OBSERVATORY 2026

Inside the dark matter of cyberspace

CYBEROO

37 Via Brigata Reggio, 42124, Reggio Emilia
Email info@cyberoo.com
www.cyberoo.com

Scientific Director: Matteo Ghiotto
Data: I-SOC & Cyberoo Incident Response Team
Scientific review: Veronica Leonardi, Luca Bonora
Content coordination: Fabio Rauzino
Editorial contributions: Andrea Coli, Andrea De Giorgio,
Barbara Sabellico, Elena Prodi, Enrico Lorenzi, Luca Benatti,
Luca Bonora, Matteo Ghiotto, Vasily Kononov
Graphic design: Gabriele Leonardi, Giorgia Gnecchi

Table of Contents

Introduction

In 2026, the greatest challenge in cybersecurity is not what we see, but what remains invisible. There is a structural gap between the perception of security and the actual extent of risk weighing on information systems: a cognitive gap, not merely a technical one.

The universe is largely influenced by dark matter, a component that neither emits nor reflects light and cannot be observed directly, but whose presence is inferred from the gravitational effects it exerts on galaxies and cosmic structures.

Similarly, in cyberspace, most of the risk is not evident, is not measured, and yet influences every decision. This cyber “dark matter” includes gaps in control mechanisms, obsolete systems, low risk awareness, unmapped exposure surfaces, inadequately managed tokens and APIs (Application Programming Interfaces), and non-human entities operating outside the field of view of security systems. This is no minor detail: it affects the entire operational balance of the organization.

The paradox is simple: as long as we don’t see a threat, we pretend it doesn’t exist. The same happens to companies: if there are no alarms, blockages, or obvious incidents, everything seems under control.

But it is precisely that silence that increases the risk. Invisible problems end up at the bottom of the priority list; the board dismisses them as minor details; budgets are cut; and small cracks turn into serious breaches. Cybersecurity, which thrives on prevention rather than spectacle, takes a back seat.

Until the moment something truly breaks. Incidents, data breaches, and operational disruptions have become a daily reality for organizations across Europe. According to the European Union Agency for Cybersecurity (ENISA) Threat Landscape, thousands of cybersecurity incidents are reported every year across the EU. The growing scale, sophistication, and frequency of these incidents highlight how cybersecurity events are no longer exceptional occurrences but an ongoing operational risk for European organizations.

Although the data speaks for itself, we delude ourselves into thinking we have the situation under control by relying on partial metrics and models that prioritize what is known (the “known knowns” of the Rumsfeld matrix), systematically overlooking what lies outside the scope of observation or appears anomalous.
But it is precisely from these anomalies that we intuit what we do not see and that could become a powerful risk multiplier. Making hidden risks visible is not a technical detail: it is essential for protecting the company.

	Known	**Unknown**
Known	Risks and information already identified and measurable.	Risks recognized but not yet quantified.
Unknown	Information available but not used in decision-making processes.	Risks outside the mental model or scope of observation.

The Rumsfeld Matrix

The 2026 edition of the Cyberoo Observatory has a clear objective: to shed light on the dark matter of cybersecurity. Not to eliminate it (it is inherent to complex systems), but to measure its effects and reduce its impact through visibility, prioritization, and repeatable decisions. Only by recognizing that the greatest risk lies in what we cannot see will we be able to ensure that evolution does not transform the invisible into a tangible factor of instability.

In the following chapters, we translate this vision into concrete analyses, studies of impacts on processes, and actionable strategies. What do you say, shall we fire up the engines? Let's go.

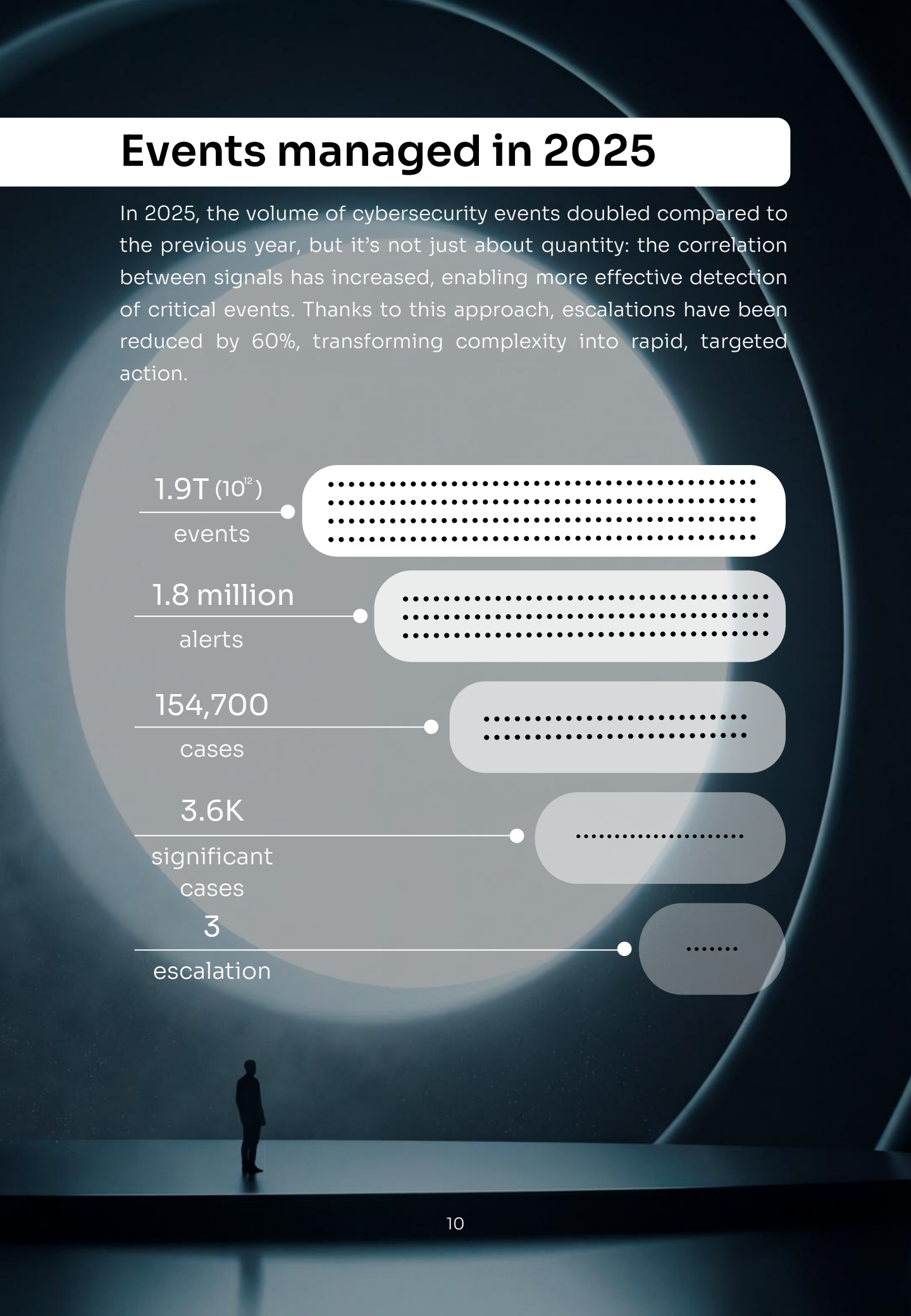
Research Methodology

Our analyses and findings are based on a quantitative and qualitative approach that integrates internal sources, such as data collected by our Cyber Security Suite, with external sources to ensure maximum data accuracy. Information is gathered through the continuous monitoring of trillions of security events managed by our proprietary systems across a sample of 700 European midsize customers, using advanced correlation to identify critical patterns and reduce statistical noise. The data is automatically organized and evaluated using models that indicate which vulnerabilities are most urgent.

It's not just about numbers: they reveal tensions, patterns, automatisms, and recurring systemic errors. Vulnerabilities and incidents are classified according to CVE (Common Vulnerabilities and Exposures) standards and mapped to the MITRE ATT&CK framework to ensure methodological consistency. Finally, the data undergoes quality control and manual review by senior analysts, with accuracy metrics and resolution times calculated based on representative statistical samples.

Events managed in 2025

In 2025, the volume of cybersecurity events doubled compared to the previous year, but it's not just about quantity: the correlation between signals has increased, enabling more effective detection of critical events. Thanks to this approach, escalations have been reduced by 60%, transforming complexity into rapid, targeted action.



The numbers clearly illustrate the intensity of cyber activity and criminal operations: our I-SOC processed 1.935×10^{12} events equivalent to over 61,358 events per second across a sample of 700 customers. This means that in one hour, across the entire monitored perimeter, we processed approximately 220.9 million security events.

If not managed with the right processes and procedures, these events could prove to be the prelude to a serious cyber incident.

Out of this nearly astronomical volume, only 1.8 million became alerts, just 0.000093% of the total, a sign of how effectively the automated pipeline filters out noise and isolates what warrants attention. Of these, 154,760 turned into actual cases (about 8.6%), while only 3,650 were truly relevant, less than 3% of the total cases. The final figure, 3 escalations,, best illustrates the maturity of the process: less than 0.1% of relevant cases required urgent intervention by our Incident Response Team.

Average time to resolution of a non-incident event:

1 hour and 48 minutes

This metric measures the time taken by our I-SOC to handle "non-incident" events that is, activities detected at an early stage of the cyber kill chain, identified before they cause tangible and visible damage to the company. The value of 1 hour and 48 minutes indicates highly optimized triage and correlation processes, with automation and playbooks that shorten the decision-making cycle and help prevent escalation.

Average resolution of an incident:

4.1 days

Average

Cyberoo

21 days

World

World

Incidents, which are inherently more complex, require structured containment, eradication, and recovery efforts. The average resolution time recorded by our Incident Response Team is 4.1 days, compared to the global average of 21 days. This figure highlights how critical rapid coordination and remediation are to limiting the impact of incidents. Short response times reduce the likelihood of propagation, limit operational and economic damage, and directly impact an organization's ability to maintain continuity in the face of critical events.

Most Common Attacks in 2025

When we look back on 2025, one thing stands out: attackers called the shots, while the rest of the world tried to keep up with them.

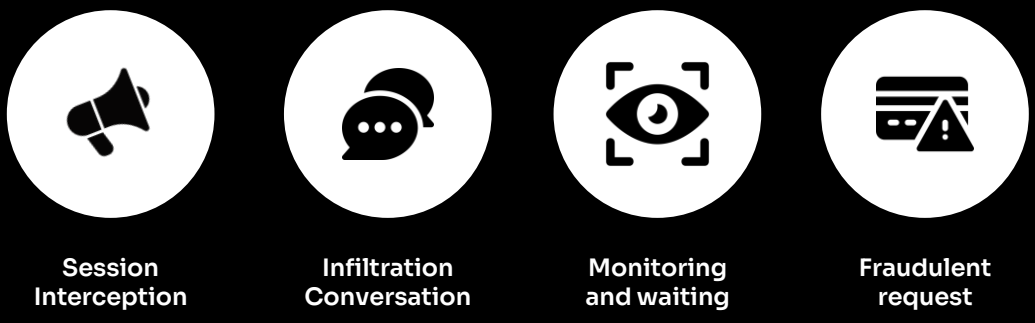
The three main attacking vectors are not “technical categories”; they illustrate how attackers have learned to move with surgical precision within complex systems. The data emerging from our analysis is clear:

Let’s take a closer look at these three types of attacks.

1) BEC & MFA bypass – 40%

Business Email Compromise (BEC) is no longer just a “simple email scam”. By 2025, it has evolved into a highly sophisticated impersonation attack, orchestrated using advanced AiTM (Adversary in the Middle) techniques that allow attackers to intercept valid sessions and bypass many types of MFA (Multi-Factor Authentication) considered “secure” just a few years ago.
What we see is almost always the same pattern:

- the attacker steals or intercepts a session,
- sneaks into the corporate conversation as a colleague,
- waits for the right moment,
- and strikes with a perfectly plausible request.



Why does it work? It doesn’t exploit a technical vulnerability, but a human one: trust in internal processes. And when the attacker “walks in the front door” using someone else’s credentials, defense tools become blind (read the in-depth analysis on p. 71).

2) Exploitation of vulnerabilities – 35%

Speed was the defining factor in 2025. A CVE would be published, and within hours, it had already been turned into a working exploit. This turned patching into a race against time, rather than a technical procedure. Exposed systems such as firewalls, VPNs, and edge appliances were hit with a speed we had never seen before. Attackers know that these devices are often critical, exposed to the Internet, difficult to update quickly, and full of active sessions to steal.

When an exploit hits, the effect is immediate: full access, rapid lateral movement, and bypassing of traditional controls. For this reason, you cannot protect a complex infrastructure with monthly patching cycles; instead, you need rapid mitigation mechanisms, exposure control, and immediate session revocation—capable of reducing the impact in the first few hours, when the damage spreads the fastest.

3) Private infostealers – 25%

For years, infostealers were considered “second-rate spyware” because they do not cause immediate, visible damage to infected systems, but operate silently, limiting themselves to collecting and stealing specific information, which they then send to servers controlled by the attackers to fuel fraudulent activities, unauthorized access, and subsequent attack campaigns. By 2025, this perception has changed drastically. We have seen a proliferation of custom versions, tailored for specific campaigns, capable of stealing:

- password,
- API token,
- session cookies,
- cloud credentials,
- already authenticated OAuth sessions.

A single exfiltration can open all the doors. Cybercriminals don't need to encrypt anything or compromise internal servers: they simply use what they've stolen to become a legitimate user, invisible to monitoring systems. And that is precisely the danger: infostealers create silent access, often weeks or months before the actual attack.

It certainly wasn't the year attackers used "new, futuristic technologies." It was the year they learned to better leverage what already worked: identity-based attacks, rapid exploitation of vulnerabilities, systematic use of stolen credentials, and psychological techniques amplified by generative AI.

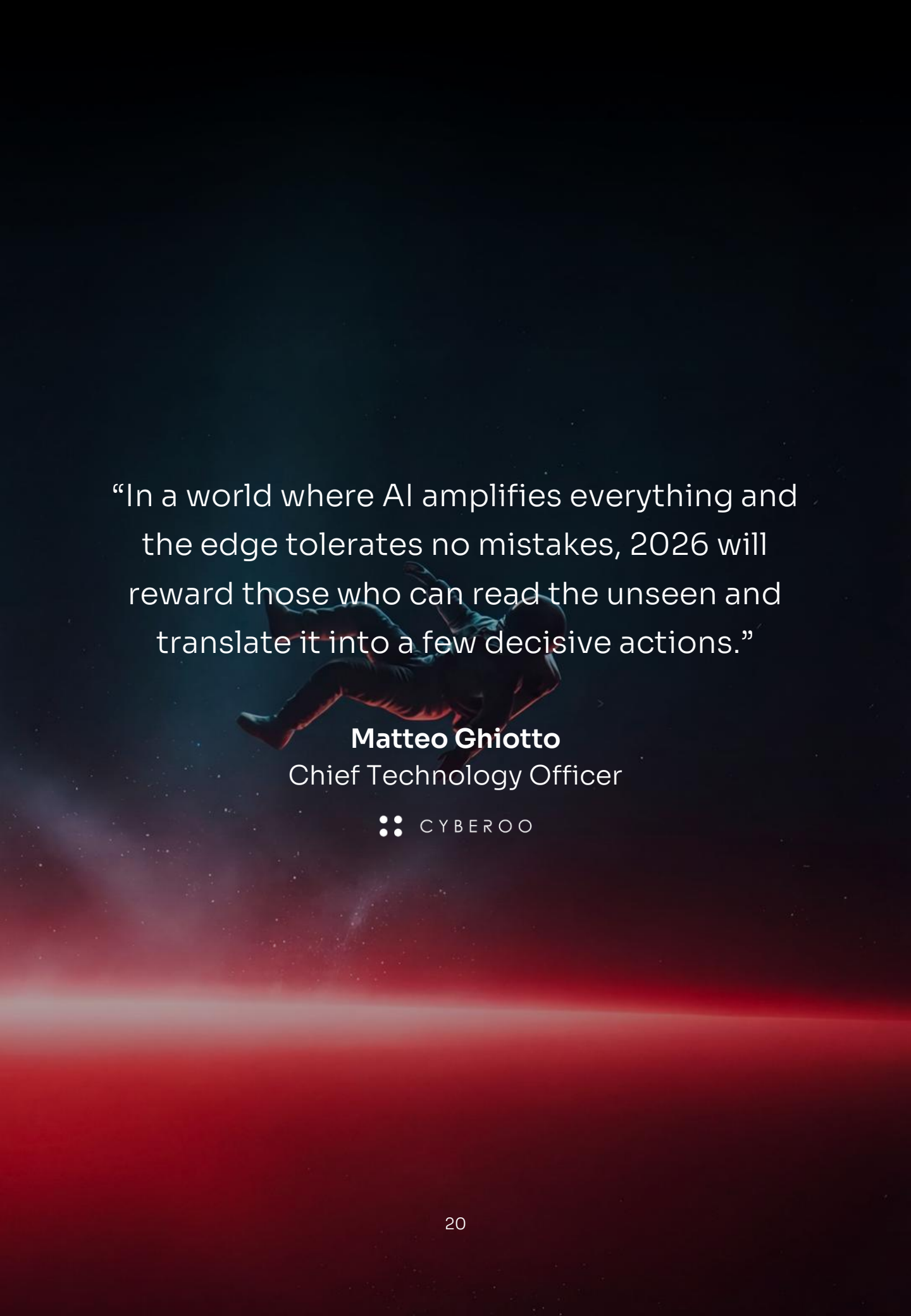
The message emerging from the data is unequivocal: cybersecurity is no longer a matter of tools, but of timing and identity. Attackers are staying one step ahead. We must learn to play strategically.

1,300 reports
submitted to the relevant authorities regarding phishing, spam, and fraud.

Trends in 2025

2025 wasn’t just a year of evolution: it was a watershed moment.
We all felt it: the speed at which threats transformed forced us to reconfigure the very way we view risk.
Traditional defenses, based on linear predictions, were overwhelmed by an ecosystem that moves faster than many companies can react.

On the following pages, we’ll explore the key trends our I-SOC has identified in the field.



“In a world where AI amplifies everything and the edge tolerates no mistakes, 2026 will reward those who can read the unseen and translate it into a few decisive actions.”

Matteo Ghiotto
Chief Technology Officer

CYBEROO

1) Exploits and Stolen Credentials

In 2025, we saw that one of the most common attack vectors was the direct exploitation of vulnerabilities, including the dreaded zero-day vulnerabilities—that is, previously unknown flaws for which no patches were available at the time of discovery. We observed this clearly in the activities of the most aggressive groups, which were able to turn a newly disclosed vulnerability into a working weapon in a matter of hours, not days.

At the same time, the market for stolen credentials has exploded: sessions, OAuth tokens, and passwords intercepted via infostealers or AiTM (Adversary-in-the-Middle) attacks have become the preferred currency for infiltrating a network undetected. When an attacker obtains a valid session, they don't need malware: they behave like a legitimate user. This is what makes the situation critical: the risk shifts from the perimeter to identity, and many companies are not prepared for this change.

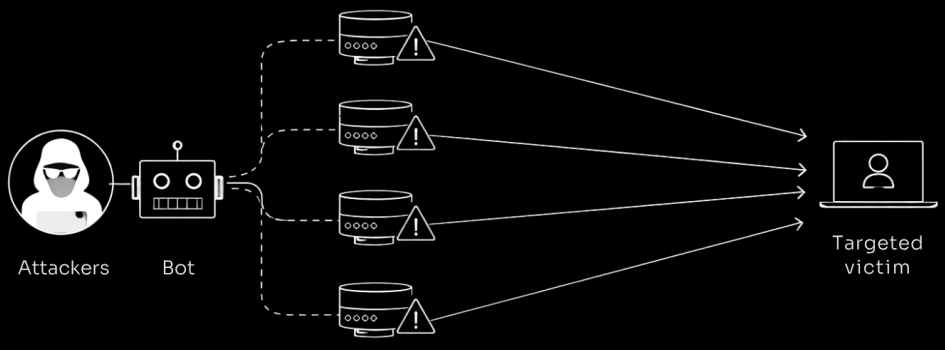
Most common passwords

(public sources)

123456
Password
Admin
Qwerty
abc123
Iloveyou
letmein
Welcome
000000

2) DDoS (Distributed Denial of Service)

Recent years have also marked a renewed surge in distributed denial-of-service (DDoS) attacks across Europe. According to the ENISA Threat Landscape, DDoS attacks represent one of the most prevalent cyber-threat categories in the European threat landscape. Not because they cause structural “damage”, but because they have become perfect distraction tools: they saturate the attention of defenses, while simultaneously triggering a reconnaissance phase or a quieter infiltration.



More and more criminal groups are using DDoS attacks as “planned cover”, coordinating them with credential-theft campaigns and automated lateral movements. These are short, pulsed, almost surgical attacks, often orchestrated by botnets powered by compromised IoT devices or by DDoS-as-a-Service offerings that can be purchased for just a few dozen euros. The most interesting thing is that today, DDoS attacks aren’t used to block, they’re used to hide.

3) AI-powered phishing

The real revolution has occurred in phishing. We've encountered flawless emails, perfectly contextualized, with AI-generated text and, increasingly, accompanied by deepfake voice messages or videos that are difficult to distinguish from an authentic message. These are no longer the classic error-ridden emails. They are credible, persuasive requests, often "digitally" signed by trusted internal figures. We have observed campaigns in which the attacker does not simply send a message but constructs a full narrative sequence: low-impact initial emails, consistent follow-ups over time, deepfake voice calls that reinforce urgency, and even the use of compromised accounts or newly registered look-alike domains to boost credibility. In some cases, messages are adapted in real time based on the victim's responses, leveraging generative models linked to OSINT (Open Source Intelligence) data, corporate organizational charts, and data from previous breaches.

This has brought to the forefront a truth we all know: cybersecurity is more human than we think (read the in-depth analysis on p. 78), and awareness plays a fundamental role in corporate security, especially when manipulation is so sophisticated. For this reason, it is essential to invest in training and security procedures such as double-checking. In a scenario where social engineering becomes adaptive, persistent, and increasingly automated, human behavior remains the final deciding factor that determines the outcome of the attack.

38,654 domains

suspected domains
detected in 2025.

4) AI-powered attacks

2025 was the first year we saw generative AI used on an industrial scale by attackers: in creating tailored content, automating reconnaissance operations, generating attack variants, and even building actual automated playbooks.

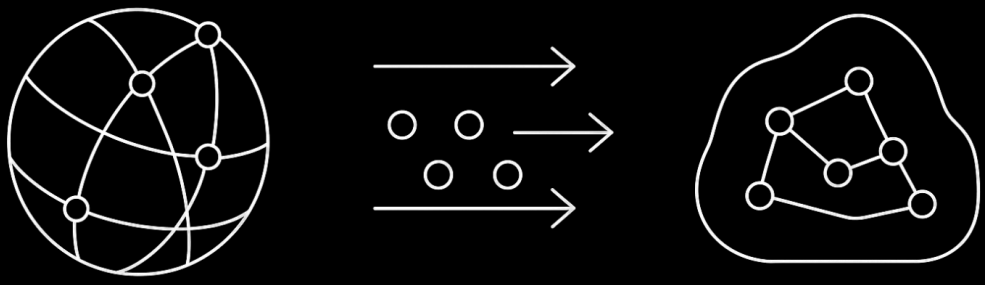
TTPs (Tactics, Techniques, Procedures) were replicated and improved at a pace that put pressure on every traditional defense model. The challenge is no longer just protecting systems, but protecting the models and ecosystems surrounding them: from Agentic AI to orchestration layers, from application access interfaces to integration mechanisms with external data and services.

This type of attack targets the “mind” of AI systems—that is, the way they reason and make decisions. For example, prompt injection (hidden or deceptive instructions in messages) attempts to make the model do things it shouldn’t; data poisoning (manipulated data to influence results) alters the quality of responses; API abuse (misuse of technical access to the system) exploits functionalities incorrectly; and data leakage (the escape or exposure of confidential information) risks the disclosure of specific data. The result is that the system can become less reliable, more vulnerable, and less secure to use in operations (we explore this topic further on page 88).

5) Pandemic Supply Chain

Chain attacks involving suppliers are now commonplace. Exposed repositories, rushed cloud migrations with poorly secured tokens and APIs, overly broad permissions in SaaS applications, or infostealers that steal credentials turn small errors into widespread compromises. In these cases, the attack is not limited to the first system hit but exploits the trust and interconnections between environments to spread rapidly: this is the dynamic known as a Supply Chain Pandemic.

A single vulnerability in a supplier or a shared component can simultaneously affect dozens of organizations. The high level of interdependence between platforms and services amplifies the impact of incidents, turning isolated events into large-scale crises. In response to this risk, the concept of Geopatriation is emerging: the trend of moving data and workloads to sovereign or regional clouds to reduce geopolitical exposure (see the in-depth analysis on p. 68), strengthen digital sovereignty, and limit the systemic effects of global compromises (Gartner, 2025).



6) Triple Extortion

Modern ransomware no longer limits itself to encrypting systems or threatening to publish stolen data. Confirming the importance of the supply chain discussed in the previous point, a further phase is emerging: Triple Extortion, in which attackers also exert pressure on the victim's partners, customers, and suppliers.

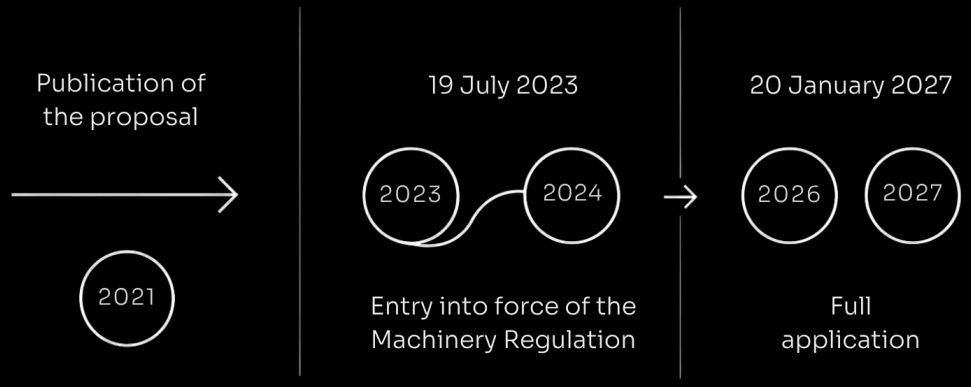
	Single Extortion	Double Extortion	Triple Extortion
Method	DDoS or direct threat	Data theft + blackmail	Encryption + theft + external pressure
Victims	Affected services	Organizations and customers	Company, users, partners
Impact	Operational disruption	Data and reputation loss	Total shutdown and systemic crisis

This means that the incident is not confined to the affected company but extends to its supply chain, turning into a crisis that involves the entire ecosystem. Even with robust backups or restored systems, the threat remains: attackers can continue to use the exfiltrated data to target third parties and amplify the damage (read the in-depth analysis on p. 83).

This evolution confirms that the operational landscape is no longer characterized by isolated attacks, but by strategies designed to exploit interdependencies, relationships, and connections. Consequently, defense must consider not only its own perimeter but the entire network of actors orbiting the organization.

7) Zero-day and edge attacks

Firewalls, VPNs, remote access appliances, and, increasingly, exposed OT devices: anything accessible from the Internet has become a prime target. Every week, multi-actor campaigns exploiting new CVEs have been observed, often targeting widely used products such as Citrix, Ivanti, Fortinet, and PAN-OS, but also industrial components and embedded systems. The attack surface is no longer limited to traditional IT: connected fleets of corporate vehicles and industrial machinery introduce new points of exposure, now explicitly addressed by the Machinery Regulation EU 2023/1230, which will apply from 20 January 2027.



The most concerning factor remains speed: the window between disclosure and exploitation is now a matter of hours. Let’s analyze in detail how these often-invisible dynamics translate into concrete operational vulnerabilities.

Most Common Vulnerabilities 2025

A perimeter that appears solid may conceal small internal cracks, nearly invisible, yet sufficient to allow entry to those who know how to exploit them. Our Observatory has recorded an inevitable phenomenon: the most exploited vulnerabilities were not the “most sophisticated” ones, but the most convenient, most widespread, and most exposed.

Those that only needed to be touched in the right place to bring down entire environments. We observe a progressive evolution of the scenarios: weaponization within a few hours, coordinated campaigns, pre-auth exploits, and targeting of core platforms. Below is the detailed analysis.

1) Edge & VPN appliances

The most intense attacks were directed at devices exposed to the Internet. Everything standing between the company and the outside world (VPNs, gateways, ADCs) became a prime target. We observed
pre-auth attacks, meaning the attacker gained access even before the system requested credentials. Some cases were particularly telling:

- **CitrixBleed 2 (CVE-2025-5777)**: allowed attackers to read the machine's memory and hijack active sessions. A vulnerability capable of putting any remote infrastructure under serious strain
- **NetScaler RCE 0-day (CVE-2025-7775)**: allowed remote code execution via an IPv6 overflow
- **Ivanti Connect Secure / Policy Secure**: was affected by two vulnerabilities actively exploited within the same timeframe, requiring patches and system integrity checks
- **Fortinet (CVE-2025-32756)**: a stack-based buffer overflow vulnerability that affected several Fortinet products, allowing unauthenticated remote attackers to execute arbitrary code or commands by sending HTTP requests with specially crafted cookie hashes.

If something is exposed on the Internet, it will always be probed, tested, and compromised. Much sooner than we expect.

“Attacks come fast:
without rapid patches and session
revocation, you’re already behind.”

Matteo Gruppi
ASOC Team Leader

CYBEROO

2) Enterprise & Core Systems

If the edge is the gateway, core platforms are the “beating heart”. 2025 has shown that attackers are not afraid to strike precisely there.

The most critical cases include:

- **SAP (CVE-2025-31324)**: a vulnerability in a system that supports finance, logistics, and the supply chain. When SAP stops, the entire company stops, making timely and prioritized corrective measures necessary
- **Microsoft SharePoint on-premises**: it was hit by a chain of zero-day exploits that allowed remote code execution via a deserialization flaw, with tools like “ToolShell” used to move rapidly between systems
- **Windows SMB Client (CVE-2025-33073)**: saw real-world activity in the wild that enabled privilege escalation and deep lateral movement, making SMB signing mandatory and reducing the use of NTLM.

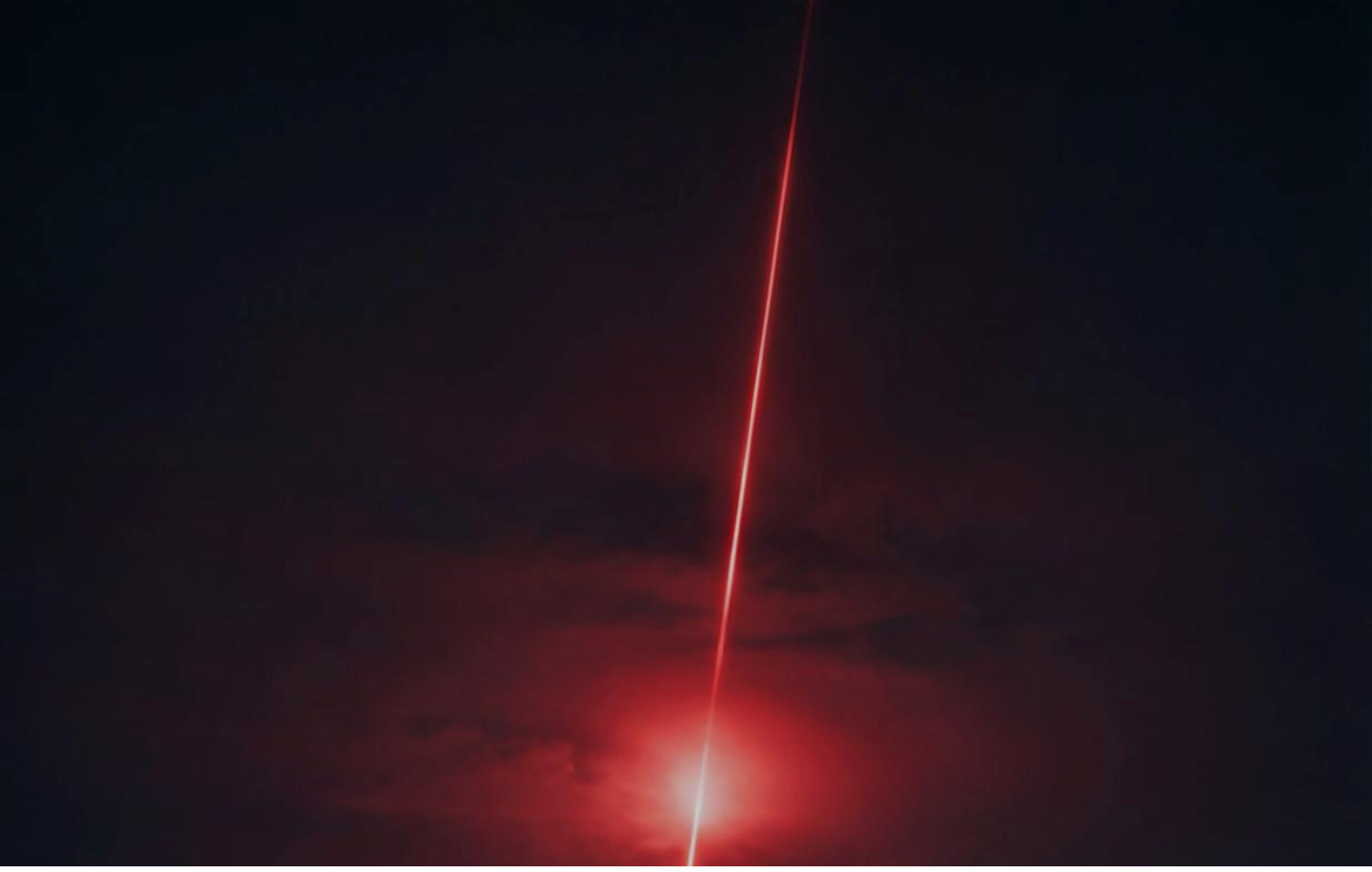
These cases were the most critical: when a core platform falls, you’re not defending a server, you’re defending the company as a living organism.

3) Virtualization & Enterprise Applications

In 2025, we saw that attackers are no longer content with targeting individual systems; they are now directly targeting the platforms that manage the entire infrastructure. There was no shortage of such cases:

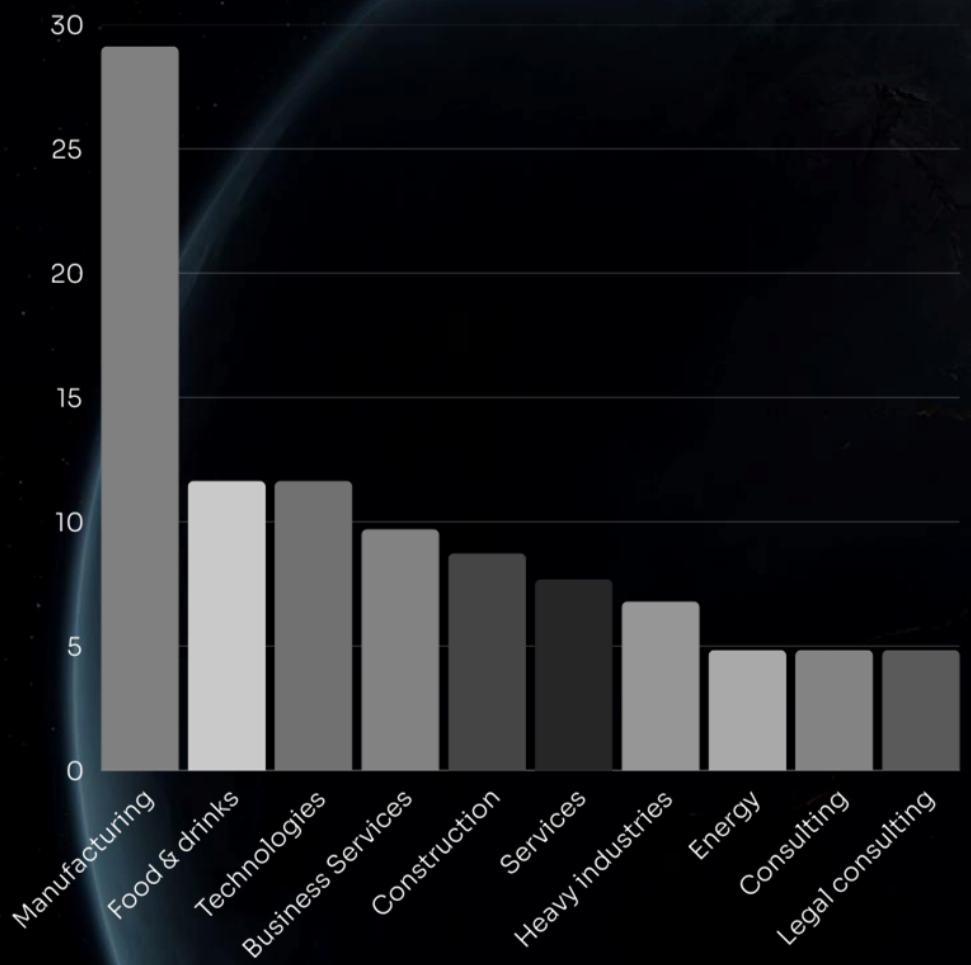
- **VMware Aria Operations / VMware Tools (CVE-2025-41244)**: This vulnerability allowed criminal hackers to gain elevated privileges within virtualized environments. In practice, when exploited, it gave them the ability to control the machine as an administrator.
- **Oracle E-Business Suite (CVE-2025-61882)**: a pre-authentication RCE via HTTP, exploited in real-world attacks. A single HTTP packet was enough to gain complete control of the instance exposed to the Internet.

When a virtualization system or a suite like Oracle E-Business is compromised, it is not a single application that falls but the entire infrastructure upon which all others rest, enabling immediate lateral movement. The most exploited vulnerabilities are not the most complex but the quickest to use, which is why managing them is no longer a one-off project: it is an ongoing program of mitigation and adaptation to the speed of the attackers.



Sectors most affected and Distribution in 2025

Sectors Most Affected in 2025



Percentage distribution of the most affected sectors in 2025

When we examine the distribution of the most frequently cited sectors among those affected in 2025, the first thing that stands out is that certain sectors appear more often in the observed cases. However, this analysis should be interpreted in relation to the weight of different sectors within the European economic landscape: a highly prevalent sector such as manufacturing, often reported among the most affected, may appear more prominent partly due to a “volume effect” linked to its widespread presence across the European industrial base.

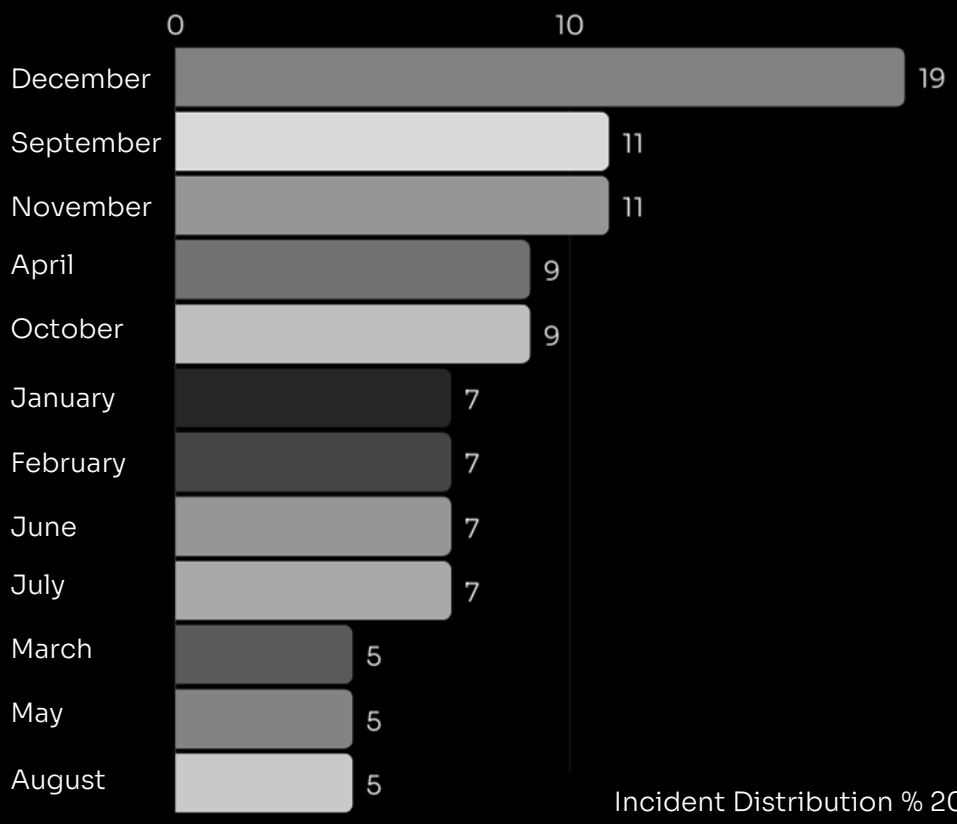
That said, every industry brings with it a specific combination of technologies, processes, and digital dependencies that, when viewed by an expert eye, can become a set of genuine “pressure points” within the cyber ecosystem.

Every disruption has a real, measurable, immediate impact, and that is exactly what attackers exploit. Next come sectors such as Food & Beverages and Technology, which, although very different from one another, share a critical element: the presence of interconnected digital processes and, in the case of technology services, a “hub” role that amplifies the effects of a compromise.

The central cluster, Business Services, Construction, and more general Services, tells yet another story: that of sectors with a high density of identities, documents, distributed workflows, and collaborations with external suppliers. In these environments, even a single compromised access point can spread rapidly, especially when IT is fragmented or managed in a hybrid mode.

Further down the list, but far from safe, we find Heavy Industries and Energy: sectors that do not face an extremely high volume of attacks but remain among the most strategically sensitive. Here, attackers’ focus is less on quantity and more on quality: critical infrastructure, legacy systems, physical-digital interdependencies, and a growing use of industrial IoT transform every vulnerability into a potential systemic risk.

Annual Distribution 2025



Incident Distribution % 2025

December is the month in which 19% of the incidents in 2025 occurred. It is likely no coincidence that this month coincides with the Christmas holidays, a time when companies are more vulnerable because they are less responsive and have fewer staff on hand. September and November are not far behind, accounting for 11% of incidents, while April and October account for 9%.

The incidents analyzed do not exclusively involve existing clients, for whom three escalations were recorded. The sample also includes non-client organizations that requested urgent support through our Incident Response service, thus allowing for a broader and more representative assessment of the operational dynamics observed.

In general, a pattern emerges that cybersecurity professionals are well acquainted with: incidents do not follow a regular trend, but tend to cluster during time windows marked by critical vulnerabilities, coordinated campaigns, or periods of heightened exposure for the industrial sector. 2025 was no exception.

What we observe is a pulsating pattern, with peaks that are not random but reflect specific events: the publication of new high-impact CVEs, the emergence of organized ransomware campaigns, the increase in infostealing activities, and the massive exploitation of edge vulnerabilities.

Another clear aspect is the persistence of risk: even in months when the volume appears low, the distribution never truly drops to negligible levels. This reflects the fact that the criminal market has now reached a state of continuous operation, with ransomware groups and emerging threat actors maintaining a constant flow of activity, often reshaped by rebranding and temporary collaborations, as in the case of the collectives formed in 2025 (as we will see in the next chapter).

Overall, the data confirms an operational reality: cybersecurity can no longer be treated as a seasonal phenomenon. The peaks tell us when attackers strike hardest, but the constant baseline reminds us that the attack surface is always active, and exploits, especially those involving stolen identities, tokens, and credentials, operate underground for weeks or months before becoming explicit incidents.

Top Threat Actors of 2025

In 2025, we saw a high concentration of activity by ransomware groups and organized cybercriminals, with trends reflecting a constantly evolving ecosystem. Some actors consolidated their presence, while others emerged following takedown operations, demonstrating just how resilient and adaptive cybercrime is.

We monitored the most significant developments, identifying the groups that had the greatest impact in Europe and globally. Let’s analyze them together.

Qilin

Qilin established itself as the most aggressive group in the second quarter of 2025, with high-impact campaigns and a strategy focused on “big game hunting” that is, targeted attacks on large organizations. Among its most notable operations, the Asahi case highlighted the group’s ability to exploit critical vulnerabilities and combine encryption techniques with data extortion, thereby increasing the pressure on victims.

Akira

Akira has distinguished itself through its widespread presence in European organizations. We are not talking about isolated attacks, but rather a constant presence, facilitated by social engineering tactics and the exploitation of stolen credentials. Its strength lies in its ability to adapt to various sectors, from manufacturing to finance, with a flexible and aggressive operational model.

INC Ransom

INC Ransom reached its peak activity in July 2025, with a series of coordinated attacks targeting critical infrastructure and large enterprises. The speed of execution and the ability to exploit known vulnerabilities made this group one of the most dangerous during the summer.

Interlock

First observed in 2024, Interlock became fully operational in 2025, prompting CISA, the FBI, HHS, and MS-ISAC to issue a joint alert on July 22. The group targets supply chains and critical infrastructure, exploiting techniques such as drive-by downloads and social engineering attacks to gain initial access, followed by data exfiltration and file encryption in a double extortion model.

SafePay

SafePay saw significant growth in the second quarter of 2025, particularly in Europe. The group has focused on targeted ransomware campaigns, using a “low-noise” approach that reduces the likelihood of detection in the early stages. This makes it particularly insidious, as victims often only discover the attack once their systems have already been compromised.

One thing is certain: threat actors are not static entities; they evolve, reorganize, and exploit every opportunity to maximize their impact (see p. 98 for an analysis of their behavior on the Dark Web). Their ability to adapt requires companies to adopt a dynamic approach to cybersecurity, based on continuous cyber threat intelligence, proactive monitoring, and resilience strategies that go beyond simply reacting to incidents (ENISA Threat Landscape, 2025).

320 new
criminal groups
identified in 2025.

New threat actors in 2025

The year 2025 saw the emergence of new groups that have reshaped the threat landscape, ranging from rebranded versions of well-known organizations to collectives formed by the merger of multiple actors employing increasingly sophisticated tactics. Let's take a look at the most significant cases.

Trinity of Chaos (ToC)

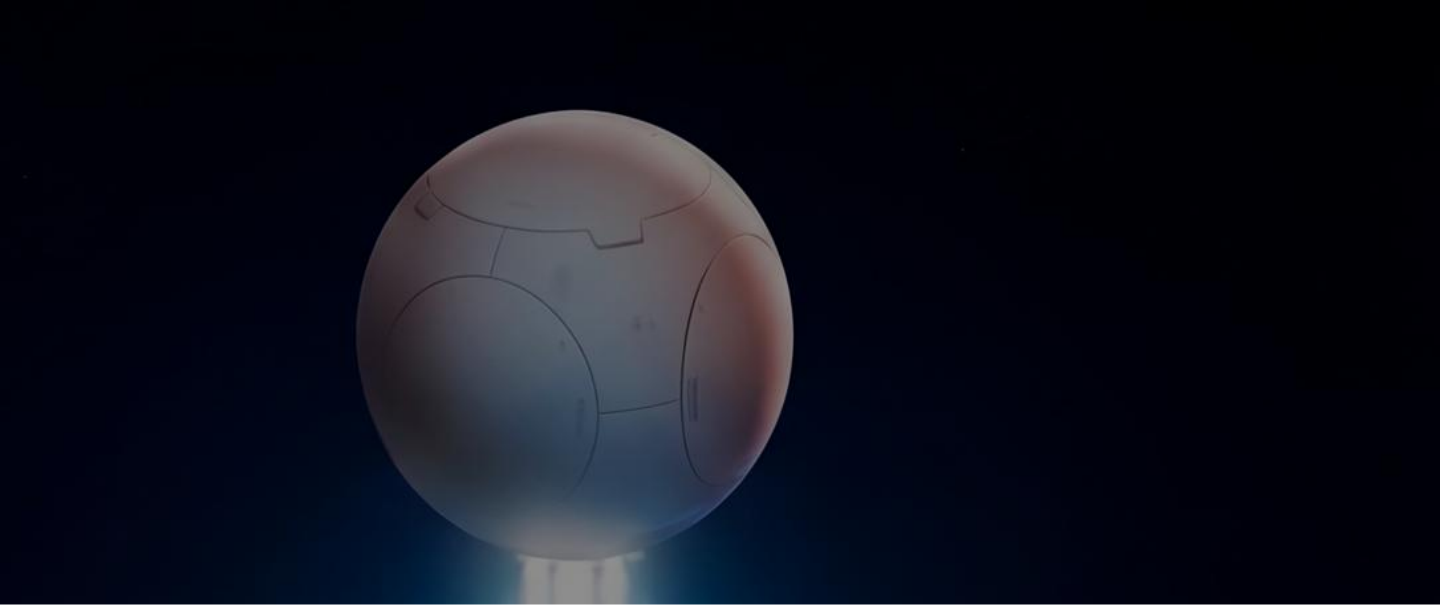
Formed between September and October 2025, ToC is a collective uniting three well-known groups: LAPSUS$, Scattered Spider, and ShinyHunters. It launched a Data Leak Site (DLS) on TOR exposing 39 companies, including major names such as Air France and Stellantis. The most exploited technique? Misconfigured Salesforce instances, which paved the way for massive data exfiltrations. This case highlights how collaboration between criminal groups can amplify the scope of attacks.

Chaos

Following the takedown of BlackSuit, Chaos emerged, likely as a rebrand of the same group, employing very similar techniques and tools. Active since February 2025, it engages in "big game hunting": targeted attacks on large companies, with initial access gained through spam, voice phishing, and remote management tools (RMM). It has continued with advanced ransomware campaigns, confirming the trend of groups changing names and fragmenting. A presence in Europe is plausible, given that BlackSuit had already targeted many EU victims.

Emerging groups exhibit a different structural dynamic: cybercrime grows not only through accumulation but also through recombination. Rapid rebranding, opportunistic cooperation, and the modular reuse of infrastructure and tactics, techniques, and procedures (TTPs) enable new actors to operate effectively before they become fully observable.

For this reason, a defense strategy based solely on tracking known groups is inherently incomplete: it becomes essential to adopt proactive threat intelligence capabilities, capable of correlating weak signals, behavioral anomalies, and emerging technical patterns before they solidify into large-scale campaigns.



Cybersecurity Challenges 2026

Let’s now turn to the new year and the challenges of the coming months. 2026 brings with it faster, more widespread, and better-coordinated threats. Attackers exploit automation, AI, and interdependencies to strike at blind spots that often remain invisible.

Cybersecurity is no longer a collection of technologies, but the ability to interpret a risk that moves faster than traditional defenses now surpassed by models that combine processes, technologies, and human expertise. The challenges that follow illustrate this new balance: an ecosystem in which seeing, understanding, and reacting become just as essential as protecting.

1) AI-powered threats & info-ops

AI is no longer merely a tool but also a multiplier for offensive operations: it reduces costs and the skills required of attackers, generating targeted phishing attacks, credible deepfakes, and realistic impersonations. Generative models produce contextualized content and automate entire stages of an attack, making it increasingly difficult to distinguish the fake from the legitimate. By 2026, two trends will emerge:

- **Agentic AI**, featuring autonomous agents capable of planning and acting, exposed to risks such as prompt injection, API abuse, data source manipulation, and excessive privileges. A compromised agent can operate with speed and persistence surpassing human capabilities; it therefore becomes essential to define governance, limits, and traceability.
- **Superagency**, or the overtaking of non-human identities over human ones: bots, APIs, services, and IoT devices become the majority and represent a new blind spot. These identities, too, can be stolen or impersonated without generating obvious signals, requiring robust rules for managing permissions and keys.

Rounding out the picture is LLMjacking, the unauthorized use of language models via key application interfaces or stolen cloud credentials, which carries risks of unexpected costs and data exfiltration. Rigorous practices are therefore needed: secure secret management, temporary keys, usage monitoring, and rate limits on APIs.

2) Compromised identities: human, non-human, invisible

Most breaches do not stem from advanced malware but from compromised identities, whether human or non-human. The former involve typical errors, such as reused passwords, impulsive clicks, and trust in authority, while Non-Human Identities (NHI) include bots, application interfaces, services, and devices that operate autonomously and often without visibility. When even a single account is compromised, the attacker can act undetected by exploiting existing access and still-valid digital credentials, or by intercepting login sessions by “man-in-the-middle” (MITM) attacks between the user and the service. In these cases, many traditional defenses are insufficient. This is why two-factor authentication, on its own, is not always enough: we need login methods that are truly fraud-resistant, we must eliminate outdated login systems, and we must use tools that detect anomalous behavior even on accounts that appear legitimate.

Compounding the problem is a crisis of trust: with deepfakes, credible voice and video clones can now be created in seconds, facilitating fraud and carefully crafted urgent requests. To defend against this, three levers are needed: behavioral biometrics (which recognizes how a person interacts), digital provenance of content (verifiable signatures and metadata for images, videos, and documents via standards like C2PA), and Zero Trust procedures, which block abuse by acting on the process with continuous verification, least privilege, and controls at critical points.

3) Edge and cloud under pressure

Attackers have shifted the focus of their operations to the edge and remote access, as these are the points most exposed to the Internet and therefore easier to exploit. In 2025, we saw vulnerabilities in gateways, VPNs, and authentication systems turn into working exploits within hours, often even before authentication took place. This demands a cultural shift: patches within hours, not weeks, and immediate revocation of potentially compromised sessions after every update.

At the same time, the growth of SaaS and the cloud has expanded the attack surface: unprotected storage, credentials, and connections exposed during migrations, overly broad permissions granted to external tools, and even the uncontrolled use of artificial intelligence services. To defend against this, cloud security must be continuous: tools are needed that show what is active, how it is configured, and how data and access flow. Added to this are system separation, traffic controls for AI services, periodic key rotation, and the search for “hidden” administrators and unauthorized shares.

The boundary is no longer just digital. Physical AI, that is, AI within industrial robots, medical devices, and OT systems—requires protecting not only data but also operational continuity and physical security.

This requires controlled architectures: zone-based segmentation, bastion hosts, and single-use jump hosts, session recording, strong MFA, and strictly regulated remote access.

The variable that determines everything today is time. Attackers are industrializing the discovery-exploitation cycle, while many companies still operate on a monthly maintenance schedule. By 2026, we'll need a fast-patch approach for exposed systems, with clear playbooks to revoke compromised tokens and sessions and perform hardening immediately after the update. It's not about "patching faster", it's about redesigning the organization to make decisions faster.

	Vulnerability or risk	Method of attack	Operational Operational
Edge and remote access	Authentication systems and gateways exposed to the Internet	Exploits Pre-authentication exploits targeting recent vulnerabilities	Unauthorized access to the corporate network
Cloud and SaaS	Exposed credentials and excessive permissions	Abuse of shared access or hidden administrative accounts	Data loss and loss of control over cloud assets
Physical AI	AI integrated into physical systems without adequate protections	AI integrated into physical systems without adequate safeguards	Risks to physical security and business continuity

4) The Year of Real Regulatory Deadlines

2026 is the year when cybersecurity truly meets European regulation: it is no longer enough to simply be “compliant”, but to demonstrate compliance through traceable processes, documented oversight, and verifiable decisions.

The Artificial Intelligence Act (EU) 2024 entered into force on 1 August 2024, but its main provisions will start to apply from 2 August 2026. From that date, mandatory requirements for high-risk AI systems and transparency obligations for deepfakes and AI-generated content will become applicable across the European Union. Companies using these systems will have to demonstrate that they manage them properly: using AI according to instructions, providing adequate data, monitoring its operation, and immediately reporting risks or incidents to the authorities. In the meantime, the Commission is publishing guidelines and codes of conduct to help organizations prepare. Essentially, by August 2026, anyone developing or using high-risk AI must be able to prove, with facts and documentation, that they have full control, traceability, and transparency over the entire lifecycle of their solutions (read the in-depth analysis on p. 91).

On the NIS2 Directive (EU) 2022/2555 front, the coming years mark the operational implementation phase across the European Union following its transposition by Member States. Organizations covered by the directive must notify significant cybersecurity incidents to the relevant national authorities and CSIRTs. The framework requires an early warning within 24 hours, an incident notification within 72 hours, and a final report within one month.

Organizations classified as essential or important entities must implement appropriate technical and organizational cybersecurity risk-management measures, as defined by the directive and further specified through national implementation frameworks across EU Member States.

The Cyber Resilience Act (EU) 2024, which entered into force on 10 December 2024, introduces for the first time EU-wide cybersecurity obligations for products with digital elements. From September 2026, manufacturers will be required to report actively exploited vulnerabilities and severe incidents without undue delay, including those affecting products already on the market. These obligations anticipate the full application of the regulation scheduled for December 2027, when comprehensive requirements on secure design, risk assessment, lifecycle management and CE conformity for digital products will fully apply across the European Union.

The framework is completed by the Digital Operational Resilience Act (EU) 2022/2554, fully applicable since 17 January 2025, which harmonizes digital resilience requirements for the European financial sector. In the following years the most demanding components become operational: advanced resilience testing, stricter oversight of critical ICT third-party providers and enhanced incident-reporting mechanisms within a significantly strengthened supervisory regime. For financial institutions, the key challenge is no longer simply achieving compliance, but demonstrating continuous resilience, effective third-party governance and the ability to maintain operations during high-impact disruptions (additional strategic details are reported on p. 95).

“Regulations set the boundaries; quality is the execution: repeatable decisions, traceable actions, suppliers under control. The rest is noise.”

Andrea Costantino
Chief Quality Officer

CYBEROO

5) Toward Post-Quantum Cryptography

Quantum computers are becoming increasingly powerful, and sooner or later, Q Day will arrive, the moment when the encryption technologies we use today will no longer be secure. When that happens, anyone who intercepted encrypted data in previous years will be able to read it without any trouble.

This is why attackers are already using the *"Harvest Now, Decrypt* Later" strategy: stealing encrypted data today that will become readable tomorrow. To protect against this, the first post-quantum encryption algorithms have been developed, designed to withstand quantum computers, but their adoption will take time and require a thorough overhaul of infrastructure.

Companies must therefore become crypto-agile, that is, able to quickly change algorithms and protocols without rewriting entire systems. There's no need to switch to post-quantum cryptography immediately, but preparation is essential: when the change comes, those lacking this flexibility risk failing to migrate in time, exposing sensitive data, critical infrastructure, and entire business processes to the risk of being decrypted, manipulated, or rendered unusable.

Cybersecurity Strategies for 2026

The year 2026 marks a clear turning point: threats are not only faster, but also smarter, automated, and capable of exploiting every weak link in the digital ecosystem.

That is why defense can no longer be a collection of tools placed side by side: it must become a living, coordinated, continuous system. In the following pages, we will explore the eight strategies a company must adopt to remain truly resilient in 2026.

1) Lack of governance: a real risk

When we step in to support organizations affected by a cyber incident, certain recurring characteristics regularly emerge: the absence of structured governance, a high level of system and network obsolescence, and insufficient investment in security processes. These elements constitute a systemic risk that affects not only the technological infrastructure but, above all, the organizational culture and management's decision-making capacity. In many cases, the Board tends to underestimate cybersecurity, viewing it as an "invisible" problem until it produces tangible effects; when an incident occurs, the organization is forced to cram into a few weeks activities and processes that would have required years of orderly implementation, with significant impacts on costs, business continuity, and reputation.

Another critical element concerns the human dimension of cybersecurity. Behavioral dynamics have a decisive influence on the outcome of many incidents: impulsivity, unchecked trust, emotional pressure, and a sense of urgency are variables that attackers systematically exploit to bypass technical controls. These mechanisms affect not only end users but also the executive and decision-making levels of the company, where hasty decisions or distorted perceptions of risk can amplify existing vulnerabilities.

For this reason, one of the strategic priorities for cybersecurity in 2026 is the adoption of robust governance, based on formalized processes, consistent investment, and targeted training programs. It is not enough to simply offer courses or issue certifications: it is necessary to address people's mental models, behaviors, and decision-making abilities, aligning skills, awareness, and accountability.

When corporate culture evolves and individuals become an active part of the protection system, the attack surface is significantly reduced. Effective governance builds an organizational wall that, if free of cracks, represents the first and most important barrier against cybercriminals.

	Problem	Impact	Priorities for 2026
Governance and Management	Absent governance and security underestimated by the Board	High costs, operational downtime and reputational damage	Structured governance and consistent investments
Infrastructure obsolescence	Outdated systems and insufficient investments	Accelerated emergency interventions	Training and skills alignment
Human factor	Weak security culture and social engineering	Exacerbation of technical vulnerabilities	Evolution of behaviors and responsibilities

2) Resilience: Cybersecurity Beyond Technology

Another belief that we must finally move beyond in 2026 is the notion that cybersecurity is purely a technological matter, where simply purchasing various tools is enough to thwart any threat. We discussed this in detail in our report “The Future of Cybersecurity,” but we’d like to revisit the topic given the importance of this paradigm. Today, attackers’ tactics, techniques, and procedures are no longer what they used to be: speed, invisibility, and automation reign supreme. Thinking that simply adding technologies is enough to feel secure no longer works.

We must combine culture and awareness with the orchestration of technologies, processes, and human expertise. It is no longer about choosing technologies but services capable of correlating signals in near-real-time, at any hour of the day, recognizing anomalous patterns, and reacting immediately within the very first seconds, especially when offices have emptied and the lights have gone out. Not only from internal threats but also external ones related to Cyber Threat Intelligence. This is what makes the difference between a contained incident and an operational disaster. The goal is not simply to reduce *Mean Time to Detect* or *Mean Time to Respond*: it is to rewrite the timeline of defense, moving from slow, sequential processes to dynamic playbooks, intelligent automation, and human teams focused on critical decisions.

This means building a system that doesn't just defend but allows the company to continue operating even while an attack is underway. This is where true resilience begins.

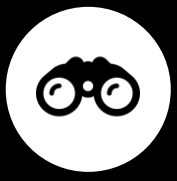
The question to start with is no longer "how well are we protected?", but "can we keep working while under attack?". This requires defining the *Minimum Viable Business*: the essential set of people, processes, applications, and data that must remain operational to allow the company to function at a minimum level even during a cyber incident.

CULTURE AND COMPETENCIES

Integrating human awareness and processes into technological orchestration.

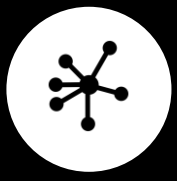
MONITORING AND 24H RESPONSE

Integrate human awareness and processes into technology orchestration.



CYBER THREAT INTELLIGENCE

Proactive monitoring of internal and external threats to prevent operational disasters.



DYNAMIC PLAYBOOKS

Replacing slow processes with intelligent automation and critical human decisions.

MINIMUM VIABLE BUSINESS

Defining the essential set of assets to ensure minimum guaranteed operability.

TRUE RESILIENCE

Building a system that allows the organisation to operate even during an attack.

“Today, it’s not technology that makes the difference, but the human factor: collaboration helps reduce real risks and build trust.”

Giuseppe Cersosimo
Head of Security Advisor Manager

CYBEROO

3) Identity-first: a top priority

Today, attacks no longer start by “breaking down a door”; they gain entry using valid credentials. We’re talking about compromised accounts, permissions already granted, or sessions that are still open. That’s why putting identity at the center isn’t a technical choice, but a matter of survival. The first step is to use multi-factor authentication that cannot be bypassed with fake messages, such as FIDO2, and to eliminate outdated systems that still allow simple attacks like repeated password attempts. In fact, attackers are increasingly exploiting access procedures considered “normal,” such as the use of codes displayed on a device or techniques where they insert themselves between the user and the service, to obtain valid authorizations without stealing passwords.

Then there is application authorization management, which must be clear and controlled: no open registration, monitored applications, verified permissions, and constant monitoring. Immediate session revocation, automatic rotation of access keys, and protections against device code-based spoofing are needed to block invisible access. Finally, there is the monitoring of identity-related threats: the only way to detect when a legitimate account behaves abnormally or when an authorized application is malicious. This is what enables the detection of hidden activities and abuses that are difficult to spot.

4) Risk-based patching: addressing vulnerabilities where it really matters

Not all vulnerabilities are the same: some are exploited immediately, while others will likely never be used. 2025 made this clear: the most serious issues hit exposed systems in a matter of moments. That's why relying solely on a technical score is no longer enough. Patches should be applied based on actual risk: if a vulnerability is already known and being exploited in attacks, has a high probability of being exploited, or affects an exposed system, it must be addressed immediately. Everything else can wait.

The most organized companies operate on two levels:

- a **fast track** for critical issues on exposed systems, with updates applied within a few hours;
- a **standard process** for internal or less urgent issues.

After the update, true effectiveness is achieved by completing remediation with critical actions: closing active sessions, changing credentials, updating keys, hardening configurations, and verifying whether there was any suspicious activity prior to the update. Without these steps, the patch risks providing only a false sense of security.

In 2026, updating doesn't just mean installing a patch, but stopping the attack, removing hidden access points, and restoring systems to a secure and controlled state. It's an ongoing process, not a simple repetitive task.

5) Continuous monitoring of cloud and SaaS

Today, data no longer resides “on-premises”: it’s in the cloud, in SaaS, and in external data flows. The idea of managing it once and then forgetting about it no longer works. A continuous program is needed—something that’s integrated into security processes every day. Truly managing the cloud and SaaS means always knowing who has access to what: applications with elevated privileges, external consents, public links, tokens and APIs that last too long, and hidden roles like shadow admins. It is in this context that SSPM (SaaS Security Posture Management) and CSPM (Cloud Security Posture Management) become indispensable: they provide constant visibility and help you keep the situation under control.

The rule is simple: fewer privileges, more segmentation. The more you limit what each application or identity can do, the less room you give a potential attacker. To protect data, you must also control what is transmitted externally: DLP solutions and egress control systems are essential to prevent specific information from ending up in external services, AI plugins, or SaaS applications used without proper governance.

This isn’t just technical common sense: it’s what regulations now require. NIS2 and DORA demand structured processes, vendor oversight, and continuous monitoring of critical dependencies. The goal isn’t to clean things up once and for all, but to immediately detect when something deviates from the expected security level.

6) Backups and Incident Response Take Center Stage

When ransomware strikes, it's not the encryption that's most frightening, it's the downtime. The only thing that determines whether a company gets back on its feet or remains crippled is an inseparable pair: robust backups + a ready-to-go incident response plan. A backup isn't just an archive—it's what allows you to get back up and running even when everything else fails. It works if it follows the 3-2-1-1-0 rule:

- 3 copies of the data,
- 2 different media,
- 1 off-site copy,
- 1 immutable or isolated copy,
- 0 errors in restore tests.

Alongside backups, you need a pre-written, tested, and internalized incident response plan: who does what, in what order, with which tools, how to communicate, how to isolate, how to restore. The difference between a company that's down for weeks and one that's back up in a few hours often comes down to this. To complete the picture, you need an MDR (Managed Detection and Response) service that monitors what's happening on hypervisors as well, and unified telemetry that connects identities, endpoints, the network, and the cloud to identify the attack chain as it unfolds.

7) Manage AI (before it manages you)

Today, artificial intelligence is a powerful accelerator: it strengthens defenses, but at the same time vastly increases attackers' capabilities. And that is precisely the point: if it isn't governed, it ends up governing us. It's not enough to "use" it: it must be kept under control. Techniques such as command spoofing, data or model poisoning, and the leakage of internal instructions clearly show that a system without rules is a real risk. Clear limits are needed: control what goes in and what comes out, continuously monitor what automated systems are doing, and define precise procedures for authorizing actions and external connections.

The situation becomes even more delicate with artificial intelligence acting autonomously through Agentic AI. These systems are no longer mere tools: they become an active part of business processes. They perform actions, make decisions, and automate tasks, often faster than humans. When a non-human agent operates with autonomous decision-making, continuous control and supervision mechanisms must be maintained. This is why it is essential to establish boundaries, objectives, and responsibilities, treating these systems as real components of the IT environment. Reference models such as AEGIS and standards like ISO/IEC 42001 (the first international standard specifically for artificial intelligence management systems) help define these rules, reducing new risks such as memory alteration and maintaining trust in intelligent systems.

8) Transparent and monitored supply chain

Today, the supply chain is so complex that trust alone is no longer enough. Too many suppliers, too much software, too many dependencies: without continuous monitoring, you're proceeding blindly. That's why it's essential to know exactly what's inside the software you use. The list of software components, known as the SBOM (Software Bill of Materials), is like an X-ray: it reveals obsolete libraries, critical dependencies, and known security issues, and is now also required in supplier management processes. When it comes to technology partners, the rules set forth by DORA come into play: maintaining a list of IT suppliers, having the right to audit them, assessing how dependent you are on a few entities, and regularly checking whether viable alternatives exist.

This isn't bureaucracy: it helps you understand how closely tied you really are to your service providers. Then there's the issue of data. If it's in the cloud or on critical online services, you must be able to control it even if it isn't physically yours. Customer-managed encryption keys serve precisely this purpose: to avoid forced dependencies, reduce unnecessary risks, and strengthen the entire supply chain. A secure supply chain is not based on promises, but on transparency, control, and constant verification. This is how you prevent an external problem from becoming your next incident.

Insights from the experts

There are topics that require analysis conducted with methodological rigor. That is why we have incorporated contributions from our specialists alongside external expertise of recognized national standing, ranging from Barbara Sabellico, a legal expert in technology regulations and speaker at TEDxModena, to Prof. Elena Prodi, a researcher at the San Raffaele University in Rome, as well as psychologist Enrico Lorenzi and neuroscience expert Andrea De Giorgio.

Each provides valuable insights.
All that remains is to “examine” them in the following pages.

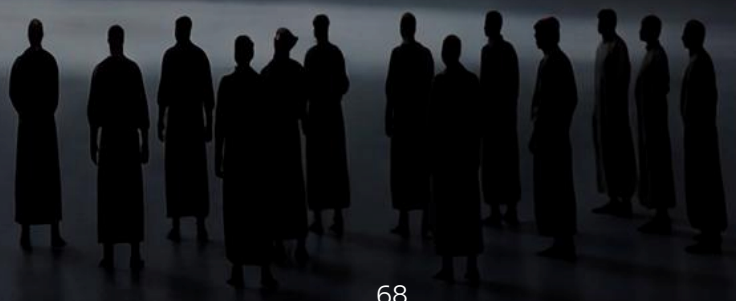
By Elena Prodi
Researcher and Lecturer, San Raffaele University, Rome

Cybersecurity Amid Geopolitical Tensions and Corporate Responsibility

Long neglected by businesses, institutions, and governments, cybersecurity is now receiving unprecedented attention. This shift is closely linked to growing geopolitical tensions, which are redefining global economic and trade balances and exposing businesses to new risks and higher levels of uncertainty. The increasing complexity and unpredictability of technological trajectories can, in fact, threaten business continuity on multiple fronts.

It is against this backdrop that the European regulatory framework on cybersecurity is being strengthened. Regulations and directives, such as those analyzed in the previous chapters, highlight the need to address a risk that increasingly stems from factors external to businesses: restrictions on technology exports, hybrid conflicts, state-sponsored cyber espionage campaigns, and vulnerabilities in global supply chains. These forms of international pressure demonstrate that cybersecurity is part of a broader process of adapting to geopolitical changes.

This new context requires companies to adopt a different mindset: it is no longer enough to protect one's own perimeter or adopt traditional technical measures.



Companies are increasingly required to consider the geopolitical nature of threats, assessing the exposure of their supply chains, their degree of dependence on strategic suppliers, and the potential impact of external political decisions.

For this reason, IT security can no longer be viewed as an incidental cost or a reactive response to incidents, but rather as an essential component of corporate strategies in an increasingly interconnected economic environment shaped by shifting international balances.

Neglecting cybersecurity means underestimating the strategic role that security and operational continuity must play as guiding principles for business action. Supply chain disruptions during the pandemic, energy crises, dependence on digital infrastructure, and instability in geographic areas crucial for the production of technological components have demonstrated how the ability to prevent and manage disruptions—including those caused by cyberattacks that reflect the climate of global instability, has become central to ensuring the day-to-day functioning of business operations. From this perspective, cybersecurity solutions become an integral part of operational resilience management, helping to preserve the stability of production and commercial processes.

The growing importance of advanced cybersecurity systems is part of a broader trend toward the expansion of advanced and digital services, which in recent years have taken on a structural role in the global economy. Even before the recent crises, international trade in services had surpassed that of goods, and within this sector, digitally delivered services, such as software, data analytics, and cybersecurity, have seen the most significant growth. Economic and geopolitical uncertainty, combined with trade tensions and the need to ensure operational continuity and stability in supply chain relationships, is further reinforcing this trend, prompting businesses and governments to invest in digital technologies and skills.

In Europe and the United States, industrial policies and incentive programs are fostering the development of these sectors, recognized as strategic levers for reducing vulnerabilities in value chains, bringing production and services closer to end markets, and supporting market diversification. In this context, cybersecurity is not merely a response to digital risks, but an essential component of the advanced services infrastructure upon which businesses' ability to operate, adapt, and compete in an increasingly complex and unstable economic environment depends.

By Andrea De Giorgio
Head of Keatrix Research & Development

Inside the Human Brain: overcoming Neurocognitive Limits to Reduce Cybersecurity Risk

The primary vulnerability underlying most cyber incidents does not lie in outdated software, weak credentials, or inadequate network configurations, but rather in an organ weighing about one and a half kilograms that consumes 20% of our basal metabolic rate.

That organ is the brain. It could be your own or that of your colleagues: any individual who, at a given moment, is processing an overload of stimuli, juggling communications between back-to-back tasks, multiple open windows, and a constant stream of incoming notifications. This is not negligence, but physiology.

If the 20th century was defined as the century of the heart, the 21st century is undoubtedly the century of the brain, because more and more disciplines recognize that cognitive and emotional processes represent the determining variable of complex behaviors.

The human factor is, in fact, involved in approximately 60% of data breaches (Verizon, 2025). Not because people are incapable, but because the human brain was not designed for cybersecurity. It was designed to ensure survival.

To understand why we click on a malicious link before thinking twice, we need to understand how the brain makes decisions. Daniel Kahneman, a Nobel laureate in economics, has described two fundamental modes: System 1, which is fast, intuitive, and automatic, and System 2, which is slow, analytical, and deliberative.

System 1 is what makes us slam on the brakes when a child crosses the street. It's also what makes us click "Verify your account now" when the email appears to come from our bank. Emotional circuits activate in milliseconds, before the prefrontal areas responsible for reasoning and control have time to intervene.

It's a perfect evolutionary shortcut for escaping a predator. Much less perfect for handling a phishing attempt at 5:30 p.m. on a Friday.

Here comes the uncomfortable part. The vast majority of employees who fall for an attack know that phishing, for example, exists. They've taken the course. They've passed the quiz. Yet, they click anyway.

The reason is that information alone is rarely enough to change behavior in a lasting way. If knowing were enough, no one would smoke, everyone would exercise, and no one would drive while looking at their phone.

From a neuroscientific perspective, the process that transforms knowledge into habitual action can be described in distinct stages: first, information is acquired (read, seen, heard) and then linked to something we already know (through examples, analogies, stories), so that the brain can encode it; after encoding it, the brain must work on it (apply it, test it, make mistakes, and correct itself); in other words, it must actively rework it to consolidate it into stable knowledge that it will be able to recall autonomously when needed. Only at this point can knowledge translate into behavior.

Many security awareness programs stop at the first two stages: they convey information and verify understanding. The result is that people “know” what they should do, but in the moment, they don’t do it. That knowledge isn’t yet ready to become behavior.

But there is an even earlier step that many training programs ignore entirely. Even before memorizing, the brain must decide that a stimulus deserves attention. If this threshold is not crossed, the learning process does not even begin.

This task falls to a distributed system that neuroscience calls the salience network.

This network functions as a continuous filter: among the thousands of stimuli that reach us every second, it selects those that are relevant (due to novelty, emotional significance, urgency, danger, or reward) and activates the cognitive resources needed to process them. Anything that does not pass through this filter simply does not enter the learning loop.

The reason why, while walking through a forest, we'll notice a snake on the ground before we even realize it is because our salience network (with our amygdala at the forefront) has recognized the threat and redirected our attention in a fraction of a second.

The problem is that a suspicious email is not a snake. It is not a stimulus for which the brain has a prepared biological response. No one is born with an instinct to protect themselves from phishing. This means that cybersecurity training faces a particularly daunting task: it must artificially construct a salience that biology does not provide, ensuring that the brain learns to treat an anomalous email with a level of attention similar to that reserved for a physical danger signal. And that is exactly why emotionally neutral content, such as slides, quizzes, and lists of rules, fails: it does not pass the salience filter; it is noise that the brain discards to save energy and time. To be effective, training must first and foremost get noticed. For the brain, this means being salient, emotionally relevant.

All of this happens precisely because “the brain is a cognitive miser; if it can’t derive value from something, it stops paying attention,” says Oren Klaff. In fact, it retains only what it deems useful, ignores what it perceives as redundant, and discards what isn’t linked to an emotion or personal meaning.

Hermann Ebbinghaus demonstrated this as early as 1885 with his famous forgetting curve: without reinforcement, we forget most information within hours or days. But there is one powerful exception: emotionally charged experiences. An event that frightens, surprises, or deeply moves us can remain etched in our memory for years.

It’s not magic; it’s neurobiology. When the amygdala is activated, it enhances memory consolidation processes in the hippocampus, strengthening the synaptic connections that make that memory lasting.

There is also a final structural problem: the one-to-many model, in which the same course is offered to hundreds of people, ignores a fundamental neuroscientific reality. Every brain is unique, shaped by its own experiences and learning history. Individual differences in attention span, working memory, and processing speed are well-documented and measurable, and they directly impact the effectiveness of any training program.

A CFO who handles sensitive financial data has a different risk profile than a production worker who accesses a shared terminal. Training them in the same way, with the same content and the same schedule, is like prescribing the same pair of glasses to all of an ophthalmologist's patients.

This is where adaptability comes into play, through technology. We're talking about systems capable of recognizing how a person learns best, which content captures their attention, which proves most effective, and which requires a change in approach—transforming training into a one-on-one process tailored to the individual's brain. In other words, technology becomes an enabler because it allows the training experience to be adapted to how each individual learns best.

All of this leads to a clear conclusion: cybersecurity, to generate real change, must align with neuroscience. It is not enough to simply add more information or increase the frequency of courses. We need to build behavioral automatisms.

And these habits are built through a precise mix: repetition spread out over time, varying contexts to generalize the behavior, emotional engagement to overcome the salience filter and activate consolidation, and adaptability to account for individual differences.

When a behavior is practiced long enough according to these principles, the brain gradually automates it: activity in the prefrontal areas, initially necessary to guide every decision, decreases, while the subcortical circuits associated with learned responses take on an increasingly important role.

It is a process analogous to the one that allows us to drive without thinking about every single movement: after thousands of repetitions in different contexts, the behavior becomes fluid and natural. Cybersecurity must aspire to this same naturalness.

The day an employee rejects a suspicious request without hesitation, reports an anomaly with the same naturalness with which they lock their front door, and checks a link before clicking without it costing them the slightest effort, that day we can say the training has worked. Not because it taught something, but because it changed the brain.

This is the contribution neuroscience can make to security awareness: transforming correct behavior from a conscious choice into an automatic response. Only then does the human factor cease to be a vulnerability and finally become part of the solution.

By Enrico Lorenzi
Psychologist and trainer

The Human Mind Under Attack: Psychology and Stress in Cyber Incidents

When discussing cyberattacks, particularly ransomware attacks, the focus tends to immediately center on financial losses, locked systems, and recovery procedures. However, it is crucial to recognize an often-overlooked truth: a cyberattack is, in every sense, an act of violence. This aggression, although directed at information systems, crashes down like a tsunami on the people working within the organization and, due to its violence, can in the worst cases amount to a true collective trauma. Companies are an ecosystem made up of resources, processes, and people (and of course much more) that coordinate with one another and influence each other.

Blocking these resources through an attack on the IT network or data literally triggers a crisis for the entire ecosystem, and those involved will experience it with great emotional turmoil. Caring for the human aspect, its psychological dimensions, during attacks is essential: it is precisely mental strength, clarity, group cohesion, and mutual trust in times of crisis that will effectively restore working conditions.

Normally, individuals navigate their daily lives (whether work or personal) following a mental framework based on stability, predictability, and security.

This framework constitutes a veritable “cognitive map” that allows us to function efficiently, reducing the brain’s energy expenditure. However, perceived safety is largely an adaptive illusion that works something like this: if yesterday was, all things considered, a normal day, and so was the day before that, and the day before that, then today will be much the same; I can rest easy.

The attacker, on the other hand, views the same scenario from the opposite perspective: a vulnerable territory, an ecosystem to exploit, a placid routine to shatter, everyone is asleep. When the attack strikes, production lines grind to a halt, computers are locked up, data becomes inaccessible, warehouses are blocked, customers are furious, data is stolen, and that cognitive map shatters, generating a perceptual and identity-shaking shock. The mind must reconstruct a new sense of reality in a suddenly hostile environment. Unprepared to do so and forced to contend with experts who are one step ahead. Not exactly a comfortable scenario.

Faced with the attack, the organization goes through well-defined emotional phases, comparable to those found in trauma processing:

1. **Denial (the initial shock):** The first reaction is a refusal to accept reality. People downplay the situation, look for a technical explanation, and try to “restart” their mental system even before the computer system.

2. **Anger and frustration:** when the inevitable becomes clear, psychological energy turns to anger. People look for someone to blame, creating interpersonal tensions and internal divisions. IT staff, in particular, may experience a strong sense of guilt and learned helplessness. Blame, stigma, and sarcasm combine into a highly disempowering mix.

3. **Bargaining:** the organization seeks cognitive and technical shortcuts, hoping for simple solutions to complex problems. It is a phase of pragmatic illusion, where internal pressure drives the need to “do something”, even in the absence of a coherent strategy. Sometimes this makes matters worse, putting other data or systems at risk.

4. **Depression and apathy:** when the extent of the damage becomes clear, a form of collective despondency sets in. A mechanism of psychological withdrawal is triggered, similar to that observed in acute post-traumatic stress disorder: cognitive slowing, loss of motivation, decision-making apathy, mistrust, a sense of helplessness, and disorientation.

5. **Acceptance (paradoxically, the riskiest phase):** when a recovery plan takes shape, a sense of relief emerges that, paradoxically, can reduce future vigilance. The illusion of “having learned the lesson” often fades over time, giving way to a new vulnerability, this time psychological and organizational: future probabilities are miscalculated, assuming it won’t happen again and that, even if it does, well, we know what to do. It is a new illusion.

Uncertainty about “if and when” the situation will end constitutes one of the most powerful stressors known to neuroscience. At the neurobiological level, prolonged activation of the hypothalamic-pituitary-adrenal (HPA) axis leads to a continuous release of cortisol, the chronic stress hormone. If this condition persists, it impairs the functions of the prefrontal cortex (the seat of rationality, problem-solving, and strategic planning) and amplifies the limbic response, dominated by fear, anger, and impulsivity.

In other words, uncertainty temporarily “disconnects” the rational brain from the emotional one. The result is systemic cognitive inefficiency: people struggle to make decisions, commit trivial errors, and lose the ability to assess priorities and risks. The entire organization enters a state of “cognitive fog”.

A cyberattack, like any traumatic event, leaves a lasting imprint not only on individuals but also on the collective identity of the organization. A “memory of trauma” is created that can resurface in the form of hypervigilance, distrust of systems, or, conversely, repression and denial. The collective psyche, like the individual one, oscillates between the desire for control and the need to forget. Only a process of organizational trauma processing, comprising psychological debriefing, ongoing training, transparent communication, and empathetic leadership can transform the wound into learning. To this must be added work on security culture by strengthening surveillance systems and transforming them into new routines.

The most mature response to an attack is not merely technical, but psychological and cultural. Resilient companies do not merely install more powerful firewalls, but train the collective mind to react to stress, tolerate uncertainty, and maintain trust and cooperation even in moments of chaos. Resilience, in this context, becomes a form of “psychological immunity,” capable of containing damage and facilitating cognitive and operational recovery. A cybersecurity plan that does not include a psychological approach to crisis is an incomplete plan. Like any trauma, digital trauma leaves scars: recognizing, understanding, and transforming them is the first step toward building a truly human security culture.

By Andrea Coli
Incident Response Manager, Cyberoo

The Paradox of Inertia: When the Certainty of an Attack Does Not Lead to Resilience

It has become an axiom in the cybersecurity sector that “it’s not a question of if, but when”. For those working in incident response (IR), this statement is not trivial or rhetorical, but rather the premise for disastrous operational scenarios. Despite widespread awareness of the inevitability of a cyber incident, this certainty does not systematically translate into concrete decisions aimed at improving security posture or increasing organizational resilience.

Real-world experience reveals just how difficult it is for IT managers to secure the necessary support for investments in risk prevention and mitigation. Requests for time and budget for fundamental activities such as system hardening, network segmentation, periodic patching, and the implementation of essential configurations and controls regularly face delaying responses such as “let’s put it in next year’s budget” or superficial underestimations like “we already have antivirus software; that’s more than enough”.

This resistance is not merely financial but is rooted in a strong resistance to cultural change. Organizational inertia persists even in more stringent regulatory contexts, such as NIS2, which mandates legally required measures and processes. Too many companies that are not digital by design struggle to grasp the logic of proactive security.

For many organizations, particularly businesses and government agencies, cybersecurity is still viewed as a “technical issue” rather than a strategic matter tied to business continuity.

Cybersecurity policies are not a technical whim but represent the first line of defense against potentially devastating impacts. Security is costly and ineffective when conceived as merely the sum of technologies to be installed, rather than as an organizational model, a strategy, and a process of continuous improvement.

In today’s landscape, attack scenarios have become significantly more sophisticated. Attack vectors have evolved, including increasingly sophisticated phishing, targeted attacks against the supply chain, and the exploitation of zero-day vulnerabilities within a matter of hours. We are witnessing a time gap of over twenty years between the evolution of threats and the defenses adopted by many organizations, which remain anchored to the idea that a simple, up-to-date antivirus and a firewall are sufficient to ensure protection.

As an Incident Response Manager, I have witnessed consequences of despair and regret that far exceed the direct financial damage. A successful ransomware attack can result in weeks of downtime, the loss of thirty years’ worth of documents, penalties resulting from the failure to provide services or supplies, severe reputational damage, GDPR violation fines, and, in extreme cases, the permanent closure of the business.

Last but not least, cyber incidents can have a direct impact on people, with employees being furloughed and, at the management level, executives facing intense pressure due to the stress caused by a complete halt in production.

A crucial, often underestimated element is the management of data breaches. Seeing specific employee data (ID cards, passports, medical records of family members) published and made publicly available poses a significant risk to individuals and requires an immediate shift in approach to the structuring of cybersecurity processes.

During the most critical phases of managing a cyber incident, management clearly understands which actions are truly indispensable: allocating the budget to eliminate obsolete systems, adopting continuous threat monitoring and response services, introducing more robust identity verification methods—at least for remote access and email—launching targeted training programs, using advanced endpoint control tools, and defining stricter security rules.

However, the problem persists: once business operations resume and the immediate “crisis” is over, there is often a return to the status quo. In many cases, the attack, the disruption of operations, and the economic and reputational damage are not enough to trigger lasting change. The correct strategic reasoning should not focus on how much it costs to protect oneself, but rather on how much it would cost to shut down again.

The literature is full of best practices, and the necessary processes exist. The difference between a company that survives an attack and one that succumbs often lies in the proper management of data resilience. While backups were not standard practice as recently as ten years ago, awareness of the issue is greater today, even though the technical and procedural solutions are sometimes questionable.

A complete backup is a lifeline for the company.

In managed ransomware attacks, companies with comprehensive and tested backups achieved a clear and relatively rapid recovery timeline. Conversely, those with compromised backups faced total darkness: a lack of essential data and services from which to recover.

Even if there is no budget for advanced solutions, such as immutable backups, or if management is not yet convinced of the need to invest, it is imperative to establish a rigorous internal procedure:

1. **define** realistic **RPO (Recovery Point Objective) and RTO (Recovery Time Objective)** that align with business needs;

2. **Store backups offline**, or at least ensure their logical and physical separation from the production environment;

3. **Periodically test** recovery procedures;

4. **Simulate attack scenarios** and ask the critical question: “Would the backups survive this scenario?”

When, despite postponed requests and denied budgets, operations come to a halt, it will be the backups that save the company and everything that depends on it. It is essential to adopt processes and solutions that protect data, work, and, above all, people. The same awareness now gained regarding the necessity of backups must be extended to cybersecurity, understood as a structured, effective, and efficient model.

By Matteo Ghiotto
Chief Technology Officer, Cyberoo

AI Agents & Model Poisoning: The Invisible War Over the Data That Fuels AI

By 2026, AI is no longer a laboratory experiment: Large Language Models (LLMs) have become the beating heart of companies, with access to specific data and the ability to act autonomously. But this evolution brings with it an uncomfortable truth: security is no longer just about code, but about the integrity of the data streams that feed AI Agents. The shift from passive chatbots to autonomous agents has opened up a significant attack surface, where vulnerabilities cannot be resolved with a simple patch. To understand where the danger lies, let's look at the OODA (Observe, Orient, Decide, Act) framework, used to analyze decision-making processes:

- **Observe**: they gather information from the web and internal databases, which may contain sensor spoofing techniques or invisible textual artifacts designed to alter the system's perception;
- **Orient**: the model's "worldview" can be corrupted by training data poisoning or semantic backdoors inserted months earlier;
- **Decide**: in agent-based systems with feedback or learning mechanisms, decision-making logic can be distorted through reward hacking, turning the decision-making process into an attack vector;
- **Act**: this is where the "blast radius" is unleashed: calls to malicious tools or data exfiltration disguised as legitimate operations.

The discovery of 2025? Experimental studies have shown that even a few hundred malicious documents, under specific training or fine-tuning conditions, can be enough to implant a backdoor, regardless of the model's scale. Once "frozen," the backdoor can remain dormant for years, activating only when it encounters the right trigger phrase.

Agents based on Retrieval-Augmented Generation (RAG) open up new attack surfaces, particularly through context and knowledge base poisoning, which operates silently and is extremely difficult to detect. One of the most insidious techniques is Indirect Prompt Injection (IPI): the attacker does not tamper with the AI's prompt, but instead poisons the sources from which the system draws information, such as documents, emails, or databases. The AI retrieves data that appears harmless but contains hidden malicious instructions.

Complicating the picture are frameworks like RIPRAG, which demonstrate how effective it is to optimize this poisoned content using Reinforcement Learning from Black-box Feedback (RLBF). The result? Just a handful of well-designed documents, even amidst millions of legitimate files, are enough to significantly influence the AI's behavior.

That's not all. There's the issue of web agent fingerprinting: many have recognizable digital fingerprints, due to automation frameworks or browsing patterns.

Malicious sites exploit this feature to employ cloaking techniques: they display a harmless page to human users, while sending hidden instructions to the AI agent that can prompt it to invoke downstream tools or workflows, leading to the execution of unauthorized actions.

The risk escalates in multi-agent systems. This is where persistent attacks come into play, such as those analyzed by the BackdoorAgent framework, which can nest within the agents' memory or tools and propagate over time. Furthermore, in multi-agent systems, agent-to-agent (A2A) communications and shared resources can create a domino effect capable of compromising the entire ecosystem. Today, the security of AI agents can no longer follow the rules of traditional models. Simply controlling content is not enough: these agents operate in dynamic environments, access external tools, and make decisions autonomously. We need Zero Trust architectures designed specifically for AI, based on an identity-first approach: management of non-human identities, separation of roles within agents, and distributed protection throughout the entire processing cycle.

Finally, another crucial point: semantic integrity. It is not enough for information to be correct; what matters is how it is interpreted by AI and what actions it can generate. In 2026, the real challenge is to build architectures capable of reliably distinguishing between information and instructions, thereby limiting the scope of AI. Only in this way can we transform autonomy from a potential risk into a lever for trust and scalability.

By Barbara Sabellico
Attorney & DPO, Legal Tech Expert

The AI Act and the Protection of Digital Rights: New Compliance Challenges

In European regulation of artificial intelligence, the AI Act represents one of the fundamental pillars for the protection of digital rights. Its goal is to balance technological innovation with the protection of privacy, non-discrimination, and citizen safety. From my perspective as a lawyer specializing in digital rights, with a focus on the GDPR, the AI Act, and cybersecurity, it is now essential to guide companies toward a proactive compliance approach capable of mitigating not only legal risks but also, and above all, the reputational risks they may face.

The gradual entry into force of the AI Act, with prohibitions taking effect in February 2025 and obligations for high-risk systems starting in August 2026, requires organizations to take action immediately. This is also because the penalties can reach up to 6% of global turnover. The regulation adopts a risk-proportionate approach, prohibiting practices deemed unacceptable, such as real-time biometric recognition in public spaces or subliminal manipulation, which directly impact privacy and human autonomy (Art. 5–6).

Particular attention is paid to systems classified as high-risk, a category that also includes various artificial intelligence applications in the field of cybersecurity, such as in surveillance systems, critical infrastructure management, or credit scoring.

In such cases, the AI Act imposes strict requirements regarding transparency, robustness, and human oversight. In practice, this means designing secure systems from the development phase onward, capable of withstanding adversarial attacks, data poisoning, and model inversion techniques. All of this takes place within the framework of the rights guaranteed by the Charter of Fundamental Rights of the European Union (Articles 8 and 21) and in constant coordination with the GDPR, particularly to reduce the risk of discriminatory bias and high-impact data processing.

The AI Act, however, cannot be interpreted in isolation. Its regulatory framework is necessarily intertwined with the NIS2 Directive, which covers essential sectors such as energy, healthcare, and finance, and with the DORA Regulation for the financial sector. These instruments introduce obligations regarding ICT risk management and incident reporting, with different timeframes: 24 hours in the case of NIS2 and up to 15 days for the AI Act. This necessitates that companies harmonize their compliance processes, avoiding overlaps or, worse, gaps in accountability.

In the case of artificial intelligence systems applied to cybersecurity, the responsibility of senior management, already provided for in Article 5 of NIS2, is also of particular importance. This responsibility now extends to accountability for any violations of digital rights. A defective or inadequately supervised AI model, for example in threat detection systems, can generate false positives with concrete impacts on the rights of defense or the privacy of data subjects.

In this regard, robust compliance also helps strengthen an organization's legal resilience by reducing its exposure to class-action lawsuits under the Directive on Representative Actions and to investigations by supervisory authorities, such as the Data Protection Authority.

From an operational standpoint, the AI Act introduces very concrete requirements. System robustness must include the ability to withstand evasion attacks, in line with the requirements of Article 32 of the GDPR. Transparency, on the other hand, entails the need to maintain auditable logs and traceability throughout the entire system lifecycle, while also ensuring the right to an explanation of automated decisions. Added to this is the obligation to report serious incidents within 15 days and the option to adhere to sector-specific codes of conduct, which provide a presumption of compliance. Taken together, these elements outline a structured and clear framework, particularly relevant for top management.

In such a context, it is important to ask what concrete actions can be taken. For CEOs and COOs, a first step is to establish a Data Protection Board with oversight functions over artificial intelligence systems, accompanied by the definition of ethical policies consistent with the guidelines of the EDPB (European Data Protection Board). It is also essential to carefully assess risk profiles, including those of a criminal nature, especially when dealing with health or financial data.

From a technical standpoint, CTOs and IT managers are required to properly classify internally used AI systems, identifying high-risk ones, particularly when artificial intelligence is integrated into security tools. The adoption of techniques such as federated learning or differential privacy can significantly help reduce data exposure and improve overall compliance.

The role of the CISO also becomes central. They are called upon to act proactively, conducting AI-specific DPIA, testing for vulnerabilities according to the frameworks defined by ENISA, and preparing incident response plans tailored to attack scenarios amplified by artificial intelligence.

To complete this picture, it is essential to provide mandatory training on digital rights for all parties involved in the use of intelligent systems, accompanied by periodic audits and gap analyses. Simulations of AI cyber scenarios and voluntary certifications can also become a tangible competitive advantage.

The AI Act should not be viewed merely as another regulatory requirement. If approached with a methodical and strategic vision, it can become a real opportunity to establish ethical leadership in the field of digital rights and strengthen the trust of customers, partners, and stakeholders.

By Luca Benatti
Product Manager, Cyberoo

From Regulatory Chaos to Super-Compliance: The True Leap Forward

2026 is not simply a new chapter in European cybersecurity; it is the moment when many organizations will discover whether what they have built over the past few years truly stands up to scrutiny. After two years experienced as a long “preparation” phase—filled with policies, documents, and initial formal adjustments—we are now entering the most challenging phase: that of concrete implementation, active monitoring, and operational responsibilities.

This is where the limitations of a fragmented approach to compliance become starkly apparent, where each regulation is managed as an isolated entity. NIS2, the AI Act, the CRA, and DORA are not separate issues: they are different facets of the same underlying requirement—namely, demonstrating control, rapid decision-making, and the ability to respond in an orderly manner under pressure. Continuing to treat them as separate checklists means accumulating costs, creating inconsistencies, and, above all, increasing the risk of operational collapse at the worst possible moment: when something actually happens.

In the European context where cybersecurity investment levels and incident exposure remains uneven across Member States, entering 2026 with fragmented compliance frameworks is not only inefficient: it becomes a multiplier of operational risk.

In fact, 2026 marks the transition from “declared” safety to verifiable safety, where it is no longer enough to have a plan; one must demonstrate the ability to execute it, even within a few hours and with incomplete information. It is in this context that the concept of super-compliance takes shape: no longer a mere sum of obligations, but a unified model that brings together processes, roles, languages, and decision-making mechanisms.

Effective super-compliance does not add bureaucracy; it reduces it, because it eliminates duplication and forces the organization to clarify who decides what, in what timeframe, and based on what evidence. Incident reporting, risk management, supplier oversight, and governance of critical technologies and new attack surfaces must speak the same language; otherwise, they remain disconnected pieces.

Without this integration, the risk is not only operational but strategic: many organizations find themselves subject to technological decisions imposed by large non-EU vendors, with infrastructure increasingly beyond their control and resources drained outward just when they are needed to strengthen security, resilience, and internal decision-making capacity.

Artificial intelligence now also fits into this framework: not as a “futuristic” concept, but as a concrete element of business processes, one that introduces new risks if not governed by clear criteria. Standards such as ISO/IEC 42001 should be viewed precisely in this light—not as yet another certification, but as an attempt to bring AI back within the established frameworks of management, accountability, and control.

As Agatha Christie’s character Hercule Poirot reminds us: “Method, order, and little gray cells: that is all one needs. One must never guess; it is a shocking and destructive habit for the faculties of logic”.

Ultimately, the message of 2026 is quite simple and for this reason often ignored: it is not the one who checks off the most requirements who wins, but the one who manages to build a coherent system, capable of functioning even when the context becomes unstable.

Super-compliance is this step toward maturity: transforming compliance from a defensive exercise into an organizational capability, because in the new European landscape, the real difference won’t be made by avoiding a penalty, but by managing to keep operating while everything around you accelerates.

By Vasily Kononov
Threat Intelligence Lead, Cyberoo

The Dark Web in 2026: The New Race Against Time in Cybersecurity

The Dark Web is only a tiny fraction of the Internet, but its impact is growing every day. In 2025, the Tor network surpassed 3 million daily users: a sign that this “shadow zone” is by no means marginal. Even when law enforcement dismantles criminal marketplaces, the system quickly regenerates itself. In this context, the real challenge is no longer preventing every attack, but how quickly we can bounce back after a blow.

Today, the Dark Web is no longer a chaotic hub of illegal activity: it has become an organized ecosystem, with internal rules, specialized roles, and even “professionals” in digital crime. Among these, cyber mercenaries stand out—small teams operating in closed, ultra-selective forums where reputation is everything and every transaction is airtight. These environments adopt mechanisms typical of regulated markets: mandatory escrow systems, active moderation, dispute arbitration, and strict rules that ensure operational stability and trust among participants.

But the real revolution is the arrival of autonomous AI agents designed to attack. We’re not talking about simple chatbots: these are systems that manage entire offensive campaigns, from reconnaissance to data exfiltration. The result? Attacks that once took days are now wrapped up in minutes. The machines race ahead, while defenses often remain slow and human.

This time gap marks the definitive shift away from prevention as the sole objective: operational resilience becomes the new metric of cyber maturity.

In the most discreet markets, initial access, zero-day exploits, and “custom” services are purchased, such as cracking administrative hashes to gain total control of an infrastructure. When a group gains access to a compromised server but fails to crack the privileged account hash, it posts a request on the forum offering a reward for its decryption.

The figures vary enormously: from small sums of $50–$200 up to offers of thousands of dollars for particularly complex hashes. Cracking on commission is entrusted to specialists equipped with multi-GPU workstations and optimized configurations, a highly technical niche that allows for transforming partial access into full control of compromised systems.

Demand is also growing for silent attacks—without ransomware—aimed solely at stealing data without leaving a trace. The real value today lies in invisible persistence, not in making a splash. And as ransomware continues to evolve, digital trust is faltering: deepfakes, advanced phishing, and hyper-targeted campaigns exploit psychology more than technology, bypassing even MFA. Identity is becoming the new security perimeter.

Ransomware-as-a-Service models are increasingly integrating with initial access brokers, while deepfakes and sophisticated phishing attacks are accelerating the crisis of digital trust, rendering many traditional controls ineffective.

What does all this mean? That the Dark Web is now a “regulated” criminal enterprise, with capabilities similar to those of APTs (Advanced Persistent Threats). No company is too small to be a target. Updated risk models, end-to-end visibility, privileged access controls, and automated responses are needed. It is no longer about preventing every intrusion, but about restoring clean systems before the attacker adapts.

Today more than ever, Dark Web-focused cyber threat intelligence and automated response are becoming essential to compete with an adversary that operates non-stop and at machine speed.

By Luca Bonora
Cybersecurity Evangelist, Cyberoo

From Data to Consciousness: The Future of Humanity Between Empowerment and Responsibility

By simultaneously considering the context of emerging technologies, advanced artificial intelligence, and human-machine integration, we can strongly argue that, in this historical period, cybersecurity lies at the intersection of a new technical and philosophical frontier where security no longer concerns only data, but the very integrity of human consciousness and digital identity. In other words, it is the intersection between transhumanism (the use of technology to overcome human limitations) and panpsychism (consciousness as a pervasive property of matter) that leads us toward a new understanding of cybersecurity itself.

In public discourse, cybersecurity is often perceived as a set of technical countermeasures. In reality, it is a socio-technical discipline: it combines technology, processes, roles, responsibilities, and governance. Today we are witnessing a significant shift that primarily alters the object of protection: not just data and infrastructure, but aspects touching on identity, decision-making capacity, and perception, extending to consciousness itself. To understand this shift, it may be helpful to refer to the theory of transhumanism. In this practical philosophical approach, AI, neurotechnologies, and human-machine integration are driving toward a model in which the human being appears increasingly measurable, modifiable, and optimizable, to the point of leading us to believe that one day, not too far in the future, AI will equal and surpass human intelligence.

The idea of a “snapshot” of the mind, which implies that if the mind were merely information, it could be copied and transferred, represents a form of extreme materialism that, regardless of its actual feasibility, influences the collective imagination and decisions regarding the development of technological products.

This opens up three scenarios, which are not mutually exclusive, and are particularly important in relation to risk management issues:

- **positive scenario**: AI and biomedicine increase autonomy and the capacity for care;
- **competitive scenario**: pressure to enhance capabilities for economic and military advantage, leading to new inequalities;
- **systemic risk scenario**: very powerful AI that is not well governed and/or a hyper-surveilled society.

At this point, the link between freedom, consent, and manipulation becomes central, if not critical, especially for those observing cyber and organizational impacts. It is necessary to understand whether the enhancement so coveted by transhumanism remains an individual choice or an implicit obligation to compete, work, and access services.

AI enables profiling and propaganda. With technologies that interact with the brain (neurotech), the risk of control and targeted influence increases further, as it directly targets cognitive and emotional states.

The crucial element is what distinguishes humans from machines: consciousness and free will. It is precisely free will that concretely allows us to understand that freedom is not merely the absence of coercion, but also the ability to choose without opaque manipulation.

Quantum physics' critique of transhumanism thus becomes relevant as a precautionary principle: reducing consciousness to computation can lead to treating the person as an upgradeable device, a conceptual error that can become an ethical error and, in turn, a security error (misplaced priorities, insufficient safeguards, opaque accountability). If the human becomes the asset, the risks are not limited to data exfiltration but extend to:

- **compromise of integrity** involving the manipulation of signals and decisions, not just data;
- **abuses of informational power** caused by extreme profiling, which can lead to discrimination and blackmail;
- **dependence on complex supply chains** involving devices, firmware, the cloud, updates, and third parties.

The picture expands further with security at the intersection of biology and digital technology, known as cyberbiosecurity, and with neuroscience applied to information technology, where biology and digital technology intertwine to the point of making biology itself a surface of attack, and where brain-computer interfaces (BCIs) can enable new forms of communication and direct control of devices.

This introduces specific risks such as violations of mental privacy, vulnerability to cyberattacks, addiction, and social control.

Furthermore, the so-called neuro-digital divide is growing, that is, the gap between protecting the mind as a privilege for those who can afford safer technologies and others, who are exposed to invasive, obsolete, or poorly regulated solutions.

In a world accelerating toward human-machine fusion, the choice is not between enthusiasm and rejection, between transhumanism and panpsychism, but between governed adoption and blind adoption. The dilemma between symbiosis and preservation becomes a matter of security. We need to understand scenarios, risks, and strategies to protect augmented identities and emerging cognitive systems by analyzing how the technological extension of the self and the potential for distributed consciousness influence digital security, privacy, and ethical responsibility.

As the boundary between neurons and bits blurs, protecting the mind means protecting freedom. This requires rules of consent, control, accountability, and verifiability commensurate with the new risks. Free will, in this context, is not a mere philosophical ornament but represents the good to be defended, around which to redefine priorities, controls, and governance, in essence, cybersecurity.

Conclusion

2026 is not a year to fear: it is the year in which we learn to observe what previously eluded us, to read the hidden part of our digital ecosystem—its dark matter. That is where most of the risk lies: forgotten or poorly managed identities, redundant data, unmapped processes, tokens left hanging, exposures that no dashboard shows but that influence every decision. The future of cybersecurity does not hinge on what we see clearly, but on what we are finally able to make visible.

The transformation we are experiencing requires a profound shift in mindset: moving from fragmentation to coherence, from the accumulation of isolated tools to integrated services and governance models, from slow processes to rapid, traceable decisions.

The organizations that will thrive will not be those with the most solutions, but those capable of building a culture and viewing their digital ecosystem as a living system: governed identities, mapped exposure surfaces, continuously monitored cloud and SaaS, assessed and monitored supply chains, remediation within hours for critical vulnerabilities, ready and tested playbooks, and MDR as the security operations center.

The dark matter of cyberspace teaches us that risk is not what we see, but what we do not yet understand. Interpreting it has now become a strategic skill: it means distinguishing the signal from the noise, anticipating the spread of a vulnerability before it becomes an incident, recognizing a valid identity that is behaving abnormally, and detecting the subtle patterns that precede an escalation.

2026 marks the end of the perimeter illusion and the beginning of the era of intelligent adaptation. It is the year when security ceases to be a barrier and becomes an organizational capability: infrastructure that reacts, identities protected before they are exploited, systems capable of sustaining operational continuity even under attack, AI models that are managed and not merely adopted. We can no longer defend the present with tools from the past: we need living processes, capable of learning, correlating, and anticipating.

The future of cybersecurity is already here. And it is not a dark future: it is a future just waiting to be illuminated with the right mindset and the right choices. A future that must be seen for what it is, without shortcuts or easy illusions, because the invisible forces governing digital systems act regardless: we cannot stop them, but we can understand them and use them to our advantage, before they drag us out of orbit.

We are Cyberoo

We are a European company specializing in cybercrime defense, with an advanced cybersecurity model that combines cutting-edge technology, proven processes, and highly specialized human expertise. Our mission is clear: to prevent cyberattacks and ensure business continuity even when the environment becomes uncertain.

We are one of the few European vendors that has been recognized for three consecutive years as a “Representative Vendor” in the Gartner® Market Guide for Managed Detection and Response Services, an achievement that confirms the strength of our approach and our ability to deliver a service that meets the highest international standards.

Our I-SOC consists of over 100 specialists across four tiers, with a dedicated Incident Response team. We operate as an extension of companies’ security functions, transforming trillions of events into a few rapid, targeted actions. In addition to our operational work, each client is supported by a Security Advisor Manager (SAM), who guides cybersecurity strategy and continuous improvement based on data, evidence, and real-world risk.

Our technological independence allows us to be completely agnostic and focused on client needs. Being present and rooted in Europe, we guarantee proximity, stability, and operational continuity in a geopolitical context that increasingly demands control and digital sovereignty.

Bibliography

CISA (2025). Known Exploited Vulnerabilities Catalog (KEV), Cybersecurity and Infrastructure Security Agency

CISA ICS (2025). ICS Advisories and Alerts

ENISA (2025). Threat Landscape Report

ENISA (2025). NIS2 Technical Implementation Guidance

European Commission (2024). The AI Act Explorer

European Parliament (2023). Cyber Resilience Act – Legislative text

European Banking Authority (EBA) (2025). DORA Technical Standards

Gartner (2025). Gartner Survey Reveals Geopolitics Will Drive 61% of CIOs and IT Leaders in Western Europe to Increase Reliance on Local Cloud Providers

IEC (2024). IEC 62443 Standard for Industrial Cybersecurity.

NIST (2024). Post-Quantum Cryptography Standardization Process

OpenAI (2025). Disrupting malicious uses of AI

Sysdig (2024). LLMjacking: Stolen Cloud Credentials Used in New AI Attack

“Illuminate what is dark within me.”

John Milton, Paradise Lost