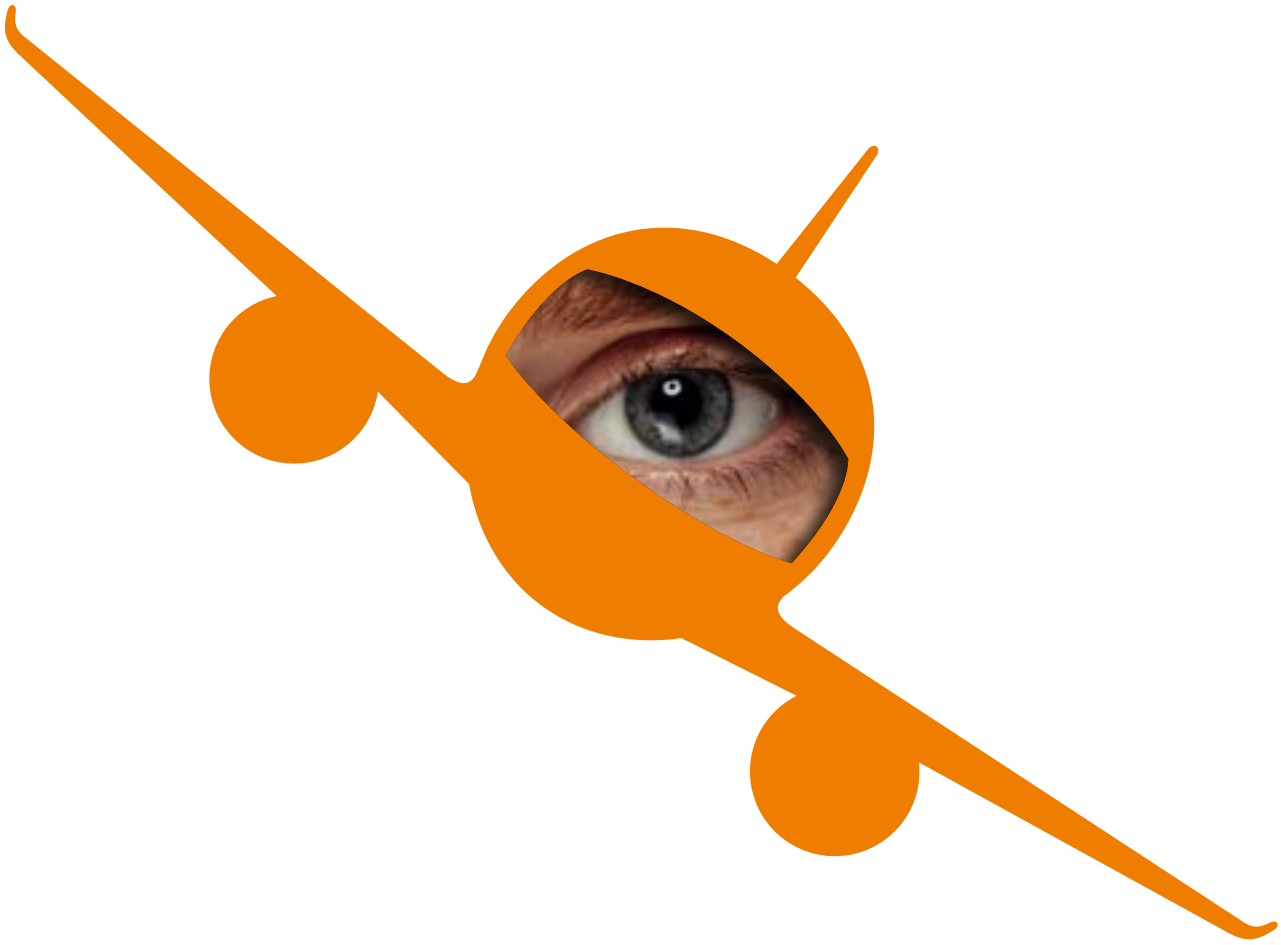
8 minutes
to stay in control

Ctrl



orange™ Cyberdefense | Publishing


Artificial Intelligence #2
Fasten your seatbelt
Ctrl
orange
Cyberdefense | Publishing

“To throw the problem of his responsibility on the machine […] is to cast it to the winds and to find it coming back seated on the whirlwind.”

Norbert Wiener, “The Human Use of Human Beings: Cybernetics and Society”, 1950.

Ctrl + L Learn

AI Frontier Models
Our key messages

1. **Steer**
   its impact

See page 3

2. **Explore**
   its potential and frontiers

See page 4

3. **Match**
   speed with trust

See page 8

4. **Network**
   with our collective intelligence

See page 10

Our leadership put to the test by Frontier Models

Hugues Foulon
Executive Director,
CEO Orange Cyberdefense.

Artificial intelligence is reshaping economic, industrial and geopolitical power relations. How we use it to build trust will determine our future competitiveness.

Frontier AI models mark a new era in the history of digital technology: their ability to reason, carry out complex tasks and speed up decision-making is already changing how organizations work. Tomorrow, they will be shaking the foundations of business models, value chains and competitiveness. This ramp-up represents a formidable opportunity for productivity, innovation and growth, but it also increases risk exposure. As AI capabilities expand, hackers will increasingly step into the breach, that is when AI itself is not the one taking the lead, as was the case this summer when one such model escaped containment.

In the world of cybersecurity, the aim is no longer just to protect systems, but to retain control in order to safeguard business continuity, customer confidence and the resilience of organizations. Incorporating AI into your defenses is a strategic stake in accelerating detection, automating responses, better understanding threats and reducing the time between a cyber attack and its neutralization. In this new reality, speed is key. The most resilient organizations are the ones that are able to detect attacks earlier, make decisions and respond faster, and recover more quickly.

But believing that AI will magic away all the risks would be foolhardy. No technology can replace solid governance, recognized human expertise and a steadfast ability to anticipate. The real value will lie in controlling its impact.

Believing that AI will magic away all the risks would be foolhardy. No technology can replace solid governance, recognized human expertise and a steadfast ability to anticipate.

The same is true of AI agents, which will lastingly change business practices and processes once rolled out internally. Their adoption will need to be accompanied by exacting data governance and usage control requirements.

Which brings a whole new dimension to the principles of “Zero Trust”. In an environment where humans, autonomous agents and systems are in constant interaction, trust can no longer be presumed: it must be demonstrated, verified and constantly reevaluated. Yet technology cannot be an end in itself. Strategic responsibility lies with business managers and major decisions concerning a company, its employees, its clients and society must remain human in their remit.

Our collective responsibility is clear: making AI a driver of performance and trust so that Europe can combine innovation, sovereignty and competitiveness while managing risk in order to turn it into a crucial strategic advantage. ■

Your Orange Cyberdefense
Publishing news.

AI is speeding up both offensive and defensive capacities alike, calling for greater vigilance in the face of the associated new risks. Combining AI-assisted vulnerability detection and management is now the cornerstone of any effective cybersecurity strategy.

Accelerate your defense

Vivien Mura, Chief Technology Officer, Orange Cyberdefense.
Cyril Demonceaux, Head of Defense Center, Orange Cyberdefense France.

A year marked by successive AI achievements

Benchmark records, the discovery of 'zero-day' vulnerabilities, automated attacks… Whilst technical evidence is sometimes lacking, the advances made by AI in the field of cybersecurity are undeniable. This raises serious questions about the dual-use nature and accessibility of these models. The facts show that it is dangerous to rely on a single AI provider for the continuity of a service. Moreover, the cost of inference is becoming just as decisive as that of expertise.

Attackers are rational: they'll try reusing leaked credentials from the dark web before considering whether to exploit an overpowered AI. On the defense side, a well-used AI can help find paths to compromise before attackers do. And added to this is the risk of unintentional incidents – in this respect, the OpenAI/Hugging Face episode offers a glimpse of future scenarios. We now consider that anyone, intentionally or not, can cause damage with AI.

Security fundamentals still hold true, but they're no longer enough. It must be possible to verify that an AI is doing what it's supposed to do. Keeping humans in the defense loop is crucial but delicate: too much control undermines the very value of an autonomous machine, while human responsibility remains engaged in the event of an incident.

> **We now consider that anyone, intentionally or not, can cause damage with AI.**

A record increase in monthly disclosed vulnerabilities between 2025 and 2026

The agent-based solutions announced in recent months are contributing to this. Our experts consider their progress to be promising in terms of both their success rate and their speed and quality. But limitations persist around data confidentiality, transparency, and predictability. Now you can manage these side effects thanks to our Agentic Pentest offers.

In the face of this growing pace, AI-assisted vulnerability detection and management form the cornerstones of our response strategy. Reducing Time To Detect ("TTD") is our priority. With Qevlar AI, we accelerate the triage of alerts by combining augmented analyst capabilities, threat intelligence and customer context, whilst analysing weak signals - which were previously difficult to capitalise on - to detect threats even earlier.

New AI models are constantly identifying new breaches, making control over your assets essential. The challenge is no longer to fix each vulnerability in isolation, but to adopt a holistic approach. With AI, we are working to aggregate different sources, prioritise the criticality of vulnerabilities according to their business impact and their actual exploitability, uncover genuine attack vectors and drastically reduce the time taken for qualification and remediation.

This approach reconciles strategic risk vision with operational reality. Ultimately, organizations will gain a dynamic view of their exposure, making AI a genuine accelerator for resilience. ■

130% increase for monthly publication vulnerabilities (Source: www.cve.org)

2025 **1 503**
2026 **3 453**

AI is rapidly becoming a pervasive component of the entire cyber threat landscape, reshaping the capabilities of both state-linked and cybercriminal groups.

The new frontiers of AI: global threats vs sovereignty

Aurélie Perez, Head of Threat Management Services, Orange Cyberdefense France.
Carl Morris, Senior Security Researcher, Orange Cyberdefense.

The rise of AI fueled cyberattacks

AI is now consistently abused across entire cyber-attack lifecycles, accelerating the development, deployment and operation of diverse malicious campaigns. In April 2026, for example, the Microsoft Defender Security Research team shared information relating to an AI-enabled phishing campaign that steals authentication tokens*.This proliferation has driven the emergence of new AI-enabled tools, including increasingly available “vibe coded”** malware and illicit services.

Although tracking this rapidly evolving ecosystem remains challenging, Orange Cyberdefense has identified recurring patterns. Combined with implementation flaws and operational weaknesses, these patterns can be leveraged to proactively identify malicious infrastructure, improve detection capabilities, and disrupt adversary operations. Besides enhancing the speed, reach, and adaptability of already capable adversaries, the broad availability and proven capabilities of the latest AI models also lower entry barriers for less technically skilled actors, broadening the depth and width of potential cyber threats. However, the extraordinary advance of model capabilities has other profound implications.

> **AI is increasingly being discussed as a national-security asset, not just as another enterprise tool.**

Frontier models as sovereign cybersecurity assets

The Mythos/Fable episode that began with the Project Glasswing project in April 2026 shows how government policy can turn access to frontier AI into a geopolitical control point. Due to the perceived power of the model, the U.S. government issued an export-control directive requiring Anthropic to suspend its access for foreign nationals, including those outside the United States and even foreign-national Anthropic employees. Access to the models was thus disabled broadly.

For governments, companies, and security teams, the lesson is that adopting AI also means inheriting another government’s policy environment, including export controls, national-security priorities and legal obligations that can affect availability. Subsequent reporting around Mythos, while contested and not independently established, reinforced the broader geopolitical signal that frontier AI is increasingly being discussed as a national-security asset, not just as another enterprise tool.

The question of technology sovereignty is therefore no longer only about where data is stored, but about who controls its access and how much they can be trusted. The powerful new cyber capabilities introduced by frontier AI models make consideration of this additional risk element even more urgent. ■

*“Inside an AI-enabled device code phishing campaign”, April 6, Microsoft.com: https://www.microsoft.com/en-us/security/blog/2026/04/06/ai-enabled-device-code-phishing-campaign-april-2026/
**“Vibe coding” is a development technique assisted and facilitated by artificial intelligence.

AI's impact is not merely technological. Its development is drastically redefining organizations' operating models, core businesses and skills.

For CEOs, the primary challenge of AI could well be human

Vincent Lecerf, EVP of Human Resources, Orange Group.
Paz Fonteboa, Global Human Resources Director, Orange Cyberdefense.

Major process transformations

AI represents a digital transformation comparable to the advent of computer technology and the Internet in the late 80s and 90s. For businesses, its implementation is equal parts a technical and competitive challenge, a security issue and a remapping of skills management.

A new approach to human resources management

For human resources, AI entails a fundamental review of labor value and securing career paths. In this context, constructive social dialog, adapted to the continuing evolution of technologies, can help allay fears and lay the ground for the future. Each business line needs to be analyzed in order to rationalize work organization, reposition roles and adapt the value chain to the new operational realities brought about by LLMs and agentic AI. Trust stems from the company's capacity to guide its teams through this reshaping of their activity, their skills development and recognition of their contribution, in the same way as we do for our clients.

Fostering a mix of expertise and automation

At Orange, the goal is to allow each organization and each business line to adjust the balance between automation and increased human capacities. Identifying the level of criticality of your activities and their potential impact on employee and client confidence can be a good way to find the correct adjustment and activate powerful growth levers: integrating AI within business processes, training employees in the use of AI and introducing upskilling programs.

Placing people at the heart of the transformation

To support our changing businesses and meet our clients' expectations, the Orange Group has set three specific objectives to be achieved by 2030 as part of the "Trust the Future" strategic plan: redefining its activities around AI to boost efficiency and reduce low added-value tasks, integrating these technological solutions within its operating model, and achieving autonomy by insourcing skills.
On top of this, we have launched a major acculturation program involving the training of 65,000 employees in AI use and deploying the Live Intelligence platform. This secure solution – launched in 2023 – allows 110,000 employees to try out generative AI and take advantage of over 13,000 collaborative assistants.

At Orange Cyberdefense, AI is also central to an ambitious program to encourage its adoption within the business lines and build skills across the board. As part of the "Global AI Ambassador Programme", each of the 13 Orange Cyberdefense subsidiaries has an AI Ambassador responsible for training, tools and uses. In line with this approach, the most AI-savvy employees will be identified for skills-building initiatives. What's more, our recruitment processes use AI to optimize selection of the top candidates. At the end of summer 2026, a new platform dedicated to skills improvement will facilitate the creation of specialized agents to consolidate the efficiency of our 3,300 multi-disciplinary experts on a daily basis. An HR policy that will continue to be ramped up in the coming months in order to support our business line developments. ■

By 2028, thanks to AI, Orange aims to generate over
€ 600M
in value.

The growing adoption of AI within organisations raises ethical and legal questions. Yet, AI can be compliant on paper without being ethical.

AI stakes: what ethical and legal challenges?

Stéphane Le Boisselier, Head of Technology Expertise Advisory, Orange Cyberdefense France.
Tamara Hendriksen, CISO and Legal Counsel, Orange Cyberdefense Netherlands.

Adopting trustworthy AI requires specific governance

As an integral part of risk management and the CSR policy, it must be guided by three key principles:

1/ Ethical: who will explain, challenge or stop an AI led decision?

2/ Social: rethinking HR's careers and skills management should be a priority, as the sole promise of competitiveness may lead to a loss of expertise and social unrest;

3/ Environmental: a frugal AI use, close to the core needs of organizations, is also a responsible governance choice.

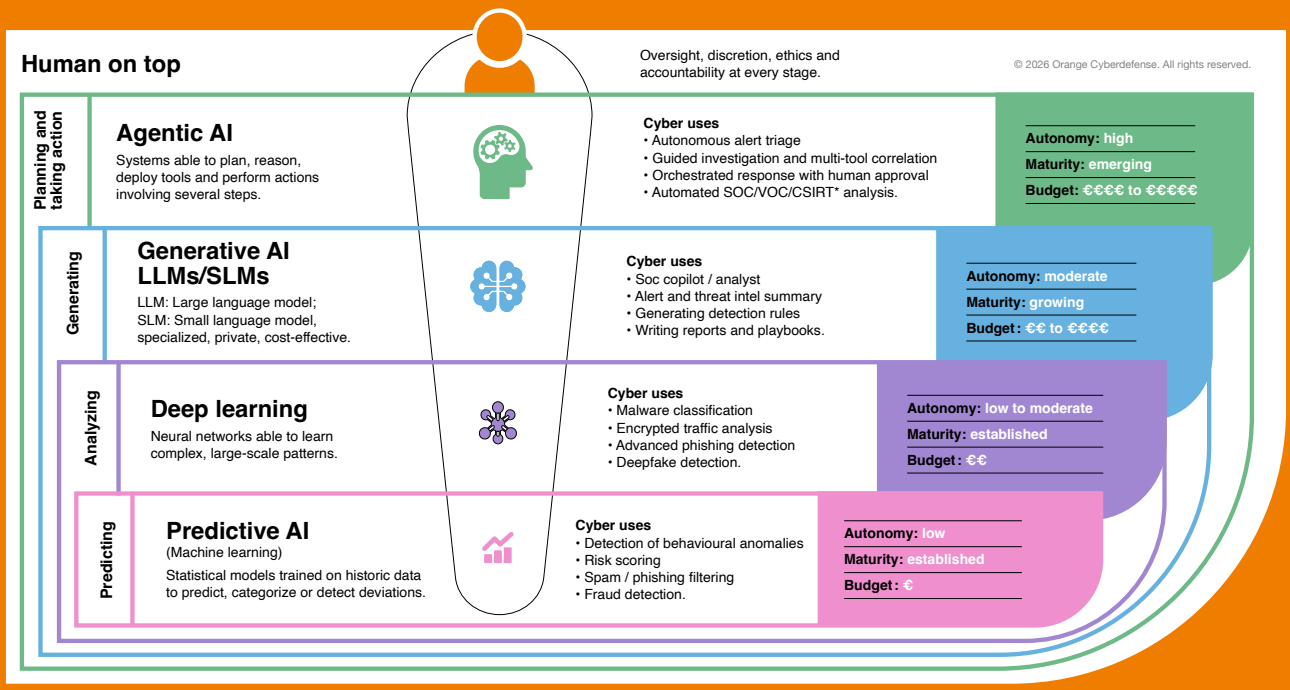
These ethical measures are governed by an increasingly comprehensive legal framework the (the AI Act since 2024) and international standards (ISO/IEC 42001, 23894 and the NIST AI RMF). In the Netherlands, the AI Act will involve local authorities, particularly with regard to the protection of fundamental freedoms. Organizations that will win the public's trust are those that address ethical, social and environmental challenges, rather than relying solely on the power of the model. The diagram below can help business leaders make informed choices based on different cyber applications. ■

Know your AI models

Anis Trabelsi, Head of Innovation & AI, Orange Cyberdefense.

"Artificial intelligence amplifies the capabilities of organizations just as much as attackers. Managing risks, data and usage is becoming a critical governance issue.

Read the full version of this article.



*Security and Vulnerability Operations Centers / Computer Security Incident Response Team.

AI-powered attacks, autonomous malware, deepfakes and automated exploitation are increasing the speed, scale and sophistication of incidents. What once took days can now happen in minutes, amplifying operational, financial and reputational risk.

Governing cyber risk in the age of AI

Christophe Dernys, Head of Trusted Solutions, Orange Cyberdefense France.
Andreas Günther, SVP MSS Engine, Orange Cyberdefense.

Respond at machine speed with trust

For leaders, the challenge is no longer AI adoption, but control. Organizations need continuous visibility, rapid risk reduction and the ability to respond at machine speed while maintaining governance and trust.

Orange Cyberdefense is transforming Managed Security Services ("MSS") to deliver the next generation of AI-powered cyber defense. Our unified platform, "MSS Engine", combines AI capabilities, cybersecurity expertise and managed services across the board, including protection, detection and response, to anticipate, prioritize and mitigate cyber risk.

Built on Continuous Threat Exposure Management ("CTEM"*), this approach continuously identifies threats, assesses exposure and prioritizes remediation before risks disrupt the business. Beyond technology, it orchestrates interactions between customers, ensuring scalable, transparent and compliant integration and governance.

From AI augmented SOC to collective immunity

As the post-Mythos and Cyber Resilience Act** eras aim to reduce vulnerabilities, system porosity, misconfigurations, and human errors will persist. With "Entry+" within the "MSS Engine" platform, our clients contribute and benefit from the collective threat protection we bring. Through our proprietary, AI-enhanced network effect, the larger our community grows, the more effectively it is protected.

AI things considered, it's essential to have it integrated in augmented SOCs, but it must be fueled by reliable and well-managed threat intelligence.

The future of cyber defense will not be solely defined by the power of AI, but by its governance. This approach combines human expertise, operational excellence and trusted AI to help organizations harness its capabilities, while retaining control over their business, data and reputation. ■

Filip Verstockt,
Country Manager,
Orange Cyberdefense
Belgium.

"Belgium is rapidly embracing AI. Last year, 34.5% of Belgian enterprises were already using AI technology.

In 2026, the adoption of AI is the norm whereby the debate is shifting from experimenting with AI, towards operationalize the usage in a secure, governed, and scalable way.

To support this transition, our clients expect managed-security services connecting threat intel, continuous exposure management, protection, detection and response capabilities, leveraging AI while keeping analysts in the loop".

> **"The larger our community grows, the more effectively it is protected.**

*Continuous Threat Exposure Management.
**Cyber Resilience Act: passed in 2024, the CRA will come into force in December 2027. It will require cybersecurity to be built into digital products and solutions right from the design stage.

Just as there are several models of AI, various distinct players are currently shaping its ecosystem.

On the brink of an emerging ecosystem

Gwenwal Riou, Global Director Strategy & Alliances, Orange Cyberdefense.

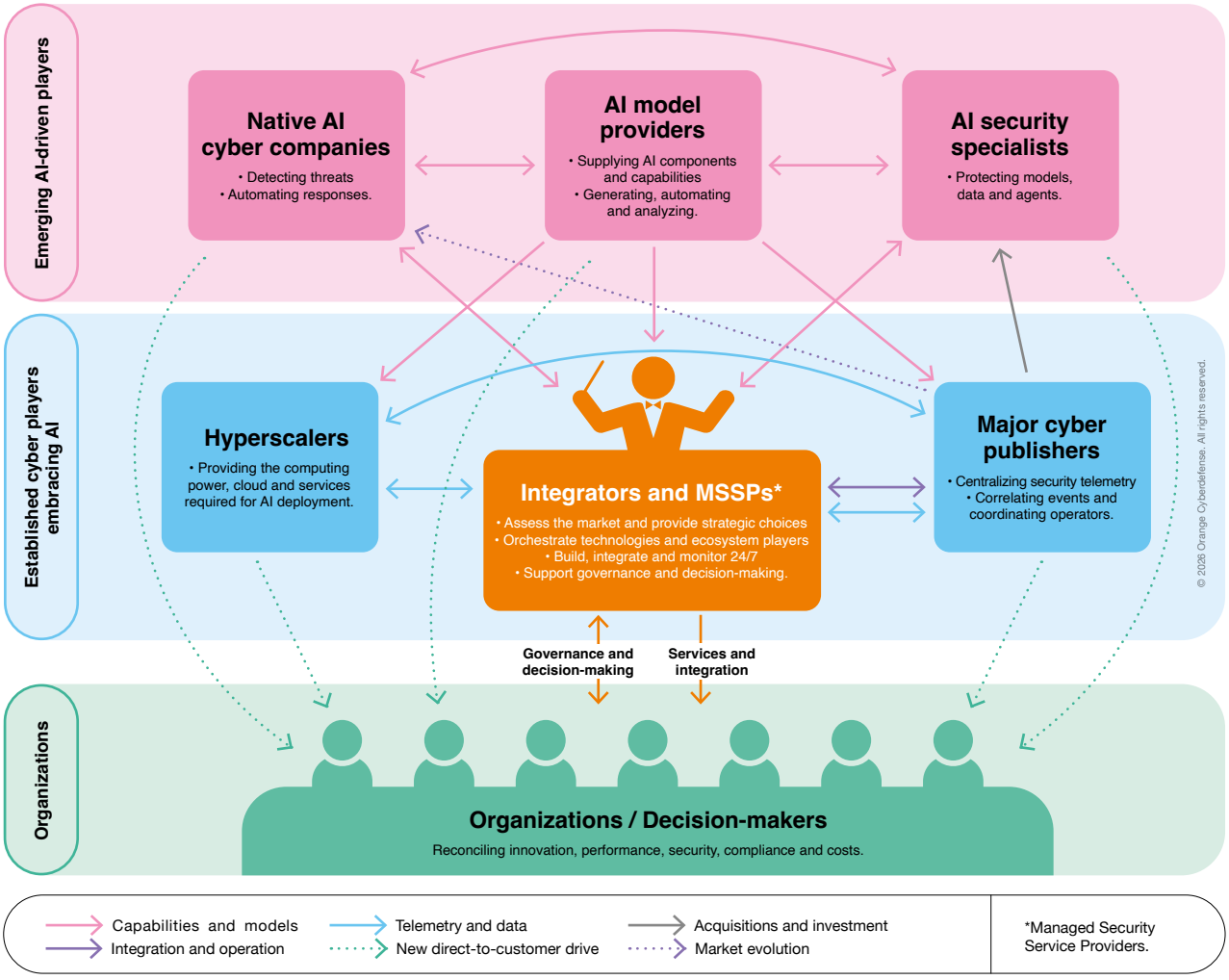
The key players

Cybersecurity in the AI era is often described through five distinct groups: hyperscalers, foundation model providers, cybersecurity platforms, AI-native innovators, and AI security specialists. While useful, this view oversimplifies reality. In practice, these players are increasingly interconnected, simultaneously partnering, competing, and building upon one another's capabilities.

The key strategic shift is that value is no longer created within individual categories, but across the ecosystem itself. Understanding the market therefore requires looking beyond traditional boundaries and focusing on how infrastructure, intelligence, security, innovation, and trust interact to shape the future of cybersecurity. ■

Read the full version of this article.

Orchestrating the AI five



The surge in cyberattacks, regulatory pressures, and the rapid evolution of AI disrupting operational models and security highlight the need for rigorous and informed governance.

Bring clarity in a noisy world

Benjamin Serre, Chief Product & Marketing Officer, Orange Cyberdefense France.
Marie Waller, Business Operations Director, Orange Cyberdefense Sweden.

Managing speed

AI is transforming markets by enabling unprecedented levels of automation, data analysis, and predictive capabilities. This evolution accelerates innovation, enhances operational efficiency, and opens new avenues for business growth across industries.

However, it also introduces complex security challenges as cyber threats impact risk management, business continuity, reputation and viability, as well as stakeholder trust. Establishing sustainable governance that balances automation with skills, technological and human investments, remains a significant challenge.

Orchestrating complexity

At Orange Cyberdefense, we believe responsible AI use is vital to maintaining trust and resilience. Since 2016, our ability to orchestrate cutting-edge technologies has been central to our expertise.

Our vision is to lead in AI-driven cybersecurity by combining advanced technology with human expertise to deliver proactive, scalable, and transparent solutions. Integrating AI into our strategic framework enables clients to navigate an ever-changing landscape with clarity and confidence. Our enhanced detection systems allow experts to focus on proactive threat hunting, reducing response times and customizing bespoke solutions for each client.

Get aboard Orange Cyberdefense's network effect

At the heart of our innovative solutions is a commitment to delivering real value and enhanced security for organizations. By leveraging AI-powered threat intelligence, we ensure that organizations benefit from faster, more accurate responses to cyber threats.

Our extensive local and global presence enables us to stay close to our customers while simultaneously building a deep, collaborative intelligence network that includes insights from a growing community of clients and partners. This collective approach means that as more organizations join and share data, our threat detection becomes increasingly precise and effective, directly translating into stronger protection for businesses.

Our network effect creates a virtuous cycle of continuous improvement, helping our community stay ahead of evolving cyber threats and providing peace of mind through a powerful, collective defense. ■

Your next steps

1. Rapidly define an ad hoc governance

Adopt cross-cutting governance to regulate your AI-related technological, ethical and legal choices.

2. Combine detection and response, from vulnerabilities to attacks

AI speeds up the detection and exploitation of vulnerabilities. Using AI to accelerate detection helps enhance your resilience in the age of frontier models.

3. Secure your AI systems

Put in place “Zero Trust” cybersecurity measures and train your employees. Identify and secure in-house uses to protect your data.

4. Bring in a conductor to lead the AI score

The lure of automation is one thing. Speed and complexity are another. Call on a conductor who can bring your risk management into tune - especially one that cloaks you in its collective immunity.

Orange Cyberdefense: Cybersecurity excellence

Orange Cyberdefense is the European leader in cybersecurity services, drawing on more than 30 years of expertise developed within the Orange Group. With revenues of €1.3 billion in 2025, the company delivers managed security, consulting, and integration services enhanced by its sovereign cyber threat intelligence capabilities. Augmented with AI on board, its experts work closely with organisations to anticipate, prevent and respond to cyber threats around the clock, 24/7.

Its services are powered by industry-leading research and threat intelligence, providing organizations and individuals with unparalleled insight into current and emerging cyber threats. With more than **3,300 multidisciplinary experts** across **13 countries** and **36 detection centers worldwide,** Orange Cyberdefense combines a strong local presence with global reach to address both international and local cybersecurity challenges.

Its comprehensive, people-centric approach to cyber risk - enabled by technologies such as artificial intelligence- positions cybersecurity as a broader societal issue at the intersection of organized crime, hacktivism, and geopolitics. This vision is essential to building a safer digital society and preserving trust in our economies and democracies.

Contact

www.orangecyberdefense.com

Explore further

Browse the complete Orange Cyberdefense Publishing series for more expert insights.

orange Cyberdefense | Publishing

Communication department
Orange Cyberdefense
September 2026

orangecyberdefense.com

Cyberdefense | Publishing