Resilience-by-Design: Tabletop Evidence on AI-Enabled Cybercrime, Coordination, and Public Trust

Gil Baram

May 2026

Abstract

Artificial intelligence is widely described as a force multiplier for cybercrime, yet scholarship and policy debate remain narrowly focused on discrete techniques such as AI-generated phishing or malicious code. This paper argues that AI-enabled cybercrime is better understood as an ecosystem-level transformation that compresses decision timelines, amplifies ambiguity, and erodes the boundaries between criminal, state-linked, and hybrid actors. The cumulative effect is sustained pressure on the institutional architecture that underpins contemporary cybersecurity governance.

The argument draws on three structured tabletop exercises conducted across the United States, Singapore, and Israel, with senior participants from government, law enforcement, critical infrastructure, technology firms, and civil society. Analysis followed reflexive thematic analysis with dual-rapporteur observational triangulation and hotwash participant validation. A consistent finding emerged across sites: failures were driven less by inadequate technical detection than by governance breakdowns under uncertainty, as participants repeatedly faced high-stakes decisions about attribution, disclosure, escalation, and public messaging before they could verify the underlying facts.

From these patterns, the paper identifies three structural deficits of the current regulatory architecture: accountability gaps, coordination deficits, and decision-right ambiguity. We argue that addressing them requires a shift from reactive incident response toward a resilience-by-design posture that builds coordination capacity, decision-right clarity, and public trust into institutions before crisis rather than improvising them under pressure. The paper contributes empirically grounded analysis to debates on AI governance, cyber norms, and the institutional conditions under which open societies can sustain public trust amid an accelerating threat landscape.

Keywords: AI-enabled cybercrime; ecosystems; tabletop exercises; accountability

1. Introduction

In 2025, the FBI's Internet Crime Complaint Center received more than 1 million complaints reporting losses exceeding $20 billion, a 24% increase from the previous year (Internet Crime Complaint Center, 2026). The scale of the problem is accelerating with AI: 90% of surveyed financial crime professionals observed an increase in AI-driven attacks, underscoring the degree to which AI capabilities are becoming embedded in criminal operations at scale (Nasdaq Verafin, 2026).

This escalation exemplifies a broader trend driven by the proliferation of artificial intelligence (AI) capabilities within the cybercriminal ecosystem. Attacks have become less costly to execute, more difficult to attribute, and are conducted at greater speed. AI tools now enable the large-scale generation of highly personalized phishing campaigns, the synthesis of convincing audio and video impersonations at minimal cost, and the automation of vulnerability exploitation that previously required specialized human expertise (Antebi et al., 2024; Kaloudi & Li, 2020; Guembe et al., 2022). Consequently, there has been an increase in the volume of cybercrime and an expansion of the perpetrator base. The emergence of Ransomware-as-a-Service (RaaS) and related crime-as-a-service (CaaS) models has broadened access to advanced capabilities, previously restricted to technically sophisticated actors. As a result, hospitals, municipalities, small businesses, and individuals have become viable targets for operators with limited expertise (Blauth et al., 2022).

A core paradox persists: research and policy often prioritize nation-state threats and worst-case technical scenarios, while common AI-driven attacks that affect ordinary users receive comparatively less governance attention. When AI-enabled attacks occur, existing frameworks often prove inadequate. Responders have limited time, may encounter fabricated evidence, and must contend with attackers who use automated networks that evade traditional attribution methods. The consequences are tangible. MacColl et al. (2024) show that ransomware inflicts cascading financial, reputational, and psychological harm, disrupting entire supply chains, healthcare systems, and local governments. Individuals uninvolved in the attack frequently suffer harm and are often unable to seek redress due to deficiencies in current accountability systems (Lusthaus, 2018; Wall, 2007).

We argue that AI-enabled cybercrime transforms the entire ecosystem by altering how threats emerge, propagate, and are addressed. It introduces speed, uncertainty, and confusion, thereby blurring the boundaries between legal actors upon whom accountability depends. To support this argument, data from three tabletop exercises (TTXs) conducted in 2024-2025 in the United States, Singapore, and

Israel revealed that while technical detection mechanisms were effective, governance structures failed under conditions of uncertainty. In all cases, participants were required to make critical decisions regarding attribution, disclosure, and escalation prior to confirming essential facts, often due to uncertainty exacerbated by AI-generated deception. Two research questions are at the heart of this paper: How do practitioners respond to AI-enabled cybercrime under operational uncertainty? What do these response patterns reveal about the structural fit between existing governance frameworks and the AI-enabled cybercrime ecosystem?

This paper makes three primary contributions. Empirically, it presents a multi-site, multi-jurisdiction TTX study that investigates AI-enabled cybercrime from a governance perspective. Conceptually, it advances the field by arguing that existing governance frameworks, which are often designed around static threat actors and discrete technical risks, are inadequate for addressing the dynamic, ambiguous, and ecosystem-driven challenges posed by AI-enabled attacks. This theoretical innovation redirects attention from optimizing technical defenses within institutional silos to developing adaptive governance structures that can manage uncertainty and interdependence in AI-mediated environments (Smith, 2025; Bohr, 2026). Methodologically, it demonstrates the value of structured TTXs as tools for studying emerging security governance challenges in contexts where real-world data are unavailable (Kävrestad et al., 2025).

The structure of this paper is as follows. Section 2 develops the framework; Section 3 sets out the TTX methodology and sites; Section 4 presents five patterns of governance failure; Section 5 synthesizes them into three structural deficits; Section 6 discusses implications and concludes.

2. Conceptual Framework: From Attack Taxonomies to Governance Questions

2.1 The Limits of the Technical Frame

The technical literature on AI-enabled cybercrime has developed taxonomies that illustrate how AI reshapes the attack surface. These include mapping techniques across kill chain phases (Kaloudi & Li, 2020), quantifying their distribution (Guembe et al., 2022), documenting the weaponization of generative models for credential synthesis and identity fabrication (Yamin et al., 2021), and proposing investigative frameworks that extend beyond existing matrices (Sarkar et al., 2023). This body of work

is empirically robust and operationally significant, providing defenders with insights into adversarial capabilities and identifying where AI offers the most substantial leverage within the attack chain.

However, this literature is scoped to threat intelligence rather than institutional response. Major frameworks, including the MITRE ATT&CK matrix (Strom et. al., 2018, MITRE Corporation, 2026) and the Lockheed Martin Cyber Kill Chain (Hutchins et al., 2011), catalog adversary behaviors to support detection and mitigation. They do not address the governance challenges these behaviors create for decision-makers: challenges that arise not from new attack vectors per se but from the conditions these vectors impose on the institutions responsible for response.

Three such conditions are systematically absent from the technical frame. The first is *accountability under synthetic deception*. When an AI-generated synthetic identity passes through multiple layers of institutional verification, including employment screening, financial authentication, and organizational access controls, the resulting failure cannot be attributed to a single technical vulnerability (Yamin et al., 2021). It implicates the verification standards of every institution involved, yet no current framework delineates how responsibility should be allocated across this chain. AI-enabled crime-as-a-service ecosystems compound this by fragmenting what was once a unified criminal act across networks of specialized providers (Blauth et al., 2022; MacColl et al., 2024).

The second condition concerns *coordination under jurisdictional ambiguity*. AI-enabled cybercrime frequently spans multiple sectors, jurisdictions, and regulatory regimes. A ransomware attack on a hospital, for instance, may activate regulations across healthcare, finance, law enforcement, and critical infrastructure, each involving distinct actors, timelines, and constraints on information sharing. This challenge predates AI but has been exacerbated by the characteristics of AI-driven attacks described in the following section.

The third condition pertains to *decision rights under uncertainty*. AI-enabled tools, including deepfakes, polymorphic malware, and AI-generated false flags weaponize ambiguity itself. Traditional cyberattacks produced forensic traces that, though sometimes incomplete, could indicate identifiable actors. AI-enabled attacks can fabricate such traces and impersonate other groups, rendering the notion of a 'responsible actor' analytically unstable and rise doubts about attribution (on the attribution and offensive cyber operations: Baram, 2023; Egloff & Smeets, 2021; Rid & Buchanan, 2015). Smith (2025), analyzing AI's effect on democratic governance, characterizes this as "persistent doubt," a sustained state of uncertainty that undermines the evidentiary basis on which decision-makers rely.

The core issue is not whether detection tools can identify AI-generated content but rather who possesses the authority to act, and on what basis, when available information cannot be fully trusted.

These three gaps, accountability, coordination, and decision rights, fall outside the intended scope of threat intelligence frameworks. Addressing them requires a conceptual vocabulary drawn from governance theory, institutional analysis, and research on organizational decision-making. Criminological scholarship has begun to delineate this boundary: the emerging field of AI governance examines how societies structure oversight, accountability, and decision-making authority in AI-shape environments (Nordström, 2022; Smith, 2025; Bohr, 2026). However, this body of work has not yet been systematically applied to the context of cybercrime.

2.2 Beyond Techniques: AI-Enabled Cybercrime as an Ecosystem-Level Transformation

Having identified three governance conditions absent from the technical frame, this section proposes a framework for understanding the structural forces that produce them. The prevailing technical perspective, which treats AI as a force multiplier for established attack types, captures certain realities but overlooks a structural dimension. The cumulative impact of AI-enabled capabilities is not an enhancement of individual techniques, but a transformation of the cybercrime ecology that challenges the foundational assumptions on which current governance systems rest. This paper identifies three interrelated transformations, each of which serves as a stressor on the institutions and frameworks tasked with managing cybercrime.

The first is *timeline compression*. AI accelerates the operational tempo of cyberattacks to a speed that surpasses the sequential decision-making processes of institutional responders. While traditional intrusion timelines allowed defenders hours or days to detect, attribute, and contain breaches, AI-assisted attack chains now reduce these windows to mere minutes (Guembe et al., 2022; Kaloudi & Li, 2020; Zero Day Clock, 2026). Institutional response mechanisms, such as legal authorization for countermeasures, inter-agency notification, executive escalation, and public communication, were not designed for this pace. Nordström (2022), analyzing AI's implications for public policy decision-making, notes that AI generates "great uncertainty" due to information scarcity and the rapid pace of events eliminates the deliberative time that institutional procedures require.

The second is *ambiguity amplification*. Deepfakes, synthetic identities, and false flags obscure attribution as an incidental byproduct and generate ambiguity as a strategic asset, turning the accountability and decision-rights gaps identified above into exploitable vulnerabilities. Whereas those governance gaps

describe what institutions struggle to resolve, ambiguity amplification describes the mechanism through which AI-enabled attackers deliberately intensify those struggles. The ambiguity produced is not a temporary information gap to be resolved through investigation but a designed feature of the attack itself.

The third is *actor boundary erosion*. The fragmented crime-as-a-service architecture has a structural consequence beyond the accountability problem it creates for legal systems: it erodes the boundaries between criminal, state-linked, and hybrid actors. When ransomware affiliates operating under one banner use infrastructure leased from another, deploy malware developed by a third, and launder proceeds through a fourth, the categories that governance systems rely on (organized crime versus state-sponsored threat versus lone actor) lose their classificatory utility. AI accelerates this erosion by enabling automated handoffs between ecosystem nodes, reducing the need for direct human coordination and making the overall architecture more fluid and less legible to investigators.

Together, these transformations systematically misalign existing governance frameworks with current cybercrime realities.

2.3 Governance Frameworks and Their Limits

The three transformations identified above expose a systematic mismatch between the architecture of AI-enabled cybercrime and the governance frameworks designed to counter it. Existing operational frameworks, including the NIST Cybersecurity Framework (CSF 2.0) (National Institute of Standards and Technology, 2024), the EU's NIS2 Directive (European Parliament and Council of the European Union, 2022), and ISO 22361's crisis management guidelines (International Organization for Standardization, 2022), share a common set of assumptions: sequential incident timelines, limited sectoral impacts, identifiable actors, and clear jurisdictional boundaries. These conditions held for the cybercrime landscape in which the frameworks were developed, and they retain utility when incidents involve attributable actors within a single jurisdiction. However, as Bohr (2026) observes, governance frameworks designed for human decision-making create "governance illusions" when outcomes result from machine coordination at speeds preventing meaningful human intervention (see Sarkar et al., 2023).

Regulatory approaches face an analogous structural limitation. Whether risk-classification-based (the EU AI Act; European Parliament and Council of the European Union, 2024; Ghosh et al., 2025), disclosure-based (the SEC's 72-hour cybersecurity incident reporting rule), or compliance-based

(sectoral guidance frameworks in Singapore and China), these approaches presuppose clearly attributable responsibility and identifiable regulated entities. These conditions do not hold for adversarial actors operating across borders (Sun et al., 2026). The architecture of AI-enabled crime is designed to distribute risk and accountability in ways that the regulatory pyramid model, in which graduated sanctions deter rational actors, cannot address.

Public-private coordination introduces an additional structural challenge. Effective governance requires rapid information exchange between private-sector organizations that hold real-time technical intelligence and government agencies that hold legal authority. The velocity of AI-enabled threats consistently outpaces the trust-building, legal authorization, and information-sharing protocols that connect them (Ghosh et al., 2025; Lin, 2025). Recent innovations, including CISA's Joint Cyber Defense Collaborative (Cybersecurity and Infrastructure Security Agency, 2025), sector-specific Information Sharing and Analysis Centers (U.S. Government Accountability Office, 2023) and ENISA's cross-sectoral threat assessments (European Union Agency for Cybersecurity, 2024), represent advances in baseline preparedness. Yet none resolves the decision-authority question that arises when multiple institutional functions activate concurrently under compressed timelines. The next sections present empirical evidence for this mismatch from three tabletop exercises across three jurisdictions.

3. Methodology

3.1 Research Design: Why Tabletop Exercises?

This study employs tabletop exercises (TTXs) as its primary research instrument. TTXs are structured, facilitated simulations that place expert participants in realistic decision-making scenarios under controlled conditions, requiring them to respond to an evolving crisis narrative through deliberation rather than operational deployment (Smith, Kollars, & Schechter, 2024). TTXs have a long tradition in security studies, emergency management, and political science, where they are used to generate structured behavioral data about decision-making under uncertainty (Reddie et. al., 2024). More recently, they have been adopted for cybersecurity governance research, where Vogt et al., (2025) describe the practice of "preparedness wargaming" – exercises designed to generate actionable governance insights through facilitated expert deliberation (Vogt et al., 2025: 186).

The choice of TTXs is deliberate. Surveys and interviews capture what respondents say they would do, not how they decide under simulated pressure. Case studies of actual incidents face severe practical

constraints: incident data is access-restricted, selectively disclosed, and biased toward resolved rather than silently failed cases (Lusthaus, 2018; MacColl et al., 2024), and post hoc reconstruction obscures the decisions considered but not taken, which is where governance failure is visible. Elite interviews remain subject to rationalization and socially desirable response patterns, particularly when failures of coordination or decision authority are involved. TTXs, by contrast, generate the urgency and uncertainty that are the phenomenon under study (Kovalsky et al., 2024) and force participants to confront the gap between planned responses and actual behavior (Kävrestad et al., 2025).

The wargaming literature has largely used TTXs for strategy, deterrence, and escalation in state-on-state contexts (Lin-Greenberg et al., 2022). This study extends the method to governance architecture: the question is not which decisions actors prefer but how existing institutional arrangements structure what is decidable under AI-enabled operational conditions. TTX data occupies a distinctive epistemological position: it is neither strictly experimental, since participants know they are in a simulation, nor purely observational, since scenarios are constructed to elicit specific decision conditions. Informed by the wargamer's trilemma framework (Reddie et al., 2024), which identifies trade-offs among analytical utility, contextual realism, and player engagement, TTX outputs are interpreted here as evidence of how participants reason within defined yet plausible settings, not as causal predictions of real-world outcomes. This aligns with the use of wargaming as a social science method for emerging security issues where real-world data are scarce (Lin-Greenberg et al., 2022).

A potential concern with TTX-based research is scenario-construction bias. This means the exercise design could embed the outcomes identified by the analysis. This study addresses that issue by separating scenario inputs from analytical outputs. The scenarios presented participants with attack conditions based on real-world incident types, like ransomware on critical infrastructure, supply chain compromise, and multi-sector cascades. They did not present governance dilemmas. The governance failure patterns in Section 4 arose from participants' behavior during deliberation: decision sequencing, coordination breakdowns across role groups, and unresolved authority disputes. These were not scripted into the scenario design. The scenario includes operational escalations (e.g., a media leak or a ransom demand), but it does not specify how participants should respond or which institutional

processes should be activated. The distinction between what happens to participants and how they respond is foundational to TTX methodology (Smith et. al., 2024; Lin-Greenberg et. al., 2022).

3.2 Data Collection: Three Sites, Three Jurisdictions

Three TTXs were held from December 2024 to December 2025 at three locations. Each examined AI-enabled cybercrime governance from a distinct institutional perspective. They all shared the research question: how do practitioners respond to AI-enabled cybercrime under operational uncertainty? The TTXs differed in format and in who took part, reflecting cross-country and cross-sector differences (see Appendices A, B).

The first exercise happened in December 2024 at a research university in the western US. About 25 people from business, law enforcement, universities, and policy attended for a day. The event started with an expert panel on current AI-enabled cybercrime threats. Next, participants participated in two rounds of simulations. In round 1, they acted as AI-using cybercriminals, either as a low-resource group or a well-resourced group close to a government. In round 2, they became private security providers handling a major cyber breach. To collect data, we observed, took notes, and later facilitated a group discussion.

The second exercise took place in October 2025 at a security policy institution in Singapore. The exercise had a panel discussion on AI-cybercrime threats, followed by a two-part simulation. In the first round, participants acted as a private company responding to a corporate breach, requiring forensic work, decision-making about responsibility, and advising clients amid uncertainty. The second round moved to a national scenario in which a government cybersecurity agency dealt with AI-generated malware targeting the energy, transport, and water sectors, along with AI-driven disinformation that caused public panic. Participants included experts from government, industry, and academia.

The third exercise was held in December 2025 as a simulation at a major international cybersecurity conference in Tel Aviv, Israel. The scenario centered on an insurance firm targeted by a supply-chain attack initiated through AI-generated deepfake social engineering directed at a supplier. The crisis escalated through four incidents: a supply-chain disconnect decision, full ransomware encryption with a 40 BTC demand, confirmation of data exfiltration, and a media leak. Participants assumed crisis management team roles: CEO, CISO/CIO, legal counsel, public relations, and business continuity. This exercise differed from the first two in format. Findings are treated accordingly in the analysis.

3.3 Data Analysis Procedures

To analyze these data from the exercises, the data sources included notes from two rapporteurs per exercise, notes from the principal investigator, post-exercise transcripts, and facilitator debrief materials from all three sites. We used reflexive thematic analysis, as described by Braun and Clarke (2006; 2019). Here, the researcher looks for themes by deeply examining the data, not by fitting it into pre-set categories. This works well for emerging research topics, such as AI-enabled cybercrime, where we do not yet know which ideas to look for (see Braun & Clarke, 2019, 594).

The analytical process proceeded in three phases. First, the PI reviewed the notes and observation records from each site independently to identify decision points, coordination breakdowns, and unresolved governance questions within each exercise. Second, patterns identified at each site were compared across exercises to distinguish site-specific dynamics from cross-jurisdictional regularities. Patterns that recurred across at least two of the three sites were treated as cross-site findings; patterns unique to a single site were reported as contextual observations. Third, the cross-site patterns were synthesized against the conceptual framework developed in Section 2 to generate the three structural deficits presented in Section 5.

Post-exercise hotwash discussions, in which participants reflected on their own deliberations, provided an initial form of participant validation, allowing the research team to test emerging observations against participants' own accounts. The cross-site convergence requirement, demanding recurrence across exercises that differed in format, jurisdiction, and participant composition, provides a structural robustness check that follows the logic of a most-different-systems design (Anckar, 2008): convergent patterns across heterogeneous contexts provide stronger evidence of structural dynamics than replication within a single design would.

Several limitations should be noted (see Appendix C). The analytical process was led by a single researcher (the principal investigator), who holds interpretive authority. Rapporteur notes, while structured by observation protocols, are not verbatim transcripts and may reflect rapporteurs' interpretive emphases. TTX data are behavioral data collected under simulated conditions; because participants know they are in a simulation, institutional inertia that would appear in real incidents may be suppressed. The Tel Aviv exercise differs in design from the other two exercises, limiting direct comparison. Its findings are treated as complementary rather than equivalent. All findings are treated as indicative of governance dynamics rather than as causal claims.

All participants agreed to take part before the exercises. They were informed that the exercises were for research purposes, that their discussions would be recorded in notes and observations, and that their names and organizations would not appear in reports. Data were stored on secure university servers for the team only. We used the Chatham House Rule, so we refer to people's comments by their roles, not by name, throughout this paper.

4. Findings

Five governance failure patterns emerged consistently across the three exercises. They are presented below with evidence from each site and a brief analytical connection to the conceptual framework developed in Section 2. These five were selected through the cross-site triangulation procedure described in Section 3.3. Five governance failure patterns emerged consistently across the three exercises. All participant remarks are cited under anonymized role labels, no individual names or institutional affiliations are disclosed.

4.1 Decisions Before Verification

Across the exercises, participants had to make high-stakes choices, like deciding on attribution, disclosure, escalation, and how to communicate publicly, before verifying critical facts. AI-enabled attacks made this harder by slowing the flow of information and actively increasing confusion through deepfakes, false flags, and fake evidence.

In the US exercise, the group faced a scenario where the criminal actor's identity could not be determined. One participant summarized the dilemma: Was the attack a nation-state-backed, purely criminal, or a mix of both? No clear answer was available, yet attribution was required before the incident could be contained. In a previous round, a task forced a cybersecurity firm to issue client guidance before finishing forensic attribution. A law enforcement participant noted that there are still no reliable tools to detect AI-augmented attacks. Even recognizing whether an attack used AI, let alone identifying the source, remains uncertain at the forensic level.

The Singapore exercise revealed a parallel challenge. In Round 1, corporate leadership demanded public guidance from the forensic team before the forensic team had determined the entry point or the scope of the compromise. Round 2 escalated the problem: AI-generated disinformation spread

claims that defense networks were compromised before authorities could verify or refute them. Participants faced a dilemma with no good option: responding immediately with unverified information risked making things worse if proved wrong, while waiting for verification risked allowing panic to grow. The decision had to be made quickly, but the needed information was not yet available.

In the Tel Aviv exercise, participants had to decide whether to pay a 40 BTC ransom while still confirming the scale of data theft. They did not know if 5TB of data had truly been stolen or what it held, all while under time pressure and facing misinformation about the data's sensitivity. This illustrates what Nordström (2022) calls decisions under great uncertainty: conditions in which AI-generated confusion shortens the time normally needed for careful decision-making.

4.2 Timeline Compression Overwhelming Institutional Process

AI reduces the operational tempo of attacks. It pushes it structurally below the threshold at which institutional response processes can function sequentially. The frameworks for notification, authorization, coordination, and escalation were designed for days-long timelines. In contrast, AI-enabled attack cycles operate on timelines measured in hours or less.

In the US exercise, an industry panel participant with 20 years of threat intelligence experience noted a trend. The average attack cycle of about five days might compress to just hours with AI-assisted coordination. This is especially true if autonomous agents are used in swarm-style configurations (Schroeder et. al., 2026). On the exploitation side, a new threat may appear within days of a new model release. The SEC's 72-hour cybersecurity incident disclosure rule was cited as structurally inadequate. It sets a notification deadline but does not provide guidance on what to report if the scope is unknown at the time of disclosure. As one panelist noted, many firms simply do not know how to respond within the required timeframe.

The Singapore exercise showed this compression at the national level. In Round 2, power disruptions were detected within one hour after the first anomaly. In that same hour, transport and water-critical infrastructure also reported issues. The national coordination architecture consisted of a central cybersecurity hub, sector-specific threat information platforms, and computer emergency response teams. It was designed for sequential incident management: one sector, one timeline, one escalation path.

In the Tel Aviv exercise, the entire crisis arc was compressed into a single simulation day. This simulated about 48 hours of real incident time. It included initial supply chain detection, ransomware

encryption, confirmation of exfiltration, a media leak, and the pay-or-refuse decision. Legal counsel's authorization, board notification demands, and regulatory disclosure obligations were all triggered at once. This simultaneous activation rendered any decision framework that depended on ordered escalation ineffective.

This pattern operationalizes Bohr's (2026) governance illusion: institutional interfaces (playbooks, escalation matrices, notification protocols) suggest sequential control, while AI-enabled attack coordination unfolds at speeds that render those interfaces non-functional.

4.3 Actor Boundary Erosion and Attribution Paralysis

AI-enabled cybercrime makes it unclear who is a responsible, attributable actor. The architecture of such crime distributes agency across automated systems, synthetic identities, and service-provider networks. This structure defeats the assumptions of existing attribution frameworks.

In the US exercise, the higher-resource criminal group's toolkit included AI-enabled attribution deception and false flag operations as standard capabilities. During the pre-exercise industry panel, a threat intelligence expert described North Korean operatives using AI-generated identities to pass background checks and secure employment at cybersecurity companies, thereby undermining the identity-verification infrastructure on which employment-based accountability rests.

The Singapore exercise highlighted a forward-looking aspect of the attribution problem. During panel deliberation, participants identified new evidentiary categories for AI-enabled crime investigation. These included AI interaction logs, prompt records, and guardrail trigger points. None of these categories is currently defined in law or standard forensic practice. Participants also observed that state attribution may be fundamentally less tractable than technical attribution, which focuses on identifying tools and techniques (see Lee, 2023). They recommended prioritizing technical attribution in early response phases. In the Round 2 false flag scenario, a foreign outlet falsely reported that defense networks had been compromised. This made correct attribution require discrediting a fabricated narrative spreading faster than any forensic process could keep up.

In the Tel Aviv exercise, a legal participant stated the core dilemma directly. It was unclear whether the incident was criminal or a national security matter. This determination directly affected the decision to pay the ransom. Payments may carry different legal implications depending on group affiliation. This determination was not resolved during the exercise. The supply chain structure created further

complexity. The insurance company suffered consequences from a breach started by its supplier. This raised unresolved questions about accountability for the initial failure.

This pattern extends Van der Wagen and Pieters' (2015) cyborg crime framework to the AI context. When AI agents execute consequential actions, such as deploying ransomware or posting exfiltrated data, they act on pre-set parameters. This occurs without specific human instruction at the moment of execution. The concept of criminal intent becomes unclear not for evidentiary but for architectural reasons.

4.4 Public-Private Coordination Failure

AI-enabled cybercrime incidents require simultaneous coordination between private-sector actors, who hold real-time technical intelligence, and government actors, who hold legal authority. Existing coordination architectures were designed for sequential and sector-specific incidents. Today, AI-speed, multi-sector cascades generate coordination failures. This occurs even in jurisdictions with well-established national cybersecurity frameworks.

In the US exercise, a mid-round inject required the CEO of a private cybersecurity firm to advise client companies while attribution and scope remained unresolved. The firm held forensic data; the government held legal authority to act on it. Neither could operate effectively alone, yet information-sharing protocols required attribution before data could be legally exchanged, creating a circular dependency.

The Singapore exercise demonstrated this failure. When transport and water critical information infrastructure were simultaneously affected in Round 2, the coordination architecture required separate notification and response tracks for each sector. Strategic partners began disconnecting from shared information networks immediately upon learning of the cascade, a rational decision that, collectively, destroyed the cooperative information environment when it was most needed. The absence of legal and communications professionals from the participant pool was visibly felt, as participants could not adequately model the legal authorization or public communication dimensions of the response.

The Tel Aviv exercise further highlighted these issues. The simultaneous activation of the CEO, CISO/CIO, legal, PR, and business continuity roles produced competing inputs without a unified decision framework. The PR function's priority – to protect institutional reputation before it is irrecoverably lost – was structurally at odds with the legal function's and the CISO's priorities (contain

and investigate before disclosing). No pre-agreed decision-right framework resolved these competing inputs under time pressure.

4.5 Crisis Communication as an Unresolved Governance Gap

AI-generated disinformation exploits the communication gap inherent in any cybercrime incident. Traditional crisis communication playbooks, such as NIST's incident response framework (NIST SP 800-61) and SEC disclosure requirements, assume that verified information will eventually become available. AI-enabled disinformation undermines this assumption by flooding the information environment with plausible falsehoods when public communication is most urgent and when responders cannot yet distinguish fabricated content from verified facts. The result is that the standard approach of waiting for confirmed information before issuing public statements becomes self-defeating.

In the US exercise, a law enforcement participant reported an increase in observed cybercrime attacks that incorporate a fear-based emotional component. An academic participant characterized end-user education as functionally equivalent to a Turing Test: ordinary users cannot reliably distinguish AI-generated phishing from legitimate communications, rendering communication-based defenses at the individual level structurally inadequate. An industry participant demonstrated this point concretely: voice cloning for deepfake audio required a one-hour investment and an eleven-dollar subscription.

The Singapore exercise demonstrated the clearest example of this pattern. In Round 2, the scenario described AI-generated fake news claiming that defense networks had been compromised went viral before authorities could verify or deny it. The resulting decision-making dilemma had no good options: correcting the narrative immediately with unverified information risked being wrong and thus amplifying the disinformation; waiting for verification allowed public panic to grow unchecked.

In the Tel Aviv exercise, a media leak was created, resulting in a simultaneous PR crisis alongside the active incident response. The PR function's operational requirement, like a rapid public statement to protect institutional reputation, was incompatible with the forensic function's requirement to maintain information control while investigating scope. This tension was structurally unresolvable under the time constraints of the scenario.

5. Analysis: Three Structural Deficits

The five patterns from Section 4 reflect three structural deficits in AI-enabled cybercrime governance. These recur across jurisdictions and institutions because they are embedded in governance frameworks rather than tied to any single organization's preparedness. This section synthesizes these empirical findings into three structural claims, using the conceptual framework from Section 2 to show why these patterns constitute a systematic governance problem rather than isolated failures.

5.1 Accountability Gaps

The first structural deficit is the inability of existing accountability frameworks to identify a responsible actor. This is especially true when AI-enabled cybercrime spreads across automated systems, service-provider networks, and synthetic identities. Two empirical patterns create this deficit: Decisions before Verification (§4.1) and Actor Boundary Erosion (§4.3).

The crime-as-a-service ecosystem in AI-enabled cybercrime breaks a once unified criminal act into specialized roles. No single actor can be identified as fully responsible for the harm. Fragmentation in responsibility has been documented in the RaaS literature for several years (Blauth et al., 2022; MacColl et al., 2024). The TTX data provides evidence that this fragmentation undermines accountability frameworks in real time. In all three exercises, participants faced attribution questions they could not resolve within operational timelines.

Legal accountability frameworks – whether criminal law, civil liability, or regulatory enforcement require an identifiable actor with identifiable intent. When AI executes a consequential action autonomously, e.g., deploying ransomware, posting exfiltrated data, or generating disinformation, on preset parameters without specific human instruction at the moment of execution, the question of mens rea (criminal intent) becomes indeterminate at the moment of harm. This is not a temporary evidentiary gap that will close as forensic capabilities mature; it is a structural feature of how AI-enabled criminal ecosystems are designed. Van der Wagen and Pieters' (2015) cyborg crime framework anticipated this condition in the botnet era of cybercrime. The TTXs data show that AI has deepened the problem by adding layers of synthetic identity, automated decision-making, and attribution deception, making the human-machine agency boundary even less tractable.

The supply chain dimension compounds this deficit further. In the Tel Aviv exercise, the Insurance company suffered consequences from a breach initiated by its supplier, raising questions about which entity was accountable for the initial failure. Even when a proximate responsible party is identified,

the chain of contractual and technical dependencies distributes accountability across entities, none of which individually bears full responsibility. The Singapore exercise revealed a forward-looking dimension: participants identified new evidentiary categories, such as AI interaction logs, prompt records, guardrail trigger points, that would be necessary to establish accountability for AI-enabled crimes but that are not yet defined in law or standard forensic practice in any of the three jurisdictions examined. The accountability gap is therefore simultaneously a legal, forensic standards, and architectural gap.

5.2 Coordination Deficits

The second structural deficit reveals why AI-enabled cybercrime is an ecosystem-level transformation, not just an accelerated form of old threats. Existing coordination architectures were built for a cybercrime environment where incidents are sequential, sector-specific, and bounded. AI-enabled operations have fundamentally reorganized this environment. This deficit arises from two empirical patterns: Timeline Compression (§4.2) and Public-Private Coordination Failure (§4.4).

Governing AI-enabled cybercrime requires real-time coordination among actors with fundamentally different authority, information, and incentives. Private-sector organizations hold real-time technical intelligence, such as forensic data, network logs, and threat indicators. Government agencies hold legal authority over law enforcement action, international attribution, and escalation. Neither can act optimally without the other. But the trust-building, legal authorization, and information-sharing protocols that connect them were built for an ecosystem in which incidents arrive one at a time, unfold at human speed, and remain within recognizable sectoral boundaries (Ghosh et al., 2025; Lin, 2025). The TTXs data show that what breaks under AI-enabled conditions is not trust or willingness to cooperate but the coordination architecture itself. This is a structural mismatch between the tempo of the threat ecosystem and the tempo of the institutional response.

The Singapore exercise provided the clearest demonstration. When energy, transport, and water critical information infrastructure were simultaneously affected, the national coordination architecture required separate notification and response tracks for each sector. The hub-and-spoke model, in which a central cybersecurity agency coordinates sector-specific responses, assumed sequential triage: one sector at a time, with the coordinating body allocating attention accordingly. AI-enabled operations collapsed that assumption by generating concurrent, cross-sector effects that overwhelmed the sequential design. More critically, strategic partners began disconnecting from shared information networks immediately upon learning of the cascade – a rational individual decision that collectively

destroyed the cooperative information environment when it was most needed. The ecosystem transformation matters here because AI-enabled operations eliminate the temporal window within which actors might otherwise negotiate their way to a cooperative equilibrium. It is not that the actors are unwilling to cooperate. Rather, the ecosystem no longer provides the conditions for cooperation to be organized.

The US exercise revealed a parallel failure at the public-private interface: a circular dependency between information sharing and legal authorization. Forensic access requires victim cooperation; attribution requires forensic access; legal authorization to act requires attribution; and the victim is managing the crisis under time pressure, with no institutional incentive to slow down for government processes. Each link in this chain was designed for sequential operation within a stable institutional environment. AI-enabled incidents render the sequence by compressing the entire chain into a timeframe shorter than any single step requires. This is the signature of an ecosystem-level problem: the failure does not reside in any one institution's performance but in the interaction architecture that connects them.

Bohr's (2026) coordination transparency framework includes interaction logging, live monitoring, intervention hooks, and boundary conditions. It is relevant to addressing this deficit. However, that framework was built for AI governance inside organizations. It was not meant for crisis coordination between sovereign actors, private entities, and international partners under these conditions.

5.3 Decision-Right Ambiguity

The third structural deficit is the absence of pre-agreed frameworks: established procedures or sets of rules for adjudicating competing decision requirements among actors with different mandates, different risk tolerances, and different information thresholds. This deficit is constructed from three empirical patterns: Decisions Before Verification (§4.1), Public-Private Coordination Failure (§4.4), and Crisis Communication (§4.5).

AI-enabled cybercrime incidents activate multiple institutional functions simultaneously, each operating under a different priority structure, in conditions where existing organizational hierarchies offer no mechanism for determining precedence. The Tel Aviv exercise provided the clearest illustration. When CEO, CISO/CIO, legal counsel, public relations, and business continuity functions were activated concurrently, their priorities proved structurally incompatible under the time constraints of the scenario: legal's obligation was to manage liability disclosure and regulatory

compliance; PR's imperative was to protect institutional reputation before irreparable damage; the CISO's priority was to contain the breach and preserve forensic evidence before any disclosure. The ransom decision crystallized this dynamic. Across all three exercises, the question of whether to pay required simultaneous input from legal (sanctions exposure if state-affiliated), technical (recovery options without payment), financial (liquidity and insurance coverage), public relations (reputational consequences), and potentially law enforcement (designation guidance). Under time pressure, no group could identify who held final authority.

The crisis communication dilemma compounds this ambiguity. In both the Singapore and US exercises, participants confronted the question of whether to issue a public statement about an ongoing incident when facts remained unverified. The absence of verification thresholds for communication, specifying what confirmation is needed before disclosure, means each incident revisits the same governance uncertainty rather than operating within established parameters.

These three structural deficits, accountability gaps, coordination deficits, and decision-right ambiguity, do not operate in isolation but interact to amplify weaknesses across the governance environment, creating conditions that AI-enabled cybercrime is architecturally positioned to exploit.

6. Conclusions

This paper argues that AI-enabled cybercrime is an ecosystem-level transformation, a structural reorganization of criminal capability. This change compresses decision timelines, amplifies ambiguity, and erodes the actor boundaries on which legal accountability depends. Drawing on data from three TTXs in the US, Singapore, and Israel, the analysis identified five recurring governance failures. These are: decisions forced before verification; institutional processes overwhelmed by compressed timelines; attribution rendered indeterminate by distributed criminal architectures; public-private coordination defeated by simultaneous cascades; and crisis communication undermined by AI-generated disinformation. These patterns reveal a single finding: the main failure is not inadequate technical detection but governance breakdown under uncertainty. Three structural deficits emerge: accountability gaps, coordination deficits, and decision-right ambiguity. These are deeply embedded in current governance frameworks and are not tied to any single organization's preparedness. That is

why these deficits appear across jurisdictions with different institutional systems and cybersecurity maturity levels.

The three structural deficits share a common origin: governance architectures designed for attributable, sequential, and jurisdictionally bounded incidents are being applied to crimes that are none of these things. Accountability models will need to evolve from single-actor liability toward distributed responsibility structures that reflect how AI-enabled crime operates, across fragmented provider networks where no single entity controls the full attack chain. Coordination systems will need to shift from sequential triage to parallel, multi-sector response architectures with pre-negotiated information-sharing protocols. Decision authority will need to be specified before crises occur. Some movement in these directions is already visible. CISA's Joint Cyber Defense Collaborative has begun developing AI-specific collaborative playbooks (Cybersecurity and Infrastructure Security Agency, 2025). NIS2 introduces supply-chain security obligations for essential entities (European Parliament and Council of the European Union, 2022). The NCSC's forward-looking threat assessments explicitly incorporate AI-enabled scenarios (NCSC, 2025). Yet these developments remain incremental adjustments within frameworks whose foundational assumptions, as the TTXs evidence demonstrates, no longer hold. The gap is not in the absence of reform efforts but in the persistence of the institutional logic they operate within: a logic of identifiable actors, sequential response, and clear jurisdictional authority that AI-enabled cybercrime has structurally undermined.

The societal implications are significant. When AI-enabled cybercrime causes harm, such as financial loss, disrupted healthcare, or compromised infrastructure, ordinary people bear the consequences. The documented governance failures mean that the institutions responsible for protecting those people cannot identify who is accountable, coordinate quickly enough, or agree on who has the authority to act. These are not technical failures. They are failures of the institutional architecture on which society depends to manage harm in AI-mediated environments.

References

Anckar, C. (2008). On the applicability of the most similar systems design and the most different systems design in comparative research. *International Journal of Social Research Methodology*, *11*(5), 389-401. https://doi.org/10.1080/13645570701401552

Antebi, S., Azulay, N., Habler, E., Ganon, B., Shabtai, A., & Elovici, Y. (2024). GPT in sheep's clothing: The risk of customized GPTs. arXiv preprint. https://doi.org/10.48550/ARXIV.2401.09075

Baram, G. (2023). A sliding scale of secrecy: Toward a better understanding of the role of publicity in offensive cyber operations. *Journal of Cyber Policy*, *7*(3), 275–293. https://doi.org/10.1080/23738871.2023.2184708

Blauth, T. F., Gstrein, O. J., & Zwitter, A. (2022). Artificial intelligence crime: An overview of malicious use and abuse of AI. *IEEE Access*, 10, 77110–77122. https://doi.org/10.1109/ACCESS.2022.3191790

Bohr, J. (2026). Coordination transparency: Governing distributed agency in AI systems. *AI & Society*. https://doi.org/10.1007/s00146-026-02853-w

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Broeders, D., De Busser, E., & Pawlak, P. (2019). *Three Tales of Attribution in Cyberspace: Criminal Law, International Law and Policy Debates*. https://scholarlypublications.universiteitleiden.nl/access/item%3A2967050/download

Cybersecurity and Infrastructure Security Agency. (2025). *JCDC AI Cybersecurity Collaboration Playbook*. U.S. Department of Homeland Security. https://www.cisa.gov/sites/default/files/2025-01/JCDC%20AI%20Playbook.pdf

Egloff, F., & Smeets, M. (2021). Publicly attributing cyber attacks: A framework. *Journal of Strategic Studies*, *46*(4), 1–32. https://doi.org/10.1080/01402390.2021.1895117

European Parliament and Council of the European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on

artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union, L 2024/1689. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

European Parliament and Council of the European Union. (2022). Directive (EU) 2022/2555 of 14 December 2022 on measures for a high common level of cybersecurity across the Union (NIS 2 Directive). *Official Journal of the European Union*, L 333, 80–152. http://data.europa.eu/eli/dir/2022/2555/oj

European Union Agency for Cybersecurity. (2024). *ENISA Threat Landscape 2024*. ENISA. https://www.enisa.europa.eu/publications/enisa-threat-landscape-2024

FBI. (2026). *Internet Crime Report 2025*. https://www.ic3.gov/AnnualReport/Reports/2025_IC3Report.pdf

Ghosh, A., Saini, A., & Barad, H. (2025). Artificial intelligence in governance: Recent trends, risks, challenges, innovative frameworks and future directions. *AI & Society*. 40(7), 5685-5707. https://doi.org/10.1007/s00146-025-02312-y

Guembe, B., Azeta, A., Misra, S., Osamor, V. C., Fernandez-Sanz, L., & Pospelova, V. (2022). The emerging threat of AI-driven cyber attacks: A review. *Applied Artificial Intelligence*, 36(1), Article 2037254. https://doi.org/10.1080/08839514.2022.2037254

Hutchins, E. M., Cloppert, M. J., & Amin, R. M. (2011). Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. *Leading Issues in Information Warfare & Security Research*, *1*(1), 80–106. https://www.lockheedmartin.com/content/dam/lockheed-martin/rms/documents/cyber/LM-White-Paper-Intel-Driven-Defense.pdf

International Organization for Standardization. (2022). ISO 22361:2022 Security and resilience Crisis management – Guidelines (1st ed.). https://www.iso.org/standard/50267.html

Kaloudi, N., & Li, J. (2020). The AI-based cyber threat landscape: A survey. *ACM Computing Surveys*, 53(1), 1-34. https://doi.org/10.1145/3372823

Kävrestad, J., Johansson, S., & Bergström, E. (2025). Using tabletop exercises to raise cybersecurity awareness of decision-makers. In G. Oliva, S. Panzieri, B. Hämmerli, F. Pascucci, & L. Faramondi (Eds.), Critical information infrastructures security: CRITIS 2024. Lecture Notes in Computer Science (Vol. 15549). Springer. https://doi.org/10.1007/978-3-031-84260-3_14

Kovalsky, M., Schechter, B. H., & Carvajal-Kim, L. R. (2024). Breaching the C-suite: Wargaming in the private sector. In F. L. Smith III, N. A. Kollars, & B. H. Schechter (Eds.), *Cyber wargaming: Research and education for security in a dangerous digital world* (pp. 151–163). Georgetown University Press.

Lee, H. (2023). Public attribution in the US government: Implications for diplomacy and norms in cyberspace. *Policy Design and Practice*, 6(2), 198-216. https://doi.org/10.1080/25741292.2023.2199964

Lin-Greenberg, E., Pauly, R. B. C., & Schneider, J. G. (2022). Wargaming for international relations research. *European Journal of International Relations*, 28(1), 83–109. https://doi.org/10.1177/13540661211064090

Lusthaus, J. (2018). *Industry of anonymity: Inside the business of cybercrime*. Harvard University Press.

MacColl, J., Hüsch, P., Mott, G., Sullivan, J., Nurse, J., Turner, S., & Pattnaik, N. (2024). *Ransomware: Victim insights on harms to individuals, organisations and society*. Royal United Services Institute for Defence and Security Studies. https://kar.kent.ac.uk/id/eprint/104628

MITRE Corporation. (2026). *MITRE ATT&CK: Enterprise Matrix* (v14). https://attack.mitre.org/

Nasdaq Verafin. (2026). *2026 Global Financial Crime Report*. Nasdaq, Inc. https://verafin.com/nasdaq-verafin-global-financial-crime-report

National Cyber Security Centre. (2025, May 7). *Impact of AI on cyber threat from now to 2027*. NCSC. https://www.ncsc.gov.uk/report/impact-ai-cyber-threat-now-2027

National Institute of Standards and Technology. (2024). *The NIST Cybersecurity Framework (CSF) 2.0* (NIST CSWP 29). https://doi.org/10.6028/NIST.CSWP.29

National Institute of Standards and Technology. (2012). *Computer Security Incident Handling Guide* (NIST Special Publication 800-61, Revision 2). U.S. Department of Commerce. https://doi.org/10.6028/NIST.SP.800-61r2

Nordström, M. (2022). AI under great uncertainty: Implications and decision strategies for public policy. *AI & Society*, 37(4), 1703-1714. https://doi.org/10.1007/s00146-021-01263-4

Reddie, A. W., Booth, R. E., Goldblum, B. L., Lakkaraju, K., & Reinhardt, J. C. (2024). Cyber wargames as synthetic data. In F. L. Smith III, N. A. Kollars, & B. H. Schechter (Eds.), *Cyber*

wargaming: Research and education for security in a dangerous digital world (21–36). Georgetown University Press.

Rid, T., & Buchanan, B. (2015). Attributing cyber attacks. *Journal of Strategic Studies*, 38(1-2), 4-37. https://doi.org/10.1080/01402390.2014.977382

Sarkar, G., Singh, H., Kumar, S., & Shukla, S. K. (2023). Tactics, techniques and procedures of cybercrime: A methodology and tool for cybercrime investigation process. *Proceedings of the 18th International Conference on Availability, Reliability and Security* (ARES '23). https://doi.org/10.1145/3600160.3605013

Schroeder, D.T., Cha, M., Baronchelli, A., Bostrom, N., Christakis, N.A., Garcia, D., Goldenberg, A., Kyrychenko, Y., Leyton-Brown, K., Lutz, N. & Marcus, G. (2026). How malicious AI swarms can threaten democracy. *Science*, *391*(6783), 354-357. https://doi.org/10.1126/science.adz1697

Smith, C. (2025). Normalizing doubt: AI, democratic confidence, and the OECD trust framework. *AI & Society*. https://doi.org/10.1007/s00146-025-02789-7

Smith, F. L., Kollars, N. A., & Schechter, B. H. (2024). Shall we play a game? Fundamentals of cyber wargaming. In F. L. Smith, N. A. Kollars, & B. H. Schechter (Eds.), *Cyber wargaming: Research and education for security in a dangerous digital world* (1–17). Georgetown University Press.

Strom, B. E., Applebaum, A., Miller, D. P., Nickels, K. C., Pennington, A. G., & Thomas, C. B. (2018). *MITRE ATT&CK: Design and philosophy* (Technical Report MTR180314). The MITRE Corporation. https://www.mitre.org/sites/default/files/2021-11/prs-19-01075-28-mitre-attack-design-and-philosophy.pdf

Sun, J., Gu, S., & Su, R. (2026). AI-empowered responsive regulation for preventing future crimes: An empirical inquiry into the regulatory pyramid to combat future crimes in China and Southeast Asia. *Asian Journal of Criminology*, 21(8). https://doi.org/10.1007/s11417-025-09477-x

U.S. Government Accountability Office. (2023). *Critical Infrastructure Protection: National Cybersecurity Strategy Needs to Address Information Sharing Performance Measures and Methods* (GAO-23-105468). https://www.gao.gov/products/gao-23-105468

U.S. Securities and Exchange Commission. (2023). *Cybersecurity Risk Management, Strategy, Governance, and Incident Disclosure*, Final Rule, Release Nos. 33-11216; 34-97989, 88 Fed. Reg. 51896 (August 4, 2023). https://www.sec.gov/rules-regulations/2023/07/s7-09-22

Van der Wagen, W., & Pieters, W. (2015). From cybercrime to cyborg crime: Botnets as hybrid criminal actor-networks. *British Journal of Criminology*, 55(3), 578–595. https://doi.org/10.1093/bjc/azv009

Vogt, J., Kollars, N., & Poznansky, M. (2025). Preparedness wargaming for critical infrastructure resilience: Taiwan digital blockade wargame. *The Cyber Defense Review*, 10(2), 181–198. https://doi.org/10.55682/cdr/qgz7-pqvc

Lin, L. (2025). Organisational challenges in US law enforcement's response to AI-driven cybercrime and deepfake fraud. *Laws*, 14(4). https://doi.org/10.3390/laws14040046

Yamin, M. M., Ullah, M., Ullah, H., & Katt, B. (2021). Weaponized AI for cyber attacks. *Journal of Information Security and Applications,* 57, 102722. https://doi.org/10.1016/j.jisa.2020.102722

Wall, D.S. (2007). *Cybercrime: The Transformation of Crime in the Information Age*. Polity Press. Series: Crime and Society. ISBN: 978-0-7456-2736-6

Zero Day Clock, 2026. https://zerodayclock.com/