Trust Boundaries under Synthetic Impersonation

Межі довіри за умов використання технологій синтетичної підміни ідентичності

Karol Jędrasiak [A]
Corresponding author: PhD in Pedagogy, e-mail: kjedrasiak@wsb.edu.pl, ORCID ID: https://orcid.org/0000-0002-2254-1030
Piotr Gawliczek [B]
PhD., Prof., e-mail piotr.gawliczek@uwm.edu.pl, ORCID ID: https://orcid.org/0000-0002-0462-1168

[A] WSB University, Dąbrowa Górnicza, Poland.
[B] University of Warmia and Mazury in Olsztyn, Poland

Кароль Єндрасяк [A]
Corresponding author: PhD, e-mail: kjedrasiak@wsb.edu.pl, ORCID ID: https://orcid.org/0000-0002-2254-1030
Пйотр Гавлічек [B]
PhD., професор, e-mail: piotr.gawliczek@uwm.edu.pl, ORCID ID: https://orcid.org/0000-0002-0462-1168

[A] Академія WSB, м. Домброва-Гурнича, Польща
[B] Вармінсько-Мазурський університет в Ольштині, Польща

Received: July 30, 2026 | Revised: September 28, 2026 | Accepted: October 31, 2026

UDC 004.056.5:004.8

DOI: https://doi.org/10.33445/sds.2026.16.5.14

Purpose. To examine how synthetic impersonation may influence privileged decision-making across organisational trust boundaries in digital infrastructure and to develop an execution-centred framework integrating verification, authorisation, execution control, and evidentiary accountability.

Methodology. A structured scenario-based analysis was conducted using 61 constructed privileged decision episodes representing managed security service providers, cloud providers, enterprise data centres, and managed IT services. Descriptive statistics and cross-table comparisons were applied to examine trust-boundary crossings, out-of-band verification, execution outcomes, evidence preservation, and decision-path reconstructability. Four supplementary exercise observations provided qualitative illustrations.

Findings. Of the 61 episodes, 46 involved organisational trust-boundary crossings. All 21 episodes without out-of-band verification were classified as unauthorised execution, whereas the 40 episodes involving verification comprised 21 stopped attacks, 10 false stoppages, and nine legitimate executions. Evidence was preserved in 49 episodes, while 40 decision pathways were fully reconstructable. These scenario-dependent relationships do not establish causal effectiveness.

Theoretical implications. The framework distinguishes communication authentication, identity assurance, authority validation, and execution approval while integrating decision-path reconstructability into cybersecurity accountability.

Practical implications. The framework provides a basis for assessing supplier-related verification procedures, privileged execution controls, and evidentiary completeness.

Originality. The study integrates established cybersecurity mechanisms into an execution-centred model of privileged decision pathways under synthetic impersonation.

Limitations. Constructed scenarios restrict external validity and do not support population-level or causal generalisation.

Article type: Research Article.

Мета роботи. Дослідити вплив синтетичної підміни ідентичності на ухвалення рішень щодо виконання привілейованих дій під час перетину організаційних меж довіри в цифровій інфраструктурі та розробити концептуальну модель, що інтегрує перевірку, авторизацію, контроль виконання й доказову підзвітність.

Методологія дослідження. Застосовано структурований сценарний аналіз 61 сконструйованого епізоду ухвалення привілейованих рішень у середовищах постачальників керованих послуг безпеки, хмарних провайдерів, корпоративних центрів обробки даних і керованих IT-сервісів. Використано описову статистику, порівняння взаємопов'язаних записів і чотири додаткові сценарні спостереження.

Результати дослідження. У 46 сценаріях зафіксовано перетин організаційних меж довіри. Усі 21 сценарій без незалежної перевірки завершилися несанкціонованим виконанням. Серед 40 сценаріїв із перевіркою зафіксовано 21 зупинену атаку, 10 випадків помилкового блокування та дев'ять випадків правомірного виконання. Доказові матеріали збережено у 49 сценаріях, а повну реконструкцію послідовності рішень забезпечено у 40. Виявлені співвідношення не доводять причинної ефективності заходів захисту.

Теоретична цінність дослідження. Розмежовано автентифікацію комунікації, підтвердження ідентичності, перевірку повноважень і дозвіл на виконання дії.

Практична цінність дослідження. Запропоновано основу для оцінювання процедур незалежної перевірки, контролю привілейованих операцій і повноти доказової інформації.

Оригінальність дослідження. Інтегровано наявні механізми кібербезпеки в модель захисту процесу ухвалення привілейованих рішень.

Обмеження дослідження. Сконструйовані сценарії не забезпечують статистичного узагальнення або доведення причинних зв'язків.

Тип статті: дослідницька стаття.

Key words: synthetic impersonation; trust boundaries; decision security; privileged execution; out-of-band verification; digital infrastructure; evidentiary reconstructability.

Ключові слова: синтетична підміна ідентичності; межі довіри; безпека ухвалення рішень; привілейовані дії; незалежна перевірка; цифрова інфраструктура; реконструйованість доказової інформації.

Introduction

Synthetic impersonation has evolved from a predominantly representational challenge into a significant operational cybersecurity concern, particularly in digital infrastructure environments

where apparently legitimate communications may initiate high-consequence decisions (Chesney & Citron, 2019; Kietzmann et al., 2020; Mirsky & Lee, 2021). Advances in generative artificial intelligence have reduced the technical and financial barriers to producing convincing synthetic voice, video, text, and multimodal communications. Consequently, AI-enabled impersonation increasingly presents risks not only to information authenticity but also to organizational decision-making, access governance, and the execution of privileged operations (CISA, 2024; ENISA, 2024; Europol, 2025).

The operational significance of synthetic impersonation arises when communication perceived as credible is treated as sufficient grounds to initiate or authorize an action. In distributed digital infrastructure, requests concerning privileged access, configuration modifications, credential resets, security exceptions, incident response, or service restoration may directly affect system integrity, availability, confidentiality, and operational continuity. Under these conditions, the central security challenge extends beyond determining whether a communication artifact is authentic. It concerns whether the claimed identity, associated authority, and requested action have been independently verified before operational execution.

This challenge becomes particularly pronounced when decision-making processes cross organizational and technical trust boundaries. Contemporary digital infrastructure relies on interconnected relationships among internal security teams, cloud service providers, managed security service providers, contractors, technology suppliers, and external operational partners. These relationships create multiple points at which communication credibility, technical authentication, identity assurance, organizational authority, and execution approval may be incorrectly treated as equivalent. A technically authenticated account, a familiar communication channel, or an apparently legitimate supplier representative does not necessarily establish that a request originates from an authorized individual or that the requested action meets the conditions required for execution.

Existing cybersecurity frameworks provide important foundations for addressing these risks. Zero Trust Architecture emphasizes explicit verification, least-privilege access, and the rejection of implicit trust assumptions (Rose et al., 2020). Security and privacy control frameworks establish requirements for authentication, authorisation, accountability, logging, and access management (Ross & Pillitteri, 2020). Digital identity guidance further distinguishes identity proofing and authentication from subsequent decisions concerning access privileges and operational authority (Temoshok et al., 2025). At the regulatory level, the NIS2 Directive, the Digital Operational Resilience Act (DORA), and the Artificial Intelligence Act contribute to the broader governance context concerning cybersecurity risk management, third-party dependencies, operational resilience, and artificial-intelligence-related risks (Directive (EU) 2022/2555, 2022; Regulation (EU) 2022/2554, 2022; Regulation (EU) 2024/1689, 2024).

Nevertheless, the existence of these frameworks does not automatically ensure that their principles are consistently implemented throughout high-consequence decision pathways. In operational environments characterized by urgency, supplier dependence, delegated responsibilities, and distributed technical control, apparently credible communications may influence decisions before identity claims, authority, and execution conditions have been independently validated. Synthetic impersonation may exploit precisely these discrepancies between established security requirements and their practical implementation.

Previous research on synthetic media has primarily examined the generation, detection, forensic assessment, and broader security implications of manipulated audiovisual content (Chesney & Citron, 2019; Verdoliva, 2020; Mirsky & Lee, 2021). Although this research provides essential foundations for understanding synthetic-content threats, comparatively limited attention has been devoted to the organisational decision pathways through which plausible communications are converted into executable privileged requests. In particular, the interaction among

communication credibility, identity assurance, authority validation, supplier-mediated trust, independent verification, execution approval, and evidentiary accountability requires further systematic examination.

An additional research gap concerns the distinction between evidence preservation and the reconstructability of consequential decisions. The availability of communication records, authentication logs, or technical event data does not necessarily establish whether the complete sequence connecting a request to its verification, approval, and execution can be reconstructed. This distinction is particularly important in supplier-dependent environments, where relevant information may be distributed across organizational boundaries, communication platforms, privileged-access systems, and operational logging infrastructures.

Against this background, the aim of this study is to examine how synthetic impersonation may affect privileged decision pathways across organizational trust boundaries in digital infrastructure and to identify the verification, execution-control, governance, and evidentiary conditions associated with different decision outcomes within a structured scenario-based analytical corpus.

To achieve this aim, the study addresses three research questions:

RQ1. Under what conditions are apparently legitimate communications converted into executable privileged requests across digital trust boundaries?

RQ2. Which organizational verification and execution controls distinguish stopped attacks, unauthorized execution, legitimate execution, and false stoppage within supplier-dependent operational scenarios?

RQ3. How do evidence preservation and decision-path reconstructability influence the assessment of accountability in privileged decision-making following suspected synthetic impersonation?

The study adopts an execution-centered analytical perspective in which the primary object of protection is the decision pathway connecting communication, identity claims, authority validation, execution approval, and operational action. This perspective does not replace established identity and access management principles or Zero Trust Architecture. Instead, it examines how these principles may be operationalised at decision points where communication credibility can influence privileged execution.

The research is based on a structured corpus of 61 scenario-instantiated privileged decision episodes representing four operational archetypes: managed security service providers, cloud service providers, enterprise data centres, and managed information technology services. Four additional scenario-based exercise observations illustrate procedural weaknesses under operational pressure. The analytical design focuses on trust-boundary crossings, verification procedures, execution outcomes, evidence preservation, and decision-path reconstructability.

Because the analytical corpus consists of constructed scenarios rather than naturally occurring incident-response records, the findings are interpreted as scenario-dependent evidence concerning the internal relationships among selected decision-security mechanisms. The study does not seek to estimate the population-level prevalence of synthetic impersonation incidents or establish the real-world effectiveness of individual security controls. Instead, it provides a structured basis for examining decision-path vulnerabilities and developing propositions for subsequent empirical validation.

The scientific contribution of the study consists of three interrelated elements. First, it integrates synthetic impersonation, organizational trust boundaries, and privileged execution into a unified decision-security framework. Second, it operationalizes privileged decision episodes as socio-technical analytical units connecting communication, verification, authority, execution, and evidence. Third, it distinguishes evidence preservation from full decision-path reconstructability as separate dimensions of operational accountability.

Building on earlier decision-security research (Jędrasiak, 2026), the present study extends the analytical perspective towards supplier-dependent digital infrastructure and introduces a structured scenario-based examination of the relationships among verification procedures, trust-boundary crossings, execution outcomes, and evidentiary reconstruction.

The remainder of the article is organised as follows. The Literature Review examines existing research on synthetic impersonation, digital identity, cybersecurity governance, and evidentiary accountability. The Conceptual Framework develops the execution-centred decision-security model. The Materials and Methods and Analytical Strategy sections describe the scenario corpus, analytical variables, and methodological procedures. The Results section presents the scenario-based findings, followed by their interpretation in the Discussion. The final sections address research limitations and conclusions.

Literature Review

Research on synthetic impersonation has developed at the intersection of artificial intelligence, digital forensics, cybersecurity, and organisational risk management. Early studies primarily examined the implications of synthetic audiovisual content for information authenticity, privacy, institutional trust, and national security. Chesney and Citron (2019) identified the potential of deepfake technologies to undermine established assumptions regarding the reliability of audiovisual evidence, while Verdoliva (2020) systematised the technical challenges associated with detecting manipulated media. Mirsky and Lee (2021) subsequently provided a comprehensive analysis of deepfake generation and detection techniques, demonstrating the growing sophistication of synthetic-content technologies and the corresponding limitations of forensic identification methods.

These studies established an important foundation for understanding synthetic media as a cybersecurity threat. However, their principal analytical focus remained on the characteristics of manipulated artefacts, their detection, and their potential informational consequences. Such approaches are necessary for identifying suspicious content but do not fully explain the organisational mechanisms through which apparently credible communications may influence high-consequence operational decisions.

Subsequent research has expanded the analysis of synthetic impersonation across different communication modalities. Studies addressing audio-based impersonation, audiovisual manipulation, image artefacts, deep-network representations, and hybrid detection approaches have demonstrated that synthetic-content threats are not confined to a single technological medium (Buchwald & Jędrasiak, 2024; Jędrasiak, 2024a, 2024b; Jędrasiak & Bijoch, 2025; Jędrasiak & Wolański, 2023, 2024). These contributions are particularly relevant because contemporary organisational communication increasingly combines voice, video, text, and digital collaboration platforms.

Nevertheless, improvements in synthetic-content detection do not necessarily translate into equivalent improvements in operational decision security. Even when suspicious content is identified, organisations must determine whether the associated request is legitimate, whether the claimed identity has been adequately verified, whether the requester possesses the required authority, and whether the requested action satisfies established execution conditions. Conversely, a communication that does not exhibit detectable synthetic characteristics may still constitute an unauthorised or malicious operational request.

This distinction indicates that media authenticity and decision legitimacy should be treated as related but analytically separate dimensions of cybersecurity.

Recent threat assessments have increasingly recognised the operational implications of AI-enabled impersonation and social engineering. Reports published by cybersecurity agencies and law-enforcement organisations identify the exploitation of organisational authority, familiar communication channels, and established business relationships as relevant components of

contemporary cyber threats (CISA, 2024; ENISA, 2024; Europol, 2025; FBI, 2025; INTERPOL, 2024, 2025, 2026). These assessments broaden the understanding of synthetic impersonation beyond manipulated media and highlight the importance of organisational processes through which communications are interpreted and acted upon.

However, threat reporting primarily contributes descriptive evidence concerning emerging risks, attack patterns, and operational vulnerabilities. It does not necessarily provide a systematic analytical model explaining how communication credibility, identity assurance, authority validation, and execution approval interact within a privileged decision pathway. This limitation is particularly relevant in distributed digital infrastructure, where decisions frequently involve multiple organisational actors and technical environments.

Contemporary digital infrastructure depends on interconnected relationships among internal security teams, cloud service providers, managed security service providers, technology suppliers, contractors, and external operational partners. These relationships create distributed trust boundaries across which communication, identity claims, delegated authority, and operational instructions may be transmitted. The resulting security environment cannot be adequately described solely through the traditional distinction between internal and external network perimeters.

Cybersecurity supply-chain research and institutional guidance recognise third-party dependencies as integral components of organisational risk management. Boyens et al. (2024) emphasise the importance of cybersecurity supply-chain risk management practices, including supplier assessment, accountability, monitoring, and the management of security requirements across organisational relationships. Such principles are particularly relevant when external entities have technical access, participate in incident response, or initiate requests that affect privileged operations.

The regulatory environment reinforces the significance of these dependencies. Directive (EU) 2022/2555 establishes cybersecurity risk-management requirements, including supply-chain security, while Regulation (EU) 2022/2554 addresses digital operational resilience and third-party risk in information and communication technology within the financial sector. Regulation (EU) 2024/1689 contributes to the broader governance framework for artificial intelligence and associated risks. Although these instruments differ in scope and regulatory purpose, they collectively demonstrate the increasing importance of organisational accountability, third-party governance, and risk-based security management.

Nevertheless, compliance-oriented frameworks do not automatically establish how individual privileged requests should be evaluated when identity claims, delegated authority, urgency, and supplier relationships intersect. A contractual relationship with a trusted supplier does not, by itself, establish the legitimacy of every request originating from that supplier's communication environment. Similarly, recognised organisational roles do not necessarily confer unrestricted authority to initiate technically consequential actions.

Established cybersecurity architectures provide essential mechanisms for addressing these problems. Zero Trust Architecture rejects implicit trust assumptions and requires explicit verification and least-privilege access when resources are accessed (Rose et al., 2020). Security and privacy control frameworks establish requirements concerning authentication, authorisation, access control, logging, accountability, and the management of privileged operations (Ross & Pillitteri, 2020). Digital identity guidance further distinguishes identity proofing and authentication from subsequent decisions concerning access rights and operational authority (Temoshok et al., 2025).

These frameworks provide a substantial theoretical and practical foundation for securing digital infrastructure. Accordingly, synthetic impersonation should not be interpreted as a threat that invalidates existing Identity and Access Management or Zero Trust principles. Instead, it highlights the importance of consistently applying those principles throughout the organisational decision process, particularly where communication credibility may be incorrectly interpreted as evidence of operational authority.

Four distinct assurance dimensions are relevant in this context. Authentication establishes whether an account, session, or technical interaction satisfies specified authentication requirements. Identity assurance concerns the confidence associated with the claimed identity of the communicating individual. Authority validation determines whether the individual is permitted to request a particular action under the applicable organisational conditions. Execution approval determines whether the procedural and technical requirements for implementing the action have been met.

Although these dimensions are conceptually related, they are not interchangeable. A technically authenticated session does not necessarily establish that a particular privileged request is authorised. Similarly, a verified individual may possess legitimate organisational credentials without having authority to request a specific configuration change, security exception, or emergency intervention.

This distinction is particularly significant under synthetic impersonation because an attacker may exploit inconsistencies among communication credibility, identity verification, organisational authority, and execution controls. The operational vulnerability therefore arises not exclusively from the sophistication of synthetic content but from the possibility that apparently legitimate communication may be accepted as sufficient justification for privileged action.

Research concerning incident handling and security accountability introduces an additional dimension of the problem. Cichonski et al. (2012) emphasise the importance of structured incident-response procedures, while Ross and Pillitteri (2020) establish broader control requirements relevant to auditability, logging, and accountability. These contributions demonstrate that effective cybersecurity governance requires not only preventive and detective controls but also the preservation of information necessary to investigate security-relevant events.

In distributed digital infrastructure, however, relevant evidence may be fragmented across communication platforms, identity-management systems, privileged-access tools, supplier environments, service-management applications, and technical logs. The preservation of individual records does not necessarily ensure that the complete decision pathway can subsequently be reconstructed.

For analytical purposes, evidence preservation and decision-path reconstructability should therefore be distinguished. Evidence preservation concerns the retention of records or artefacts relevant to a particular event. Decision-path reconstructability concerns the ability to establish a coherent sequence connecting the initiating communication, claimed identity, verification procedures, authority validation, approval decisions, technical execution, and resulting operational outcome.

This distinction is particularly relevant in synthetic impersonation scenarios because the evidentiary record must support an assessment not only of what technical action occurred but also of why the request was considered legitimate and which organisational controls were applied or bypassed.

The reviewed literature consequently reveals three interconnected areas requiring further analytical integration.

First, synthetic-media research provides substantial knowledge concerning the generation, detection, and forensic assessment of manipulated content, but its principal focus does not consistently extend to the organisational conditions under which communication credibility becomes operational authority.

Second, established cybersecurity and digital identity frameworks define essential requirements for authentication, authorisation, access management, and accountability. However, the interaction of these mechanisms within supplier-dependent, high-consequence decision pathways requires further operational analysis, particularly under conditions of urgency and potentially deceptive communication.

Third, incident-response and security-control literature recognises the importance of evidence preservation and auditability, but the distinction between retaining individual records and

reconstructing the complete authorisation and execution pathway remains insufficiently integrated into research specifically addressing synthetic impersonation.

The present study addresses these intersections through an execution-centred decision-security perspective. Building on earlier research on decision security in high-trust environments (Jędrasiak, 2026), it examines privileged decision episodes as socio-technical units that connect communication, identity claims, authority validation, independent verification, execution outcomes, and evidentiary accountability.

The proposed analytical perspective does not seek to replace established synthetic-media detection methods, Identity and Access Management mechanisms, or Zero Trust Architecture. Rather, it integrates these existing foundations into a decision-path framework focused on the organisational conditions under which apparently credible communications may result in privileged execution.

Within this framework, trust boundaries are understood as points at which communication, identity claims, delegated authority, or approval decisions move between different organisational or technical assurance contexts. Their security significance depends on whether the relevant verification and execution controls remain effective when requests cross those boundaries.

The literature review therefore provides the conceptual foundation for examining three related questions: how plausible communications become actionable privileged requests, how organisational verification and execution controls are associated with different scenario outcomes, and how evidence preservation and decision-path reconstructability contribute to the assessment of operational accountability.

These relationships are subsequently examined through a structured scenario-based analytical corpus. Because the scenarios are constructed rather than derived from a representative sample of naturally occurring incidents, the analysis is intended to clarify decision-security mechanisms and generate propositions for further empirical validation rather than establish population-level causal effects.

Conceptual Framework

The conceptual framework of this study is based on the proposition that synthetic impersonation becomes an operational cybersecurity threat when apparently credible communication influences the authorisation or execution of high-consequence actions. Within distributed digital infrastructure, the principal security concern is therefore not limited to the authenticity of an individual communication artefact. It extends to the integrity of the decision pathway through which identity claims, organisational authority, verification procedures, and execution permissions are translated into privileged operational actions.

This perspective is consistent with established cybersecurity principles concerning explicit verification, least-privilege access, accountability, and the separation of authentication from authorisation (Rose et al., 2020; Ross & Pillitteri, 2020). However, the present framework places particular emphasis on the organisational decision points at which technically legitimate communications may be incorrectly interpreted as sufficient grounds for privileged execution.

The framework distinguishes four interrelated but analytically separate assurance dimensions: authentication, identity assurance, authority validation, and execution approval.

Authentication refers to the process of verifying that a digital entity, account, session, or communication endpoint satisfies specified authentication requirements. Successful authentication provides evidence concerning the validity of the authentication process but does not independently establish the legitimacy of every request transmitted through the authenticated environment.

Identity assurance concerns the degree of confidence that the individual initiating or communicating a request is the person whose identity is claimed. In synthetic impersonation scenarios, this dimension is particularly important because generated voice, video, or text may

reproduce characteristics associated with a legitimate individual without establishing that individual's actual participation in the communication.

Authority validation determines whether the identified individual possesses the organisational, contractual, or delegated authority required to request a specific privileged action. Authority must be assessed in relation to the requested operation, the applicable organisational rules, and the operational circumstances. A legitimate identity does not necessarily imply unrestricted authority to initiate technical changes.

Execution approval concerns the final determination that the requested action satisfies the procedural, organisational, and technical conditions required for implementation. This dimension may include independent verification, confirmation by designated approvers, dual-control procedures, and compliance with established change-management requirements.

These dimensions form a sequential analytical structure, although their practical implementation may involve parallel or repeated verification activities. Their separation is necessary because the successful completion of one assurance procedure does not automatically establish the validity of subsequent decisions.

A technically authenticated communication channel may transmit a request originating from an impersonator. A correctly identified individual may lack authority to initiate the requested action. Similarly, an individual with the relevant authority may submit a request that cannot be executed until additional procedural or technical conditions are met.

Consequently, the legitimacy of privileged execution depends on the combined adequacy of the applicable assurance procedures rather than on the apparent credibility of the initiating communication alone.

Trust boundaries and privileged decision pathways

The concept of a trust boundary provides the organisational and technical foundation for the proposed framework.

For the purposes of this study, a trust boundary is defined as a point at which a communication, identity claim, authority claim, or approval decision moves between different organisational or technical assurance contexts and becomes capable of influencing a consequential operational action.

Such boundaries may arise between internal organisational units, shift-based operational teams, service desks and infrastructure administrators, cloud service providers and customers, managed security service providers and client organisations, or contractors and the systems they are authorised to support.

Unlike a conventional network perimeter, a trust boundary in this framework is not necessarily associated with a physical or technical separation. It may also exist within an organisational decision process when responsibility for identity verification, authority confirmation, approval, or execution passes from one actor to another.

This interpretation reflects the distributed nature of contemporary digital infrastructure, in which operational authority and technical capabilities are frequently delegated across organisational relationships (Boyens et al., 2024).

A privileged decision pathway is understood as the sequence of communication, verification, authorisation, approval, and execution activities associated with a request capable of materially affecting the security or operation of a digital system.

Privileged actions may include elevation of administrative access, firewall configuration changes, routing modifications, identity and access management exceptions, enabling remote administration, backup restoration, and urgent interventions outside established change-management windows.

The defining characteristic of such actions is their capacity to produce consequential changes in system security, availability, access, configuration, or operational continuity.

Within this framework, synthetic impersonation may influence the decision pathway by introducing a misleading identity claim, an apparently authoritative instruction, or a plausible operational justification for bypassing established verification procedures.

The resulting risk is associated with the possibility that communication credibility may be treated as a substitute for independently verified authority.

Execution-centred decision security model

The proposed execution-centered decision security model conceptualizes the transition from an initiating communication to an operational outcome as a sequence of interconnected decision stages.

The first stage involves receiving a communication requesting a privileged action. The request may be transmitted via telephone, Voice over Internet Protocol, email, collaborative platforms, video communication, or service management systems.

At this stage, the communication may contain claims concerning the requester's identity, organisational role, operational urgency, or relationship with a recognised supplier.

The second stage involves identifying whether the request crosses an organisational or technical trust boundary. Such a crossing occurs when the request moves between assurance contexts or requires one organisational actor to rely on identity, authority, or approval information originating from another actor or system.

The third stage concerns the application of verification and authorisation controls. These controls may include callback procedures, verification through an independent communication channel, confirmation against an approved contact inventory, validation of designated approvers, and assessment of the requester's authority.

Independent verification is particularly important where the initiating communication may itself be compromised or synthetically generated. Verification performed exclusively through the original communication channel may reproduce the very trust assumptions that made the suspicious request credible.

The fourth stage involves the execution decision. At this point, the organisation determines whether the request satisfies the conditions necessary for implementation or whether execution should be suspended, rejected, or escalated for further verification.

The fifth stage concerns the operational outcome. Within the scenario-based analytical design, four outcome categories are distinguished: Attack Stopped, False Stoppage, Legitimate Execution, and Unauthorized Execution.

Attack Stopped refers to a scenario in which an illegitimate request is prevented from producing the intended privileged action.

False Stoppage refers to a scenario in which a legitimate request is incorrectly blocked or prevented from proceeding.

Legitimate Execution refers to a scenario in which an authorised request is executed in accordance with the applicable requirements.

Unauthorized Execution refers to a scenario in which a privileged action is performed without satisfying the required conditions of identity assurance, authority validation, or execution approval.

These categories distinguish security failures from operational availability consequences. In particular, preventing unauthorised execution is not the only relevant performance objective, because excessive or incorrectly applied controls may also prevent legitimate high-consequence operations.

The final component of the framework concerns evidence preservation and decision-path reconstructability.

Evidence preservation refers to the retention of relevant communication records, verification results, approval information, execution logs, and other materials necessary for subsequent examination.

Decision-path reconstructability refers to the ability to establish a sufficiently complete and coherent sequence connecting the initiating request with the verification procedures, authority assessments, approval decisions, and final operational outcome.

Although these concepts are related, they represent different analytical dimensions. The existence of preserved records does not necessarily establish that the complete decision pathway can be reconstructed.

For example, a technical log may confirm that a configuration change occurred without documenting the identity claimed by the requester, the source used for callback verification, or the organisational authority under which the action was approved.

The evidentiary component therefore extends the framework beyond immediate execution security towards post-incident accountability, auditability, and organisational learning.

The proposed execution-centred decision security framework, including the principal verification stages, execution outcomes, and evidentiary accountability mechanisms, is presented in Figure 1.

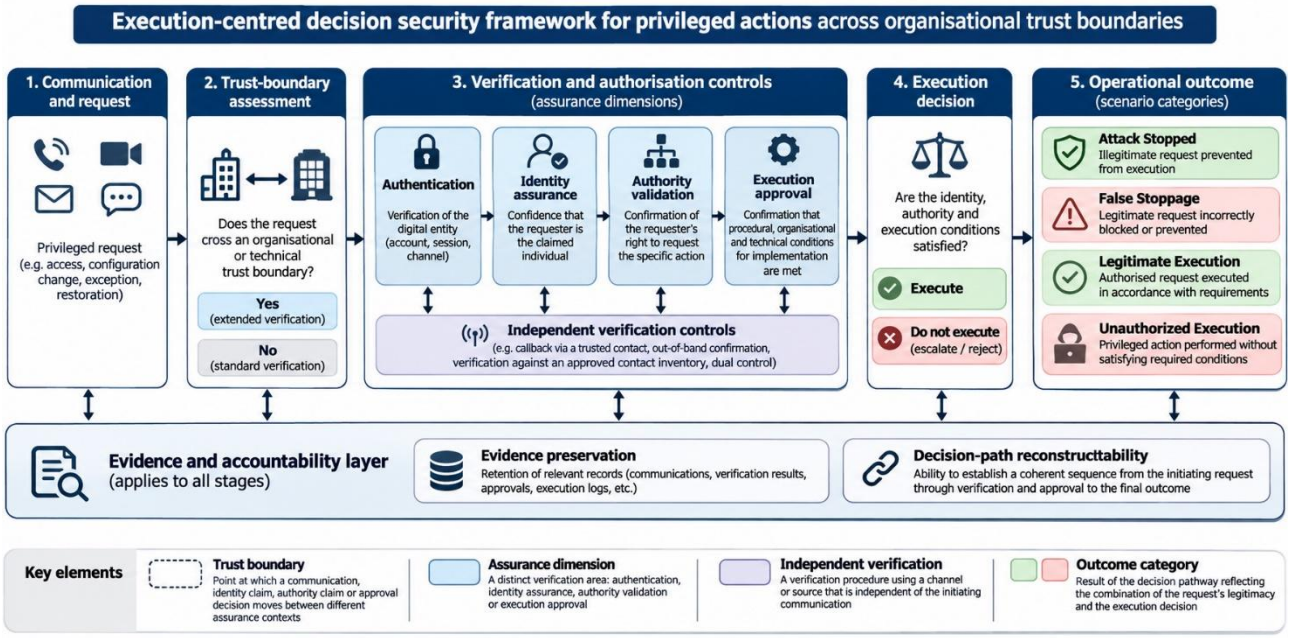


Figure 1. Execution-centred decision security framework for privileged actions across organisational trust boundaries

Source: Developed by the authors.

Analytical relationships within the proposed framework

The framework identifies three principal relationships for examination within the scenario-based analytical corpus.

The first concerns the relationship between trust-boundary crossings and outcomes of privileged decisions. Requests crossing organisational boundaries may involve additional dependencies on supplier relationships, delegated authority, communication channels, and verification procedures.

However, crossing a trust boundary is not inherently equivalent to a security failure. The analytical significance of the boundary depends on the controls applied when information or authority moves between assurance contexts.

The second relationship concerns independent verification and execution outcomes. Out-of-band verification provides a mechanism for challenging identity and authority claims through a channel or procedure that does not rely exclusively on the initiating communication.

Within the proposed model, such verification is treated as an execution-control mechanism rather than as a method for determining the technical authenticity of synthetic media.

Nevertheless, the effectiveness of independent verification depends on its implementation. A callback performed using contact information supplied within a suspicious request may fail to provide meaningful independence. Similarly, verification against outdated contact inventories or incorrectly assigned approvers may introduce additional procedural weaknesses.

Accordingly, the analytical relevance of out-of-band verification depends not only on whether a verification action occurred but also on whether the procedure was sufficiently independent and appropriate to the requested privileged operation.

The third relationship concerns evidence preservation and decision-path reconstructability. The framework assumes that the retention of individual records and the ability to reconstruct an entire decision sequence are related but distinct characteristics of operational accountability.

This distinction permits the analysis of situations in which evidence exists but remains insufficient to establish the sequence of identity claims, authority assessments, approval decisions, and execution actions.

Together, these relationships provide the conceptual basis for examining how apparently legitimate communications may influence privileged decisions across organisational trust boundaries.

The framework does not assume that any individual control is universally sufficient to prevent unauthorised execution. Instead, it treats decision security as a property of the combined verification, authorisation, execution, and evidentiary mechanisms applied throughout the decision pathway.

Relationship to existing cybersecurity frameworks

The proposed model should be understood as an analytical extension of established cybersecurity and digital identity principles rather than as a replacement for them.

Zero Trust Architecture provides the general requirement that access decisions should not depend on implicit trust associated with network location or organisational familiarity (Rose et al., 2020). Identity and Access Management mechanisms establish processes for managing digital identities, authentication, authorisation, and access privileges. Security-control frameworks further address accountability, logging, monitoring, and privileged operations (Ross & Pillitteri, 2020).

The contribution of the execution-centred perspective lies in examining how these established mechanisms interact within a complete organisational decision pathway, particularly when communication credibility, operational urgency, supplier relationships, and delegated authority may influence privileged execution.

This perspective shifts the unit of analysis from an isolated communication artefact or authenticated session towards the sequence of socio-technical decisions through which a request becomes actionable.

It also introduces evidentiary reconstructability as a distinct analytical consideration, allowing the study to examine whether an organisation can subsequently demonstrate how a privileged action was requested, verified, authorised, and executed.

The proposed framework therefore provides a structured basis for analysing decision-security mechanisms in distributed digital infrastructure under conditions of synthetic impersonation.

Its empirical application in the present study is limited to a constructed scenario-based corpus. Consequently, the framework is used to examine relationships among selected decision variables and to identify potential control weaknesses rather than to establish the population-level effectiveness of specific cybersecurity mechanisms.

The following sections describe how these conceptual elements are operationalised through scenario-instantiated privileged decision episodes, verification records, trust-boundary characteristics, execution outcomes, and evidentiary indicators.

Materials and Methods

Research design

This study employed a structured, scenario-based analytical design to examine privileged decision-making under synthetic impersonation pressure across organisational and technical trust boundaries in digital infrastructure environments. The principal unit of analysis was a privileged decision episode, defined as a sequence of interconnected communication, verification, authorisation, execution, and evidentiary activities associated with a request capable of producing a consequential operational action.

The research design adopted an execution-centred perspective consistent with the conceptual approach developed in earlier research on decision security in high-trust sectors (Jędrasiak, 2026). Rather than evaluating synthetic audio, video, or text artefacts in isolation, the study examined the organisational conditions under which apparently credible communications could influence privileged decisions.

The analytical material consisted of a bounded corpus of 61 anonymised, scenario-instantiated privileged decision episodes representing four operational archetypes. These episodes were constructed for analytical comparison and were not extracted from a representative sample of naturally occurring cybersecurity incidents. Accordingly, the study was designed to investigate relationships among predefined decision-security characteristics rather than estimate incident prevalence or establish the real-world effectiveness of individual controls.

Four additional scenario-based exercise observations were incorporated as a supplementary analytical layer to illustrate procedural weaknesses under operational pressure. These observations were considered separately from the 61-episode corpus and were not treated as additional independent cases in the principal descriptive calculations.

Scenario corpus and operational environments

The main analytical corpus comprised 61 privileged decision episodes distributed across four operational archetypes:

- Managed security service provider (MSSP) environments: 16 episodes.
- Cloud-provider environments: 15 episodes.
- Enterprise data-centre environments: 15 episodes.
- Managed information technology (IT) service environments: 15 episodes.

The distribution was intended to provide variation across organisational arrangements and operational dependencies while retaining a comparable analytical structure. It was not designed to reproduce the actual distribution of synthetic impersonation incidents across industries or infrastructure sectors.

The selected archetypes represent environments in which privileged requests may originate from internal personnel, external suppliers, managed service providers, or other actors operating within established technical and organisational relationships.

Each scenario represented a decision pathway initiated by an apparently credible communication and associated with a requested action affecting access privileges, system configuration, service availability, operational continuity, or another consequential process.

The corpus included both internal decision pathways and episodes involving external service relationships. In 46 of the 61 episodes, the scenario involved a supplier or external service dependency and a crossing of an organisational trust boundary.

This distribution reflects the study's deliberate emphasis on supplier-dependent operational environments. Consequently, the proportion of cross-boundary episodes is a characteristic of the constructed corpus and should not be interpreted as an estimate of the prevalence of such events in operational practice.

Communication channels and privileged actions

The communication layer represented six channel types: telephone, Voice over Internet Protocol (VoIP), Slack, email, Microsoft Teams video, and ticketing systems.

These channels included synchronous and asynchronous communication modes and were selected to represent different mechanisms through which privileged requests may be transmitted and perceived as credible.

The privileged change layer included several categories of consequential operational actions:

- Firewall rule modifications.
- Temporary administrative access.
- Routing or Domain Name System (DNS) changes.
- Identity and Access Management (IAM) exceptions.
- Virtual Private Network (VPN) configuration changes.
- Backup restoration.
- Secure financial-routing changes affecting wire-transfer paths.

The inclusion of both technical infrastructure operations and decision-sensitive financial processes reflected the broader analytical focus on the conversion of apparently legitimate requests into consequential actions across trust boundaries.

Synthetic-media fidelity was not manipulated as an independent experimental variable. Instead, the scenarios assumed a functional plausibility threshold: each initiating communication was considered sufficiently credible to receive ordinary operational attention.

This assumption allowed the analysis to focus on subsequent verification, authority validation, execution decisions, and evidentiary procedures. However, it also limits the study's ability to examine how variations in synthetic-content quality, latency, visual artefacts, voice similarity, or contextual sophistication affect the likelihood of successful impersonation.

Data structure and analytical variables

Each privileged decision episode was represented through a set of correlated records linked by a common episode identifier.

The analytical structure comprised eight interconnected data layers:

1. A message layer describing the initiating communication.
2. A privileged change layer describing the requested operational action.
3. A verification layer recording identity, authority, and procedural checks.
4. A trust-boundary layer describing organisational and technical dependencies.
5. A callback and approver inventory layer representing designated verification contacts and approval arrangements.
6. A normalised logging layer containing technical and procedural records.
7. An evidence preservation layer describing retained evidentiary materials.
8. An exercise observation layer containing supplementary scenario-based observations.

The episode identifier served as the principal correlation key, enabling the reconstruction of relationships among initiating communications, verification procedures, approval conditions, execution outcomes, and retained records.

The principal analytical variables concerned trust-boundary crossing, out-of-band verification, execution outcome, evidence preservation, and decision-path reconstructability.

Trust-boundary crossing was examined in relation to whether a request involved movement between organisational or technical assurance contexts, including supplier-dependent decision pathways.

Out-of-band verification concerned whether a verification procedure was performed through a channel or source separate from the initiating communication. The verification records also captured whether the communication channel was technically valid, whether identity and

organisational role were checked, whether authority was assessed, and whether a designated approver was verified where applicable.

These dimensions were treated separately because successful technical authentication does not necessarily establish the legitimacy of an individual request or the authority required to initiate a privileged action.

Execution outcomes were classified into four categories: *Attack Stopped*, *Unauthorized Execution*, *False Stoppage*, and *Legitimate Execution*.

Attack Stopped represented an illegitimate request that did not result in the intended privileged execution.

Unauthorized Execution represented a privileged action performed without satisfying the required authorisation or execution conditions.

False Stoppage represented a legitimate request that was incorrectly blocked.

Legitimate Execution represented an authorised request executed in accordance with the applicable requirements.

The distinction among these outcomes allowed the analytical framework to consider both security failures and the operational consequences of incorrectly blocking legitimate requests.

Trust-boundary governance and supplier dependencies

The trust-boundary layer was used to examine how supplier relationships and delegated operational responsibilities influenced the structure of privileged decision pathways.

The dataset included information concerning external entity types, contractual logging obligations, and classified boundary risk levels.

These characteristics were relevant because privileged operations in supplier-dependent environments may involve multiple organisations, communication platforms, approval arrangements, and technical control systems.

The analytical approach was informed by cybersecurity supply-chain risk management principles and relevant European regulatory requirements concerning third-party dependencies and operational resilience (Boyens et al., 2024; Directive (EU) 2022/2555, 2022; Regulation (EU) 2022/2554, 2022).

However, the presence of a supplier relationship was not treated as evidence of a security failure. Instead, supplier dependence and trust-boundary crossing were examined as contextual characteristics requiring appropriate verification, authority validation, and execution controls.

Evidence preservation and decision-path reconstructability

The study incorporated an evidentiary dimension to examine whether privileged decision pathways could be reconstructed following the execution or blocking of a request.

The normalised logging layer contained 183 technical and procedural records associated with the 61 decision episodes, corresponding to an average of three records per episode.

The evidence register recorded the types of preserved materials, the application of immutability controls, and the usefulness of the retained evidentiary package for subsequent examination.

Evidence preservation and decision-path reconstructability were treated as distinct analytical characteristics.

Evidence preservation referred to the availability of retained records relevant to a decision episode.

Decision-path reconstructability referred to the ability to establish a coherent sequence connecting the initiating communication, identity and authority checks, approval decisions, technical execution, and operational outcome.

This distinction was necessary because the preservation of individual records does not automatically ensure that the complete sequence of organisational and technical decisions can be reconstructed.

The analysis therefore considered whether the retained information was sufficient to determine who initiated a request, which verification procedures were performed, who approved the action, and what technical operations followed.

Supplementary exercise observations

Four additional scenario-based exercise observations were incorporated to illustrate procedural vulnerabilities associated with operational urgency, supplier-related pressure, and uncertainty concerning designated authority or escalation.

These observations served a supplementary interpretative purpose. They were not included in the principal 61-episode denominator and were not used to estimate the frequency of operational failures.

Their inclusion allowed the study to consider procedural weaknesses that may arise when decision-makers must assess apparently credible requests under time constraints.

The exercise observations were interpreted as illustrative scenario evidence rather than independent validation of the relationships identified within the main analytical corpus.

Scope and methodological limitations

The research design was intended to support structured analysis of privileged decision pathways under predefined synthetic impersonation conditions.

Its principal methodological strength lies in the use of correlated records connecting communication, verification, trust-boundary characteristics, execution outcomes, and evidentiary information within a common analytical structure.

Nevertheless, several limitations follow directly from the scenario-based design.

First, the constructed corpus does not represent a random or statistically representative sample of operational cybersecurity incidents. The observed distributions therefore describe the analytical scenarios and cannot be generalised to real-world incident frequencies.

Second, the study did not experimentally manipulate synthetic-media fidelity. Consequently, it cannot determine how changes in the technical quality of impersonation affect verification behaviour or execution outcomes.

Third, the relationships identified between verification procedures and execution outcomes may depend on the assumptions used to construct the scenarios and assign their characteristics. Such relationships should therefore not be interpreted as independent evidence of causal effectiveness without further validation.

Fourth, the four supplementary exercise observations provide illustrative material but do not independently establish the external validity of the analytical corpus.

Accordingly, the findings are interpreted as scenario-dependent observations concerning the structure and internal relationships of privileged decision-security mechanisms. Further research using independently observed operational cases, controlled exercises, or experimentally varied verification conditions would be required to evaluate the generalisability and practical effectiveness of the proposed framework.

Analytical Strategy and Validity Safeguards

The analytical strategy combined descriptive statistics, structured cross-table comparison, and mechanism-oriented interpretation to examine privileged decision pathways within the constructed scenario corpus.

The analysis proceeded in four consecutive stages.

First, the analytical records were examined for structural consistency at the episode level. Unique episode identifiers were used to correlate information across the communication, privileged change, verification, trust-boundary, logging, and evidence preservation layers.

Second, descriptive analysis was used to characterise the distribution of scenarios across organisational archetypes, communication channels, privileged action categories, trust-boundary

conditions, and execution outcomes. Counts and percentages were calculated using the 61-episode analytical corpus as the denominator unless otherwise specified.

Third, structured comparisons were conducted to examine relationships among trust-boundary crossing, out-of-band verification, and execution outcomes. These comparisons were intended to identify patterns within the constructed scenarios rather than estimate causal effects or population-level associations.

Fourth, evidence preservation and decision-path reconstructability were assessed by examining whether the available records supported the reconstruction of the initiating request, verification procedures, approval sequence, and resulting technical action.

The analysis maintained a conceptual distinction among authentication, identity assurance, authority validation, and execution approval. A technically valid communication channel was not treated as sufficient evidence of the legitimacy of a privileged request. Similarly, successful identity verification was not considered equivalent to authorisation for a specific operational action.

To support analytical consistency, information from the different record layers was compared at the episode level. Communication records, verification activities, privileged change records, technical logs, and evidence preservation entries were examined together to assess whether the recorded decision pathway was internally coherent.

This cross-table comparison strengthened the traceability of scenario information and reduced dependence on individual summary indicators. However, because the records belonged to the same constructed analytical corpus, their consistency did not constitute independent empirical validation.

The interpretation of results was therefore subject to three safeguards.

First, descriptive proportions were interpreted exclusively within the boundaries of the constructed corpus. Second, relationships between verification procedures and outcomes were not interpreted as evidence of causal effectiveness without independent validation. Third, findings concerning evidence preservation and reconstructability were distinguished from claims about the actual performance of operational incident-response systems.

The analytical strategy was informed by established principles of incident handling, auditability, access control, and digital identity assurance (Cichonski et al., 2012; Ross & Pillitteri, 2020; Temoshok et al., 2025).

Overall, the adopted approach provides a structured basis for examining decision-security mechanisms under predefined synthetic impersonation conditions while recognising the limitations associated with scenario construction, outcome classification, and external validity.

Results

The analysis comprised 61 constructed privileged decision episodes representing four operational archetypes: managed security service providers (16 episodes), cloud providers (15 episodes), enterprise data centres (15 episodes), and managed IT services (15 episodes). The distribution provided comparable representation of the selected environments and supported descriptive examination of decision-security characteristics across different organisational settings.

The scenario corpus was designed to represent operational conditions involving consequential privileged requests. High or critical urgency was assigned to 51 of the 61 episodes (83.6%). This distribution reflects the deliberate inclusion of time-sensitive decision conditions and should not be interpreted as evidence that a corresponding proportion of real-world synthetic impersonation incidents occurs under high urgency.

The scenarios included six communication channels: telephone, Voice over Internet Protocol, Slack, email, Microsoft Teams video, and ticketing systems. These channels represented different communication environments through which apparently credible requests could be transmitted. The corpus therefore enabled examination of privileged decision pathways without restricting the analysis to a single communication medium.

Execution outcomes and verification procedures

The distribution of execution outcomes is presented in Table 1.

Table 1. Distribution of execution outcomes in the scenario corpus

Execution outcome	Number of episodes	Share, %
Attack Stopped	21	34.4
Unauthorized Execution	21	34.4
False Stoppage	10	16.4
Legitimate Execution	9	14.8
Total	**61**	**100.0**

Source: *Calculated from the scenario corpus.*

The scenario outcomes were distributed between successful prevention of illegitimate requests, unauthorised execution, incorrect blocking of legitimate requests, and legitimate execution. This classification enabled the analysis to distinguish security-related outcomes from operational consequences associated with the incorrect rejection of authorised requests.

The verification records indicated that the communication channel was technically valid in all 61 episodes. However, out-of-band verification was performed in only 40 episodes (65.6%).

The relationship between out-of-band verification and execution outcomes was particularly pronounced within the constructed corpus. All 21 episodes without out-of-band verification were classified as Unauthorized Execution. The remaining 40 episodes, in which out-of-band verification was performed, were distributed among Attack Stopped (21 episodes), False Stoppage (10 episodes), and Legitimate Execution (9 episodes).

Table 2. Out-of-band verification and execution outcomes

Execution outcome	Verification performed	Verification not performed	Total
Attack Stopped	21	0	21
Unauthorized Execution	0	21	21
False Stoppage	10	0	10
Legitimate Execution	9	0	9
Total	**40**	**21**	**61**

Source: *Calculated from the scenario corpus.*

The observed distribution demonstrates complete separation between the verification categories and unauthorised execution outcomes in the analysed scenarios. Nevertheless, because the episodes were constructed, this relationship may reflect the scenario-generation and outcome-assignment procedures. It therefore cannot be interpreted as independent empirical proof that out-of-band verification eliminates unauthorised execution in operational environments.

The results also demonstrate the importance of distinguishing technical channel validity from request legitimacy. Although the communication channel was classified as technically valid in every episode, the final outcomes differed substantially. Within the analytical framework, technical channel validity was therefore insufficient to determine whether a privileged request should be executed.

Trust-boundary characteristics and supplier governance

Organisational trust-boundary crossings were recorded in 46 of the 61 episodes (75.4%). These episodes involved supplier relationships, contractors, or other external operational dependencies.

The remaining 15 episodes did not involve a recorded organisational trust-boundary crossing under the adopted classification.

The high proportion of cross-boundary episodes reflects the analytical emphasis on supplier-dependent digital infrastructure. It does not establish that such dependencies account for a comparable proportion of actual synthetic impersonation incidents.

The trust-boundary register also included risk classifications and contractual logging characteristics.

Of the 61 records, 58 (95.1%) were classified as medium, high, or critical risk. Contractual logging obligations were recorded in 41 episodes (67.2%).

These characteristics indicate that the constructed corpus included numerous decision pathways with elevated assigned risk levels, while contractual logging requirements were not present in every episode.

The difference between the number of medium-to-critical risk records and the number containing contractual logging obligations identifies a governance issue for further examination. However, the available aggregate figures do not establish how many high-risk episodes lacked logging obligations, because the corresponding cross-tabulation is not reported.

Consequently, the results support a descriptive distinction between risk classification and contractual logging coverage but do not independently demonstrate that the absence of logging obligations caused unauthorised execution or reduced decision-path reconstructability.

Evidence preservation and decision-path reconstructability

Evidence preservation and decision-path reconstructability were assessed as separate analytical characteristics.

Evidentiary materials were preserved in 49 of the 61 episodes (80.3%). However, only 40 decision pathways (65.6%) were classified as fully reconstructable.

Nine episodes (14.8%) were classified as partially reconstructable, while 12 episodes (19.7%) were classified as not reconstructable.

Table 3. Evidence preservation and decision-path reconstructability

Indicator	Number of episodes	Share, %
Evidence preserved	49/61	80.3
Fully reconstructable decision paths	40/61	65.6
Partially reconstructable decision paths	9/61	14.8
Not reconstructable decision paths	12/61	19.7

Source: *Calculated from the scenario corpus.*

The three reconstructability categories are mutually exclusive and collectively account for all 61 episodes. Evidence preservation represents a separate indicator and should not be added to these categories.

The results show that the number of episodes containing preserved evidence exceeded the number classified as fully reconstructable by nine. This difference is consistent with the conceptual distinction between retaining individual records and maintaining sufficient information to reconstruct an entire privileged decision pathway.

In the scenarios with incomplete reconstructability, available records did not always establish the complete sequence connecting the initiating communication, claimed organisational role, urgency framing, verification procedures, approval conditions, and resulting privileged action.

These findings indicate that the availability of evidence and the reconstructability of decision paths should be evaluated separately. However, the aggregate data do not establish a causal relationship between specific evidence-preservation practices and reconstructability outcomes.

Supplementary exercise observations

Four supplementary scenario-based exercise observations were examined separately from the principal analytical corpus.

The observations illustrated procedural weaknesses, including delayed callback validation, bypassing independent verification under subcontractor-related urgency, incorrect blocking due to uncertainty over designated authority, and accepting a Slack message as sufficient grounds for technical execution.

These observations provide qualitative illustrations of the decision-security mechanisms represented in the conceptual framework.

They were not included in the 61-episode statistical denominator and do not constitute independent empirical validation of the relationships observed in the constructed corpus.

Summary of descriptive findings

The principal descriptive findings are summarised in Table 4.

Table 4. Summary of the principal analytical indicators

Analytical domain	Indicator	Value
Operational conditions	High or critical urgency	51/61 (83.6%)
Execution outcomes	Attack Stopped	21/61 (34.4%)
Execution outcomes	Unauthorized Execution	21/61 (34.4%)
Execution outcomes	False Stoppage	10/61 (16.4%)
Execution outcomes	Legitimate Execution	9/61 (14.8%)
Verification	Out-of-band verification performed	40/61 (65.6%)
Trust-boundary governance	Episodes crossing an organisational trust boundary	46/61 (75.4%)
Trust-boundary governance	Medium-to-critical risk classifications	58/61 (95.1%)
Trust-boundary governance	Contractual logging obligations	41/61 (67.2%)
Evidence and accountability	Evidence preserved	49/61 (80.3%)
Evidence and accountability	Fully reconstructable decision paths	40/61 (65.6%)

Source: *Calculated from the scenario corpus.*

Overall, the results characterise the distribution of verification procedures, trust-boundary conditions, execution outcomes, and evidentiary characteristics within the constructed scenario corpus.

The most distinctive analytical pattern concerns the complete separation between episodes without out-of-band verification and those with verification in relation to unauthorised execution. However, the interpretation of this pattern depends on the assumptions underlying scenario construction and outcome assignment.

The findings also demonstrate that evidence preservation and full decision-path reconstructability represent different analytical characteristics. Their distinction is relevant for evaluating whether privileged decisions can subsequently be examined in terms of the initiating request, applied verification procedures, approval conditions, and operational outcome.

These observations provide a basis for discussing decision-security mechanisms and identifying propositions for subsequent empirical validation, without establishing population-level frequencies or causal effectiveness.

Discussion

The findings of this study provide a structured basis for examining synthetic impersonation as a decision-security challenge in digital infrastructure. The central analytical proposition is that the operational significance of synthetic impersonation extends beyond the authenticity of an individual communication artefact and concerns the conditions under which apparently credible communications are translated into privileged actions.

This interpretation is consistent with previous research addressing the implications of synthetic media for institutional trust, social engineering, and organisational security (Chesney & Citron, 2019; ISACA, 2024; Pedersen et al., 2025). It also builds on the execution-centred perspective developed by Jędrasiak (2026), which connects artefact assessment with identity assurance, authorisation, execution control, and evidentiary accountability.

However, the present study differs from observational incident research because its analytical material consists of constructed scenarios. Consequently, the results support an examination of predefined decision-security mechanisms and their relationships within the scenario

corpus rather than independent conclusions concerning the prevalence, effectiveness, or causal significance of those mechanisms in operational environments.

Synthetic impersonation and the conversion of communication credibility into operational authority

The first important implication concerns the distinction between the technical validity of a communication channel and the legitimacy of a privileged request.

All 61 analysed episodes involved technically valid communication channels, while their execution outcomes differed. Within the conceptual framework, this distinction illustrates why successful channel authentication cannot be considered equivalent to identity assurance, authority validation, or execution approval.

A communication transmitted through an authenticated account, recognised platform, or familiar organisational relationship may still contain an illegitimate request. Similarly, the identification of a legitimate individual does not independently establish that the individual possesses authority to initiate a specific privileged action.

This interpretation is consistent with Zero Trust Architecture, which rejects implicit trust assumptions and emphasises explicit verification and least-privilege access (Rose et al., 2020). It also corresponds with established security-control and digital identity principles requiring appropriate distinctions among authentication, authorisation, and accountability (Ross & Pillitteri, 2020; Temoshok et al., 2025).

The contribution of the present analytical framework is therefore not the introduction of a fundamentally new authentication principle. Rather, it concerns the integration of established assurance mechanisms into a complete organisational decision pathway.

Within this perspective, the critical analytical point is the transition from communication acceptance to operational execution. This transition requires explicit consideration of the requester's identity, the claimed organisational role, the authority associated with the requested action, and the conditions under which execution may be approved.

Synthetic impersonation introduces an additional source of uncertainty into this process because communication characteristics traditionally associated with personal identity or organisational familiarity may be artificially reproduced.

The resulting security concern is not limited to whether synthetic content can be detected. It also concerns whether the organisation can prevent communication credibility from becoming an unsupported substitute for operational authority.

Independent verification and execution outcomes

The second major finding concerns the relationship between out-of-band verification and execution outcomes.

Within the constructed corpus, all 21 episodes without out-of-band verification were classified as Unauthorized Execution. Conversely, the 40 episodes involving out-of-band verification were distributed among Attack Stopped, False Stoppage, and Legitimate Execution.

This distribution constitutes complete separation between the recorded verification categories and unauthorised execution outcomes within the analysed scenarios.

Nevertheless, its interpretation requires methodological caution.

Because the scenarios were constructed rather than independently observed, the relationship may reflect assumptions incorporated into scenario design, variable assignment, or outcome classification. Without a documented procedure demonstrating that verification status and execution outcomes were assigned independently, the observed separation cannot establish that out-of-band verification causally prevented unauthorised execution.

The result is therefore best interpreted as a scenario-dependent illustration of the role assigned to independent verification within the proposed decision-security framework.

From an operational perspective, out-of-band verification remains a relevant security mechanism because it can introduce an additional assurance step independent of the initiating communication.

However, the quality of this mechanism depends on how independence is established.

For example, a callback to a telephone number provided in the suspicious message may reproduce the original trust vulnerability rather than resolve it. Similarly, verification against outdated contact inventories, incorrectly designated approvers, or poorly maintained supplier records may fail to provide the intended assurance.

Independent verification should therefore be understood as a controlled procedure involving trusted contact information, defined authority relationships, and appropriate execution conditions rather than merely as the use of a second communication channel.

This interpretation connects the proposed framework with established principles of identity assurance, privileged access management, and organisational accountability (Rose et al., 2020; Ross & Pillitteri, 2020; Temoshok et al., 2025).

The scenario findings consequently support the formulation of a proposition for future empirical testing: independently maintained and appropriately implemented verification procedures may reduce the likelihood that apparently credible but illegitimate requests result in privileged execution.

Testing this proposition would require observational or experimental data in which verification procedures and outcomes are not predetermined by the scenario-generation process.

Trust boundaries and supplier-dependent decision pathways

The third analytical dimension concerns organisational trust boundaries and supplier dependencies.

Trust-boundary crossings were represented in 46 of the 61 episodes (75.4%). This substantial representation reflects the study's emphasis on digital infrastructure environments involving managed service providers, cloud operators, contractors, and other external operational relationships.

The result should not be interpreted as evidence that approximately three-quarters of real-world synthetic impersonation incidents involve supplier-related trust boundaries.

Instead, it demonstrates that the selected analytical corpus provides extensive coverage of decision pathways in which operational responsibility, communication, identity claims, or approval authority extend beyond a single organisational assurance context.

These conditions are relevant to cybersecurity supply-chain risk management because external operational relationships may distribute verification responsibilities and technical capabilities among several organisational actors.

Boyens et al. (2024) emphasise the importance of supplier-related security requirements, accountability, and risk-management practices. Similarly, the NIS2 Directive and DORA establish regulatory requirements relevant to cybersecurity risk management and third-party dependencies within their respective scopes (Directive (EU) 2022/2555, 2022; Regulation (EU) 2022/2554, 2022).

However, neither the presence of a supplier relationship nor the crossing of an organisational boundary necessarily constitutes a security failure.

The relevant question is whether the organisation maintains appropriate assurance procedures when requests move between different operational and authority contexts.

Within the proposed framework, supplier governance therefore extends beyond contractual recognition of a service relationship. It also concerns the practical availability of trusted callback contacts, clearly designated approvers, escalation procedures, execution restrictions, and evidentiary obligations.

The scenario corpus included contractual logging obligations in 41 of the 61 episodes (67.2%), while 58 episodes (95.1%) were classified as medium, high, or critical risk.

These figures indicate that contractual logging requirements were not recorded universally within a corpus containing a substantial number of elevated-risk scenarios.

However, the aggregate data does not establish which particular risk categories lacked logging obligations. Consequently, the results cannot support a precise estimate of a contractual logging gap among medium-to-critical risk episodes without additional cross-tabulation.

This distinction is important because descriptive differences between aggregate indicators should not be interpreted as evidence of a specific governance failure unless the underlying episode-level relationship has been demonstrated.

The findings nevertheless identify supplier verification and evidentiary obligations as relevant dimensions for further empirical investigation.

Evidence preservation and decision-path reconstructability

A particularly important conceptual contribution concerns the distinction between evidence preservation and the reconstructability of privileged decision pathways.

Evidence was preserved in 49 of the 61 episodes (80.3%), whereas 40 episodes (65.6%) were classified as fully reconstructable. Nine episodes were partially reconstructable, and 12 were not reconstructable.

These distributions illustrate that retaining evidentiary materials and establishing a complete sequence of organisational and technical decisions are not equivalent analytical outcomes.

A technical log may establish that a privileged configuration change occurred without documenting the communication that initiated the request, the identity claimed by the requester, the authority relied upon, or the verification procedure used by the approving operator.

Similarly, communication records may preserve the content of an apparently legitimate request without establishing whether the required approval and execution conditions were satisfied.

The proposed framework therefore distinguishes the existence of retained records from their sufficiency for reconstructing a complete decision pathway.

This distinction is consistent with established incident-handling and accountability principles, which emphasise the importance of relevant records for investigation, response, and organisational learning (Cichonski et al., 2012; Ross & Pillitteri, 2020).

It is also relevant to supplier-dependent environments, where information required for reconstruction may be distributed among external service providers, communication systems, privileged-access platforms, and internal operational teams.

However, the scenario results do not establish that specific evidence-preservation controls caused improvements in reconstructability. Such a conclusion would require additional analysis of the relationship between particular logging practices, record completeness, and reconstruction outcomes.

The analytical contribution is more specific: evidence preservation and reconstructability should be defined and assessed separately when evaluating the accountability of privileged decisions.

From an operational governance perspective, this distinction suggests that organisations should evaluate whether available records can explain not only what action occurred but also why the action was considered legitimate, who authorised it, and which verification procedures were applied.

Security effectiveness and the operational consequences of false stoppage

The scenario outcomes also highlight the importance of evaluating both unauthorised execution and the incorrect blocking of legitimate requests.

Within the corpus, 21 episodes were classified as Unauthorized Execution, and 10 as False Stoppage.

These categories represent different operational concerns.

Unauthorized Execution indicates that an action was performed without satisfying the required execution conditions. False Stoppage indicates that a legitimate request was prevented from proceeding.

The distinction is particularly relevant in digital infrastructure environments where privileged actions may be required to restore services, mitigate incidents, correct configurations, or maintain operational continuity.

Security procedures that prevent illegitimate actions may nevertheless impose operational costs if they incorrectly block authorised interventions.

Accordingly, a decision-security framework should not be assessed exclusively through the number of suspicious requests prevented from reaching execution.

It should also consider whether legitimate requests can be verified and processed within operationally acceptable conditions.

This observation supports a balanced evaluation perspective that includes both security and service-continuity outcomes.

Nevertheless, the constructed scenario corpus does not provide an independent basis for estimating the real-world frequency, severity, or financial consequences of false stoppage.

Such assessments would require operational data concerning the duration of delays, service criticality, recovery requirements, and consequences of rejected legitimate actions.

The role of synthetic-media detection

The findings also clarify the position of synthetic-media detection within the proposed framework.

The study does not evaluate the technical performance of specific deepfake detection systems. It therefore cannot establish whether media detection is more or less effective than independent verification or other organisational controls.

Research on synthetic media demonstrates the importance of identifying manipulated audiovisual and digital content (Chesney & Citron, 2019; Mirsky & Lee, 2021; Verdoliva, 2020).

However, the detection of suspicious content and the authorisation of a privileged action address different security questions.

Media analysis concerns the characteristics and potential authenticity of the communication artefact. Decision security concerns whether the requested action is legitimate and whether its execution satisfies the required organisational and technical conditions.

These mechanisms should therefore be considered complementary rather than interchangeable.

Synthetic-content indicators may contribute to risk assessment, escalation, and verification prioritisation. However, the absence of detectable manipulation should not independently establish the legitimacy of a privileged request.

Conversely, a request identified as potentially synthetic may require further investigation without automatically establishing that the associated operational action is illegitimate.

The execution-centred perspective therefore places media-authenticity assessment within a broader decision-security process that also incorporates identity assurance, authority validation, execution approval, and evidentiary accountability.

Theoretical and methodological implications

The theoretical contribution of the study lies in integrating several established cybersecurity concepts into a common analytical structure focused on privileged decision pathways.

Existing approaches to synthetic-media security primarily address the authenticity, manipulation, detection, and consequences of generated content. Identity and Access Management and Zero Trust Architecture address authentication, authorisation, access privileges, and explicit verification. Incident-handling and security-control frameworks address accountability, logging, and operational response.

The proposed execution-centred framework connects these dimensions by treating a privileged decision episode as a socio-technical sequence extending from the initiating communication to the final execution outcome and its evidentiary reconstruction.

This integration offers an analytical perspective for examining where communication credibility may be incorrectly converted into operational authority.

The methodological contribution concerns the representation of privileged decision episodes through correlated records describing communication, verification, trust-boundary characteristics, execution outcomes, and evidentiary conditions.

Such a structure enables systematic descriptive comparisons and may support the development of future observational or experimental studies.

However, the use of constructed scenarios limits the extent to which the present results can validate the proposed model.

In particular, the perfect separation observed between verification status and unauthorised execution should be treated as a methodological issue requiring explicit explanation rather than as conclusive evidence of control effectiveness.

The framework's scientific value therefore depends not only on the internal coherence of its conceptual structure but also on its subsequent testing against independently observed operational decisions.

Practical implications for digital infrastructure governance

The findings provide a basis for identifying operational controls and performance indicators that may be considered in future implementation and evaluation.

Relevant measures include the proportion of high-impact privileged requests subjected to independent verification, the completeness and currency of trusted callback inventories, the availability of designated approvers, the documentation of supplier-related verification responsibilities, and the reconstructability of privileged decision pathways.

Execution outcomes should also be evaluated in a manner that distinguishes unauthorised execution, successful prevention of illegitimate requests, legitimate execution, and false stoppage.

These indicators are consistent with the broader need to translate cybersecurity governance principles into observable and auditable operational procedures (Jędrasiak & Gawliczek, 2025; Ross & Pillitteri, 2020).

Nevertheless, the present study does not establish validated performance thresholds for these measures.

Organisations would therefore need to determine appropriate implementation requirements through risk assessment, operational testing, and consideration of sector-specific constraints.

Particular attention should be given to supplier-dependent privileged decision pathways because they may involve fragmented responsibility for identity verification, approval, technical execution, and evidence retention.

The practical implication is not that every request crossing an organisational boundary requires an identical procedure. Rather, verification and execution controls should be proportionate to the consequences of the requested action and the assurance available concerning the requester's identity and authority.

Synthesis of the findings

Taken together, the scenario-based findings support three principal analytical observations.

First, technical communication validity and privileged request legitimacy represent different assurance dimensions. The execution-centred framework provides a structured means of examining the transition from communication acceptance to operational action.

Second, independent verification is an important component of the proposed decision-security model. However, the complete separation between verification categories and unauthorised execution within the constructed corpus cannot independently demonstrate causal effectiveness.

Third, evidence preservation and full decision-path reconstructability are distinct characteristics of operational accountability. Their separation is particularly relevant in supplier-dependent environments where communication, approval, execution, and logging responsibilities may be distributed across organisational boundaries.

These observations reinforce the relevance of an execution-centred approach to synthetic impersonation while defining the limits of the present study's conclusions.

Further validation using independently observed operational incidents, controlled exercises, and systematically varied verification conditions is necessary to establish whether the proposed framework improves decision-security outcomes in real-world digital infrastructure environments.

Limitations

This study has several methodological limitations that should be considered when interpreting its findings and assessing their applicability to operational cybersecurity environments.

First, the analytical corpus consisted of 61 constructed, scenario-instantiated privileged decision episodes rather than a representative sample of naturally occurring cybersecurity incidents. The scenarios were designed to examine selected organisational and technical characteristics of privileged decision-making under synthetic impersonation pressure. Consequently, the observed distributions describe the constructed corpus and cannot be generalised to the actual frequency of synthetic impersonation incidents, trust-boundary crossings, or execution failures in digital infrastructure.

Second, the relationships identified between verification procedures and execution outcomes may be influenced by the assumptions used during scenario construction. In particular, all 21 episodes without out-of-band verification were classified as Unauthorized Execution, while none of the 40 episodes involving such verification resulted in that outcome. This complete separation may reflect the scenario-generation or outcome-assignment rules rather than an independently observed relationship. Accordingly, the findings do not establish that out-of-band verification causally prevents unauthorised execution or that its effectiveness can be generalised to operational environments.

Third, synthetic-media fidelity was not manipulated as an independent experimental variable. The scenarios assumed that initiating communications were sufficiently plausible to receive operational attention. Therefore, the study cannot determine how variations in voice similarity, audiovisual quality, linguistic sophistication, contextual credibility, or other synthetic-content characteristics influence decision-making or verification outcomes.

Fourth, the analytical corpus represented four selected operational archetypes: managed security service providers, cloud providers, enterprise data centres, and managed IT services. Although these environments reflect relevant forms of supplier dependence and privileged operational activity, they do not encompass the full diversity of digital infrastructure arrangements, regulatory conditions, organisational cultures, or sector-specific decision processes.

Fifth, the classification of trust-boundary characteristics, risk levels, execution outcomes, and evidentiary reconstructability depends on the definitions and coding procedures applied within the analytical framework. Without independent coding, documented inter-rater agreement, or external validation, the consistency of these classifications cannot be interpreted as evidence of their general applicability.

Sixth, evidence preservation and decision-path reconstructability were examined through scenario-associated records rather than through independently collected forensic material from real incidents. Consequently, the study demonstrates how these characteristics can be represented analytically but does not establish the practical performance of particular logging architectures, evidence-management procedures, or forensic reconstruction methods.

Seventh, the four supplementary exercise observations were used for qualitative illustration and were not incorporated into the principal statistical denominator. Their limited number and

scenario-based character prevent their interpretation as independent confirmation of the patterns identified in the main corpus.

Finally, the study did not evaluate the financial costs, operational delays, implementation complexity, or comparative effectiveness of alternative verification procedures. Although the classification of False Stoppage recognises the potential operational consequences of incorrectly blocking legitimate requests, the analysis does not quantify these consequences or establish an optimal balance between security assurance and operational continuity.

Future research should therefore employ independently observed operational incidents, controlled exercises, or experimental designs in which verification conditions and execution outcomes are determined separately. Additional priorities include transparent scenario-generation protocols, independent coding procedures, systematic variation of synthetic-content characteristics, and the evaluation of decision-security controls across different organisational and supplier environments.

Such research would provide a stronger basis for assessing the external validity, causal effectiveness, and operational applicability of the proposed execution-centred decision security framework.

Conclusion

This study examined synthetic impersonation as a decision-security challenge in digital infrastructure, focusing on the organisational and technical conditions under which apparently credible communications may influence privileged execution across trust boundaries.

The research developed an execution-centred analytical framework integrating communication credibility, authentication, identity assurance, authority validation, independent verification, execution approval, and evidentiary accountability. The framework was applied to 61 constructed privileged decision episodes representing managed security service providers, cloud providers, enterprise data centres, and managed IT services.

Regarding the first research question, the analysis demonstrated the conceptual importance of distinguishing communication-channel validity from the legitimacy of a privileged request. Although all analysed episodes involved technically valid communication channels, their execution outcomes differed. This finding illustrates why communication authenticity, identity assurance, authority validation, and execution approval should be treated as separate components of a privileged decision pathway.

Regarding the second research question, the scenario corpus exhibited complete separation between the absence of out-of-band verification and unauthorised execution. All 21 episodes without such verification were classified as Unauthorized Execution, whereas the 40 episodes involving verification were distributed among Attack Stopped, False Stoppage, and Legitimate Execution. However, because the corpus was constructed, this relationship may reflect the assumptions underlying scenario generation and outcome assignment. It therefore cannot be interpreted as independent empirical evidence that out-of-band verification eliminates unauthorised execution in operational environments.

Regarding the third research question, the analysis identified a distinction between evidence preservation and full decision-path reconstructability. Evidence was preserved in 49 episodes, while only 40 decision pathways were classified as fully reconstructable. This difference highlights the analytical importance of assessing whether retained records establish the complete sequence connecting the initiating communication, verification procedures, authority validation, execution approval, and operational outcome.

The principal scientific contribution of the study lies in integrating established cybersecurity and digital identity principles into a unified execution-centred framework for examining privileged decision pathways under synthetic impersonation pressure. The framework extends the analytical

focus beyond the detection of manipulated communication artefacts towards the organisational processes through which apparently credible requests may become actionable.

The study also identifies decision-path reconstructability as a distinct dimension of operational accountability, complementing conventional evidence-preservation and logging considerations.

From a practical perspective, the proposed framework provides a basis for examining independent verification procedures, supplier-related authority arrangements, privileged execution controls, and the completeness of evidentiary records. Its application may support the development of operational assessment procedures and performance indicators, although their effectiveness requires independent empirical validation.

The findings are subject to the limitations of the constructed scenario-based design. They do not establish the population-level prevalence of synthetic impersonation incidents or the causal effectiveness of particular security controls.

Further research should therefore test the proposed framework using independently observed operational cases, controlled exercises, and systematically varied verification conditions. Particular attention should be given to the relationship between verification independence, supplier governance, false stoppage, unauthorised execution, and decision-path reconstructability.

Overall, the study supports an execution-centred perspective in which the security of privileged operations depends not solely on the apparent authenticity of communications but on the integrity, verification, authorisation, and accountability of the complete decision pathway.

Funding

To be completed by the authors. Specify the funding source and grant number, or confirm that the research received no external funding.

Competing interests

The authors declare that they have no competing interests.

Use of Artificial Intelligence

Artificial intelligence tools were used solely for technical text checking.

References

Bohacek, M., & Farid, H. (2022). Protecting world leaders against deep fakes using facial, gestural, and vocal mannerisms. *Proceedings of the National Academy of Sciences, 119*(48), e2216035119. https://doi.org/10.1073/pnas.2216035119

Boyens, J., Smith, A., Bartol, N., Winkler, K., Holbrook, A., & Fallon, M. (2024). *Cybersecurity supply chain risk management practices for systems and organizations* (NIST Special Publication 800-161 Rev. 1, Update 1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-161r1-upd1

Buchwald, P., & Jędrasiak, K. (2025). A method to protect users from fake news disinformation content using hybrid evaluation methods based on artificial intelligence solutions and user experience. *Safety & Fire Technology, 66*(2), 6–23. https://doi.org/10.12845/sft.66.2.2025.1

Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review, 107*, 1753–1820. https://doi.org/10.15779/Z38RV0D15J

Cichonski, P. R., Millar, T., Grance, T., & Scarfone, K. (2012). *Computer security incident handling guide* (NIST Special Publication 800-61 Rev. 2). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-61r2

Cybersecurity and Infrastructure Security Agency. (2024). *Risk in focus: Generative AI and the 2024 election cycle*. https://www.cisa.gov/sites/default/files/2024-05/Consolidated_Risk_in_Focus_Gen_AI_ElectionsV2_508c.pdf

Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union. (2022). *Official Journal of the European Union*. https://eur-lex.europa.eu/eli/dir/2022/2555/oj/eng

European Union Agency for Cybersecurity. (2022). *Remote identity proofing: Attacks and countermeasures*. https://www.enisa.europa.eu/publications/remote-identity-proofing-attacks-countermeasures

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

European Union Agency for Cybersecurity. (2025). *ENISA sectorial threat landscape: Public administration*. https://doi.org/10.2824/4606183

Europol. (2025). *The changing DNA of serious and organised crime: EU-SOCTA 2025*. https://www.europol.europa.eu/publications-events/main-reports

Federal Bureau of Investigation. (2025, May 15). *Senior U.S. officials impersonated in malicious messaging campaign*. https://www.fbi.gov/investigate/cyber/alerts/2025/senior-us-officials-impersonated-in-malicious-messaging-campaign

INTERPOL. (2024). *Beyond illusions*. https://www.interpol.int/content/download/21179/file/BEYOND%20ILLUSIONS_Report_2024.pdf

INTERPOL. (2025). *Africa cyberthreat assessment report 2025*. https://www.interpol.int/content/download/23222/file/2025%20Africa%20Cyberthreat%20Assessment%20Report.pdf

INTERPOL. (2026). *Asia and South Pacific cyber threat assessment report 2025/2026*. https://www.interpol.int/content/download/24327/file/CYBER_ASP%20Cyber%20Threat%20Assessment%20Report_2025_2026_v4.pdf

ISACA. (2024). *Examining authentication in the deepfake era*. https://www.isaca.org/resources/white-papers/2024/examining-authentication-in-the-deepfake-era

Jędrasiak, K. (2025a). Analiza cyberzagrożeń współczesnych kanałów komunikacyjnych ze szczególnym uwzględnieniem zagrożeń typu DeepFake. In *Horyzonty sztucznej inteligencji a Przemysł 5.0* (pp. 37–59). Wydawnictwo Naukowe Akademii WSB. https://bazawiedzy.wsb.edu.pl/info/article/AWSB4fa8c261dcf04dfd9c1d3fbceed5c15d/

Jędrasiak, K. (2024). Audio stream analysis for deep fake threat identification. *Civitas et Lex, 1*(41), 21–35. https://doi.org/10.31648/cetl.9684

Jędrasiak, K. (2025b). Analiza cech jakości obrazu oraz artefaktów przetwarzania jako fundament detekcji deepfake. *Safety & Fire Technology, 66*(2), 168–201. https://doi.org/10.12845/sft.66.2.2025.10

Jędrasiak, K. (2026). Decision security against synthetic impersonation in high trust sectors. *European Cybersecurity Journal, 11*(1), 53–66. [Publication not independently verified.]

Jędrasiak, K., & Bijoch, J. (2025). Deep network representations as reliable indicators of synthetic content in audiovisual and clinical contexts. *Communications of International Proceedings, IBIMA 46*. [Publication not independently verified.]

Jędrasiak, K., & Gawliczek, P. (2025). Implementing artificial intelligence in data-driven enterprises through a ten-phase framework based on multiple case studies. *Social Development and Security, 15*(5), 17–34. https://doi.org/10.33445/sds.2025.15.5.2

Wolański, R., & Jędrasiak, K. (2023). Audio-video analysis method of public speaking videos to detect deepfake threat. *Safety & Fire Technology, 62*(2), 172–180. https://doi.org/10.12845/sft.62.2.2023.10

Jędrasiak, K., & Wolański, R. (2024). An image analysis algorithm for detecting modified content on the internet against cybercrime: A technique for estimating the probability of modification. *Safety & Fire Technology, 63*(1), 88–94. https://doi.org/10.12845/sft.63.1.2024.7

Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? *Business Horizons, 63*(2), 135–146. https://doi.org/10.1016/j.bushor.2019.11.006

Mirsky, Y., & Lee, W. (2021). The creation and detection of deepfakes: A survey. *ACM Computing Surveys, 54*(1), 1–41. https://doi.org/10.1145/3425780

Pedersen, K. T., Pepke, L., Stærmose, T., Papaioannou, M., Choudhary, G., & Dragoni, N. (2025). Deepfake-driven social engineering: Threats, detection techniques, and defensive strategies in corporate environments. *Journal of Cybersecurity and Privacy, 5*(2), Article 18. https://doi.org/10.3390/jcp5020018

Regulation (EU) 2022/2554 of the European Parliament and of the Council of 14 December 2022 on digital operational resilience for the financial sector. (2022). *Official Journal of the European Union*. https://eur-lex.europa.eu/eli/reg/2022/2554/oj/eng

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence. (2024). *Official Journal of the European Union*. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). *Zero trust architecture* (NIST Special Publication 800-207). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-207

Joint Task Force. (2020). *Security and privacy controls for information systems and organizations* (NIST Special Publication 800-53 Rev. 5). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-53r5

Temoshok, D., Abruzzi, C., Choong, Y.-Y., Fenton, J. L., Galluzzo, R., LaSalle, C., Lefkovitz, N., Regenscheid, A., & Vachino, M. (2025). *Digital identity guidelines: Identity proofing and enrollment* (NIST Special Publication 800-63A-4). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-63a-4

Verdoliva, L. (2020). Media forensics and deepfakes: An overview. *IEEE Journal of Selected Topics in Signal Processing, 14*(5), 910–932. https://doi.org/10.1109/JSTSP.2020.3002101

Verizon. (2025). *2025 data breach investigations report*. https://www.verizon.com/business/resources/Tc28/reports/2025-dbir-data-breach-investigations-report.pdf

This is an open access journal and all published articles are licensed under a Creative Commons «Attribution» 4.0.