative interviews provide this depth, but they do not show how common the patterns are across European or UK organisations. The study therefore offers detailed, context-based explanations rather than statistical generalisation. A later mixed-method study could test how widely the patterns transfer while retaining the interview evidence needed to explain them.

4.4 Axiology and Researcher Positionality

Axiology examines the role of values in research. Terms such as 'responsible', 'efficient', 'strategic' and 'human-centred' are not neutral, which matters in AI research. The study identifies tensions between efficiency and human involvement, experimentation and governance, and innovation and control. AI adoption is therefore judged through professional and organisational values as well as technical performance. These values also shape which problems researchers study and what they accept as successful or responsible practice.

Researcher positionality also matters because the researcher-participant relationship shapes an elite interview. The authors are established scholars in strategic and corporate communication, which supports informed questions and analysis. Their knowledge may also shape their views of professional development and legitimate communication practice. Olmos-Vega et al. (2023)

define reflexivity as reflection on how researchers' identities, assumptions and relationships affect the knowledge they create. This matters for DBA research because senior executives may describe AI adoption positively, while researchers may give management views more weight than employee or stakeholder views. Participants' senior status may also affect which questions can be asked and which problems are disclosed.

This topic also contains competing values. A communication department may measure AI adoption through efficiency, speed and cost savings. Stakeholders may judge the same practice through transparency, fairness, authenticity and accountability. The study shows internal organisational tensions, but a wider assessment of responsible communication would include external stakeholder perspectives. This would reduce the risk of treating management priorities as the only valid standard.

4.5 Reliability, Validity, Trustworthiness and Bias

The study has several methodological strengths. Its sample of 30 CCOs from the 300 largest European corporations provides detailed evidence from decision-makers whose strategic knowledge is hard to reach through ordinary employee samples. They can discuss organisational priorities, governance boundaries and implementation decisions across communication functions. Participants from several European organisations also reduce the risk that the findings reflect one unusual company. The sample is well suited to the strategic level of the research question, although senior access cannot replace evidence about operational practice.

Elite interviews also create quality issues. Aguinis and Solarino (2019) argue that qualitative research with elite participants needs transparency because access, sampling and interpretation affect credibility and replicability. Senior executives may not know exactly how employees use AI and may present governance or responsibility in a positive way. Their accounts are valuable senior perspectives, but they may not fully describe implementation. Replicability here should mean a transparent and defensible process, not an expectation that another interview study will produce identical accounts.

Sociotechnical systems theory and the participants' seniority and experience support credibility (Lincoln and Guba, 1985). The range of large European corporations supports transferability across similar firms but limits relevance to small and medium-sized enterprises, public bodies and third-sector organisations. Transferability also depends on enough contextual detail for readers to judge whether another setting is comparable. Dependability and confirmability rely on a clear record of interviews and analysis. Qualitative researchers do not need to make identical coding decisions, but readers need an audit trail that shows how interpretations were developed and tested.

The sample is both a strength and a weakness. It provides a rare senior-level view but gives priority to the most influential organisational group. Communication employees, technical specialists, regulators and external stakeholders are excluded. A wider design could compare how CCOs, staff who use AI tools and stakeholders who receive AI-supported communication

understand the same governance issues. This data-source triangulation could reveal conflicts between formal policy, daily tool use and stakeholder experience that a CCO-only study cannot observe.

Qualitative research focuses on the credibility of interpretations and their fit with the evidence rather than statistical construct validity. Internal validity is limited by self-reported executive accounts without documents or observation. External validity is restricted to large European corporations; in qualitative terms, these are issues of credibility and transferability. Possible biases include selection bias, social-desirability bias, elite-informant bias and researcher interpretation. Reflexive records, analysis of negative cases, multiple coders and evidence from employees, documents and stakeholders would improve credibility, dependability and confirmability (Aguinis and Solarino, 2019; Olmos-Vega et al., 2023). These additions would not remove interpretation, but they would make the basis for each claim easier to assess.

4.6 Ethical Considerations

The ethical issues operate at two levels. First, the research design requires informed participation and confidentiality. Informed consent should explain data use, quotation, withdrawal and the limits of anonymity. CCOs from very large European corporations might be identified from their sector, role and organisational practices even when company names are removed. They may also discuss trade secrets, weak governance or internal projects. Careful anonymisation, secure data handling and cautious use of identifiable quotations are therefore essential.

Second, the topic creates ethical issues. The study identifies governance, technical integration and continuing human engagement as central to Responsible AI. Responsible communication also affects privacy, authorship, transparency, misinformation, bias and stakeholder selection. Bowen (2024) gives communication professionals two ethical duties: using AI responsibly and advising organisations about wider AI impacts that may become public issues. The study supports this strategic role, but its executive-centred design cannot fully examine stakeholder rights or unequal effects. Ethical assessment therefore needs evidence from groups that may bear risks without controlling the system.

4.7 Critical Appraisal and Alternative Philosophical Approach

Buhmann et al. (2026) make a useful contribution by challenging overstated claims about rapid AI transformation. From a sociotechnical perspective, access to technology does not transform an organisation by itself. The senior European sample is relevant to strategic communication research and to a DBA focus on management practice. The study supports the view that Responsible AI is an organisational capability developed through governance, learning, technical integration and continuing human judgement, not simply through generative AI tools. Its value lies in explaining limits and conditions, not in claiming universal outcomes.

The study also has useful limits. Its interviews capture practice at one point in a fast-changing context. They cannot show how AI use develops over time, whether governance prevents harm or whether AI-supported communication affects long-term stakeholder trust. Large corporations may also have more technical, legal and communication resources than other organisations. Executive statements cannot confirm whether policies are followed in daily work. Longitudinal and observational evidence would be needed to assess change and implementation.

Future DBA research would be stronger as a critical-realist qualitative multiple-case study. It could combine interpretive interviews with organisational documents, communication policies, AI governance frameworks, workflow evidence and stakeholder accounts to identify deeper mechanisms. Critical realism would examine whether Responsible AI communication works, what makes it work, for whom, in which organisational and regulatory settings, and with which unintended consequences (Fletcher, 2017). Case selection should include contrasting sectors, organisation sizes and governance settings in the UK. This comparison would show which mechanisms depend on context and which may transfer.

The selected study provides a philosophical and methodological link to the proposed DBA research. Its interpretive, qualitative evidence shows the value of managerial meaning, while its sociotechnical approach suggests a realist interest in structures and mechanisms. A future study could use an explicit critical-realist ontology, interpretive epistemology and reflexive, value-aware axiology. This combination would make the philosophical assumptions clear, support organisational explanation and retain the ethical and stakeholder concerns needed for a Responsible AI-enabled corporate communication framework.

5. Discussion and Contribution

5.1 Integrated Discussion

The literature indicates that Responsible AI in corporate communication should be treated as a strategic organisational capability, not only as a technology. Strategic communication aligns purposeful communication with organisational goals. AI creates value when it works with professional judgement, organisational learning and governance (Zerfass et al., 2018; Buhmann et al., 2026). This capability links technical resources with roles, routines and relationships. The key management question is therefore how communication departments can use AI to improve monitoring and response without weakening evidence, accountability or stakeholder relationships.

Misinformation research supports this conclusion. Mehta et al. (2021) describe an organisational response as detection, verification, response selection and evaluation. Lu and Jin (2024) show that correction timing, detail and message form affect outcomes. AI may support early monitoring, evidence searches and response drafting, but it cannot replace contextual judgement because misinformation differs in intent, seriousness and stakeholder impact. Speed is useful only when evidence is checked and people retain final responsibility. This makes response authority and verification standards management controls rather than optional communication tasks.

5.2 Philosophy and Research Practice

Research philosophy shapes what can be known about this topic. Positivist and post-positivist experiments provide evidence on message effects, trust signals and response timing, but often separate variables from organisational structures. Interpretive studies explain managerial meaning, uncertainty and professional practice, but provide less support for statistical generalisation. Buhmann et al. (2026) use a sociotechnical, interpretive approach to show interactions between technology, tasks, structures and professional roles. A critical-realist position connects visible communication outcomes with underlying processes such as governance, culture, professional norms and regulation (Fletcher, 2017). Methodological pluralism is therefore useful: experiments can test effects, while qualitative cases can explain how and why those effects arise in organisations.

5.3 Practical and Theoretical Contribution

The practical contribution is an evidence-based sequence for Responsible AI-enabled communication. It starts with Responsible AI governance and ethical communication. These support AI-enabled monitoring and verification, accountable response authorisation, transparent stakeholder communication, stakeholder engagement, and post-response evaluation and learning. Together, the activities can protect stakeholder trust, organisational reputation and long-term

organisational resilience. The sequence is an initial framework rather than a tested model and provides a basis for further DBA research. In practice, feedback may move between stages, so the framework should be treated as connected controls rather than a fixed linear process.

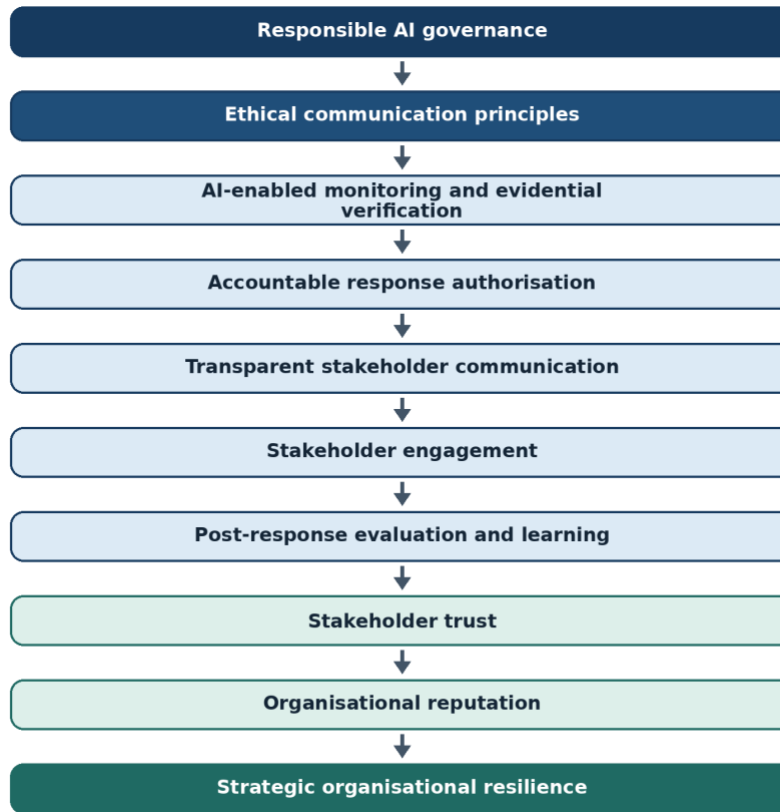


Figure 3. Proposed Responsible AI-enabled corporate communication framework.

Source: Author's synthesis of the reviewed literature.

Governance comes first because AI-enabled communication depends on more than individual practitioners. Mäntymäki et al. (2022) place AI governance within corporate and IT governance. Papagiannidis, Mikalef and Conboy (2025) identify structural, relational and procedural mechanisms for Responsible AI. Communication teams need clear decision rights, cross-functional coordination, written verification procedures, data controls and escalation routes before misinformation appears. Horneber (2025) warns that formal policies can become separated from practice. Governance is credible only when it is applied.

The theoretical contribution combines views that are often studied separately. Strategic Communication explains why AI use should support organisational objectives and relationships. Misinformation and crisis-communication research guide the identification of and response to information threats. Stakeholder Theory identifies the interests involved, and Legitimacy Theory explains how stakeholders judge organisational actions. Responsible AI governance turns ethical principles into practice. Park and Yoon's (2024) findings on transparency and trust also show that

perceptions of a technology can shape perceptions of its organisation, linking AI governance with corporate legitimacy.

In the UK, Responsible AI should form part of communication capability rather than a separate technical compliance exercise. It is not yet clear whether the proposed relationships work equally well across sectors and stakeholder groups. Future DBA research can test and improve the framework through qualitative multiple-case studies of UK organisations, including real or simulated misinformation incidents, governance arrangements, professional judgement and stakeholder expectations. Relevant outcomes include response quality, stakeholder trust, organisational legitimacy, policy-practice alignment and learning after an incident.

6. Conclusion

This report examined how Responsible AI can support corporate communication when UK organisations face digital misinformation and seek to maintain stakeholder trust. The literature shows that misinformation is not only a communication issue. It is an organisational risk that can damage reputation, legitimacy and stakeholder relationships. Generative AI increases this risk by creating synthetic content, although it can also help organisations monitor, understand and respond to information more quickly. Its value therefore depends on organisational capability rather than access to a tool.

Technology alone is not the answer. AI-supported communication needs sound governance, ethical communication, evidence checks, human judgement and clear accountability. Crisis-communication, stakeholder and organisational-legitimacy research also shows that a fast response does not guarantee trust. Stakeholders judge what an organisation says and whether the process behind the message is transparent, fair, credible and open to challenge.

The philosophical and methodological analysis shows why research context matters. Experiments explain message effects and transparency, while qualitative research can reveal organisational structures, professional views and governance limits that shape AI adoption. The selected study shows uneven adoption of AI in corporate communication and a strong dependence on learning, governance and continuing human engagement. Critical realism is suitable for future research because it links visible communication outcomes with the organisational mechanisms that produce them. A multiple-case design can also compare management claims with documents, workflows and stakeholder experience.

This report presents Responsible AI as a strategic communication capability rather than only a technology or compliance requirement. The initial framework links Responsible AI governance and ethical communication with AI-enabled monitoring and verification, accountable response authorisation, transparency, stakeholder engagement, trust, reputation and organisational resilience. These relationships are not yet well established in UK organisations. Future multiple-case research should test and adapt the framework across UK settings and identify the conditions under which it improves responsible communication.

References

- Aguinis, H. and Solarino, A.M. (2019) 'Transparency and replicability in qualitative research: The case of interviews with elite informants', *Strategic Management Journal*, 40(8), pp. 1291–1315. DOI: 10.1002/smj.3015.
- Akbarighatar, P. (2025) 'Operationalizing responsible AI principles through responsible AI capabilities', *AI and Ethics*, 5, pp. 1787–1801. DOI: 10.1007/s43681-024-00524-4.
- Amazeen, M.A. (2024) 'The misinformation recognition and response model: an emerging theoretical framework for investigating antecedents to and consequences of misinformation recognition', *Human Communication Research*, 50(2), pp. 218–229. DOI: 10.1093/hcr/hqad040.
- Berente, N. et al. (2021) 'Managing artificial intelligence', *MIS Quarterly*, 45(3), pp. 1433–1450. DOI: 10.25300/MISQ/2021/16274.
- Bigus, J., Hua, K.P.M. and Raithel, S. (2024) 'Definitions and measures of corporate reputation in accounting and management: commonalities, differences, and future research', *Accounting and Business Research*, 54(3), pp. 304–336. DOI: 10.1080/00014788.2022.2149458.
- Boman, C.D. (2021) 'Examining characteristics of prebunking strategies to overcome PR disinformation attacks', *Public Relations Review*, 47(5), 102105. DOI: 10.1016/j.pubrev.2021.102105.
- Boman, C.D. and Schneider, E.J. (2021) 'Finding an antidote: Testing the use of proactive crisis strategies to protect organizations from astroturf attacks', *Public Relations Review*, 47(1), 102004. DOI: 10.1016/j.pubrev.2020.102004.
- Bowen, S.A. (2024) "'If it can be done, it will be done": AI ethical standards and a dual role for public relations', *Public Relations Review*, 50(5), article 102513. DOI: 10.1016/j.pubrev.2024.102513.
- Broda, E. and Strömbäck, J. (2024) 'Misinformation, disinformation, and fake news: lessons from an interdisciplinary, systematic literature review', *Annals of the International Communication Association*, 48(2), pp. 139–166. DOI: 10.1080/23808985.2024.2323736.
- Buhmann, A. et al. (2026) "'We're right in the foothills": the adoption of artificial intelligence in corporate communication departments', *Corporate Communications: An International Journal*, advance online publication, pp. 1–21. DOI: 10.1108/CCIJ-09-2025-0296.
- Christensen, E. and Andersson, R. (2025) 'To hasten slowly: The prudence of slow AI implementation in public relations', *Public Relations Review*, 51(2), 102557. DOI: 10.1016/j.pubrev.2025.102557.
- Coombs, W.T. (2022) 'Situational Crisis Communication Theory (SCCT): Refining and clarifying a cognitive-based theory of crisis communication', in Coombs, W.T. and Holladay, S.J. (eds.) *The Handbook of Crisis Communication*. 2nd edn. Hoboken, NJ: Wiley, pp. 193–204. DOI: 10.1002/9781119678953.ch14.

- Cornelissen, J. (2023) *Corporate Communication: A Guide to Theory and Practice*. 7th edn. London: SAGE.
- Corsi, G., Marino, B. and Wong, W. (2024) 'The spread of synthetic media on X', *HKS Misinformation Review*, 5(3), pp. 1–19. DOI: 10.37016/mr-2020-140.
- Department for Science, Innovation and Technology (2023) *A Pro-Innovation Approach to AI Regulation*. Available at: <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper> (Accessed: 18 August 2026).
- DiMaggio, P.J. and Powell, W.W. (1983) 'The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields', *American Sociological Review*, 48(2), pp. 147–160. DOI: 10.2307/2095101.
- Dong, C. and van den Berg, M. (2025) 'Revisiting PR professionalism and ethics in the generative AI era through PR practitioners' insights', *Public Relations Review*, 51(3), 102582. DOI: 10.1016/j.pubrev.2025.102582.
- Fletcher, A.J. (2017) 'Applying critical realism in qualitative research: Methodology meets method', *International Journal of Social Research Methodology*, 20(2), pp. 181–194. DOI: 10.1080/13645579.2016.1144401.
- Freeman, R.E. (1984) *Strategic Management: A Stakeholder Approach*. Boston, MA: Pitman.
- Germinder, L.-A. and Capizzo, L. (2025) 'A strategic communication practitioner imperative: Contextualizing Responsible AI as part of responsible advocacy', *International Journal of Strategic Communication*, 19(2), pp. 340–356. DOI: 10.1080/1553118X.2024.2430959.
- Giannitsas, D., Sun, R. and Baptista, J. (2026) 'In artificial intelligence (AI) we (dis)trust? Navigating institutional pressures for automation and augmentation in the implementation of AI in organizations', *Information and Organization*, 36(1), 100609. DOI: 10.1016/j.infoandorg.2026.100609.
- Hadley, E., Blatecky, A. and Comfort, M. (2025) 'Investigating algorithm review boards for organizational responsible artificial intelligence governance', *AI and Ethics*, 5, pp. 2485–2495. DOI: 10.1007/s43681-024-00574-8.
- Hillo, J., Vento, I. and Erkkilä, T. (2025) 'Algorithmic governance: Experimental evidence on citizens' and public administrators' legitimacy perceptions of automated decision-making', *Public Administration*, advance online publication, pp. 1–16. DOI: 10.1111/padm.70028.
- Hollebeek, L.D. et al. (2024) 'Consumers' online brand-related misinformation engagement: a weapons of influence perspective', *Journal of Strategic Marketing*, 32(8), pp. 1415–1434. DOI: 10.1080/0965254X.2024.2380657.
- Horneber, D. (2025) 'Understanding the implementation of Responsible Artificial Intelligence in organizations: A neo-institutional theory perspective', *Communications of the Association for Information Systems*, 57, pp. 185–218. DOI: 10.17705/1CAIS.05708.
- Information Commissioner's Office (2023) *Guidance on AI and Data Protection*. Available at: <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/> (Accessed: 18 August 2026).

- Jahng, M.R. (2021) 'Is fake news the new social media crisis? Examining the public evaluation of crisis management for corporate organizations targeted in fake news', *International Journal of Strategic Communication*, 15(1), pp. 18–36. DOI: 10.1080/1553118X.2020.1848842.
- Jiang, Y.N., Cheng, Y. and Wang, Y. (2025) 'Leveraging AI in CSR: How social-oriented chatbots influence Chinese consumers' supportive actions via dialogic communication', *Public Relations Review*, 51(1), 102541. DOI: 10.1016/j.pubrev.2025.102541.
- Jin, Y. et al. (2020) 'The effects of corrective communication and employee backup on the effectiveness of fighting crisis misinformation', *Public Relations Review*, 46(3), 101910. DOI: 10.1016/j.pubrev.2020.101910.
- Jöhnk, J., Weißert, M. and Wyrski, K. (2021) 'Ready or Not, AI Comes—An Interview Study of Organizational AI Readiness Factors', *Business & Information Systems Engineering*, 63(1), pp. 5–20. DOI: 10.1007/s12599-020-00676-7.
- Kraus, S. et al. (2022) 'Literature reviews as independent studies: guidelines for academic practice', *Review of Managerial Science*, 16(8), pp. 2577–2595. DOI: 10.1007/s11846-022-00588-8.
- Laskin, A.V. and D'Agostino, G. (2024) 'The Delphi Panel investigation of artificial intelligence in investor relations', *Public Relations Review*, 50(4), 102489. DOI: 10.1016/j.pubrev.2024.102489.
- Lincoln, Y.S. and Guba, E.G. (1985) *Naturalistic Inquiry*. Beverly Hills, CA: SAGE.
- Lu, X. and 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), 102514. DOI: 10.1016/j.pubrev.2024.102514.
- Mahajan, R. et al. (2023) 'Stakeholder theory', *Journal of Business Research*, 166, 114104. DOI: 10.1016/j.jbusres.2023.114104.
- Martin, K. and Waldman, A. (2023) 'Are algorithmic decisions legitimate? The effect of process and outcomes on perceptions of legitimacy of AI decisions', *Journal of Business Ethics*, 183(3), pp. 653–670. DOI: 10.1007/s10551-021-05032-7.
- Mehta, A.M. et al. (2021) 'A process view of crisis misinformation: How public relations professionals detect, manage, and evaluate crisis misinformation', *Public Relations Review*, 47(2), 102040. DOI: 10.1016/j.pubrev.2021.102040.
- Men, L.R. et al. (2023) 'Shaping corporate character via chatbot social conversation: Impact on organization-public relational outcomes', *Public Relations Review*, 49(5), 102385. DOI: 10.1016/j.pubrev.2023.102385.
- Mäntymäki, M. et al. (2022) 'Defining organizational AI governance', *AI and Ethics*, 2(4), pp. 603–609. DOI: 10.1007/s43681-022-00143-x.
- Mökander, J. and Floridi, L. (2023) 'Operationalising AI governance through ethics-based auditing: an industry case study', *AI and Ethics*, 3, pp. 451–468. DOI: 10.1007/s43681-022-00171-7.

- National Institute of Standards and Technology (2023) *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. NIST AI 100-1. Gaithersburg, MD: NIST. DOI: 10.6028/NIST.AI.100-1.
- Ofcom (2024a) *A deep dive into deepfakes that demean, defraud and disinform*. Available at: <https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/deepfakes-demean-defraud-disinform> (Accessed: 7 August 2026).
- Ofcom (2024b) *Understanding misinformation: an exploration of UK adults' behaviour and attitudes*. Available at: <https://www.ofcom.org.uk/media-use-and-attitudes/attitudes-to-news/understanding-misinformation-an-exploration-of-uk-adults-behaviour-and-attitudes> (Accessed: 7 August 2026).
- Oh, J. and Ki, E.-J. (2024) 'Can we build a relationship through artificial intelligence (AI)? Understanding the impact of AI on organization-public relationships', *Public Relations Review*, 50(4), 102469. DOI: 10.1016/j.pubrev.2024.102469.
- Olmos-Vega, F.M. et al. (2023) 'A practical guide to reflexivity in qualitative research: AMEE Guide No. 149', *Medical Teacher*, 45(3), pp. 241–251. DOI: 10.1080/0142159X.2022.2057287.
- Papagiannidis, E., Mikalef, P. and Conboy, K. (2025) 'Responsible artificial intelligence governance: a review and research framework', *The Journal of Strategic Information Systems*, 34(2), article 101885. DOI: 10.1016/j.jsis.2024.101885.
- Park, K. and Yoon, H.Y. (2024) 'Beyond the code: The impact of AI algorithm transparency signaling on user trust and relational satisfaction', *Public Relations Review*, 50(5), 102507. DOI: 10.1016/j.pubrev.2024.102507.
- Park, K. and 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, 1160. DOI: 10.1057/s41599-025-05116-z.
- Park, S. and Nan, X. (2026) 'Generative AI and misinformation: a scoping review of the role of generative AI in the generation, detection, mitigation, and impact of misinformation', *AI & Society*, 41, pp. 1501–1515. DOI: 10.1007/s00146-025-02620-3.
- Paul, J. and Criado, A.R. (2020) 'The art of writing literature review: What do we know and what do we need to know?', *International Business Review*, 29(4), article 101717. DOI: 10.1016/j.ibusrev.2020.101717.
- Sauer, P.C. and Seuring, S. (2023) 'How to conduct systematic literature reviews in management research: a guide in 6 steps and 14 decisions', *Review of Managerial Science*, 17(5), pp. 1899–1933. DOI: 10.1007/s11846-023-00668-3.
- Saunders, M.N.K., Lewis, P. and Thornhill, A. (2023) *Research Methods for Business Students*. 9th edn. Harlow: Pearson.
- Suchman, M.C. (1995) 'Managing legitimacy: Strategic and institutional approaches', *Academy of Management Review*, 20(3), pp. 571–610. DOI: 10.5465/amr.1995.9508080331.
- Uysal, N. and Deng, T. (2025) 'The new brave world of public relations: The impact of organizational dynamics on adoption of AI', *International Journal of Strategic Communication*, 19(2), pp. 303–323. DOI: 10.1080/1553118X.2024.2447019.

- Wang, S. et al. (2026) 'How rationale and process transparency shape perceived legitimacy in AI-assisted decisions: Experimental evidence from China and the United States', *Government Information Quarterly*, 43(2), 102141. DOI: 10.1016/j.giq.2026.102141.
- Wardle, C. and Derakhshan, H. (2017) *Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making*. Available at: <https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html> (Accessed: 18 August 2026).
- Yue, C.A. et al. (2024a) 'Artificial intelligence for internal communication: Strategies, challenges, and implications', *Public Relations Review*, 50(5), 102515. DOI: 10.1016/j.pubrev.2024.102515.
- Yue, C.A. et al. (2024b) 'Public relations meets artificial intelligence: Assessing utilization and outcomes', *Journal of Public Relations Research*, 36(6), pp. 513–534. DOI: 10.1080/1062726X.2024.2400622.
- Zerfass, A. et al. (2018) 'Strategic communication: Defining the field and its contribution to research and practice', *International Journal of Strategic Communication*, 12(4), pp. 487–505. DOI: 10.1080/1553118X.2018.1493485.
- Zhang, N. et al. (2024) 'How organization response influences public sentiment and behavior in online crises: The role of response strategies, tenor of discourse and executives', *Public Relations Review*, 50(1), 102395. DOI: 10.1016/j.pubrev.2023.102395.