International Journal of Computational Innovations, Intelligent Systems and AI
July 2026, Volume: **2**, N.º **2** | Pages 7–25
doi: 10.64439/cisai.v2i2.43
e-ISSN : 3084-7907

ARTICLE

Modern cybersecurity architecture for fraud prevention in administrative services

Arquitectura moderna de ciberseguridad para prevención de fraudes en servicios administrativos

Enrique Castellares Cuya* · **Jose Rengifo Espina**

Facultad de Ingeniería, Universidad Tecnológica del Perú, Lima, Perú

*Corresponding Author: 1221978@utp.edu.pe

Submitted: 01 June 2026 Accepted: 14 July 2026 Published online: 15 July 2026

Abstract

As service organizations advance in their digital modernization, security does not always keep pace, generating vulnerabilities that conventional protection schemes fail to address. To tackle this problem, the present work proposes a cybersecurity architecture oriented toward fraud prevention in a service sector company in Lima, Peru, whose design is grounded in the documentary analysis of 385 technical incident records. This analysis revealed that only 54.3% of vulnerabilities were addressed annually, with monthly fluctuations between 44.4% and 62.5%, while the effective response rate did not exceed 34%, reaching lows of 19% during periods of peak operational load. These findings, combined with the concentration of unresolved incidents in malware, unknown threats, and ransomware, expose structural failures in the prioritization and closure of security events. In response to this diagnosis, the proposed architecture integrates identity and access management, advanced threat protection, cloud security, regulatory compliance, and extended analytics powered by artificial intelligence, forming a defense-in-depth capable of reducing residual exposure and sustaining a robust anti-fraud response in digitalized administrative environments.

Keywords: Computer security, risk management, threat detection, operational continuity, artificial intelligence

Resumen

A medida que las organizaciones de servicios avanzan en su modernización digital, la seguridad no siempre crece al mismo ritmo, generando vulnerabilidades que los esquemas convencionales de protección no logran cubrir. Para abordar este problema, el presente trabajo propone una arquitectura de ciberseguridad orientada a la prevención de fraudes en una empresa del sector servicios de Lima, Perú, cuyo diseño se fundamenta en el análisis documental de 385 registros técnicos de incidentes. Dicho análisis evidenció que solo el 54,3% de las vulnerabilidades fueron atendidas con oscilaciones mensuales entre 44,4% y 62,5%. Mientras que el índice de respuesta efectiva no superó el 34%, con mínimos del 19% en los períodos de mayor carga operativa. Estos resultados, junto con la concentración de incidentes no resueltos en malware, amenazas desconocidas y ransomware, revelan fallas en la priorización y el cierre de eventos de seguridad. Ante este diagnóstico, la arquitectura propuesta articula gestión de identidades y accesos, protección avanzada contra amenazas, seguridad en entornos cloud, cumplimiento normativo y analítica extendida potenciada con inteligencia artificial, conformando una defensa en profundidad capaz de reducir la exposición residual y sostener una respuesta antifraude robusta en entornos administrativos digitalizados.

Palabras clave: Seguridad informática, gestión de riesgos, detección de amenazas, continuidad operacional, inteligencia artificial

Cite as: E. Castellares Cuya and J. Rengifo Espina, "Modern cybersecurity architecture for fraud prevention in administrative services," *CISAI*, vol. 2, no. 2, pp. 7–25, July 2026, doi: 10.64439/cisai.v2i2.43.

1. Introduction

In recent years, digitalization has gone from being a competitive advantage to becoming an operational condition for most organizations. The massive incorporation of cloud services, transactional platforms, and computer systems has brought with it a considerable expansion of the attack surface, exposing organizations to increasingly sophisticated threats. In this scenario, computer fraud has ceased to be a peripheral risk and has become a threat that compromises availability, operational continuity, institutional reputation, and user trust, with financial consequences of growing magnitude. Some recent cases help to gauge the real impact of these attacks. Houghton [1] documents the attack on the British Library, with estimated losses of between six and seven million pounds sterling, while Wadesango et al. [2] report that cyberattacks against Zimbabwean institutions generate annual losses of close to 1.8 billion dollars. In the Peruvian context, León [3] records growth of more than 26% in economic cybercrime and more than 30,000 cases of electronic fraud in the first quarter of 2024, with a year-on-year increase of 25%. These figures reveal that no sector or organizational scale is exempt from exposure, including administrative service companies with supporting technological infrastructure.

Faced with this panorama, the advance of knowledge in this field shows valuable, albeit partial, contributions. From a technological perspective, Wahid and Hassini [4] demonstrate that automation and artificial intelligence not only contribute to threat detection, but also help with patterns of digital fraud by endowing it with greater speed, adaptability, and evasion capacity. However, this finding takes on particular relevance when one considers that attackers exploit these same technologies to make their methods more sophisticated, which is why Abu-Khadrah [5] warns that current attacks combine identity theft, malware, credential exploitation, and configuration abuse in a coordinated manner, which makes them difficult to contain through isolated controls. Both contributions agree in pointing out that the complexity of digital fraud requires equally integrated responses, capable of operating simultaneously across multiple layers of the organizational infrastructure. From an executive and organizational standpoint, the literature complements this approach with equally relevant elements. For example, Agyepong et al. [6] underscore that measuring performance in security operations centers is a necessary condition for sustaining an effective response over time, since without clear indicators, gaps tend to become normalized and invisible to management. Added to this is the warning from Yaseen [7], who points out that technology adoption only generates real value when accompanied by user trust, which implies that any security architecture must also consider the human and cultural dimension. Consistent with this, Cevallos et al. [8] associate cyber resilience with sustained organizational capabilities, thereby positioning security as an institutional attribute rather than an isolated technical function. Despite these advances, an applied gap persists regarding how to translate attack records into a measurable anti-fraud architecture, especially in service organizations that have scattered tools but lack unified indicators, defined control domains, and integrated implementation sequences.

To address this gap, the present article provides a modern cybersecurity architecture oriented toward fraud prevention in administrative services, grounded in the analysis of technical incident records documented during 2025 at a company in Lima, Peru, which reported a high rate of attacks, partial interruptions to its operations, and estimated losses of 120 thousand soles. This context justifies the need for a structured response and offers an empirical basis for formulating a verifiable solution. Hereafter, the article is structured into six sections. Section 2, literature review, develops the conceptual and empirical foundations underlying the proposal. Section 3, case study, characterizes the organizational context that motivates the proposal. Section 4, methodology, describes the study's approach, the documentary source, the variables, the indicators, and the analytical procedure. Section 5, results, presents the vulnerability diagnosis, the performance of the protection system, and the proposed architecture in response to the gaps identified. Section 6, discussion, interprets the findings in light of the literature and specifies their implications for cybersecurity management. Section 7, conclusions, synthesizes the main contribution, the study's limitations, and the future validation route.

2. Literature review

This section examines the theoretical foundations underlying the study's architectural proposal. It is organized into three axes: fraud detection and technology adoption, security operations and their performance indicators, and cyber resilience supported by artificial intelligence and governance. Based on this review, the consolidated advances in the discipline are identified, as well as the applied gap that justifies the contribution of the present work.

2.1 Fraud detection and technology adoption

The understanding of digital fraud has undergone a significant transformation in recent years, moving from approaches based on static rules toward hybrid systems capable of combining supervised learning, behavioral analytics, and institutional trust criteria. In this regard, Wahid and Hassini [4] highlight the value of hybrid frameworks with artificial intelligence for transactional platforms, demonstrating that combining multiple layers of analysis significantly improves detection capacity. Along the same lines, Afriyie et al. [9] show that supervised algorithms achieve greater accuracy when trained with consistent patterns, underscoring the importance of having quality historical records as the basis for any solution. However, technology alone does not guarantee results if users do not perceive transparency, fairness, and reliability in the systems, an aspect that Yaseen [7] identifies as decisive for the success of any implementation. In the context of administrative services, this tension between technical capacity and organizational acceptance is especially relevant, because the architecture must protect sensitive assets while sustaining operational continuity, traceability, and auditable decisions. Recognizing this duality is the starting point, but it is not sufficient without mechanisms that allow systematic measurement, prioritization, and response, an aspect addressed in the following axis.

2.2 Security operations and performance indicators

Identifying threats is only part of the problem, since the other, equally critical, part is the capacity to manage them efficiently and consistently over time. Along these lines, Agyepong et al. [6] propose evaluating the performance of analysts and operations centers through indicators that go beyond simply counting alerts, recognizing that the real effectiveness of a security team is not reflected in the volume of events processed but in the quality and timeliness of its response. Complementing this perspective, Forsberg and Frantti [10] highlight the importance of evaluating vulnerabilities by considering detection, prioritization, and remediation as interdependent stages, such that a failure in any one of them compromises the final outcome. For their part, Umer et al. [11] link mitigation efficiency with deterrence mechanisms and technological trust, adding an institutional dimension to operational management. These contributions reveal that an organization can have advanced tools and still fail if it does not integrate and govern them with comparable indicators and defined processes. The challenge, therefore, does not consist of accumulating technological solutions, but of building an architecture in which events, identities, assets, controls, and responses flow with low friction and high visibility. This level of integration, however, requires something more than well-measured processes, since it also demands that the organization be capable of anticipating, adapting, and recovering, which places the discussion within the realm of cyber resilience addressed in the next subsection.

2.3 Cyber resilience, artificial intelligence, and governance

Efficient operational management opens the way to a more strategic dimension of the problem, namely organizations' capacity to anticipate risks, absorb impacts, and learn from each incident. Within this framework, Nandy et al. [12] discuss the use of predictive models to anticipate risks before they materialize, representing a paradigmatic shift from traditional reactive approaches. Consistent with this, Moura et al. [13] analyze the role of artificial intelligence in incident detection and response, while Raj [14] links modern architectures with adaptive controls and extended monitoring capable of dynamically adjusting to the threat landscape. This orientation converges with contemporary security approaches that integrate cloud protection, identity management, continuous monitoring, data security, and regulatory compliance into a cohesive framework. However, automation without clear prioritization criteria can produce alert fatigue or opaque decisions, which calls for explicit governance of security processes. Despite these advances, the

literature reviewed addresses each dimension independently, without offering a proposal that articulates detection, operation, and resilience into an integrated, measurable, and replicable architecture for service organizations. It is precisely this gap that the present study seeks to address, proposing a sociotechnical architecture that articulates advanced tools, measurable processes, clear roles, training, periodic auditing, and continuous improvement as a verifiable response to the problem of fraud in digitalized administrative environments.

Therefore, the foregoing shows that organizational cybersecurity has matured from ad hoc detection toward an integrated vision that combines technology, operation, and governance. However, this theoretical maturity contrasts with the reality of many service organizations, where capabilities remain fragmented and indicators nonexistent. It is from that gap that the present study builds its proposal, translating the foundations reviewed into a measurable architecture adapted to the operational context analyzed, whose basis is described below in the methodology section.

3. Case study

The organization analyzed is a Lima-based company that provides logistical, financial, accounting, technological, and administrative support services to a group of companies in the Peruvian automotive sector. Its entire operation depends on digital platforms, enterprise management systems, and cloud services, which means that any failure in its technological infrastructure is not confined to a single area, but spreads immediately to all the companies in the group, paralyzing processes, blocking operations, and generating losses that accumulate with every hour of downtime. This dependence makes the organization a particularly attractive target for attackers, who can exploit any breach with a far-reaching cascading effect.

Despite this exposure, the level of protection is not commensurate with the risk it faces. The available records documented more than 600 attack attempts, 20 partial service interruptions, and estimated economic losses of 120 thousand soles, making clear that the situation required a deeper and more structured response. With this scenario as a starting point, the present study set out to go beyond recording what happened, in order to understand the underlying causes of these gaps and translate them into a cybersecurity architecture that not only corrects what failed, but also offers a solid basis for progressively reducing risk exposure.

4. Methodology

This section explains how the study was carried out, what information was used, and how it was analyzed to arrive at the architectural proposal. It is organized into four subsections describing the study's approach, how the data were selected, the indicators that were calculated, and the procedure followed to obtain the results.

4.1 Approach and unit of analysis

In order to propose a cybersecurity solution that works in practice, it is first necessary to understand what is failing. For this reason, the starting point of the study was to review the records left by the organization's security system throughout 2025. Each time the system detected an attack attempt, an alert, or a security problem, it generated a record with information on what happened, when it happened, and whether or not it was resolved. Based on this analysis, it was possible to precisely identify the weakest points of the system and determine the elements the proposal should address, so that it responds to what is actually happening rather than to hypothetical scenarios.

4.2 Sample, criteria, and instruments

This subsection describes the origin of the data used in the study, the criteria a record had to meet to be included in the analysis, and the instrument used for its extraction. Since security systems continuously generate records of everything that happens in the infrastructure, not all of them contain sufficient information to be rigorously analyzed, since some are incomplete, others are duplicates, and others simply do not provide useful data. For this reason, clear selection criteria were defined to allow work with only the most complete and reliable records, including those with the date and time of the event, type of attack,

severity level, resolution status, and other fields key to the analysis. After this review and cleaning process, 385 valid records remained, distributed across the twelve months of 2025. To extract the information in an orderly and consistent manner, a structured extraction form was used, designed from the format already employed by the organization's own security team, which facilitated data collection and ensured that all data were comparable with one another. Table 1 illustrates a concrete example of what a valid record looked like once extracted using this form.

Tabla 1. Example of a record included in the analysis

Field	Example record
Date and time	14/03/2025 – 02:47
Attack type	Malware
Severity level	High
Confidence level	94%
Agent version	6.2.1
Analysis category	Endpoint
Mitigation status	Not attended

4.3 Variables and indicators

With the records cleaned, the next step was to define how to measure the actual state of cybersecurity. To this end, two indicators were calculated. The first is the addressed-vulnerability rate, which answers a specific question: of all the security problems detected, how many were actually resolved. If this percentage is low, it means that many threats remain unaddressed, leaving the organization exposed. This indicator is calculated as shown in Equation 1.

$$\text{AVR}(\%) = \frac{\text{Vulnerabilities addressed}}{\text{Total vulnerabilities}} \times 100 \tag{1}$$

The second indicator is the response index, which measures how many of the alerts generated by the system were correctly managed, that is, investigated, classified, and adequately resolved. This indicator shows whether the security team not only detects problems but is also capable of resolving them, and is expressed through Equation 2.

$$\text{RI}(\%) = \frac{\text{Alerts correctly processed}}{\text{Total alerts}} \times 100 \tag{2}$$

Both indicators were calculated month by month and then summarized into an annual average and a monthly variation range. Taken together, they make it possible to build a diagnostic picture of the actual state of cybersecurity in the organization, overcoming the limitation of traditional detection metrics, which record what happens but do not indicate whether it was resolved or how effectively. In this sense, the two indicators are articulated so that one bounds the magnitude of unresolved exposure while the other reveals whether the security team has the capacity to close the full incident-management cycle. This duality is therefore what makes it possible to draw a relevant distinction between an organization that detects threats but does not neutralize them and one that responds reactively but without the systematicity needed to sustain a robust security posture, a distinction that is decisive for precisely guiding the components of the proposed architecture. Under this premise, the next subsection addresses the analysis procedure and the ethical aspects that have guided the research.

4.4 Analysis procedure and ethical aspects

Once the indicators had been defined and the records selected, the analysis was carried out in four sequential steps that made it possible to move from raw data to diagnostic conclusions. In the first step, all records were reviewed to discard those with incomplete or repeated information, thereby ensuring that the subsequent analysis rested on reliable data. In the second step, the valid events were grouped by month,

by threat type, by severity level, and by resolution status, which made it possible to clearly identify which types of attack recurred most frequently and which systematically went unresolved. In the third step, the two indicators were calculated month by month, providing a complete view of how the security situation evolved throughout the year. In the fourth and final step, each problem identified was related to a specific area of the cybersecurity architecture, thereby establishing a direct connection between what failed and what the proposed architecture seeks to correct. Regarding ethical aspects, the analysis was carried out on aggregated records that do not contain identifiable personal data, and therefore did not involve any intervention on individuals. The findings obtained from this procedure are presented in the following section.

5. Results

This section presents the findings obtained from the analysis of the 385 technical security records documented during 2025. It is organized into three subsections corresponding to the study's three analytical axes: the vulnerability of administrative services to attack attempts, the performance of the existing protection system, and the cybersecurity architecture proposed in response to the gaps identified. The results are expressed through descriptive statistics, counts, percentages, and ranges, since the study did not formulate hypotheses or apply inferential contrast tests.

5.1 Vulnerability of administrative services

This subsection characterizes the organization's exposure to computer threats during 2025, based on the indicators calculated from the technical incident records. Over the course of the year, 385 security events were documented, with a monthly frequency ranging between 14 and 101 attacks. The month with the highest volume was November, with 101 events recorded, while the months with the lowest activity were September and October, with 14 and 18 attacks respectively. Figure 1 presents the monthly distribution of these events.

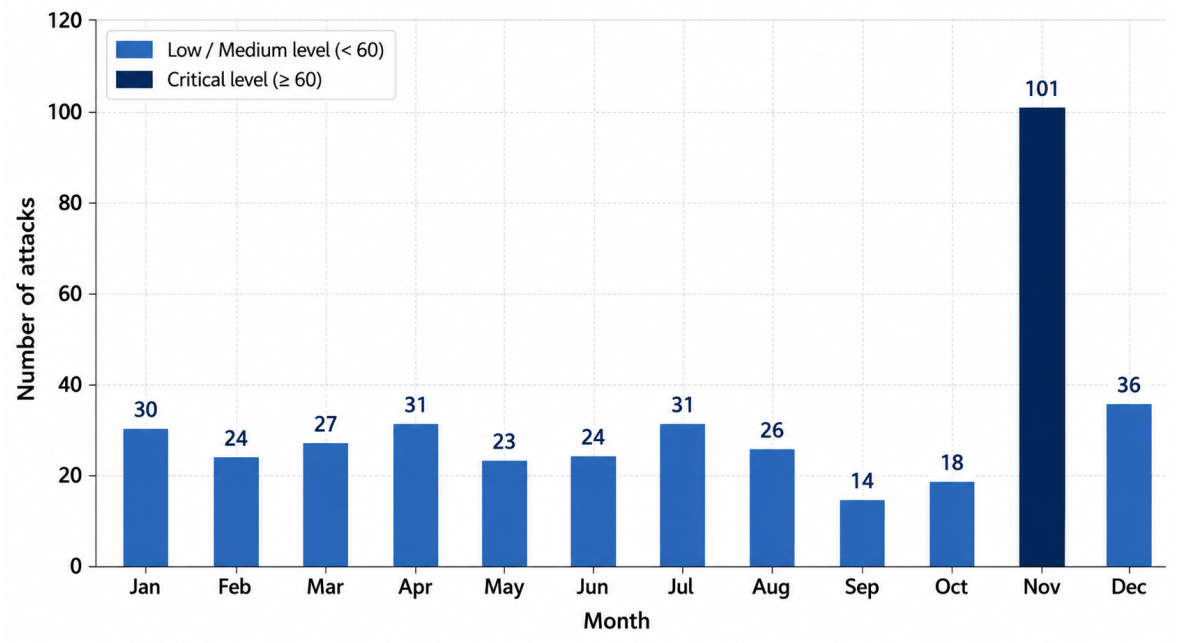

Figura 1. Monthly number of attacks recorded in 2025

Classification by monthly frequency range made it possible to identify three levels of criticality. Eight months fell within a range of 10 to 30 attacks and were classified as low level, three months fell between 30 and 60 attacks and corresponded to the medium level, and one month exceeded 60 attacks, reaching the critical level, associated with the increase in activity typical of the year-end closing period. This distribution shows that extreme exposure was concentrated in a specific period, although the presence of attacks was constant throughout the year.

Regarding threat type, unknown and packed-type attacks predominated at the lower criticality levels, malware occurred recurrently at the medium and high levels with 54 events recorded at high criticality, and ransomware was concentrated at the critical level with 30 documented attacks. Figure 2 summarizes the distribution by type and criticality level.

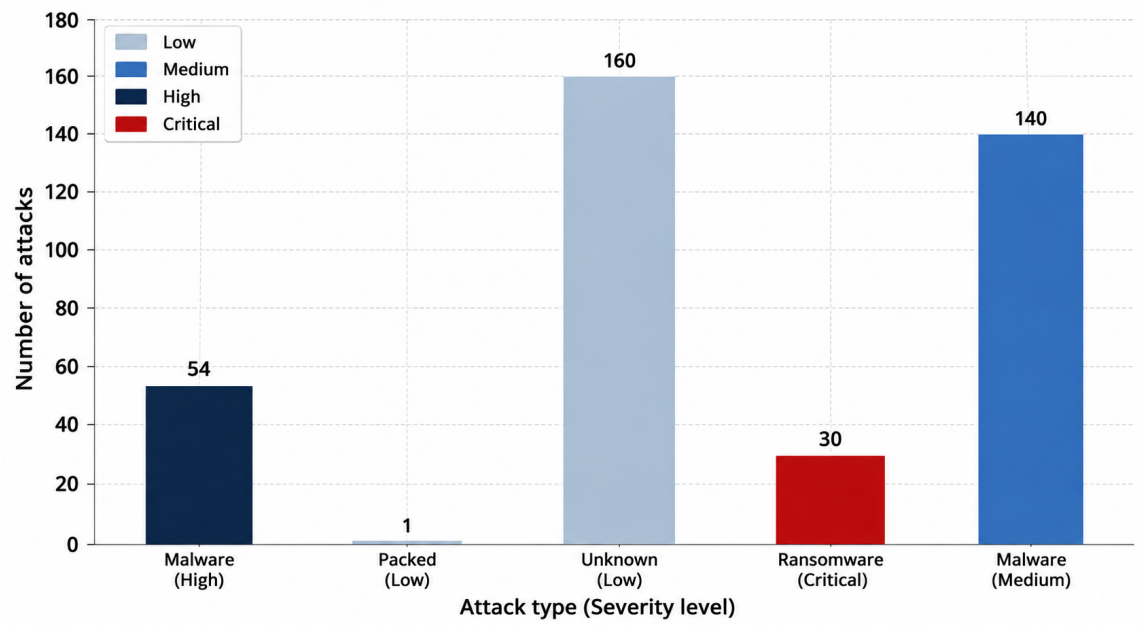

Figura 2. Analysis by attack type and criticality level

The quarterly analysis showed that high- or critical-criticality attacks ranged between 3 and 5 events in each of the first three quarters, while in the fourth quarter that range widened significantly, to between 12 and 32 events, evidencing a heavier load of severe threats in the final stage of the year. Figure 3 illustrates the evolution of each criticality level across the four quarters.

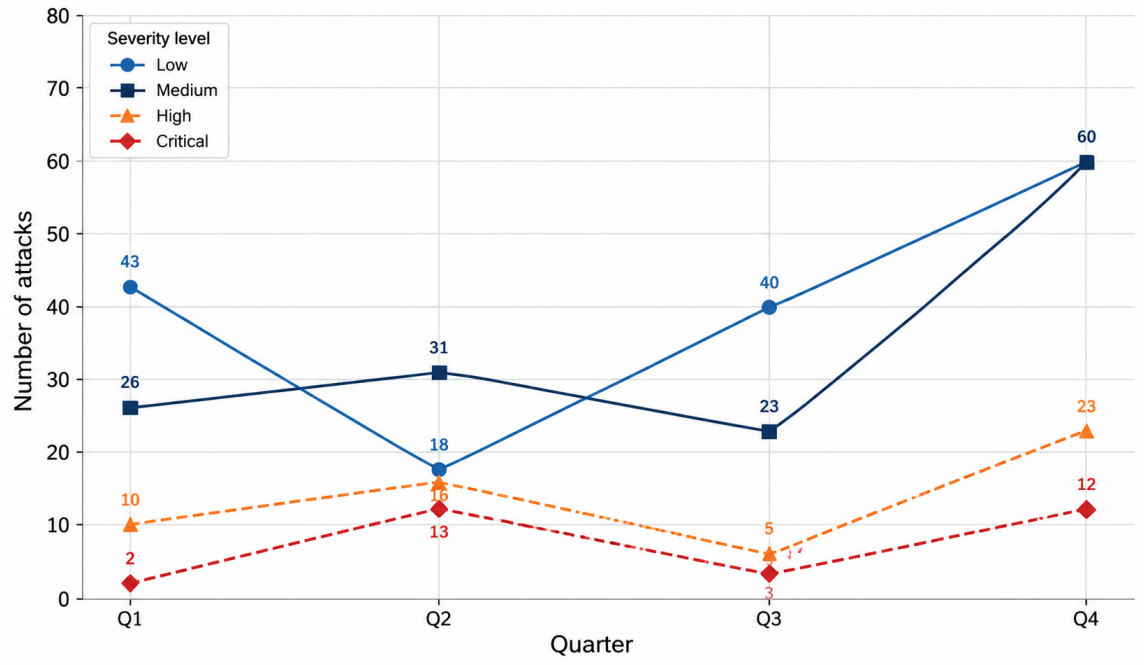

Figura 3. Quarterly analysis of attacks by criticality level

The addressed-vulnerability rate, calculated with 209 vulnerabilities addressed out of 385 total records, reached an annual value of 54.3%, with a monthly variation between 44.4% and 62.5%. This means that practically half of the events recorded were not resolved, leaving the organization with a residual exposure sustained throughout the entire period. Figure 4 shows the monthly evolution of this indicator.

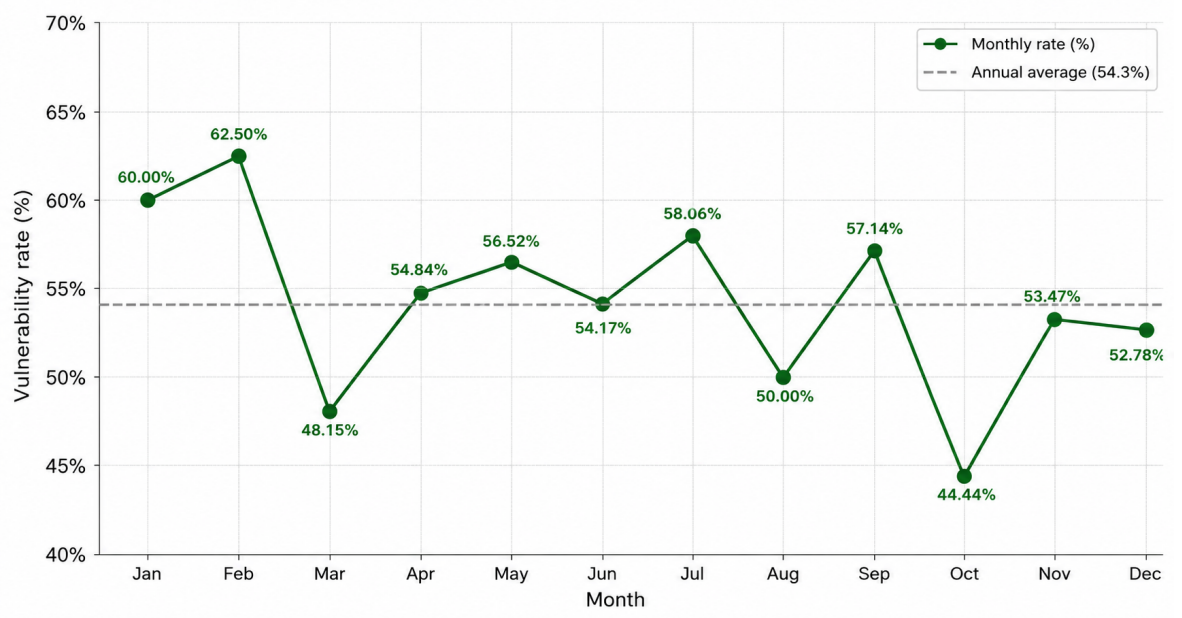

Figura 4. Monthly evolution of the addressed-vulnerability rate (%) in 2025

5.2 Performance of the protection system

Once the exposure had been characterized, the next step was to evaluate the capacity of the existing protection system to deal with these events, using mitigation status and the timeliness of the responses recorded as a reference. Unmitigated attacks ranged between 6 and 47 cases per month, attacks with deficient attention between 1 and 18 cases per month, and attacks with optimal attention between 0 and 41 cases per month. November concentrated the highest number of unmitigated attacks (47), while March and July recorded no optimal attention at all. Figure 5 presents the monthly evolution of these three mitigation states.

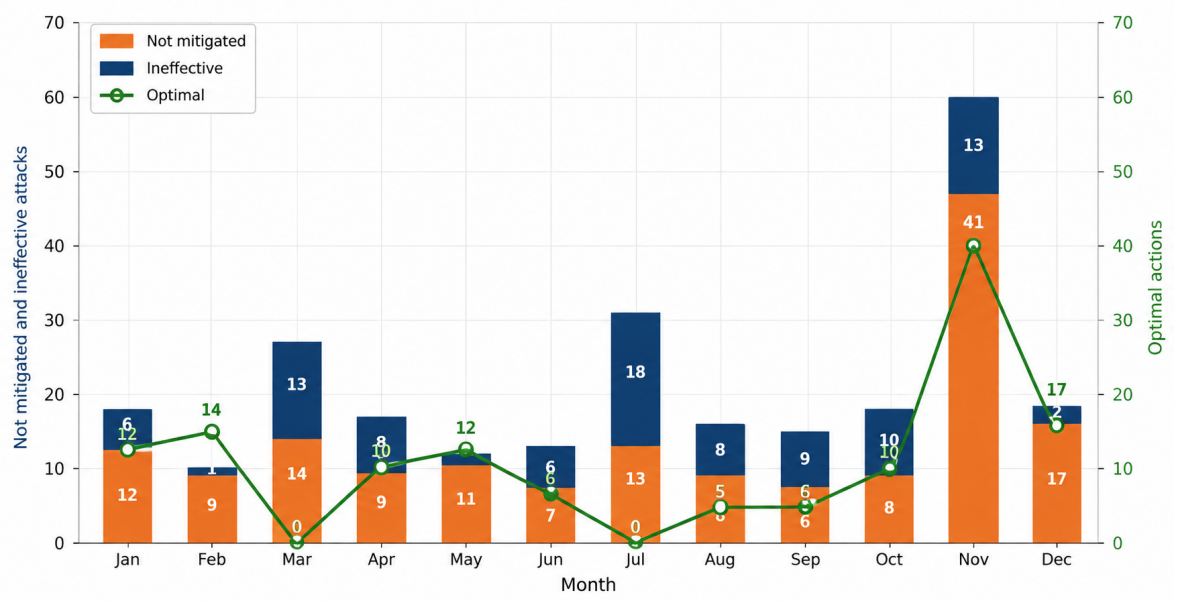

Figura 5. Number of attacks by mitigation level

Likewise, analyzing optimal attentions by criticality level reveals not only how many incidents were adequately resolved but also which types of threats concentrated the team's greatest effective response capacity during the period, as illustrated in Figure 6. Within this framework, critical cases attended within the optimal timeframe ranged between 0 and 5 per month, a range that does not reflect operational stability but rather the scant frequency with which the most severe incidents were adequately managed, confirming that it was precisely the highest-risk events that most frequently escaped timely response capacity. Added to this, high-risk cases showed a slightly wider range, between 0 and 8 optimal attentions per month, like-

wise remaining at low values relative to the total volume recorded in that category. Medium-criticality cases, for their part, showed the greatest dispersion, with values between 0 and 17 optimal attentions and a notable peak in November that coincided with the year's highest attack volume, suggesting a reactive mobilization of resources under volume pressure rather than a sustained structural improvement. This pattern therefore evidences a resource allocation that prioritized volume over severity, a limitation that the architecture proposed in the following section seeks to correct through prioritization mechanisms based on actual criticality.

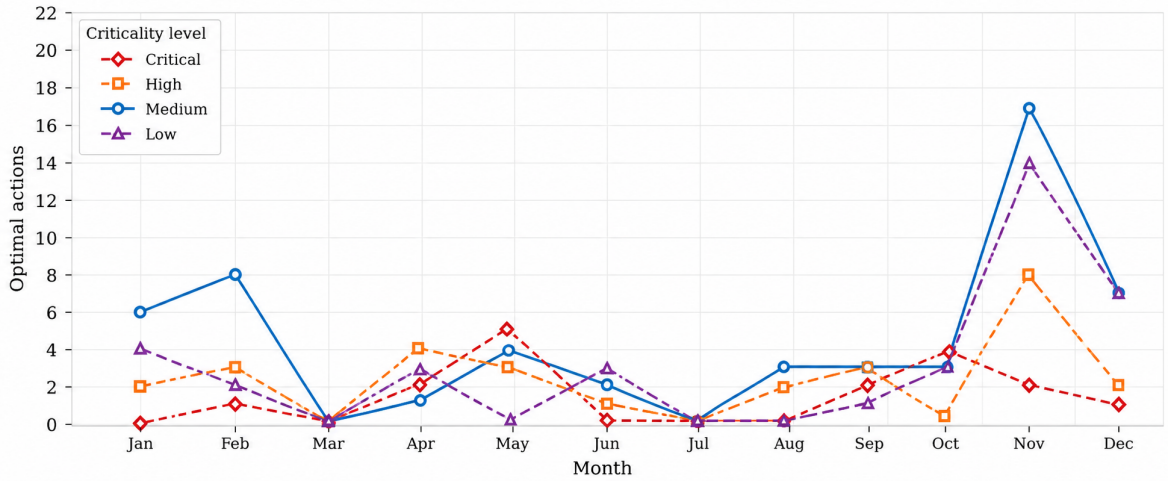

Figura 6. Optimal attentions by criticality level

Now, when examining the composition by attack type among events optimally mitigated, considerable variability is observed, reflecting both the changing nature of the threats received and the system's uneven capacity to respond to each category at different times of the year, as shown in Figure 7. In this context, malware accounted for between 7% and 33% of monthly optimal attentions, ransomware exhibited the widest range of variation, from 6% to 78%, and attacks of unknown origin ranged between 10% and 54%, indicating that no category stably dominated the composition of effective attentions over the period. Added to this, March and July recorded no optimal attention at all, meaning that all events received in those months fell outside the adequate resolution threshold. Given this scenario, the absolute response gaps observed in those periods cannot be attributed to a specific type of threat but rather to general limitations in the security architecture, which reinforces the need for a structural, rather than merely reactive, intervention.

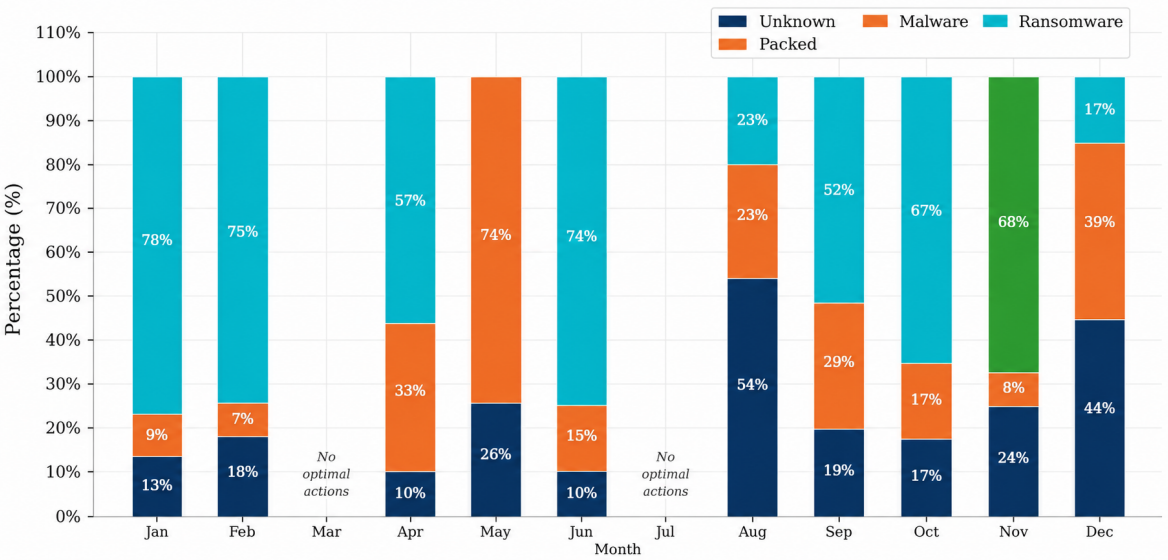

Figura 7. Monthly percentage distribution of attack type in optimal attentions

Continuing along this same line, examining the total number of attacks that did not receive effective attention by threat type reveals a worrying concentration in the categories of greatest potential impact, where 90 corresponded to malware classified as high criticality due to the risk of inoperability it represents for administrative services, 73 to threats of unknown origin classified as medium criticality given their high potential for causing loss or exposure of information, and 13 to ransomware classified as low criticality since it is a threat that could be partially contained before causing irreversible damage. Given this distribution, the fact that the largest volume of unaddressed events corresponds precisely to high-criticality malware indicates that the current architecture fails not only in quantity but in selectivity, leaving unresolved the incidents with the greatest operational consequences. This deficit becomes even more acute when examining the internal geographic distribution of unaddressed attacks, since the commercial area concentrated 70 events of unknown origin and 38 of malware, accumulating the largest volume in the organization due to its constant exposure to multiple digital communication channels, positioning it as the most vulnerable and priority front for the intervention of the proposed architecture, as evidenced in Figure 8.

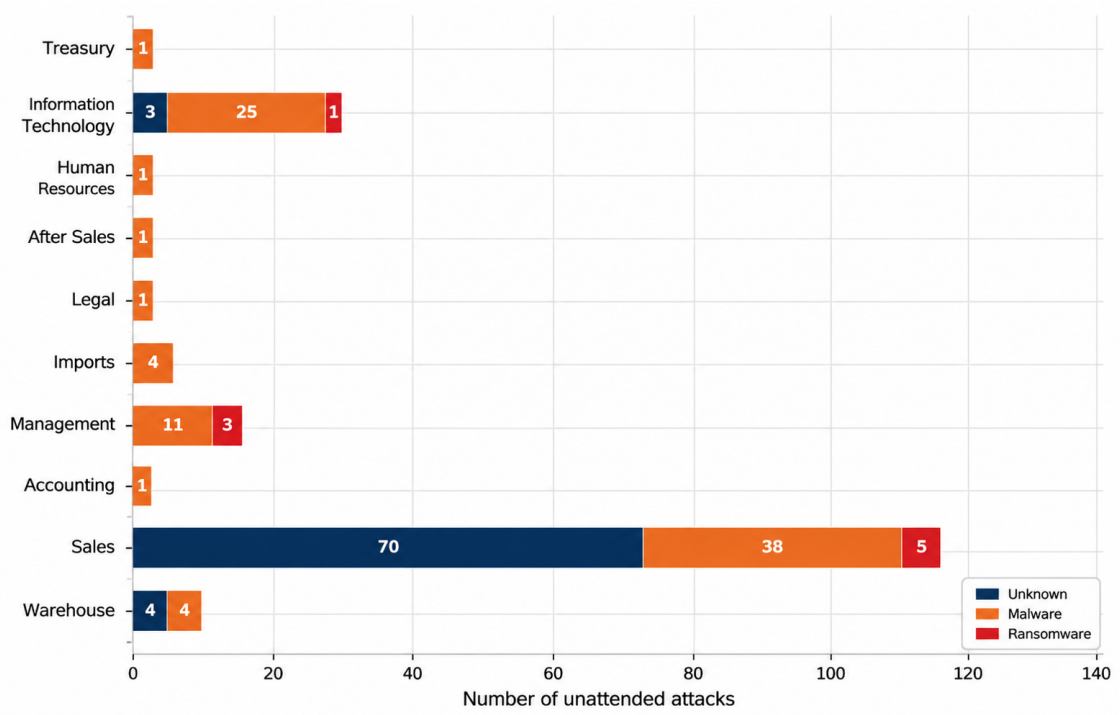

Figura 8. Unaddressed attacks by area and threat type

To complement the foregoing, Figure 9 presents the monthly evolution of this indicator.

5.3 Proposed modern cybersecurity architecture

The findings presented in the preceding subsections reveal an addressed-vulnerability rate of 54.3 % and a response index of 34 %, figures that are not isolated but reflect a security architecture that lacks integration, prioritization, and effective closure capacity. Given this diagnosis, the solution cannot be limited to reinforcing existing controls, but requires a structural response. With this purpose, a modern cybersecurity architecture was designed, aimed at reducing residual exposure, improving the performance of the protection system, and strengthening computer-fraud prevention in digitalized administrative environments.

The architecture is organized into six functional domains that operate in an interconnected manner and are articulated around a central core of advanced detection and artificial intelligence, as illustrated in Figure 10. The first domain, identity and access management, controls who can access each system, with what privileges, and under what conditions, implementing multi-factor authentication, single sign-on, and role-based access control to reduce the risk of unauthorized access, which was a recurring source of incidents in the diagnosis.

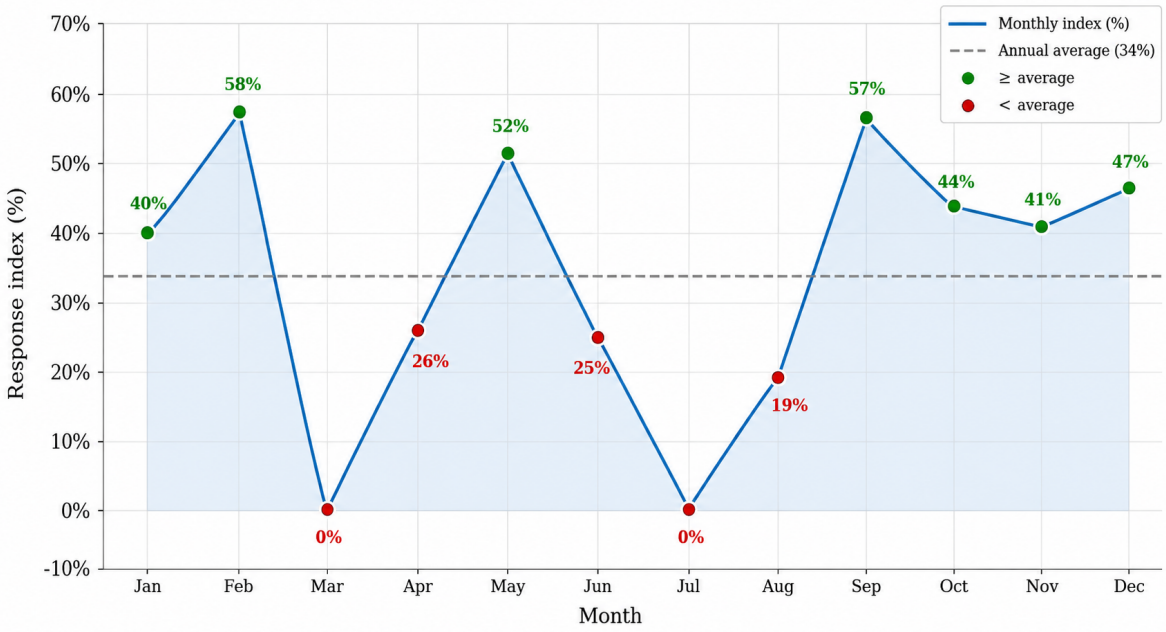

Figura 9. Monthly evolution of the response index (%) in 2025

The second domain, threat protection, detects, contains, and neutralizes malicious events such as malware and ransomware before they can cause damage, relying on endpoint detection and response, network traffic filtering, and real-time antimalware engines. The third domain, cloud security, protects the ERP services and digital platforms that sustain the operation, ensuring correct configurations and visibility over applications deployed in cloud environments through cloud access security brokers and configuration-compliance verification.

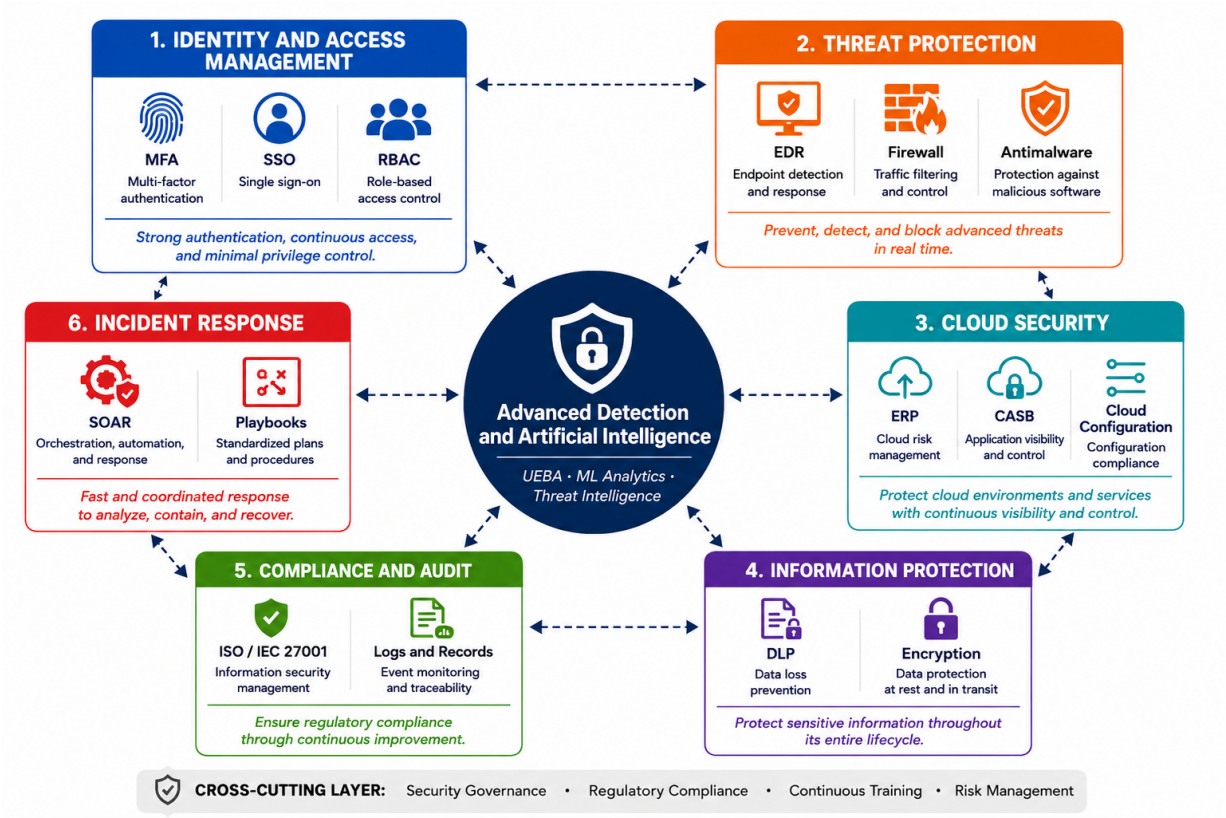

Figura 10. Proposed modern cybersecurity architecture

As for the fourth domain, information protection classifies data according to its sensitivity level and establishes leak-prevention and encryption controls both at rest and in transit, an especially critical risk in organizations that handle financial and accounting information. For its part, the fifth domain, compliance and audit, periodically verifies that controls are functioning, documents evidence through monitoring logs, and generates the corrective actions needed, aligned with standards such as ISO 27001. Added to this is the sixth domain, incident response, which ensures that every event is handled under a defined and traceable protocol through orchestration, automation, and standardized playbooks, closing the cycle that the current architecture left open. Transversally, a layer of security governance, regulatory compliance, ongoing training, and risk management underpins the operation of all domains, recognizing that security is both a technical and an organizational problem.

The value of this architecture lies not merely in the existence of the six domains but in the way they interact through the advanced-detection core, turning what in the current architecture was a collection of isolated controls into a cohesive, self-adjusting response system. In this sense, Figure 11 illustrates the operational flow followed by an incident from its detection to its resolution and the resulting learning, showing that each stage does not operate independently but feeds and conditions the next. When the system identifies an event through sensors, logs, or EDR agents, it is normalized and filtered through parsing, triage, and deduplication before being sent to the artificial-intelligence core, where it is correlated with other events to determine whether it forms part of a broader attack pattern rather than an isolated incident. Based on this correlation, the event is prioritized according to its actual criticality level and an automated response is triggered through SOAR and playbooks if the case allows, or escalated to the corresponding analyst when the complexity of the incident requires human judgment, thereby ensuring that response resources are allocated based on severity rather than order of arrival.

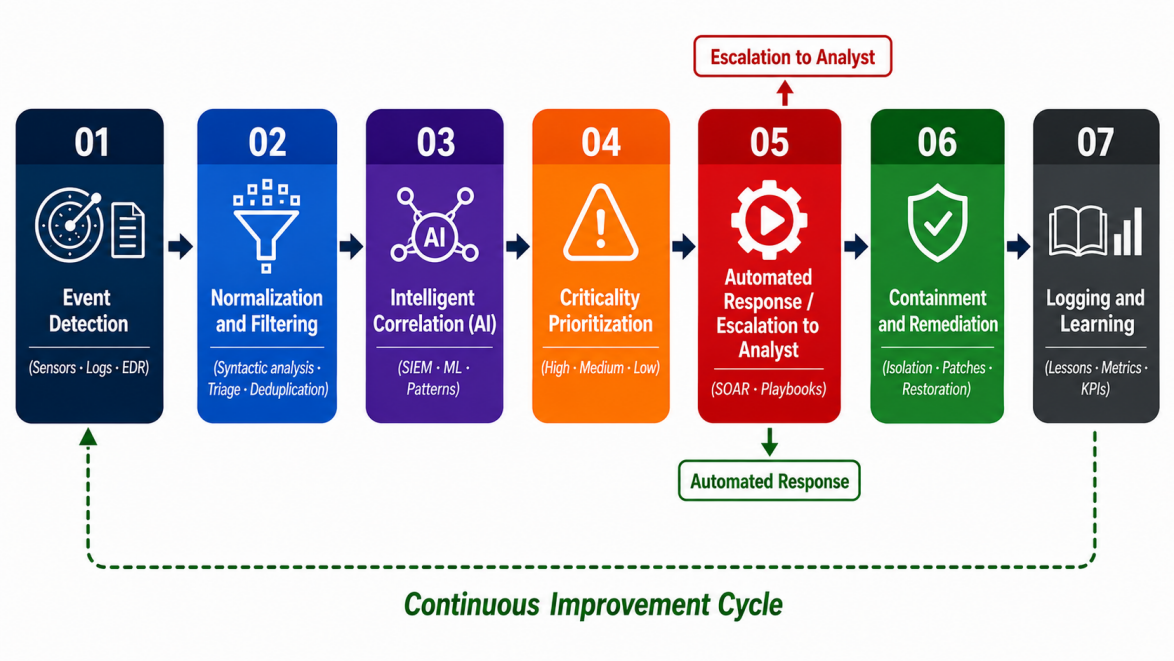

Figura 11. Operational flow of incident detection, prioritization, and response

Once the incident has been contained and remediated through isolation, patching, or rollback, the system does not conclude the cycle there; instead, it records the lessons learned and updates performance metrics to improve the detection of similar events in the future, closing a continuous-improvement loop that the current architecture does not contemplate. This capacity for cumulative learning is precisely what distinguishes a reactive architecture from an adaptive one, since every incident managed becomes input for fine-tuning detection thresholds, adjusting correlation rules, and progressively reducing response times. This design directly addresses the central weakness identified in the diagnosis, where the 54.3 % addressed-vulnerability rate and the 34 % response index were not the result of a lack of resources but of the absence of a systematic process connecting detection with effective response, criticality-based prioritization, and structured recording of lessons learned—three dimensions that this flow integrates explicitly and measur-

ably.

Now, to put this architecture into practice, its implementation is organized into three progressive phases over twelve months, as shown in Figure 12. The first phase, called foundations and protection and spanning months 1 through 4, installs identity, access, and basic protection controls, establishes security policies, and carries out the asset inventory and risk assessment that serve as the basis for the following phases, with the goal of raising the addressed-vulnerability rate above 65 % and the response index above 40 %. Building on that base, the second phase, detection and automation, developed between months 5 and 8, connects the domains to one another through the artificial-intelligence core, activates event correlation, deploys cloud protection and automated response, pursuing an addressed-vulnerability rate above 75 % and a response index above 55 %. Based on the progress consolidated in that stage, the third phase, optimization and continuous improvement, spanning months 9 through 12, reinforces the process through periodic audits, staff training, dynamic rule adjustment, and architecture review, with the aim that the addressed-vulnerability rate exceed 85 % and the response index reach 70 %. This staggered rollout ensures that each phase builds on the achievements of the previous one and that the indicators serve as an objective signal of progress throughout the entire transformation process.

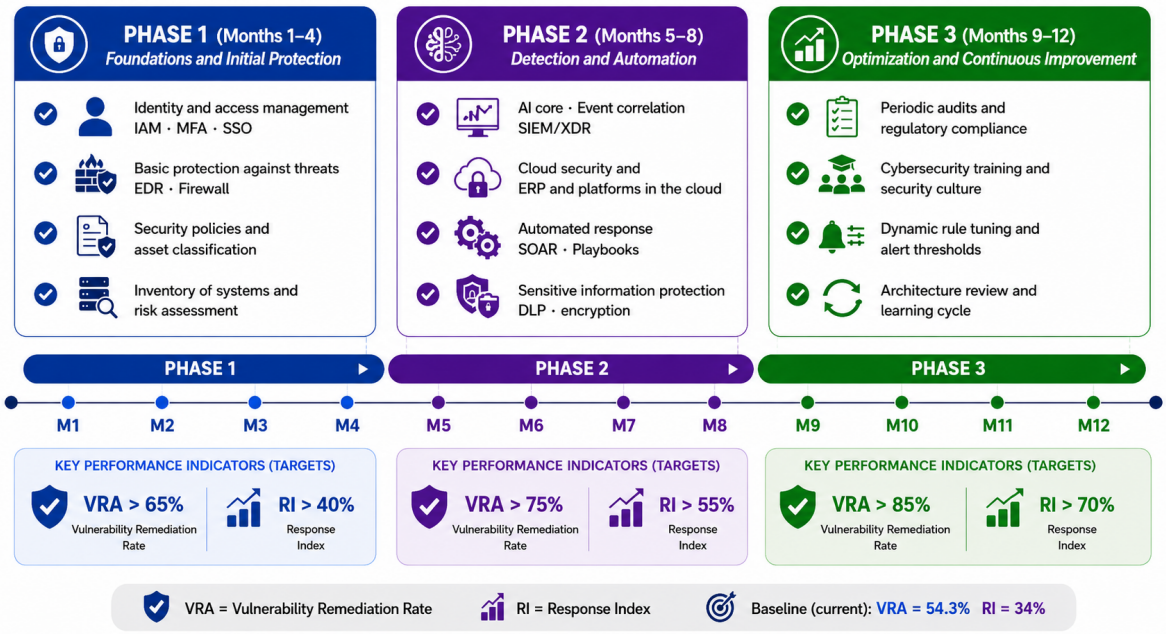

Figura 12. Roadmap for implementing the proposed architecture

Together with the conceptual foundations already described, the soundness of a cybersecurity architecture depends not only on the coherence among its components but also on the empirical backing that justifies each design decision. In this sense, the body of studies underlying this proposal covers a broad spectrum of dimensions, from intelligent anomaly detection and statistical pattern analysis to social engineering, user awareness, and the link between security risk and business processes. Each reference was selected for its direct correspondence with one of the gaps identified in the diagnosis, ensuring that the proposal does not constitute a generic solution but a response calibrated to the organization's specific limitations.

In this way, the architecture not only relies on consolidated evidence but also establishes an explicit bridge between the analysis findings and the design decisions adopted. Some studies support the selection of indicators and the quantitative reading of anomalous patterns, others justify the integration of awareness mechanisms and a security culture, and others directly connect operational risk with digitalized administrative processes. Taken together, these references confirm that each component of the proposed architecture is grounded in prior research rather than arbitrary design criteria, a relationship synthesized in Table 2.

Tabla 2. Complementary studies underlying the architectural proposal

Author	Contribution axis	Use in the manuscript
Zioviris et al. [15]	Intelligent detection	Supports the use of sequential models for fraud.
Zogaj et al. [16]	Statistical analysis	Reinforces the quantitative reading of anomalous patterns.
Vanini et al. [17]	Risk management	Connects payment anomalies with anti-fraud management.
Ukwandu et al. [18]	Critical infrastructure	Shows operational impacts in digitalized sectors.
Torres Gamarra [19]	Structural gaps	Contextualizes regional threats, legislation, and gaps.
Szczepaniuk et al. [20]	Intelligent systems	Provides criteria for security in connected environments.
Santillán et al. [21]	Comprehensive review	Underpins the multi-domain view of cybersecurity.
Rodrigues et al. [22]	E-commerce	Relates digital fraud to operational prevention.
Pittoli et al. [23]	Architectures	Supports the need for context-adapted designs.
Pauta Ayabaca et al. [24]	Emerging technologies	Links technology adoption with preventive capabilities.
Orosco-Fabian et al. [25]	Bibliometrics	Situates regional cybersecurity research.
Schroeder et al. [26]	Measurement	Supports the selection of security indicators.
Awobelem George [27]	Digital governance	Reinforces security as an institutional capability.
Khan et al. [28]	Anomaly detection	Supports analytics for unknown events.
Hernández Aros et al. [29]	Financial fraud	Connects machine learning with fraud detection.
Fueres Quilumbango et al. [30]	Social engineering	Integrates users and training into the architecture.
Chaudhary et al. [31]	Awareness	Justifies measuring security-culture programs.
Rosado et al. [32]	Business processes	Connects security risk with administrative processes.
Bennouk et al. [33]	Vulnerabilities	Supports automated detection and management of flaws.
Alshomrani et al. [34]	Explainable AI	Provides transparency for early detection.

Based on this theoretical framework, each domain of the architecture was associated with a central reference that underpins both its design and its operational function within the integrated protection system. This correspondence between academic evidence and architectural component ensures that design decisions respond to criteria validated in the literature and not solely to technological considerations, as detailed in Table 3.

When examining the domains, it becomes clear that their design does not follow an additive logic but rather an architecture of interdependencies in which each component reinforces and conditions the next. For example, the identity and access domain constitutes the entry point of the entire protection chain, since without rigorous control of privileges and authentication, any subsequent mechanism operates on a compromised basis. Building on that control, the threat-protection domain extends coverage toward endpoints and the network, on the assumption that the malware and unknown-activity incidents identified in the diagnosis cannot be contained solely at the perimeter. At that same operational level, cloud security complements this coverage by protecting the ERP environments and digital services that underpin the organization's administrative operation. On that technical basis, the information-protection domain adds a custody layer specifically oriented toward sensitive data, whose exposure or loss represents the greatest fraud risk in financial and accounting contexts. Added to this is the compliance and audit domain, which closes the control cycle by periodically verifying that the preceding mechanisms function as intended and generating the evidence needed for accountability. In this way, the advanced detection and artificial intelligence domain operates transversally across all the others, correlating events, prioritizing alerts, and learning from historical patterns so that the system improves its response capacity with each incident

managed.

Tabla 3. Domains and components of the proposed architecture

Reference	Domain	Main function
Raj [14]	Identity and access	Control privileges, multi-factor authentication, conditional access, and high-risk accounts.
Umer et al. [11]	Threat protection	Detect, contain, and respond to malware, unknown events, and malicious activity on endpoints and the network.
Cevallos Ramos et al. [8]	Cloud security	Protect ERP services, configurations, workloads, and policies deployed in cloud environments.
Yaseen [7]	Information protection	Classify, safeguard, and monitor sensitive data to reduce leakage, manipulation, or loss.
Agyepong et al. [6]	Compliance and audit	Periodically verify controls, evidence, accountable parties, gaps, and corrective actions.
Moura et al. [13]	Advanced detection and AI	Correlate events, prioritize alerts, automate response, and learn from historical patterns.

6. Discussion

Having presented the results and the architectural proposal, it is now necessary to interpret their scope, contrast them with the available literature, and delimit the implications for organizational management. Digitalized administrative services concentrate financial, logistical, and documentary processes that, owing to their interdependence and volume, represent a persistent target for organizational fraud. In this context, an organization's capacity to detect threats and turn them into timely responses does not depend solely on available technology, but on the coherence among its detection, prioritization, and operational closure processes. The results of this study reveal precisely a disconnection between these two capabilities. The addressed-vulnerability rate of 54.3% indicates that the system identifies more than half of the risk events, but the 34% response index shows that only a fraction of those alerts translates into traceable, timely, and verifiable actions. This gap does not constitute a specific technical deficit, but rather reflects an insufficiency in the organizational security architecture, where detection and operational management function as decoupled processes, which widens exposure to identity theft, data manipulation, and service interruption.

This interpretation aligns with the warning from Agyepong et al. [6], who point out that maturity in security operations is not measured by the volume of alerts generated, but by the quality of the decisions they produce. In a complementary manner, Forsberg and Frantti [10] argue that effective vulnerability management requires detection, prioritization, and remediation to form a measurable, closed cycle. Under these criteria, the indicators used capture the initial stages of the cycle but not its closure, since no data are available on mean containment time, residual severity, recurrence, or cost per incident. This qualification does not invalidate the findings, but rather clarifies their scope. The results are diagnostic of operational activity rather than of systemic maturity, and should therefore be read as a starting point for a broader evaluation.

The digital-fraud dimension introduces an additional layer of complexity that deserves particular attention. Afriyie et al. [9] show that detection models based on machine learning critically depend on the consistency of the training data and the stability of operational patterns. Abu-Khadrah [5] underscores that digital fraud methods evolve quickly enough to render static rule-based systems obsolete. Projected onto the context studied, both arguments indicate that an effective response architecture cannot be limited to automating pre-existing controls or increasing the volume of alerts. Its value lies in integrating data,

context, prioritization, and explainability to reduce false negatives, mitigate alert fatigue, and eliminate opaque decisions. The absence of these capabilities in the current system is consistent with the observed gap between detection and operational closure.

Building on this diagnostic basis, the proposed architecture responds to gaps that span multiple organizational-control domains. Raj [14] promotes modern frameworks with extended monitoring and adaptive response. Moura et al. [13] show that artificial intelligence strengthens proactive detection when integrated with governed processes. Cevallos Ramos et al. [8] link these approaches to organizational cyber resilience. Within this framework, identity, cloud, data protection, compliance, audit, XDR, and artificial intelligence must operate as interdependent layers. When one fails, the others provide containment, traceability, and learning. This redundancy pursues a concrete objective, namely preventing an isolated event from escalating into operational fraud or affecting administrative continuity in environments with distributed users and cloud-shared data. The proposal is architectural rather than experimental, so its practical validity remains contingent on a controlled implementation with comparable indicators before and after the intervention.

Translating this into management decisions requires recognizing that security is a core operational capability rather than a technical support function. The organizational costs associated with the gap detected include service interruptions, information restoration, overload of the infrastructure team, and erosion of institutional trust. For this reason, implementing the proposed architecture demands defined accountable parties, a timeline, a budget, ongoing training, and periodic audits. It also requires dashboards that connect technical indicators with business continuity, residual risk, and verifiable response times. Impact should be measured against a pre-implementation baseline using the same indicators employed in this diagnosis, a condition that will allow results to be rigorously compared and prevent attributing improvements that have not been empirically verified.

Finally, the study's limitations should be read as delimiters of its contribution rather than as objections to its relevance. The case analyzed corresponds to a single organization, so the transferability of the findings depends on the degree of similarity in context, size, digital maturity, and resource availability. The design is descriptive rather than experimental, so it identifies gaps and justifies an architecture but does not establish causality or prove effectiveness. The available records cover a specific operational period, and their quality is conditioned by the corporate classification platform. These conditions precisely define the next step, which consists of validating the architecture through a pre- and post-implementation longitudinal design with comparable indicators, including residual risk, operational continuity, and actual anti-fraud response capacity.

7. Conclusions

This study provides evidence that security gaps in digitalized administrative environments do not stem from isolated technological shortcomings but from the absence of integration between detection and effective response. Across 385 technical records documented during 2025, the addressed-vulnerability rate did not exceed 54.3% and the effective response index stood at 34%, meaning that almost half of the events detected passed through the system without generating a traceable or verifiable action. These values do not reflect insufficient technical capacity but rather an architecture that lacks the integration needed to close the full cycle between threat detection and its operational resolution. In response to this condition, the architecture designed introduces a deliberate logic of interdependence in which identity and access management, advanced threat protection, cloud security, safeguarding of sensitive information, regulatory compliance, continuous auditing, extended analytics, proactive detection through artificial intelligence, staff training, and institutionalized continuous improvement together form a system in which the failure of one domain does not produce a collapse but a containment, because the remaining domains sustain traceability and response. This property, absent from the current architecture, is what determines the difference between a system that detects and one that protects.

It should be noted that the architecture proposed in this work constitutes a design grounded in diagnostic evidence and theoretical support, but its validation is not part of the scope of this study. That validation represents the most immediate and necessary line of future work, and should be carried out through follow-up

studies measuring the evolution of the same indicators after implementation, incorporating complementary metrics such as containment time, residual severity, and recurrence rate. Subsequently extending this analysis to organizations of different sizes and sectors would make it possible to examine the transferability of the findings and contribute to the development of comparable reference frameworks for cybersecurity in digitalized administrative environments.

Author Contributions. The authors confirm their contribution to the article as follows: conceptualization, E.C.C.; methodology, J.R.E. and E.C.C.; investigation, E.C.C. and J.R.E.; formal analysis, J.R.E.; writing – original draft preparation, E.C.C.; writing – review and editing, E.C.C.; visualization, J.R.E. and E.C.C. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgments. The authors thank the Universidad Tecnológica del Perú and the specialists who facilitated the operational understanding of the case, while maintaining the confidentiality of sensitive data.

Funding. This study was self-funded by the authors.

Informed Consent Statement. Not applicable. The study used anonymized organizational technical records and did not involve the direct recruitment of human participants.

Data Availability Statement. The operational data belong to the company studied. The aggregated matrices, coding criteria, and instruments can reasonably be requested from the authors, subject to confidentiality restrictions.

Conflicts of Interest. The authors declare no conflicts of interest.

Statement on the Use of Artificial Intelligence. The authors used generative artificial intelligence tools exclusively to improve the clarity and style of the text. The analysis, interpretation, and conclusions are the exclusive intellectual responsibility of the authors, who validated and approved the final version.

References

[1] F. Houghton, M. Winterburn, and K. Oakley, "Rhysida Ransomware Attack on the British Library: Prioritisation, Expertise, and Funding Issues," *Information Technology and Libraries*, vol. 44, no. 1, Mar. 2025. doi: 10.5860/ital.v44i1.17112.

[2] N. Wadesango and E. Maveneka, "Cyberthreats and their impact on financial integrity: Evaluating the effectiveness of local authorities cybersecurity policies in preventing and detecting fraud," *Corporate Law and Governance Review*, vol. 7, no. 2, p. 32, Apr. 2025. doi: 10.22495/clgrv7i2p3.

[3] V. H. Che Leon Antinori, "Influence of Economic Cybercrime in Peru," *SCIENDO*, vol. 27, no. 4, pp. 415–418, Oct. 2024. doi: 10.17268/sciendo.2024.071.

[4] D. F. Wahid and E. Hassini, "An augmented AI-based hybrid fraud detection framework for invoicing platforms," *Applied Intelligence*, vol. 54, no. 2, pp. 1297–1310, Jan. 2024. doi: 10.1007/s10489-023-05223-x.

[5] A. Abu-Khadrah, S. Al-Washmi, A. Mohd Ali, and M. Jarrah, "Cyber-fraud detection methodology by using machine learning algorithms," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 15, no. 4, p. 3949, Aug. 2025. doi: 10.11591/ijece.v15i4.pp3949-3956.

[6] E. Agyepong, Y. Cherdantseva, P. Reinecke, and P. Burnap, "A systematic method for measuring the performance of a cyber security operations centre analyst," *Computers & Security*, vol. 124, p. 102959, Jan. 2023. doi: 10.1016/j.cose.2022.102959.

[7] H. Yaseen and A. Al-Amarneh, "Adoption of Artificial Intelligence-Driven Fraud Detection in Banking: The Role of Trust, Transparency, and Fairness Perception in Financial Institutions in the United Arab Emirates and Qatar," *Journal of Risk and Financial Management*, vol. 18, no. 4, p. 217, Apr. 2025. doi: 10.3390/jrfm18040217.

[8] C. D. R. Cevallos Ramos, F. F. Navarrete Chavez, F. R. Marquez Sanay, and M. P. Andrade Romero, "Strengthening cyber resilience in universities using artificial intelligence for proactive threat detectio," *Data and Metadata*, vol. 4, p. 1109, Jul. 2025. doi: 10.56294/dm20251109.

[9] J. K. Afriyie, K. Tawiah, W. A. Pels, S. Addai-Henne, H. A. Dwamena, E. O. Owiredu, S. A. Ayeh, and J. Eshun, "A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions," *Decision Analytics Journal*, vol. 6, p. 100163, Mar. 2023. doi: 10.1016/j.dajour.2023.100163.

[10] J. Forsberg and T. Frantti, "Technical performance metrics of a security operations center," *Computers & Security*, vol. 135, p. 103529, Dec. 2023. doi: 10.1016/j.cose.2023.103529.

[11] M. A. Umer, E. G. Belay, and L. B. Gouveia, "Fortifying Industry 4.0: Internet of Things Security in Cloud Manufacturing through Artificial Intelligence and Provenance Blockchain—A Thematic Literature Review," *Sci*, vol. 6, no. 3, p. 51, Sep. 2024. doi: 10.3390/sci6030051.

[12] M. Nandy and A. Dubey, "Applications of artificial intelligence to strengthening cyber security for threat detection and response," in *AIP Conference Proceedings*, Bangkok, Thailand, 2026, p. 020029. doi: 10.1063/5.0298670.

[13] L. Moura, A. Barcaui, and R. Payer, "AI and Financial Fraud Prevention: Mapping the Trends and Challenges Through a Bibliometric Lens," *Journal of Risk and Financial Management*, vol. 18, no. 6, p. 323, Jun. 2025. doi: 10.3390/jrfm18060323.

[14] A. Raj, V. Narayan, V. Muskan, A. Sani, P. Sharma, and S. S. Sarma, "Modern ransomware: Evolution, methodology, attack model, prevention and mitigation using multitiered approach," *SECURITY AND PRIVACY*, vol. 7, no. 6, p. e436, Nov. 2024. doi: 10.1002/spy2.436.

[15] G. Zioviris, K. Kolomvatsos, and G. Stamoulis, "An intelligent sequential fraud detection model based on deep learning," *The Journal of Supercomputing*, vol. 80, no. 10, pp. 14 824–14 847, Jul. 2024. doi: 10.1007/s11227-024-06030-y.

[16] G. Zogaj, F. Ismaili, E. Idrizi, and A. Luma, "Statistical Analysis of Unique Web Application Vulnerabilities: A Quantitative Assessment of Scanning Tool Efficiency," *SEEU Review*, vol. 20, no. 1, pp. 136–152, Jun. 2025. doi: 10.2478/seeur-2025-0021.

[17] P. Vanini, S. Rossi, E. Zvizdic, and T. Domenig, "Online payment fraud: from anomaly detection to risk management," *Financial Innovation*, vol. 9, no. 1, p. 66, Mar. 2023. doi: 10.1186/s40854-023-00470-w.

[18] E. Ukwandu, M. A. Ben-Farah, H. Hindy, M. Bures, R. Atkinson, C. Tachtatzis, I. Andonovic, and X. Bellekens, "Cyber-Security Challenges in Aviation Industry: A Review of Current and Future Trends," *Information*, vol. 13, no. 3, p. 146, Mar. 2022. doi: 10.3390/info13030146.

[19] N. Torres Gamarra and M. Zuniga Carnero, "Panorama actual de la ciberseguridad: amenazas, legislacion y brechas estructurales desde una revision sistematica," *Zenodo*, Jun. 2025. doi: 10.5281/ZENODO.15605545.

[20] E. K. Szczepaniuk and H. Szczepaniuk, "Cybersecurity of Smart Grids: Requirements, Threats, and Countermeasures," *Energies*, vol. 18, no. 18, p. 5017, Sep. 2025. doi: 10.3390/en18185017.

[21] H. Santillan, J. A. Arevalo Satan, and P. Wong, "A Comprehensive Analysis of Cybersecurity Infrastructure in Academic Environments," *Ingenieria*, vol. 35, no. 1, pp. 11–23, Oct. 2024. doi: 10.15517/ri.v35i1.60075.

[22] V. F. Rodrigues, L. M. Policarpo, D. E. Da Silveira, R. Da Rosa Righi, C. A. Da Costa, J. L. V. Barbosa, R. S. Antunes, R. Scorsatto, and T. Arcot, "Fraud detection and prevention in e-commerce: A systematic literature review," *Electronic Commerce Research and Applications*, vol. 56, p. 101207, Nov. 2022. doi: 10.1016/j.elerap.2022.101207.

[23] P. Pittoli, P. David, and T. Noel, "Security architectures in constrained environments: A survey," *Ad Hoc Networks*, vol. 79, pp. 112–132, Oct. 2018. doi: 10.1016/j.adhoc.2018.06.001.

[24] S. L. Pauta Ayabaca and J. E. Alvarez Gavilanes, "Uso y necesidad de Tecnologias Emergentes en las empresas cuencanas para el fortalecimiento academico," *Pacha. Revista de Estudios Contemporaneos del Sur Global*, vol. 3, no. 9, p. e210123, Sep. 2022. doi: 10.46652/pacha.v3i9.123.

[25] J. R. Orosco-Fabian, "Ciberseguridad en educacion superior: una revision bibliometrica," *Revista Digital de Investigacion en Docencia Universitaria*, vol. 18, no. 2, p. e1933, Jul. 2024. doi: 10.19083/ridu.2024.1933.

[26] K. Schroeder, H. Trinh, and V. Y. Pillitteri, "Measurement guide for information security volume 1 : identifying and selecting measures," National Institute of Standards and Technology (U.S.), Gaithersburg, MD, Tech. Rep. NIST SP 800-55v1, Dec. 2024. doi: 10.6028/NIST.SP.800-55v1.

[27] Awobelem A. George, Akeem Olakunle Ogundipe, and Aminat Bolaji Bello, "Cybersecurity In healthcare systems: safeguarding electronic health records (EHRs) and medical devices against emerging cyber threats," *World Journal of Advanced Research and Reviews*, vol. 25, no. 2, pp. 2249–2262, Feb. 2025. doi: 10.30574/wjarr.2025.25.2.0592.

[28] W. Khan and N. Ebrahim, "ANOGAT-Sparse-TL: A hybrid framework combining sparsification and graph attention for anomaly detection in attributed networks using the optimized loss function incorporating the Twersky loss for improved robustness," *Knowledge-Based Systems*, vol. 311, p. 113144, Feb. 2025. doi: 10.1016/j.knosys.2025.113144.

[29] L. Hernandez Aros, L. X. Bustamante Molano, F. Gutierrez-Portela, J. J. Moreno Hernandez, and M. S. Rodriguez Barrero, "Financial fraud detection through the application of machine learning techniques: a literature review," *Humanities and Social Sciences Communications*, vol. 11, no. 1, p. 1130, Sep. 2024. doi: 10.1057/s41599-024-03606-0.

[30] E. A. Fueres Quilumbango, R. O. Andrade Paredes, and M. V. Yamba Yugsi, "Proteccion contra ataques de ingenieria social: analisis cualitativo de estrategias, tecnologias y patrones especificos," *Religacion*, vol. 10, no. 44, p. e2501310, Oct. 2024. doi: 10.46652/rgn.v10i44.1310.

[31] S. Chaudhary, V. Gkioulos, and S. Katsikas, "Developing metrics to assess the effectiveness of cybersecurity awareness program," *Journal of Cybersecurity*, vol. 8, no. 1, p. tyac006, Jan. 2022. doi: 10.1093/cybsec/tyac006.

[32] D. G. Rosado, L. E. Sanchez, A. J. Varela-Vaca, A. Santos-Olmo, M. T. Gomez-Lopez, R. M. Gasca, and E. Fernandez-Medina, "Enabling security risk assessment and management for business process models," *Journal of Information Security and Applications*, vol. 84, p. 103829, Aug. 2024. doi: 10.1016/j.jisa.2024.103829.

[33] K. Bennouk, N. Ait Aali, Y. El Bouzekri El Idrissi, B. Sebai, A. Z. Faroukhi, and D. Mahouachi, "A Comprehensive Review and Assessment of Cybersecurity Vulnerability Detection Methodologies," *Journal of Cybersecurity and Privacy*, vol. 4, no. 4, pp. 853–908, Oct. 2024. doi: 10.3390/jcp4040040.

[34] M. Alshomrani, A. Albeshri, A. A. Alsulami, and B. Alturki, "An Explainable Hybrid CNNTransformer Architecture for Visual Malware Classification," *Sensors*, vol. 25, no. 15, p. 4581, Jul. 2025. doi: 10.3390/s25154581.