Exploring organisations' preparedness to address malicious digital influence operations: A professional perspective*

Michael Thorn[1,*]

[1] *University of Portsmouth, School of Criminology and Criminal Justice, Portsmouth, United Kingdom*

Abstract

Malicious digital influence operations are now recognised as a strategic challenge for states, democratic institutions and civil society. Less attention has been given to how private-sector organisations understand and prepare for such activity as an organisational risk. This paper examines senior professional perspectives on malicious digital influence operations affecting commercial organisations. Drawing on 14 semi-structured interviews with participants across different organisational contexts, the study explores how professionals perceive the threat, distinguish it from ordinary criticism or online discourse, allocate responsibility, and assess organisational preparedness. The study applies framework-informed thematic analysis, following an abductive, hybrid inductive-deductive approach, structured around four domains: organisational awareness, perceived risk and impact, responsibility and ownership, and preparedness and response capability. Participants often recognise elements of malicious influence activity but commonly interpret them through existing business categories such as reputation, customer trust, complaints, activism, fraud, security, compliance or crisis management. The paper argues that private-sector protection is constrained less by complete absence of awareness than by conceptual ambiguity, fragmented ownership, weak escalation pathways and limited rehearsal of integrated response capability. It names a central asymmetry: organisations that are fluent in projecting influence outward, through legitimate marketing and public relations, remain comparatively unprepared to receive it, that is, to recognise and defend against the same persuasive techniques when they are turned against the organisation through deception, coordination and coercion. The paper reframes this defensive capability as an enterprise governance task, not individually as a communications, legal or cyber responsibility.

Keywords

digital influence operations, disinformation, misinformation, information warfare, security, organisational preparedness, socio-technical systems, enterprise risk governance, private sector

CEUR-WS.org/Vol-4334/paper27.pdf

* *The 12th International Conference on Socio-Technical Perspectives in IS (STPIS'26), September 16–17, 2026, L'Aquila, Italy.*

[1,*] Corresponding author.

Michael.Thorn@port.ac.uk (M. Thorn)

0009-0005-0886-2208 (M. Thorn)

© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

published 2026-11-19

1. Introduction: Malicious Digital Influence Operations as an Emerging Business Challenge

Malicious digital influence operations present a business challenge that private-sector organisations are not yet organised to meet. Although the language of influence operations is most often associated with elections, geopolitical competition and state-led interference [4], the underlying mechanisms are not confined to the political domain. They belong to a longer lineage of propaganda, the organised attempt to shape perception, emotion and behaviour in the service of an interest, now operating through digital and largely computational means [23]. Even the most common labels are unstable: "fake news" is applied inconsistently, its meaning shaped by media framing as much as by any settled definition [19]. In organisational settings disinformation is often treated narrowly as false information, when the practical challenge is broader. Commercial organisations may face fabricated claims but also misleading amplification, selective use of true information, coordinated distortion, synthetic content, covert sponsorship, hostile narrative construction and targeted harassment, the span Wardle and Derakhshan [16] taxonomise as information disorder, distinguished by harm and intent, not truth value alone (Table 1). The most recent organisationally focused review warns that misinformation is "particularly dangerous for businesses" [14, p.135] and situates its rise against the World Economic Forum's identification of mis- and disinformation as a leading near-term global risk [14], showing that the threat now registers in elite risk assessment even as scholarship on its organisational dimension remains thin.

For businesses the distinction is practical: malicious digital influence operations can cause harm even where the material circulated is not wholly false, because a genuine document, historic incident, customer complaint, executive comment or operational failure may be reframed, decontextualised or amplified in ways that distort stakeholder understanding. Organisations must therefore distinguish between misinformation, disinformation, malinformation, ordinary criticism, activism, customer dissatisfaction and deliberate attempts to manipulate perception. The intent-based distinction [2] is the starting point for this task but, as the findings show, is not by itself sufficient to capture the maliciousness organisations must in practice respond to. The distinction is also practical: the label attached to an incident shapes whether it is escalated as a communications problem, legal matter, customer complaint, fraud issue, cyber incident, security concern, reputational risk or enterprise-level threat.

Private-sector exposure is acute because many organisations depend on intangible forms of confidence. A company's commercial position may rest on consumer trust, investor assurance, public legitimacy, employee confidence, supply-chain credibility, regulatory trust or licence to operate. These assets can be affected before an organisation has established the factual basis of an allegation or identified whether the activity is organic, coordinated, sponsored or malicious. The speed of digital amplification means that organisational harm may occur before clear evidence thresholds are met.

This grey area is a large part of the problem. Malicious influence activity does not always present as a discrete incident; it may appear first as customer dissatisfaction, a viral post, repeated allegations, negative reviews, media enquiries, coordinated tagging, staff harassment, investor concern or stakeholder pressure. Organisations may then respond through existing functional categories, communications, legal, compliance, security, customer service,

marketing, risk or crisis management, each seeing part of the problem while none owns the full threat picture.

This paper examines malicious digital influence operations as an organisational preparedness issue. It explores how senior professionals understand the threat, how they distinguish it from adjacent phenomena such as disinformation or ordinary reputational challenge, where they believe responsibility currently rests, and how prepared they consider organisations to be. The paper takes its question from the wider doctoral study: To what extent are organisations protected from malicious digital influence operations?

The paper asks: How do senior professionals across different organisational contexts perceive, identify, allocate responsibility for, and assess preparedness for malicious digital influence operations affecting private-sector organisations?

Drawing on 14 semi-structured interviews, the paper argues that private-sector protection is constrained by organisational interpretation as much as by technical detection challenges. Senior professionals often recognise fragments of the threat through existing business categories, including disinformation, reputation, customer trust, activism, fraud, compliance and crisis management. Preparedness remains weakened where terminology is unstable, early-warning signals are dispersed, ownership is fragmented, and response mechanisms are insufficiently rehearsed.

Table 1. *Key terms as used in this paper.*

Term	Working definition	Reference
Misinformation	False information shared without intent to cause harm	[16]
Disinformation	False information created or shared deliberately to cause harm; the field's dominant, falsity-centred construct	[1, 16]
Malinformation	Genuine information deployed to cause harm, for example selectively or out of context	[16]
Information disorder	Umbrella term for the three forms above, distinguished by harm and intent, not truth value alone	[16]
Fake news	An unstable public label, its meaning shaped by media framing; not used analytically in this paper	[19]
Information warfare	The organised, state-linked end of influence activity; used by participants, not as this paper's construct	[24]
Malicious digital influence operation	This paper's organising construct: coordinated digital activity, spanning the forms above, that seeks to shape stakeholder perception and behaviour against an organisation's interests through deception, coordination or coercion	[16, 23]

2. Literature Positioning: From Democratic Threat to Organisational Preparedness

This review deliberately draws on professionally published books alongside peer-reviewed scholarship, a methodological choice serving the study's dual contribution to academic knowledge and professional practice. Senior practitioners do not form their understanding from peer-reviewed literature alone; the categories, vocabulary and worked examples available to that community circulate largely through trade and professional books. Peer-reviewed empirical and review work supports the evidentiary claims, while professional and practitioner texts supply conceptual framing, typology and illustration, with the limits of each made explicit at the point of use.

Research on misinformation, disinformation, and fake news has grown rapidly, but its focus remains overwhelmingly on political and societal domains. A systematic review of 1,261 peer-reviewed articles published between 2010 and 2021 found that more than half the entire dataset (647 articles) was published in 2021 alone [1]. Within this literature, disinformation is distinguished from misinformation primarily by intent: disinformation is false information deliberately created or disseminated to cause a form of harm, whereas misinformation is false but not intended to cause harm [2].

This body of research offers little to commercial organisations specifically. The review's own search terms excluded "propaganda" entirely, because the term generated too many false positives [1], leaving the corpus without the propaganda and computational-propaganda literature closest to the operational core of digital influence operations. Broda and Strömbäck themselves identify the meso level of impact, effects on organisations, interest groups and social movements, as one of the most under-researched areas across the field [1]; and even that meso-level attention concentrates on civil-society and political organisations, not commercial firms. The SAGE Handbook of Propaganda [23], which that review excludes, is organised overwhelmingly around state, military, electoral and extremist propaganda and gives private industry little more than a single chapter, and that chapter frames the corporation as a producer of propaganda through public relations; the corporation as target is absent. Olejnik [24] confirms the orientation, framing the subject principally in political, diplomatic and military terms, with the commercial organisation appearing only incidentally.

The field's political orientation was built in from the start. Freelon and Wells [4] trace disinformation's emergence as a distinct research object to the 2016 US presidential election and its aftermath; prior scholarly work consisted of isolated historical case studies that did not reference one another or build a cumulative research programme. The subsequent growth was rapid and concentrated: of all disinformation-titled publications indexed between 2010 and 2019, more than 70 per cent appeared after 2016 [4]. This growth was driven entirely by concern for democratic and electoral integrity – the special issue these authors introduce addresses state-sponsored campaigns, partisan media ecosystems, and voter belief formation, with no reference to commercial organisations as either targets or units of analysis. The pattern across all three reviews considered here is consistent: a field that expanded explosively in response to a specific political shock, with research infrastructure, definitions and case material built around that shock's political consequences.

The field's most recent systematic mapping shows that this concentration has continued. A scoping review of sixty-four peer-reviewed studies published between 2021 and 2024 covering

generative artificial intelligence and disinformation [7] organises the entire corpus into six thematic areas: political disinformation and propaganda, scientific disinformation, fact-checking, journalism and the media, media literacy and education, and deepfakes. None of the six admits the commercial organisation as a target or a unit of analysis; firms surface only in a single passing remark that disinformation can damage the reputation of individuals and organisations. The arrival of generative artificial intelligence – the development most likely to have redirected scholarly attention – has reproduced the field's political and societal centre of gravity, and although the review names the capacity to fabricate and disseminate synthetic content a "profound governance gap" [7, p.10], it continues to locate that gap at the level of democracy and the public sphere, leaving the private-sector dimension as under-addressed in 2024 as it was in the corpus Broda and Strömbäck mapped to 2021.

The absence appears again when the literature is approached from the opposite direction, organisational risk governance looking for threat categories. A bibliometric analysis of 463 enterprise security risk management articles published between 1986 and 2019 maps the field's full thematic structure: cybersecurity, risk management, security risk management, information systems, and enterprise risk management [3]. Even within this firm-level, governance-focused literature, where security risk is treated explicitly as a board-level, cross-functional concern, no thematic category addresses information manipulation, hostile narrative construction, or influence activity. The seven operational risk types formalised under the Basel II framework – internal and external fraud, employment practices, customer and product practices, physical asset damage, business interruption, and system failure – also contain no direct equivalent [3]. Cybersecurity, the nearest adjacent theme, is treated throughout as a systems-intrusion problem with no content-integrity dimension. This confirms from the governance side that malicious digital influence operations currently fall outside the formal risk categories most commercial organisations use to structure their protection.

A close reading of one rigorous account of organisational preparedness, to see which risks it includes, gives the same result. Dahmen's study of enterprise risk management and organisational resilience is commercial in setting and organisational in unit of analysis, explicitly concerned with how firms anticipate, cope with and adapt to "completely novel" and "severe" crisis events [5]. It conceptualises the organisation as a system of "both technical and social components" [5] and defines resilience in terms under which a coordinated hostile narrative campaign would qualify as a foreseeable novel threat. Yet across a multidisciplinary review and an empirical study of forty-four senior executives, the crises populating the account are the COVID-19 pandemic, the 2008 financial crisis, large-scale fraud, infrastructure interruption, a flood and a terrorist attack; information manipulation and hostile narratives appear nowhere among the novel risks a resilient firm is expected to anticipate. Where the bibliometric evidence [3] establishes the absence structurally, Dahmen demonstrates it substantively, from inside a paper that had every opportunity to include such threats and did not.

One recent exception is itself narrowly framed. Akhtar et al. [8] offer a post-2024 empirical, firm-level study of disinformation as an organisational risk: a structural-equation model drawn from 209 managers in China's maritime logistics sector, in which "disinformation control" significantly moderates the relationship between generative-AI adoption and global supply-chain resilience. Its construct of disinformation is operational: disinformation figures as noise that distorts market signals and degrades the data feeding AI-enabled decision systems, and the

coordinated, intentional and reputational dimensions central to malicious digital influence operations fall outside its scope. What it supports is the limited claim that disinformation is beginning to be modelled as a measurable firm-level risk to operational resilience; what it cannot support is any claim that organisations recognise, govern or rehearse responses to influence operations as a distinct adversarial threat. The narrowing has precedent: Petratos [13], one of the earliest studies to register misleading information as a business-level threat, framed it explicitly as "an emerging cyber risk" [13, p.763], admitting the organisational threat to the risk register only once recoded in the vocabulary of cyber defence.

The formal cybersecurity standards through which commercial organisations learn to recognise and categorise threats show the same absence. A comparative assessment of authoritative cybersecurity reference works – including NIST, CISA, CNSS, ISACA, SANS, the IETF and the FFIEC – finds that with the single exception of ENISA's 2021 Threat Landscape, none lists disinformation among its recognised cyber threats, though disinformation maps cleanly onto conventional threat criteria [11]. The authors call this a "taxonomical correction" yet to be made [11]; the implication for the organisational reader is larger: the formal threat language that firms inherit from the cybersecurity profession, and through which their security, risk and audit functions interpret what an "incident" is, contains no slot for malicious influence activity. We adopt their diagnosis of the absence while questioning their prescription: absorbing influence operations into the cybersecurity remit risks reproducing the very technocentric framing that the findings and Discussion identify as part of the preparedness problem.

A similar gap exists within the information-management discipline itself. Stone and colleagues [29] document an organisational "dark side" of information, namely the deliberate or unconscious falsification, concealment and biasing of information by an organisation's own managers and boards. They note how readily coordinated oversight fails even where the relevant functions exist – such as financial institutions with compliance, quality-assurance, risk, internal-audit and investigative roles, yet still struggling to bring them together to manage risk. Their focus is information mismanaged inside the firm, and their analysis is conceptual. This study examines hostile influence aimed at the firm from outside, using empirical evidence. Because the discipline's most developed study of organisational "misinformation" looks inward at the firm's own conduct instead of outward at an adversary, it reinforces from a further direction the absence this review documents.

Rare empirical evidence of this conceptual absence in practice comes from a longitudinal Spanish content analysis of 630 fact-checks published between 2023 and the second quarter of 2025, covering 549 organisational entities across 730 mentions [10]. Disinformation targeting organisations is persistent and strategic, with retail and banking the most extensively affected commercial sectors and Spain's tax authority, postal service and traffic agency the most intensively targeted institutions [10]. More directly relevant to preparedness, the analysis documents an organisational response in only 52.6 per cent of fact-checked cases, and characterises those responses as "predominantly reactive, fragmented, and focused on social media and corporate websites, with limited implementation of preventive strategies" [10, p.343]. The underlying records are the published fact-checks of four Spanish organisations, capturing incidents visible to public verification, and the construct studied is weighted toward fraud-adjacent activity, corporate identity impersonation, deepfake-enabled scams, phishing under institutional cover, so the evidence does not reach internal organisational practice or the full

scope of malicious digital influence operations [10]. Even within those bounds, the convergence with the present study is striking: half of affected organisations show no public response at all, and the responding half do so without an integrated capability.

Further evidence comes from the professional domain. Wolff's 2026 book on disinformation attacks on businesses, issued in a management series for practitioners and prefaced by a former head of Germany's federal intelligence service, sets out the organisational threat picture missing from the academic literature [12]. Its core distinction separates the disinformation attacker from the cybercriminal: where the latter penetrates information systems, the former targets the trust, perception and behaviour of a company's stakeholders, seeking reputational and operational damage in place of ransom [12, ch. 9]. From this distinction Wolff develops a typology of perpetrators, from state or state-affiliated actors and competitors to disgruntled insiders, investors and short sellers, hacktivists and activists, and a schema of motives spanning financial gain, competitive advantage, ideological objectives and revenge [12, ch. 9]. The supporting cases are concrete and commercial, from banks withdrawing credit from a targeted firm on the strength of an orchestrated campaign to state-sponsored activity migrating into the economic sphere during the COVID-19 vaccine race [12, chs. 8–9]. As a practitioner monograph it is cited here for the threat typology and illustrative cases it assembles; it offers no measure of prevalence or effect. That the most developed organisational account appears in a professional management series instead of peer-reviewed scholarship is itself a symptom of the neglect this paper describes.

Another professional account with a similar concern comes from Forest [17], who documents the rise in supply of "digital influence mercenaries"; actors paid to manipulate the perceptions and behaviour of target audiences on behalf of clients: hostile influence can now be bought as a service.

These strands converge on a single conclusion from independent directions. Disinformation literature largely looks outward to democratic and societal harm and barely registers the commercial organisation, a pattern the most recent generative-AI-era mapping reproduces without correcting [1, 7]. The firm-level risk literature contains no thematic or formal category for information manipulation, whether mapped structurally across the field [3] or read substantively inside a rigorous account of organisational resilience [5]. The authoritative cybersecurity reference works through which firms inherit their formal threat language make no provision for influence activity [11], and where the private sector does engage information threats in recent empirical work it does so through the vocabulary of operational noise and data integrity [8], with the most current organisational-response evidence showing that even responding firms do so reactively and without integrated capability [10]. Malicious digital influence operations sit in the interval between two mature bodies of scholarship – neither fully a communication problem nor fully a governance problem – and this interstitial position leaves commercial organisations without an established conceptual category, an agreed vocabulary, or a formal risk factor through which to recognise, own and prepare for them. The present study addresses this gap empirically.

3. Research Design and Methodology

This paper is based on a qualitative, exploratory study consisting of 14 semi-structured interviews with senior professionals operating across private-sector and commercially relevant

organisational contexts. The study is interpretive in orientation, reflecting the research objective of understanding how participants make sense of malicious digital influence operations as an emerging business challenge. The purpose is to examine how senior professionals perceive the threat, identify its limits, allocate responsibility and assess organisational preparedness, not to quantify prevalence across industry.

Participants were recruited purposively through the researcher's professional and academic networks. Selection required senior standing, executive, board, director or national-level responsibility, in role remits spanning security, risk and corporate affairs, communications, technology and business leadership, each with direct professional exposure to influence-related issues. The fourteen participants span sectors including technology and enterprise software, cyber security, national security and critical national infrastructure, policing, media and public relations, aviation, strategic communications, energy, consumer goods and financial services, and two combine academic posts with board or advisory practice. Participants were drawn from networks spanning the United Kingdom, the European Union, the United States and Canada, reflecting the study's focus on private-sector organisations in Western regulatory, media and market environments, where open information flows leave firms directly exposed within a common governance frame. Eleven are based in the United Kingdom, two in the United States and one operates internationally from the UK; operating footprints range from UK-national to global, including Asia-Pacific, and most organisations are large or multinational, alongside two medium-sized firms and one specialist consultancy. Interviews were conducted between August 2025 and February 2026, and each opened with a standard working definition of a digital influence operation.

Semi-structured interviews were selected because the phenomenon is conceptually unsettled and unevenly recognised across professional domains; a structured survey alone would risk imposing terminology participants may not use. The interview format allowed a consistent set of core questions across all participants, while permitting limited follow-up where clarification was required. Participants often discussed adjacent concepts such as disinformation, misinformation, hostile narratives, reputational risk, fraud, activism and crisis response.

The interview schedule examined baseline awareness, terminology, perceived seriousness, the line between legitimate and malicious activity, issue visibility, organisational impact, current and ideal lines of responsibility, resourcing, guidance, detection indicators and future direction. This structure served both the wider doctoral research question and this paper's focus.

All transcripts were analysed thematically, following an abductive, hybrid inductive-deductive approach in which an a priori organising structure and inductively developed themes were combined [30, 31]. The analysis began with familiarisation through repeated reading of each transcript and production of an analytic memo for each participant. Initial deductive coding domains were derived from the research question, interview schedule and study objectives. These domains were then refined through inductive coding to capture meanings, examples and distinctions emerging from participant accounts. NVivo was used to manage transcript coding, coded extract retrieval, code refinement and comparison across participants.

The analysis proceeded in five stages. First, transcripts were imported into NVivo and reviewed for familiarisation. Second, first-cycle coding was applied using the four-domain analytical framework. Third, the codebook was refined through consolidation of overlapping codes, clarification of code definitions and identification of inductive codes. Fourth, candidate themes were developed by reviewing coded extracts, analytic memos and cross-participant

patterns. Fifth, a cross-case framework matrix [32] was used to examine how patterns varied by organisational context, participant function, seniority, geography and perceived level of preparedness.

Rigour was maintained through participant identification codes, quote traceability, reflexive memoing and a documented audit trail of coding decisions. The study received ethical approval from the University of Portsmouth Ethics Committee. Informed consent was obtained from every participant before their interview, and participants were advised of their right to withdraw at any point up to the anonymisation of their data. Interviews were audio-recorded and transcribed in full, then pseudonymised at the point of transcription and referred to throughout by participant code (P001–P014), so that no personal or identifying data are analysed or reported in the findings. Data were stored securely in accordance with University of Portsmouth policy and applicable data-protection law; the recordings, transcripts, NVivo project and analytic memos form the audit trail behind the findings. Attention was given throughout to negative and divergent cases. Findings are presented as analytically grounded professional perspectives, not as statistically generalisable claims about all private-sector organisations.

4. Analytical Framework

The analytical framework was designed to connect the interview data directly to the study's central question: whether organisations are protected from malicious digital influence operations. It organises the analysis around four linked domains: organisational awareness, perceived risk and impact, responsibility and ownership, and preparedness and response capability. These domains follow a practical logic of protection: recognise the threat and distinguish its character, understand what may be harmed, determine who should act, and assess whether sufficient capability exists to respond.

Organisational awareness addresses both how participants understand the seriousness, organisational relevance and likely future direction of malicious digital influence operations, and how they distinguish such activity from adjacent phenomena. The first aspect concerns whether the issue is viewed as a strategic risk, a communications problem, a customer trust issue, a security concern, or a phenomenon largely associated with politics and state actors. The second concerns how participants separate malicious influence activity from disinformation, misinformation, malinformation, ordinary criticism, activism, fraud or routine online discourse. Recognition and discrimination form a single domain because they prove inseparable in practice: the point at which a professional registers that something is happening is also the point at which they categorise it, and misclassification at that moment shapes the entire downstream response.

Perceived risk and impact examines what participants perceive to be exposed when malicious influence activity succeeds. This includes trust, reputation, revenue, investor assurance, regulatory scrutiny, employee wellbeing, operational disruption, stakeholder confidence, supply-chain credibility and licence to operate. The emphasis is on how information activity becomes commercial, operational or governance harm.

Responsibility and ownership considers how organisations detect, escalate, own and respond to such activity. It includes current and preferred lines of responsibility, frontline warning

signals, senior decision-making, and the relationship between communications, legal, compliance, security, risk, customer-facing teams, operations and digital functions.

Preparedness and response capability assesses whether organisations have the people, processes, tools, training, escalation routes, external guidance and rehearsed response capability required to detect and manage malicious digital influence operations. This moves the analysis beyond awareness towards practical organisational readiness.

Together, these domains provide the coding spine for the analysis. They are sufficiently structured to enable systematic comparison across interviews, while retaining flexibility for inductive insights to emerge through NVivo coding. The study examines an emerging and unevenly recognised threat category, in which participants may describe similar phenomena through different professional vocabularies.

The four domains are read through socio-technical systems theory. Trist and Bamforth's founding insight, that work systems fail when their social and technical elements are optimised separately [25], Orlikowski's account of technology and organisation as shaping each other [26], and Baxter and Sommerville's diagnosis of failure at the seams between weakly connected teams and processes [9] together supply the expectation the findings then examine: that protection against influence operations depends on joint optimisation of the social and technical subsystems.

5. Findings: Senior Professional Perspectives on Organisational Preparedness

The four themes below were developed through an abductive, hybrid inductive-deductive thematic analysis. The four analytic domains provided a deductive organising structure, while the themes and sub-themes were generated inductively within and across that structure through NVivo coding and cross-case comparison of coded extracts. Each of the fourteen accounts was coded against every theme, so the prevalence figures reported below rest on a systematic case-by-case assessment. Their status is interpretive and descriptive: a count records how widely a construction recurred within these fourteen accounts and claims nothing beyond them. The analytical weight is placed on how participants framed the threat and reasoned about it, with the counts locating each reading within the dataset; consistent with the interpretive design, no frequency is treated as evidence of prevalence beyond the participants studied. That refinement produced the themes that follow: the vocabulary problem of 5.1, the definitional line of 5.2, the transmission account of 5.3, and the single integrated-preparedness deficit of 5.4. The analysis supports a more specific proposition than simple unawareness: that recognition is genuine but partial, that what is recognised is routinely absorbed into whatever risk category the observer already owns, and that the organisational distance between where the threat is first seen and where it can be acted upon is largely unbridged. The binding constraint is readiness, and readiness fails at the seams between functions, not within them.

One framing recurred among the most experienced participants and organises the themes that follow. Organisations are fluent in the production of influence through marketing and public relations, and largely illiterate in its reception: they lobby, advertise, brief journalists and manage stakeholders as routine, yet are weak at recognising the same environment turned against them. As a participant strategic influence practitioner put it, large corporates "know how to project a message, but they're weaker at recognising when someone else is shaping the

environment around them. That's the gap" (P010). This asymmetry is the connective thread through the conceptual, functional and operational problems documented below.

5.1. Recognition without a shared vocabulary: organisational awareness

All fourteen participants recognised the behaviours the study describes, whether from professional experience or once the phenomenon was set out, and all fourteen expressed that recognition through a vocabulary inherited from their own professional or sectoral world; no shared vocabulary was available. Specialist participants reached for the registers of national security and defence. Nine participants noted disinformation as a term being used within their workplace. The academic with a counter-threat background treated information warfare as the lasting category, while noting that its natural home is "national security, not commercial space" (P003). The cyber-security executive named misinformation and disinformation as dominant and tied them to their elevation on the World Economic Forum risk register (P002). A consultant with an information-operations background framed the threat for boards as a threat to "reputation management" (P004).

Commercial participants reached instead for the registers of their own operations. The energy executive was explicit that the formal term digital influence operations would not survive contact with his executive committee: "we wouldn't call it DIO around the table. We'd call it a hostile narrative, activist pressure, reputational heat, licence-to-operate problem. But it's the same damn animal" (P012). In fast-moving consumer goods the native vocabulary was "product scare, brand issue, rumours, fake reviews, or misinfo", and information warfare, the group communications director observed, would make colleagues think they had "been watching a world war two documentary" (P013). In financial services the same phenomenon arrived as "customer noise, complaint trend, social media heat... WhatsApp rumour... agent feedback" (P014), while a public-relations consultant described comparable activity simply as "negative marketing" (P008). Seven of the fourteen went further and explicitly framed the formal term as ill-fitting for a commercial setting, associating "digital influence operation" with government, politics or the military, not with their own organisation (P003, P005, P010, P011, P012, P013, P014).

The variation has practical consequences: each participant assimilated the phenomenon into the risk category their function already owns, so that the same incident is constituted as a different kind of problem, and described in those terms, depending on who first encounters it. Four participants made this consequence explicit, linking the vocabulary problem directly to fragmented ownership (P004, P010, P011, P014). The strategic influence practitioner pinned it down: where public relations frames the issue as a media matter, legal as defamation, cyber as out of scope and risk as an unplaceable register entry, "you've already lost time", because "the language problem becomes an ownership problem" (P010). The financial-services director saw the same mechanism from the customer-facing edge, where complaint-volume measurement could miss coordinated activity "under different topic, and maybe too late" (P014). Vocabulary, in this reading, is the first point at which detection either converges on a shared object or fragments into competing departmental ones.

A complication arises within the dataset: the specialist vocabulary that participants with security and military backgrounds imported, information warfare, hybrid threats, coordinated inauthentic behaviour, is itself a poor fit for the commercial object, importing a national-security frame onto a firm whose exposure runs through customers, retailers, regulators and

licence to operate. The effect is to substitute one ill-fitting register for another. The technology executive described the discomfort: an adviser used the phrase information warfare and "everyone kind of flinched... People don't want to sit in a tech board meeting and say, we're facing information warfare. They feel silly" (P011).

Divergent cases bound the claim. Prior awareness was low or latent in five of the fourteen accounts. They included senior commercial leaders in energy, consumer goods and financial services who recognised the behaviours instantly but had no framing for them (P001, P005, P012, P013, P014). The least exposed participant, a technology company founder, rated personal awareness at three or four out of ten and had never regarded the firm as a plausible target (P001). This is consistent with the theme, not contrary to it: where there is no felt exposure, there is no occasion to develop or borrow a vocabulary, and the phenomenon remains latent. Across the fourteen accounts, recognition follows lived encounter far more closely than sector.

5.2. Maliciousness beyond falsehood: perceived risk

Participants did not, on the whole, define malicious digital influence operations by the truth or falsity of the material involved. Ten of the fourteen located maliciousness primarily in conduct and effect, that is, in intent, the identity and provenance of the source, the methods used, proportionality, persistence and foreseeable harm, with the truth of a statement secondary (P002, P003, P005, P006, P007, P010, P011, P012, P013, P014). Eight articulated an explicitly multi-dimensional judgement to that effect (P003, P004, P005, P006, P007, P010, P011, P014). The strategic influence practitioner set out the fullest version, naming "intent, identity, method, proportionality and harm" and rejecting the truth-versus-lie distinction as "too simple, and too crude" (P010); a serving police commissioner added two recurring practical thresholds, that the line is crossed where activity leaves protected speech for threats, incitement or intrusion, and where a campaign persists after the target has sought to have it stop (P007). Others compressed the same judgement: "targeting to harm" (P011); "hostile is different, it is when the purpose not to solve, purpose is to damage" (P014).

The practical implication is that truthful material is routinely used as a weapon, and participants found attacks built on true material harder to respond to than outright fabrication, because the material itself cannot be denied. Seven of the fourteen volunteered a concrete example of true material weaponised (P005, P009, P010, P011, P012, P013, P014), and the examples span every commercial sector in the dataset. In aviation, an accurate video of an honest frontline mistake was amplified through coordinated sharing to portray the company as taking sides in a geopolitical conflict, which the participant judged "not actually misinformation, but... malicious" (P009). In energy, a genuine oil spill belonging to another operator was attached to the participant's firm by association: "Was it our spill? No. Did that matter online? Hell no" (P012). In technology, a board member's real conference remark was edited to imply pride in surveillance, so that "his words were real, but... the meaning they created was so false" (P011). These accounts occupy what the study frames as the malinformation register: the harm is produced by selection, sequencing and association applied to true material, so the activity is invisible to any organisational check that asks only whether a statement is false.

Not every participant took this view, and the exceptions follow a pattern of their own. Three participants placed maliciousness primarily in falsity, forming a truth-centred pole (P001, P004, P009). The start-up founder argued that "the truth is the truth. I don't think that's malicious" (P001); the aviation communications director reasoned explicitly from defamation principles,

treating truth as the ultimate defence, even while supplying one of the strongest malinformation examples in the set (P009); and the information-operations consultant, though accepting that truthful information can be weaponised, remained "not sure" whether true-but-hostile activity crosses the line (P004). One further participant came at the question through the difficulty of distinguishing satire from attack, without committing to either pole (P008). The distribution of the poles is the informative part. Those who weighted conduct and effect were, without exception, closest to live operational consequence: the practitioner who runs influence programmes, the executive whose firm had been targeted, and the energy and financial-services leaders who deal with the consequences as they happen. The three at the truth-centred pole reasoned from detached or explicitly legal vantages. This bears on practice because organisational escalation rules are typically written in the legalistic register, and criteria calibrated to falsity will systematically under-recognise the malinformation activity that the operationally exposed participants identified as the harder and more common case.

Participants gave an equally strong warning in the opposite direction, and it forms part of the finding. Five participants warned, unprompted, against treating all criticism as hostile (P004, P007, P010, P012, P013). The energy executive insisted that "some of it's fair... if we screw up, people should come after us", and cautioned against letting firms use the word malicious "as a get-out-of-jail card" (P012). For the strategic influence practitioner the point reduced to a working rule: "not every criticism is an attack. All viral posts aren't coordinated. But patterns matter" (P010). The finding is that the operative distinction is multi-dimensional, and not a licence to read hostility everywhere, that intent and provenance are frequently concealed at the moment a response decision must already be taken, and that a protection model resting on fact-checking alone will fail on the malinformation cases participants found most damaging.

5.3. Reputation as the gateway to enterprise harm: perceived impact

All fourteen participants named reputation as the primary harm, and seven treated it explicitly as a channel into wider enterprise consequence, a reading the analysis follows (P003, P005, P010, P011, P012, P013, P014). Reputation transmits hostile narrative activity into material enterprise consequence, and the consequence is sector-specific. The strategic influence practitioner stated it directly, "reputation is the transmission mechanism", before mapping destinations: investor confidence for a listed company, customer and retailer confidence for a consumer brand, licence to operate for an energy company, recruitment and regulatory attention for a technology firm, government confidence for a defence supplier (P010). The energy executive arrived at the same structure from inside his own sector, placing licence to operate at the centre and reputation as its feeder: "Licence to operate. Start there. Reputation matters because it feeds it" (P012). Dislike, in his account, became consequential only when it "turns into a permit delay, that's money... if it makes a bank put another review in the way, that's capital" (P012).

Three features of how the damage spreads distinguish this threat from the reputational risks organisations already manage. First, it cascades: eight participants described the damage spreading well beyond a single reputational hit (P001, P005, P007, P010, P011, P012, P013, P014). The financial-services director traced a chain from customer confidence to retention, frontline pressure, agent confidence, regulatory attention and employee wellbeing (P014); a public-sector director summarised it as touching "money, process, people" until "I don't think anything survives not being touched by a nasty campaign" (P005). Second, the cascade frequently runs

quietly, which defeats conventional detection; the three participants who described this most directly located the early damage in withdrawal and hesitation, with no visible complaint to detect (P011, P012, P014). In financial services, customers "don't always shout, they quietly stop using app... delay renewal, ask agent for cancellation", so that in the dashboard "you see later, not same day, but damage already start" (P014).

Third, and least anticipated by participants before they had experienced it, the gateway can open onto the physical safety and wellbeing of named individuals. Eight participants identified acute impact on named individuals as a distinct harm (P002, P007, P008, P010, P011, P012, P013, P014), and in four of these the harm extended explicitly to physical threat (P007, P008, P011, P012). The police commissioner located a growing share of harm in the movement from online to physical, through doxing, swatting and threats against officials and employees, in a feedback loop of online grievance, physical act and online approval (P007). The technology executive found the campaign most plainly hostile where it named junior, non-public staff as "enablers of genocide", which he characterised as intimidation whatever the language used (P011). At this point, as the strategic influence practitioner noted, the organisation acquires "a duty of care issue, not just a comms issue" (P010), and the impact of malicious influence on employee wellbeing emerged across the fourteen accounts as consistently underestimated.

Together, these accounts reframe what protection means: an organisation is protected to the extent that it can prevent narrative disturbance from propagating through reputation into customer behaviour, investor confidence, regulatory attention, staff safety, operational cost and licence to operate. On this evidence the firms most exposed were those whose product is itself trust, in financial services, consumer goods and the social licence of the energy sector, where the harm propagated fastest and most quietly. Capability did not confer immunity, a point five participants made directly (P002, P004, P010, P011, P012). As the energy executive put it, strong lawyers, strong public relations and good intelligence people do not settle the matter, because "you can still get your ass handed to you if the story runs faster than the facts" (P012).

5.4. Functions exist; integrated preparedness does not: responsibility, ownership and preparedness

Preparedness is the strongest theme in the data and gives the most direct answer to the research question. Twelve of the fourteen participants described organisations as possessing relevant functions in abundance while lacking any capability that connects them, so that the threat passes through the gaps between detection, interpretation, ownership and response (P002, P003, P004, P005, P006, P008, P009, P010, P011, P012, P013, P014). The financial-services director gave the theme its sharpest formulation. Their firm ran dedicated teams for customer success, service, digital product, the agency channel, complaints, compliance, risk, communications and IT security, yet "online influence problem goes across all, and that is where we are weak", because "who owns?... It depends who is loudest, maybe". Their conclusion located the failure in architecture: "Coordination is not bad people, it is bad structure" (P014). The strategic influence practitioner described the same condition in larger organisations as "dedicated teams for pieces of the problem, but not for the whole problem" (P010). Four mechanisms recur across the dataset and together specify the missing capability.

The first is the structural separation of detection from authority. In all fourteen accounts, first detection lies at the organisational periphery while the authority to act lies at the centre: the first signal reached customer success and agents in financial services (P014), customer care and

social teams in consumer goods (P013), the social-accounts manager in the targeted technology firm (P011), and the public before the company itself in the national-security consultant's experience (P006). Five participants named the failure that follows, that the periphery sees first but cannot escalate because no threshold or route exists (P008, P010, P011, P013, P014). The strategic influence practitioner drew the implication plainly: "a junior social media manager may see the first signals of what becomes a board level crisis" (P010). The consumer-goods director warned that "you can't give someone a national brand issue and a Canva login and wish them luck" (P013). Detection capacity in these accounts is genuine and structurally misplaced relative to the authority required to use it.

The second mechanism is the mismatch between the organisation's rehearsed incident architecture and the shape of this threat, named by five participants (P003, P010, P012, P013, P014). Each distinguished bounded, attributable, technically defined incidents from ambiguous, contested-meaning ones. The energy executive named the distinction directly: for a "clear crisis, we're pretty good... everybody knows the drill", but the hard case is "slow heat", a story bubbling up "bit by bit" from posts to local news to an NGO to a regulator (P012). The strategic influence practitioner generalised the pattern, that organisations "rehearse cyber incidents... product recalls. But they rarely rehearse a coordinated narrative attack that sits across all of those domains" (P010). A threat defined by ambiguity and reach across domains meets an incident architecture built for bounded, single-domain events.

The third mechanism is the distortion introduced by single-function ownership, which mirrors the assimilation problem identified in the first theme and was named directly by three participants (P010, P011, P013). Where ownership defaults to one function, the problem is refracted into that function's register and the other dimensions are lost. The strategic influence practitioner set out the full distortion: "If it sits only in comms, it becomes reputation. If it sits only in cyber, it becomes technical. If it sits only in legal, it becomes evidential. If it sits only in security, it becomes threat-led but may miss brand dynamics" (P010). The technology executive identified placing the problem with communications alone as the core error, because "sometimes the answer is not a public statement": customer reassurance, internal awareness, evidence preservation or deliberate silence may serve instead (P011). Whoever owns the response, on this evidence, determines what kind of problem it is taken to be, which is why the ten participants who addressed ideal ownership converged on senior, cross-functional ownership in place of any single contributing function, most often enterprise risk or resilience with one owner of sufficient weight to convene across functions (P002, P003, P004, P005, P008, P010, P011, P012, P013, P014).

The fourth mechanism is that the signal degrades as it travels from periphery to centre, and that the organisation's tempo cannot match the threat's; five participants described one or both aspects (P008, P011, P012, P013, P014). A head-office English summary, the financial-services director explained, flattens local-language customer sentiment, so that material that is "very sarcastic or angry" reaches the centre as "negative customer sentiment", which "sounds too soft", with the result that "if senior people don't feel the heat, they don't move fast" (P014); the technology executive found board language accurate but "a bit weak" relative to a reality that needed stating "plainly": "someone is trying to make your company radioactive" (P011). The consumer-goods director added the tempo mismatch: "rumours move in minutes. Companies move through approvals... by the time we've agreed the commas in a statement, the internet's

already built a bonfire" (P013). Preparedness, on this reading, is partly a matter of pre-authorised decision rights and rehearsed escalation, alongside capability and monitoring.

Two findings that span the themes support the central claim. First, monitoring is consistently mistaken for preparedness, a distinction four participants drew explicitly (P006, P009, P010, P014): firms "are buying monitoring tools, but monitoring is not preparedness. Seeing the wave coming is not the same as knowing what to do when it hits" (P010); the aviation director described a mature suite of monitoring tools, twice-daily cuttings, alerts and social listening, behind which sat no in-house ability to determine whether an image was fabricated or to attribute a campaign (P009); and "department exist is not same as readiness" (P014). Second, twelve of the fourteen could identify no authoritative external framework to fall back on (P002, P003, P004, P005, P006, P007, P009, P010, P011, P012, P013, P014). There is "no Bible on this" (P003); a targeted executive could find nothing that answered "what do you do Monday morning... first 24 hours. First 72 hours" (P011); "if we don't know what good looks like, we just react" (P014). The immaturity is compounded because organisations conceal these incidents, out of embarrassment in the technology case (P011) and out of operational secrecy in the energy case, with no reason to "hand the other side our operating manual" (P012), so the conditions for the peer learning that allowed cyber-security practice to mature are largely absent.

The theme therefore resolves the paper's question without a binary answer. The evidence does not support the conclusion that organisations are either protected or unprotected. Most have relevant departments, monitoring tools, crisis procedures and capable professionals. The difficulty is that malicious digital influence operations cut across the divisions those structures presuppose, crossing from customer issue to communications problem to legal exposure to staff safety to investor confidence before attribution is clear or a decision threshold has been met. The technology executive's reflection on his own targeted firm states the underlying condition plainly: the firm was "right on a lot of the facts", and "that did not solve the problem" (P011). Protection, on the evidence of these accounts, depends on the integration that connects the functions already present: a shared vocabulary that constitutes the threat as a single object, escalation thresholds that connect periphery to centre, interpretive and evidence-handling capacity alongside monitoring, pre-authorised decision rights matched to the threat's tempo, senior cross-functional ownership, and rehearsed response. The dataset identifies the absence of that integration as the locus of private-sector vulnerability, with awareness and monitoring technology already in place.

6. Discussion: The Private-Sector Preparedness Gap

The findings suggest that the private sector's exposure to malicious digital influence operations is not explained by an absence of awareness. Senior professionals across the dataset recognised many of the relevant behaviours. The larger problem is that these behaviours are not consistently integrated into a coherent organisational risk category. This produces a preparedness gap that is conceptual, functional and operational.

At the conceptual level, malicious digital influence operations are frequently absorbed into adjacent language. Participants rendered the issue variously as disinformation, information warfare, customer noise or licence-to-operate risk. Several participants showed that the most damaging activity may not be wholly false: true material used selectively, genuine events attached to the wrong organisation, partial evidence amplified without context, or private

information deployed coercively. Organisations therefore need a wider assessment model than truth-versus-falsehood alone.

At the functional level, the analysis indicates that malicious influence activity cuts across organisational divisions. It may be first detected by customer service, social media teams, frontline staff or local country teams. Yet the consequences may require legal judgement, security assessment, regulatory engagement or executive decision-making. This creates a mismatch between where signals appear and where authority lies. The customer-success account was clear that complaints, outages and data breaches have established processes, whereas rumours spreading and customer confidence declining lack clear ownership.

At the operational level, preparedness is weakened by the evidence-action gap. Organisations may need to act before attribution, intent, sponsorship or coordination can be proven. Legal teams may hesitate to use the term disinformation without evidence of deliberateness. Communications teams may seek rapid rebuttal, while risk and security teams may require a more complete threat picture. The difficulty is compounded by the absence of disinformation from the formal cybersecurity threat taxonomies that ordinarily structure how security, risk and audit functions classify incidents and authorise response [11]: a security team operating under standards that never name influence activity is unlikely to treat narrative disruption as the same class of problem as a phishing campaign, even where the operational logic is structurally similar. The gap is acute because digital influence activity can move faster than corporate verification: in FMCG, product fear may spread while batch records are still being reviewed.

The findings therefore challenge the assumption that organisational protection can be inferred from the existence of relevant departments. A company may have communications, legal, cyber, risk, public affairs, customer service, compliance and security functions, yet still lack integrated preparedness. The missing capability is an integrated operating model that links vocabulary, early-warning indicators, escalation thresholds, evidence handling, senior ownership and rehearsed cross-functional response.

The private-sector preparedness gap is therefore best understood as an enterprise governance problem: malicious digital influence operations exceed any single corporate function. Communications alone cannot contain them, because they may affect security, regulation, investor confidence and staff wellbeing; legal alone cannot, because lawful or partially true material can still be used coercively; cyber alone cannot, because there may be no system compromise; and customer teams alone cannot, because customer anxiety may signal a wider hostile narrative. The same governance question, who is responsible, recurs within the socio-technical IS community: Kennedy's fourteen-interview study of insider risk finds the same dispersal of responsibility across departmental silos and concludes that effective protection requires converged, enterprise-level ownership supported by a governance model, the counterpart of the single senior cross-functional owner this study's participants converged on [33]. Wolff draws the same line from the practitioner side: boards which locate the problem in the political sphere or treat it as a matter for IT misclassify a threat whose object is the social subsystem [12, chs. 8–9]. Protection depends on connecting these perspectives quickly enough to interpret the threat before disturbance becomes enterprise harm. The reputation-management literature treats corporate reputation both as a managed asset shaped by technology-mediated stakeholder engagement [22] and as a target of the deliberate degradation Shiraev, Keohane, Icks and Samoilenko [21] examine as character assassination. Reputation

operates as the gateway through which hostile narrative activity becomes material enterprise consequence, an exposure falling heaviest on trust-dependent sectors, with independent evidence identifying retail and banking as the commercial sectors most extensively affected [10], the sectors most prominent in this study's accounts of trust-pathway harm. The crisis-management literature reinforces the point from the response side: Akhgar, Staniforth and Waddington [18] document social media as both the channel through which organisations must detect emerging crises and the medium through which they are expected to respond, with the information processing behind timely response itself degraded under crisis conditions.

These observations can be drawn together through the socio-technical theory set out in Section 4. Trist and Bamforth's study of mechanised coal-mining showed that imposing a technological system on a social structure without regard to their interdependence degrades both, and required that "a social as well as a technological whole" be brought into existence [25, p.37]; Orlikowski adds that technology "can only condition social practices" and cannot determine them [26, p.411], so a technical countermeasure cannot of itself secure an organisational outcome. The field's own sociotechnical turn makes the same case for information threats: Sanfilippo, Zhu and Yang characterise misinformation as "sociotechnically produced" and requiring a correspondingly sociotechnical response [15, p.289], and Zhu and Yang frame it as a composite of human actors, ICT and information requiring jointly engaged countermeasures, though their framework takes the state as principal agent, leaving the firm's own preparedness unexamined [28]. Malicious digital influence operations are, in these terms, a predominantly social-subsystem phenomenon, acting upon perception, narrative and stakeholder trust; yet where organisations engage them at all, they tend to do so through technical or single-function instruments that attend to the technical subsystem while leaving the social dimension comparatively unmanaged. Malatji, Von Solms and Marnewick identify exactly this exposure within the security domain, where technical over-emphasis opens a corresponding social-side gap [6], and Mahmood, Chadhar and Firmin extend the perspective into organisational cybersecurity; their object is cyber defence, so no direct evidence on influence operations follows, yet the direction favours enlarging the social subsystem over further technical control [27].

Baxter and Sommerville locate the deeper cause of complex-system failure in organisation, not technology. Large systems fail, they argue, because those responsible do not recognise the social and organisational complexity of the environment into which the system is deployed [9, p.10], and because the separate teams and processes that ought to be coordinated are typically only weakly connected [9, p.11]. This matches the fragmentation documented in the findings: detection placed with customer-facing and frontline functions, interpretation dispersed across communications, legal and security, and authority residing with executives who receive signals late. The financial-services participant's formulation that coordination failure "is not bad people, it is bad structure" (P014) is, in socio-technical terms, a statement that the social and technical subsystems are not being jointly optimised. The fragmentation is the mechanism through which the threat succeeds.

The socio-technical reading also reframes the preparedness deficit as a capability problem. Dahmen's three-phase account of organisational resilience, the capacity to anticipate, to cope and to adapt [5], read with the socio-technical emphasis on cross-functional integration, specifies what an adequately prepared organisation would possess: an anticipatory capability that admits hostile narrative activity as a foreseeable risk; a coping capability that connects

detection to interpretation and authority in near real time; and an adaptive, learning capability that converts each incident into revised thresholds and rehearsed response. Baxter and Sommerville's proposed remedy, bridging activities that integrate otherwise separate processes [9], describes the integrated operating model the findings show to be absent: shared vocabulary, agreed escalation thresholds, evidence-handling protocols, senior ownership and rehearsed cross-functional response. The private-sector preparedness gap is, at root, a failure of socio-technical integration: a failure to optimise jointly the social and technical subsystems through which a hostile narrative must be detected, interpreted, owned and answered.

The implication for practice is that organisations should develop influence-risk preparedness as a cross-functional capability. Few organisations need a large specialist team for this; what every organisation needs is a shared vocabulary, defined escalation routes, agreed ownership, frontline reporting mechanisms, monitoring of narrative indicators, local-language interpretation where relevant, and scenario exercises built around coordinated narrative attacks as well as cyber incidents and conventional crises. Culloty and Suiter [20] model disinformation as a process spanning bad actors, platforms, audiences and countermeasures: a response addressed to only one dimension leaves the rest unmanaged. For higher-exposure sectors, including energy, technology, finance, aviation, FMCG and critical infrastructure, malicious digital influence operations should be treated as a foreseeable business risk instead of an exceptional communications event.

7. Contribution, Limitations and Future Research

This paper makes three contributions. First, it extends analysis of malicious digital influence operations beyond the dominant focus on state, electoral and democratic harms by examining the private sector as an organisational risk environment. The findings suggest that commercial organisations may encounter malicious influence activity through customer trust, product confidence, staff targeting, investor concern, regulatory scrutiny and licence-to-operate challenges.

Second, the paper contributes an organisational preparedness perspective. Senior professionals often recognise elements of the threat but not always as a distinct risk category, interpreting the issue as disinformation, reputation risk, product scare, activist pressure, fraud or hostile narrative. This conceptual fragmentation feeds directly into escalation, ownership and response, and the finding that "the language problem becomes an ownership problem" is central to the paper's contribution (P010).

Third, the paper identifies integrated preparedness as the key protection gap. Relevant functions often exist but may not be connected through shared terminology, agreed thresholds, evidence-handling processes, senior ownership or rehearsed response. The financial-services account captured this by distinguishing departmental presence from operational readiness (P014). Protection from malicious digital influence operations is, on this evidence, an enterprise governance capability above any singular communications, legal, cyber or security task. Independent quantitative evidence reinforces the pattern: the Spanish content analysis documents an organisational response in only 52.6 per cent of 630 fact-checked cases, predominantly reactive and fragmented [10]. The convergence of interview-based and content-analytic evidence from independent contexts indicates a structural integrated-preparedness deficit.

The sample is purposive and qualitative, so the findings are not statistically generalisable across the private sector; participants provided professional perceptions and experience, not independently verified incident records. The breadth of sectors strengthens the exploratory value of the study while limiting precise like-for-like comparison between industries, and the terminology itself remains unsettled.

Given the interpretive design, a note on reflexivity is warranted. The researcher is a practitioner in security intelligence and informational risk, and this professional proximity aided access and rapport and informed the reading of participant accounts, while bringing the risk of importing prior assumptions. This was mitigated through reflexive memoing, attention to divergent and negative cases, and an audit trail linking each interpretation back to coded extracts.

The qualitative themes established through this fourteen-participant study are designed to serve as the structural variables and hypotheses for a large-scale, cross-sector questionnaire instrument in the next phase of the research. That phase will examine whether the patterns identified here appear across a broader population of organisations. Further work should also examine sector-specific preparedness models for trust-dependent sectors such as finance, healthcare, technology, FMCG, aviation, energy and other critical infrastructure-adjacent industries. A later step is a preparedness maturity model covering detection, interpretation, escalation, governance, response and rehearsal.

8. Conclusion

This paper argues that private-sector protection from malicious digital influence operations remains uneven and underdeveloped. The challenge is organisational as much as technological or communicative; framed in socio-technical terms, it is a failure of joint optimisation between the social subsystem on which malicious influence operates and the technical subsystems through which organisations defend [6, 9]. The same craft that organisations use to project influence outward, through marketing and public relations, is being turned against them by adversaries who borrow its mechanics, framing, amplification, targeting and narrative construction, and deploy them deceptively, coercively and at times unlawfully. Recognising that the discipline of shaping perception is now exercised against the organisation as well as by it is the step that moves protection from malicious digital influence from a peripheral communications concern to a core enterprise capability.

Acknowledgements

The author thanks the fourteen senior professionals who participated in this study for their time and candour.

Declaration on Generative AI

During the preparation of this work, the author used Grammarly and Claude (Anthropic) to check grammar and spelling. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the publication's content.

References

[1] E. Broda, J. Strömbäck, Misinformation, disinformation, and fake news: lessons from an interdisciplinary, systematic literature review. Annals of the International Communication Association 48(2) (2024) 139–166. doi:10.1080/23808985.2024.2323736.

[2] C. Wardle, Information disorder: the essential glossary, 2018. (As cited in [1], pp. 144–145.)

[3] J. Marquez-Tejon, M. Jimenez-Partearroyo, D. Benito-Osorio, Security as a key contributor to organisational resilience: a bibliometric analysis of enterprise security risk management. Security Journal 35 (2022) 600–627. doi:10.1057/s41284-021-00292-4.

[4] D. Freelon, C. Wells, Disinformation as political communication. Political Communication 37(2) (2020) 145–156. doi:10.1080/10584609.2020.1723755.

[5] P. Dahmen, Organizational resilience as a key property of enterprise risk management in response to novel and severe crisis events. Risk Management and Insurance Review 26(2) (2023) 203–245. doi:10.1111/rmir.12245.

[6] M. Malatji, S. Von Solms, A. Marnewick, Socio-technical systems cybersecurity framework. Information & Computer Security 27(2) (2019) 233–272. doi:10.1108/ICS-03-2018-0031.

[7] A. López-Borrull, C. Lopezosa, Mapping the impact of generative AI on disinformation: insights from a scoping review. Publications 13(3) (2025) 33. doi:10.3390/publications13030033.

[8] P. Akhtar, T. Papadopoulos, H. Khan, A.M. Ghouri, S. Shamim, A. Ashraf, Risk, generative AI, disinformation control, global supply chains, and social impacts. Technological Forecasting & Social Change 226 (2026) 124609. doi:10.1016/j.techfore.2026.124609.

[9] G. Baxter, I. Sommerville, Socio-technical systems: from design methods to systems engineering. Interacting with Computers 23(1) (2011) 4–17. doi:10.1016/j.intcom.2010.07.003.

[10] L. Rodríguez-Fernández, I. Robles Marquez, A. M. Rodríguez de Austria Giménez de Aragón, Disinformation targeting companies and institutions and their response strategies: a longitudinal analysis of fact-checking in Spain (2023–2025). Revista de Comunicación 25(1) (2026) 343–362. doi:10.26441/RC25.1-2026-4171.

[11] K. M. Caramancion, Y. Li, E. Dubois, E. S. Jung, The missing case of disinformation from the cybersecurity risk continuum: a comparative assessment of disinformation with other cyber threats. Data 7(4) (2022) 49. doi:10.3390/data7040049.

[12] U. Wolff, Disinformation Attacks on Businesses: The Dark Business of Fake News & Co. and How to Combat It. Management for Professionals. Springer Gabler, Wiesbaden (2026). ISBN 978-3-658-51406-8.

[13] P. N. Petratos, Misinformation, disinformation, and fake news: cyber risks to business. Business Horizons 64(6) (2021) 763–774. doi:10.1016/j.bushor.2021.07.012.

[14] A. Malhotra, N. Evans, J. Gao, J. T. Du, C. Zheng, The issues caused by misinformation—how workers and organizations deal with it: a systematic literature review. Journal of the Association for Information Science and Technology 77(1) (2026) 133–162. doi:10.1002/asi.25016.

[15] M. R. Sanfilippo, X. A. Zhu, S. Yang, Sociotechnical governance of misinformation: an Annual Review of Information Science and Technology (ARIST) paper. Journal of the Association for Information Science and Technology 76(1) (2025) 289–325. doi:10.1002/asi.24953.

[16] C. Wardle, H. Derakhshan, Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making. Council of Europe report DGI(2017)09. Council of Europe, Strasbourg (2017).
[17] J. J. F. Forest, Digital Influence Mercenaries: Profits and Power Through Information Warfare. Naval Institute Press, Annapolis, MD (2022). ISBN 978-1-68247-722-9.
[18] B. Akhgar, A. Staniforth, D. Waddington (Eds.), Application of Social Media in Crisis Management. Advanced Sciences and Technologies for Security Applications. Springer, Cham (2017). doi:10.1007/978-3-319-52419-1.
[19] S. Boudana, E. Segev, Fake news makes the news: definitions and framing of fake news in mainstream media. Journalism Practice 20(2) (2026) 679–698. doi:10.1080/17512786.2024.2379898.
[20] E. Culloty, J. Suiter, Disinformation and Manipulation in Digital Media: Information Pathologies. Routledge Focus on Communication and Society. Routledge, London (2021).
[21] E. B. Shiraev, J. Keohane, M. Icks, S. A. Samoilenko, Character Assassination and Reputation Management: Theory and Applications. Routledge, New York (2022). ISBN 978-1-138-60917-4.
[22] T. C. Melewar, C. Dennis, P. Foroudi (Eds.), Building Corporate Identity, Image and Reputation in the Digital Era. Routledge Studies in Marketing. Routledge, London (2022).
[23] P. Baines, N. O'Shaughnessy, N. Snow (Eds.), The SAGE Handbook of Propaganda. SAGE Publications, London (2020).
[24] L. Olejnik, Propaganda: From Disinformation and Influence to Operations and Information Warfare. CRC Press, Boca Raton, FL (2025). ISBN 978-1-032-80225-1.
[25] E. L. Trist, K. W. Bamforth, Some social and psychological consequences of the longwall method of coal-getting. Human Relations 4(1) (1951) 3–38.
[26] W. J. Orlikowski, The duality of technology: rethinking the concept of technology in organizations. Organization Science 3(3) (1992) 398–427.
[27] S. Mahmood, M. Chadhar, S. Firmin, Addressing cybersecurity challenges in times of crisis: extending the sociotechnical systems perspective. Applied Sciences 14(24) (2024) 11610. doi:10.3390/app142411610.
[28] X. Zhu, S. Yang, Toward a Sociotechnical Framework for Misinformation Policy Analysis. In: Yang, S., Zhu, X., Fichman, P. (Eds.), The Usage and Impact of ICTs During the COVID-19 Pandemic. Routledge (2023) 11–45. doi:10.4324/9781003231769-3.
[29] M. Stone, E. Aravopoulou, G. Evans, E. Aldhaen, B. D. Parnell, From information mismanagement to misinformation — the dark side of information management. The Bottom Line 32(1) (2019) 47–70. doi:10.1108/BL-09-2018-0043.
[30] V. Braun, V. Clarke, Using thematic analysis in psychology. Qualitative Research in Psychology 3(2) (2006) 77–101. doi:10.1191/1478088706qp063oa.
[31] J. Fereday, E. Muir-Cochrane, Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. International Journal of Qualitative Methods 5(1) (2006) 80–92. doi:10.1177/160940690600500107.
[32] N. K. Gale, G. Heath, E. Cameron, S. Rashid, S. Redwood, Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Medical Research Methodology 13 (2013) 117. doi:10.1186/1471-2288-13-117.
[33] R. Kennedy, Managing insider risk 'entities' as the route to security convergence. In: Proceedings of the 11th International Workshop on Socio-Technical Perspectives in IS

(STPIS'25), Skopje, North Macedonia. CEUR Workshop Proceedings, vol. 4134. CEUR-WS.org (2025). Available at: https://ceur-ws.org/Vol-4134/paper3.pdf.