sopra steria next

Hack to the Future:

From Cyber Defense to Cognitive and Operational Resilience

When time-to-exploit is measured in days and intelligence becomes programmable, cybersecurity evolves into a discipline of speed, control and trust.

CIO Compass Series

Executive *Summary*

Cybersecurity is part of a context of ongoing state confrontation, illustrated both by the war in Ukraine and by recent clashes between the United States, Israel and Iran, where cyberattacks accompanied strikes in order to disrupt command and defence systems.

Cyber operations, which are hybrid and continuous, now target operational resilience as much as confidence and decision-making quality, through disinformation, the compromise of supply chains and the exploitation of intelligent systems.

In this context, European regulations (NIS2, DORA, AI Act) reflect a requirement for demonstrable resilience under constant pressure, while dependence on global platforms reinforces sovereignty issues.

Alongside this tense situation, cybersecurity is undergoing a profound transformation. On the one hand, the speed of attacks has collapsed. Critical vulnerabilities are now exploited in a matter of days, sometimes hours, leaving little room for delayed reactions or calendar-driven remediation. On the other hand, enterprise information systems are becoming increasingly intelligent. Through APIs, cloud platforms, data services and AI agents, they no longer simply execute predefined processes: they interpret context, recommend actions and, in some cases, act autonomously.

> *Cybersecurity must now achieve three objectives at the same time:* **protect an increasingly fragmented and exposed digital estate, ensure operational continuity under continuous pressure, and secure the behaviour of intelligent systems powered by Large Language Models (LLMs) and agentic AI."**

In this new reality, traditional cyber strategies (largely built around perimeter protection, periodic controls, and human-only response) are no longer sufficient. Cybersecurity must now achieve three objectives at the same time: protect an increasingly fragmented and exposed digital estate, ensure operational continuity under continuous pressure, and secure the behaviour of intelligent systems powered by Large Language Models (LLMs) and agentic AI.

This evolution marks a decisive shift. Cybersecurity is no longer a defensive function operating in the background. It becomes a core operating capability, anchoring cyber and cognitive resilience by default.

What must now be made explicit

- **If patching remains calendar-driven, organisations are structurally late.**
- **If LLM usage is not observable, Shadow AI already exists.**
- **If autonomous systems can act without approval thresholds, resilience is assumed and not proven.**

The cyber arbitrations CIOs *can no longer avoid*

Behind every cyber decision now lie four unavoidable tensions that CIOs must arbitrate explicitly:

→ Speed *VS* Control.

Time-to-exploit is measured in days, sometimes hours. Governance models must evolve to protect at this speed without paralysing delivery.

→ Autonomy *VS* Accountability.

Agentic AI accelerates detection and response, but accountability must always remain human and explicit.

→ Innovation *VS* Sovereignty.

LLMs unlock rapid value creation while introducing new dependencies. Architecture, not policy, is the only sustainable lever to preserve choice and control.

→ Coverage *VS* Cost.

Operate-by-default cannot mean tool sprawl. Resilience requires prioritisation, not accumulation.

These arbitrations frame all that follows.

A splintered digital estate *under continuous pressure*

The enterprise IT landscape is no longer a coherent, centralised system. It is a distributed ecosystem combining SaaS platforms, hybrid and multi-cloud infrastructures, APIs and data services, complex software supply chains, operational technologies and, increasingly, AI-driven decision layers.

Field evidence across industries is remarkably consistent. Identity compromise remains the primary entry point for attackers. APIs and third-party dependencies are among the fastest-growing attack vectors. And the average time-to-exploit for critical vulnerabilities is now counted in days.

In such an environment, speed and coverage of basic controls are no longer hygiene factors. They have become board-level performance indicators. A delayed patch, an unmanaged API or an unmonitored identity flow is no longer a technical issue; it is an operational risk with direct business consequences.

Cybersecurity: Making AI a strategic ally

The cyber threat landscape is entering a new phase. Generative AI is lowering the barrier to entry for attackers, enabling deepfakes, automated phishing, and more sophisticated malware. Hybrid infrastructures and complex ecosystems further expand the attack surface.

But the same technologies can radically strengthen cyber defense. AI and agentic systems make it possible to analyze massive volumes of signals, detect anomalies in real time, automate threat response, and continuously test system resilience.

At Sopra Steria, we hold three key convictions:

1. **AI must become a strategic ally for cyber defense**
   AI-powered detection, automation, and predictive capabilities are essential to operate securely at scale.

2. **Security must be embedded by design**
   Cybersecurity, governance, and ethical safeguards must be integrated from the earliest stages of AI development and deployment.

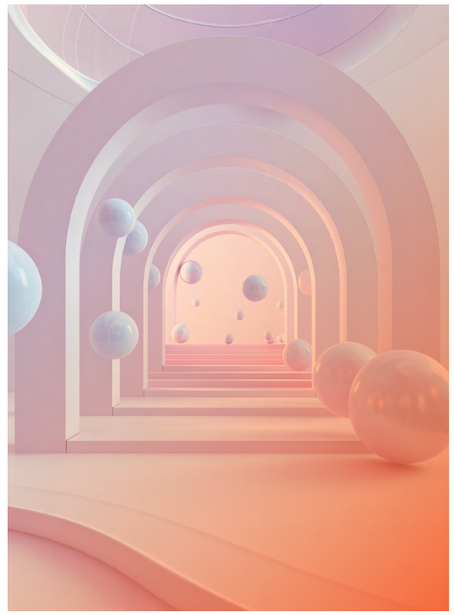
3. **Human expertise remains central**
   AI amplifies security teams but does not replace them. The combination of human judgment and intelligent automation is what creates resilient systems.

Cybersecurity is no longer only about protection.

It is about **building trusted digital ecosystems where innovation can scale securely.**

A threat model that has *fundamentally shifted*

The nature of cyber threats has evolved as much as their speed. Attackers no longer rely primarily on sophisticated zero-day exploits. Instead, they favour scalable abuse: stolen identities, legitimate remote administration tools, exposed APIs and upstream software components. The effort required is low, the impact potentially massive.



At the same time, compromise increasingly occurs before systems even reach production. Open-source dependencies, CI/CD pipelines and SaaS integrations have become preferred targets. When organisations lose control over their software supply chain, they lose control over risk itself.

With the rise of LLMs and agentic AI, a new dimension is added. Attacks no longer focus solely on breaking systems, but on manipulating behaviour. Through prompt injection, indirect instructions hidden in data or subtle interference with agentic workflows, systems continue to run, yet their decisions can no longer be trusted.

> **“With the rise of LLMs and agentic AI, a new dimension is added. Attacks no longer focus solely on breaking systems, *but on manipulating behaviour.* ”**

LLMs and the emergence *of cognitive risk*

Large Language Models introduce a fundamentally new category of cyber risk. They interpret context, generate content, recommend actions and, increasingly, trigger execution through agents. In many cases, there is no data breach and no system outage. Yet business outcomes are compromised: a financial insight is hallucinated, a compliance rule is misinterpreted, or an HR recommendation embeds unintended bias.

This is the essence of cognitive risk. Its cost is not measured in downtime, but in eroded trust, flawed decisions and reputational damage. A single well-placed hallucination in a sensitive process can outweigh the impact of a traditional infrastructure incident. Agentic AI amplifies this risk further. By connecting reasoning to action, it transforms local errors into systemic failures. Cybersecurity must therefore evolve to protect not only systems and data, but decision-making itself.

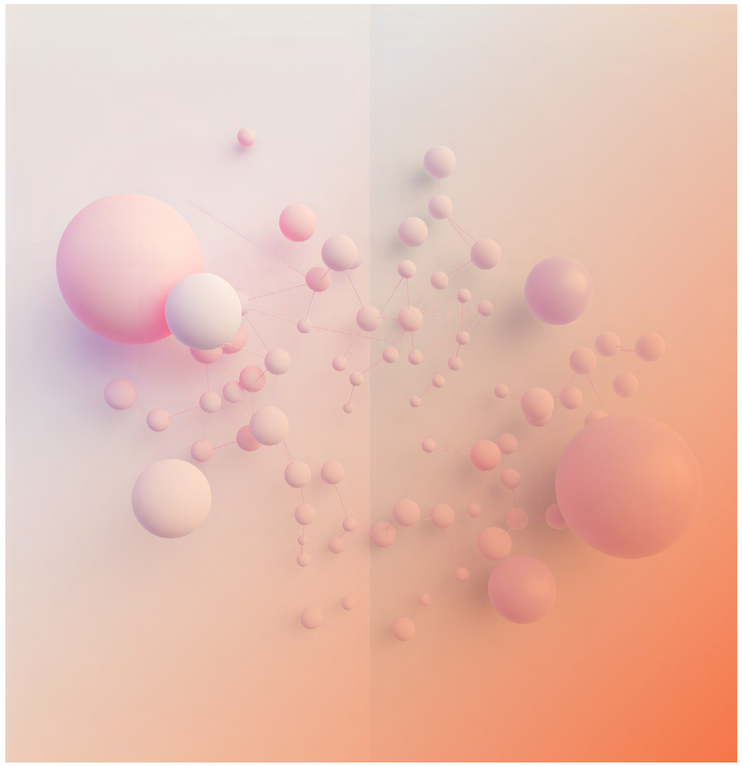
> **The 'operate by default' approach, decision security (cognitive risk) and controlled autonomy via agentic AI should be positioned not as maturity options, but *as a strategic posture of defence and continuity in an environment of sustained conflict."***

From "Secure by Design" *to "Operate by Default"*

Most cyber incidents do not occur at design time. They occur in production, under pressure, across live systems. Resilient organisations recognise this reality and adopt an operate-by-default posture. Controls are no longer assumed to be effective; they are continuously monitored. Gaps are not discovered during audits; they are detected in real time. Accountability is not implicit; it is assigned, measured and enforced through clear Service Level Objective cognitives (SLOs). This operational discipline applies across identity, endpoints, APIs, cloud configurations, software supply chains and AI systems, measured by clear SLOs. For example:

Domain	*Example SLO*
Identity & Access	*Time to disable credentials for a departing employee: < 1 hour*
Endpoint Security	*EDR agent health coverage: ≥ 99.5% of active devices*
Vulnerability Mgmt	*Time to patch critical internet-facing vulnerabilities: ≤ 5 days*
AI Model Security	*Model monitoring and input validation service uptime: ≥ 99.9%*

In this model, cybersecurity moves away from compliance-driven reporting and becomes an operational capability, fully aligned with product thinking and Post-CAPEX governance models.



Agentic AI *as a Controlled Resilience Accelerator*

Agentic AI is often misunderstood. It does not mean uncontrolled autonomy or the replacement of security teams. Instead, it refers to the deployment of small, purpose-built agents capable of perceiving signals, reasoning within defined boundaries and acting through approved tools, with human arbitration for high-impact decisions. Agentic AI is a primary mechanism to achieve an "operate-by-default" posture at machine speed.

When properly governed, agentic AI reduces response time, absorbs alert overload and standardises reaction quality. It enables cybersecurity to operate at machine speed where appropriate, while preserving human accountability where it matters most.

When poorly governed, however, it amplifies failure. The differentiator is not the sophistication of algorithms, but the clarity of architecture and governance.

Securing LLMs and Agentic Systems: *Foundational Principles*

Securing intelligent systems requires a disciplined approach.

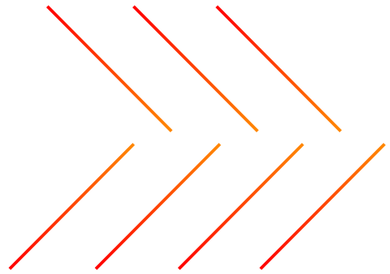
LLMs must be treated as untrusted components, regardless of whether they are proprietary or open-source. Security cannot be assumed within the model; it must be built around it, through sandboxing, policy constraints and continuous monitoring.

A strict separation between reasoning, authority and execution is essential. Models may reason, systems may execute, but humans must always arbitrate. Combining unrestricted data access, decision authority and execution within the same trust boundary is structurally unsafe.

Static rules are insufficient for probabilistic systems. Adaptive guardrails, such as contextual input validation, runtime policy enforcement and behavioral anomaly detection are required to manage risk dynamically. For instance, an adaptive guardrail would not just block keywords, but assess the user's role, the data's sensitivity, and the query's context to permit or deny an action, such as allowing an HR manager to query salary data but blocking the same query from a junior analyst.

Models themselves must be governed like critical infrastructure. Maintaining an inventory, classifying use cases by risk, tracking versions and evaluations becomes mandatory. In this context, a Model Bill of Materials (MBOM) is emerging as a core cyber artefact.

Finally, agentic AI must also be used defensively, under strict control. Automated containment is appropriate for low-risk actions, while human approval remains essential for high-impact decisions. The objective is controlled autonomy, not blind automation.



Avoiding the Most Common *Anti-Patterns*

Certain practices now create a dangerous illusion of security. Push-based or OTP multi-factor authentication, for example, is increasingly bypassed and no longer sufficient for critical flows, which must migrate toward passkeys and FIDO2. APIs that are not inventoried remain invisible and therefore unprotected. Legitimate remote administration tools blur the line between administrator and attacker unless they are treated as monitored exceptions. And uncontrolled use of public LLMs creates silent data and intellectual property leakage through Shadow AI.

These anti-patterns must be explicitly identified and retired.

Regulation, Sovereignty and Trust *as Design Constraints*

European regulations such as NIS2, DORA and the AI Act are not administrative exercises. They formalise a clear expectation: provable resilience. Leading organisations do not approach regulation defensively. They use it as a design constraint to regain control over critical technologies, reduce dependency and lock-in, and industrialise trust as an operational capability.

LLM security sits at the intersection of cyber, data, legal and business governance. Fragmentation is no longer an option.

Cybersecurity *in the Post-CAPEX Age*

As with cloud and AI, cybersecurity economics are shifting. The focus moves from tool accumulation to measurable outcomes, from inventories to demonstrable risk reduction. High-performing CIOs increasingly run cybersecurity as a portfolio of platform products: identity, endpoints, APIs, software supply chains, detection and response, and AI & Agent Governance. Each has a clear owner, explicit service levels and measurable impact.

For the AI & Agent Governance platform, key SLOs would include:

- **Prompt Injection Detection Rate:** ***> 95%***
- **Detection & Response Latency for Model Abuse:** ***< 1 minute***
- **Time to Onboard New Model with Full Guardrails:** ***< 1 day***

Cybersecurity becomes a value-preserving system rather than a sunk cost.

The Compass *Perspective*

Cybersecurity no longer protects systems alone. It protects operations, decisions and trust. LLMs and agentic AI do not weaken cyber strategies; they force them to mature. The CIOs who succeed will be those who prioritise what attackers touch first, secure intelligence without slowing innovation, and transform cyber resilience into a structural advantage for the business.
Over the coming years, AI-assisted social engineering will intensify, software supply chains will remain under pressure, and regulatory scrutiny on AI behaviour will increase. At the same time, agentic defence capabilities will become standard in mature organisations.

The differentiator will not be tooling. It will be operational discipline, architectural clarity and governance.

In the age of agentic systems, cybersecurity is no longer a shield. It is a control system for intelligence. The future of cybersecurity is not just automated. It is agentic, governed and resilient by default.

Contacts

Julien MASSON
Group Head of Tech Advisory
→ julien.masson2@soprasterianext.com

Contributors of our European Network

Olivier BROUSSEAU
Architecture & Tech
→ olivier.brousseau@soprasterianext.com

Isabelle PONS
Data & AI
→ isabelle.pons@soprasterianext.com

Valérie THIBLET
Data & AI
→ valerie.thiblet@soprasterianext.com

Axelle ROESCH
CIO Advisory
→ axelle.roesch@soprasterianext.com

Thomas OTTO
CIO Advisory
→ thomas.otto@soprasteria.com

Philipp MAIER
CIO Advisory
→ philipp.maier@soprasteria.com

Simon KAYE
Architecture & Tech
→ simon.kaye@soprasteria.com

Ray BAKER
Architecture & Tech
→ ray.baker@soprasteria.com

Andy WHITEHURST
Architecture & Tech
→ andy.whitehurst@soprasteria.com

Eirik HASLESTAD
CIO Advisory
→ eirik.haslestad@soprasteria.com

Jahn-Fredrik SJOVIK
CIO Advisory
→ jahn-fredrik.sjovik@soprasteria.com

Tor NESET
Architecture & Tech
→ tor.neset@soprasteria.com

Marius SANDBU
Data & AI
→ marius.sandbu@soprasteria.com

Robert HEXSPOOR
Architecture & Tech
→ robert.hexspoor@soprasteria.com

Hugo HERMSEN
CIO Advisory
→ hugo.hermsen@soprasteria.com

Juan Antonio MARTINEZ SANCHEZ
Architecture & Tech
→ juanantonio.martinezsanchez@soprasterianext.com

Mauro PALMARINI
CIO Advisory
→ mauro.palmarini@soprasterianext.com

Conception: Spintank, photos credits: GettyImages ©