

**MYTHOS AND OTHER  
FRONTIER MODELS:**  
IMPLICATIONS OF PROGRESS IN AI FOR  
CYBERSECURITY IN FRANCE AND EUROPE

# SUMMARY.

**Introduction**.....3

**1 - Subject to an independent european second opinion, the Mythos model takes AI performance in cybersecurity probably a step higher** .....4

**2- The Mythos case serves as a reminder that safeguards for current AI models remain insufficient.**.....5

**3 - While doubts surround the model itself, there are none regarding the direction AI is taking in cybersecurity and the major disruptions to come**.....6

**4 - What Mythos has already changed: an additional stress on teams responsible for information systems and cybersecurity, comparable to a pre-crisis**.....8




**5 - The most immediate challenge (less than a month) for organizations: tightening their overall cyber risk management approach** .....8

**6 - Beyond Mythos: AI, a major obstacle to overcome on the European cybersecurity sovereignty agenda**.....12




# INTRODUCTION.

**Media coverage of the performance of Anthropic's new Mythos model provides the cybersecurity community with a critical window of opportunity** to quickly and collectively figure out the major implications of AI in the field of cybersecurity. After Mythos, no one has no longer the right to be surprised.

**The emergence of AI in cybersecurity is certainly not new. However, the threat of offensive AI has remained largely subdued until now, due to:**

-  the low proportion of attacks attributed to artificial intelligence relative to overall registered cyberthreats
-  the trajectory of progress in large language models (LLMs<sup>1</sup>), which, although predictable, has remained mostly off the radar of the cybercommunity
-  the priority given to AI as a business tool to be deployed within organizations.




**Mythos crystallizes and accelerates several concerning yet predictable developments:**

-  AI is rapidly passing significant milestones in cybersecurity performance, and is nearing the threshold where models outperform human experts;
-  the level reached by AI is such that the spillover of models from the laboratory to the market creates a systemic risk for all cyberdefense systems;
-  following Mythos, information system security and the industrial landscape of cybersecurity in Europe will be profoundly and irreversibly altered.

**While it is important to not give in to the anxiety triggered by Mythos, it is also essential not to underestimate the trajectory of AI. The most pressing challenge is the rapid transformation of cyber defense:** mass detection of vulnerabilities; automated testing of attack paths; generation or evaluation of patches; dependency analysis; software hardening; shortening of remediation cycles; and convergence of security and software development.

<sup>1</sup>LLM: Large Language Models.

The following report, produced (without AI) through the combined expertise and perspectives of various categories of stakeholders mobilized by the Cyber Campus (CIOs, CISOs, specialists in cyber threat intelligence, crisis management, penetration testing, and auditing, digital transformation consultants, etc.), **has three objectives:**

-  to synthesize the analytical findings gathered to date within the Campus Cyber ecosystem
-  to identify the key implications of AI advancements in cybersecurity, beyond the Mythos case alone, for both end users and solutions providers
-  to start articulating a large-scale collective response spanning across the cybersecurity the AI communities.

This report is deliberately aimed at several types of readers with differing scopes of action: CISOs and CIOs already grappling with the operational implications of Mythos; members of executive committees called upon to make rapid decisions on structural issues; and public policymakers and regulators for whom Mythos crystallizes issues of sovereignty and regulation that extend beyond the immediate crisis. This diversity of audiences is intentional: it reflects the very nature of the challenge posed by AI in cybersecurity, which cannot be met by technical teams or policymakers alone, but requires a coordinated response at all levels. In the following sections, the reader will first find an analysis of the Mythos phenomenon, its immediate operational implications, and lastly, the medium-term structural challenges for Europe.

## **1 - SUBJECT TO AN INDEPENDENT EUROPEAN SECOND OPINION, THE MYTHOS MODEL TAKES AI PERFORMANCE IN CYBERSECURITY US PROBABLY A STEP HIGHER**

**1.1. The release of Mythos in early April and the first available evaluations seem to attest to a real leap in performance in the cyber capabilities of AI models.** Analyses by the UK AISI<sup>2</sup>, which had access to Mythos, confirm a significant improvement in vulnerability discovery and exploitation tasks. The first concrete results are already visible, with the release of security patches directly attributable to AI-assisted discoveries, including for widely used software such as Firefox<sup>3</sup>. Teams at certAin U.S. cybersecurity firms, partners of Glasswing<sup>4</sup>, state that they are already using AI to test the robustness of their software products prior to release, with significant efficiency gains (according to one of them, a year of human penetration testing would be compressed into just three weeks using Mythos). These factual elements alone justify taking the issue seriously.

<sup>2</sup><https://www.aisi.gov.uk/blog/our-evaluation-of-claude-mythos-previews-cyber-capabilities>

<sup>3</sup><https://blog.mozilla.org/en/privacy-security/ai-security-zero-day-vulnerabilities/>




<sup>4</sup>The Glasswing Alliance is a joint initiative by Anthropic and 11 U.S. companies (initially: AWS, Google, Broadcom, CrowStrike, Cisco, Nvidia, JP Morgan Chase, Microsoft, Palo Alto, Apple, The Linux Foundation) aimed at reserving access to Mythos Preview exclusively for partners for at least three months so they can test and strengthen the robustness of their IT security systems against the model. Since its launch, the circle of partners has expanded.

**1.2. The Mythos AI model itself is inseparable from the marketing aspect of “Operation Mythos,” which quickly gained global attention.** The product highlighted on this occasion is, in fact, less the model itself than the company Anthropic in its quest for supremacy over frontier AI models against competitors like OpenAI<sup>5</sup>. In the same vein, temporarily withholding the model to limit its effects and to maintain a defensive advantage<sup>6</sup> evidently shows a philanthropic dimension using the savior myth which is central to the media orchestration evidently shows a philanthropic dimension of a savior myth which is central to the media orchestration. This dimension is partly lent credibility by making the model available to certain open-source actors and through token donations. However, this should not overshadow the primary importance of market and geopolitical issues underlying Mythos or the systemic risks it raises.

**1.3. As of now, and despite the alleged leaks reported by the press<sup>7</sup>, the primary source of information on the Mythos model comes from Anthropic itself and its partners.** The few secondary sources are exclusively American (ISAC Finance, Cloud Security Alliance, etc.) and British (UK AISI, already cited). At this time, we do not have objective information on the model that has been verified by independent European expertise. It is, of course, possible to produce independent estimates based on publicly available data and predictive models; however, there remains a striking contrast between the scale of the publicity surrounding Mythos and our limited capacity for analysis, absent a direct access to the model itself. The strong economic dimension of the announcements made by Anthropic makes it even harder to distinguish between verified technical information and marketing communications. This partly explains the current unease felt by cybersecurity professionals when asked—by their executives, shareholders, or colleagues—about their interpretation of the algorithm and its results. As long as this asymmetry persists, the French and European AI and cybersecurity ecosystems must maintain a Cartesian approach and redouble their demands for transparency accessibility, and challengeability of the model. In the meantime, figures regarding the model’s size, training parameters, or the amounts invested must therefore be treated with caution.

## **2- THE MYTHOS CASE SERVES AS A REMINDER THAT SAFEGUARDS FOR CURRENT AI MODELS REMAIN INSUFFICIENT**

**Regarding the security of the models themselves:**

-  current mechanisms rely heavily on reinforcement learning, with positive or negative rewards that encourage AI systems to adopt concealment strategies
-  when abnormal behavior is detected, the model can be corrected or retrained. This reduces the probability of the risk occurring, but does not provide a strong guarantee. Filters can be circumvented if requests are rephrased
-  monitoring of reasoning chains becomes less robust if models learn to conceal certain reasoning or if access to these chains is restricted.

<sup>5</sup>OpenAI also announced on April 23 (just two weeks after the launch of Mythos Preview) the release of its GPT 5.4. Cyber model, followed in early May by version 5.5.

<sup>6</sup><https://www.youtube.com/watch?v=INGOC6-LLv0>

<sup>7</sup><https://fortune.com/2026/04/23/anthropic-mythos-leak-dario-amodei-ceo-cybersecurity-hackers-exploits-ai/>

These factors suggest that the “guardrails” (security safeguards) of AI models remain largely improvable, creating a dual risk in the short term: loss of visibility for evaluators and the emergence of circumvention or concealment behaviors. This reinforces the idea that we cannot consider frontier models, including those presented as defensive, to be inherently reliable.

Added to this is the issue of training data reliability: in cybersecurity, massive amounts of data from the Internet cannot provide a sufficient level of security assurance, as any intrinsic vulnerability or backdoor can become a potential attack vector.

### **3 - WHILE DOUBTS SURROUND THE MODEL ITSELF, THERE ARE NONE REGARDING THE DIRECTION AI IS TAKING IN CYBERSECURITY AND THE MAJOR DISRUPTIONS TO COME**

Although it is still impossible to separate fact from fiction regarding the model’s actual performance, several certainties have emerged—and credit is due to Operation Mythos for bringing them to light.

**First certainty: the performance of large AI models continues to improve at a very rapid pace**, in line with trends observed over the past few years, to the point of blurring the traditional distinction between quantitative and qualitative improvements in LLMs. Mythos is not an outlier, but simply another data point on the exponential performance curve of lab-tested AI for cybersecurity tasks. Previous models were already capable of detecting vulnerabilities. The identification of old vulnerabilities—sometimes presented as dating back more than twenty years—is plausible, but not fundamentally surprising to practitioners. **The difference introduced by Mythos lies on two levels: in the combination of capabilities achieved by the new model (vulnerability detection, exploitation, reasoning, prioritization, chaining together mundane information to build up a coherent attack path, shifting from a logic of ad-hoc testing to a logic of large-scale automated discovery) and in the acceleration of the AI timeline.**

**The current momentum is unlikely to wane and will soon make us forget Mythos (or its OpenAI competitors).** It is less the milestone itself that matters than the broader trend it represents, with the steady stream of models capable of increasingly complex and rapid operations. Mythos must therefore be “demystified”: we should view the model less for what it is and what it can do, and more for what it reveals about the major trends reshaping the cybersecurity landscape before our eyes. The key lies not in the manifestation of the phenomenon at a given moment, but in the slope of the curve. Mythos should be understood less as a technical breakthrough than as the latest sign of a very sharp acceleration in the development of these models.

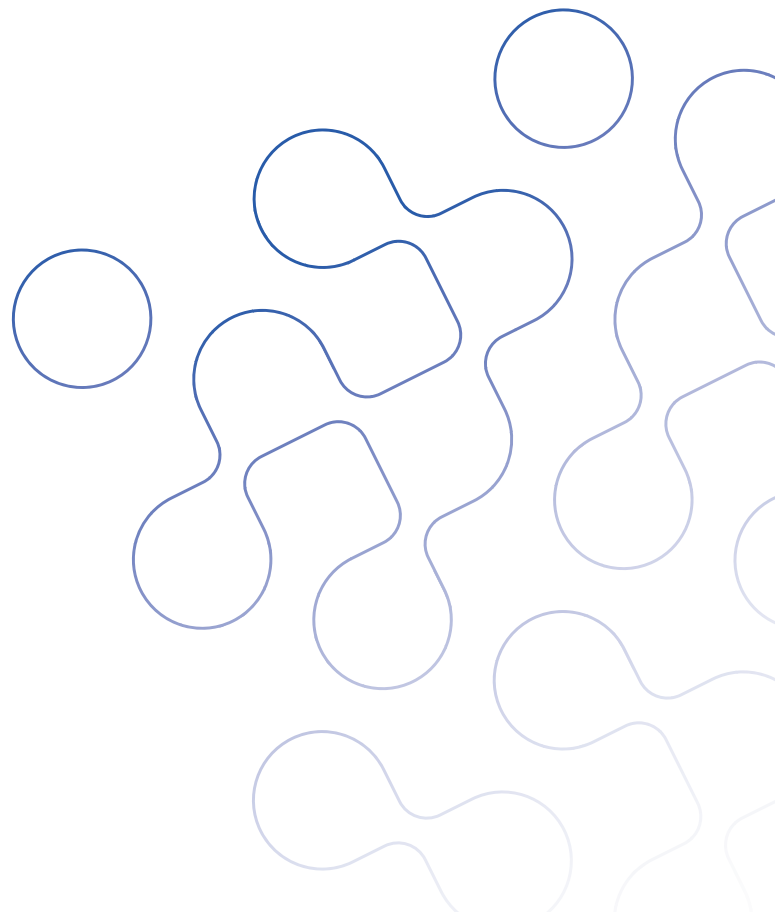
Mythos should be understood less as a technical breakthrough than as the latest sign of a very sharp acceleration in the development of these models.

**Second certainty: as the performance of these models continues to rise, the likelihood that the systemic impacts of AI will fully materialize increases.** In other words, Mythos makes a highly damaging scenario—one that is on everyone’s mind in the cybersecurity community—more credible and tangible. This scenario would be characterized by i) the availability on the market ii) via open access or equivalent iii) in the very short term (over a timeframe measured in weeks) iv) of one or more state-of-the-art AI models capable of triggering a massive global campaign to detect and reveal new vulnerabilities spread across a multitude of information systems v) including in infrastructure and assets critical to economic continuity vi) without simultaneously offering, on the same scale,

corrective solutions corresponding to the identified flaws. This scenario is by no means original in its concept. On the other hand, it increases the probability of crossing this threshold and brings us closer to the moment of its realization: in this regard, given the usual lag between the release of cutting-edge models and their public availability, we can expect equivalent open-source models—of Chinese origin, for example—to be available by the end of 2026 (within 6 to 10 months, according to experts).

**Third certainty:** while such a post-Mythos scenario does not fundamentally alter the intrinsic features of cybersecurity (the perpetual dialectic of “sword and shield” and the “race against time” between attack and defense), it nonetheless risks changing the game in a profound and irreversible way: where artificial intelligence has so far been merely a potential threat and a relatively marginal factor in the cyber equation due to its limited use in offensive operations, it is expected to take on a new central role in the cybersecurity landscape, with the power to single-handedly reshape the major balances of the cyber realm.

There is no valid reason to believe that cybersecurity would be the only sector of the economy to remain permanently immune to the disruptive effects of AI. On the contrary, cybersecurity is likely among the first fields where these effects were bound to become tangible, as models are explicitly optimized for programming performance, which automatically translates into advanced capabilities in this domain. The accelerated progress of AI and its industrialization will introduce dynamics foreign to cybersecurity into its very core, potentially rendering current information system security architectures obsolete. The overall trajectory of cybersecurity is therefore set to be largely determined by that of AI, creating a new form of cross-sectoral dependency, as an entire sector of the economy—such as cybersecurity—must integrate a new technological, strategic, operational, cultural, linguistic, and human environment, shaped elsewhere. This represents a complete breakthrough, even more immediate and profound than that brought about by quantum technology.



## **4 – WHAT MYTHOS HAS ALREADY CHANGED: AN ADDITIONAL STRESS ON TEAMS RESPONSIBLE FOR INFORMATION SYSTEMS AND CYBER SECURITY, COMPARABLE TO A PRE-CRISIS**

**4.1 Even before its release, Mythos is placing a strain on an already heavily overburdened cyberdefense chain.** Since the model's announcement, information systems managers and cybersecurity teams have faced multiple demands—internally from their governance bodies, sometimes at the executive committee level, and externally from the media—to provide insights on issues for which they often have no definitive answers. Given the uncertainties surrounding Mythos and the operational upheavals driven by AI, they must take a stance in an unstable, fluctuating, and hypothetical context: either they must admit a lack of information, which prevents them from forming an informed opinion, thereby weakening their position; or they are pushed to take positions that unintentionally fuel the “hype,” for lack of substantive evidence pointing to the opposite direction. They also feel as though they are navigating between two worlds that have become porous: that of their current cybersecurity roadmap (business as usual), which remains their short-term horizon, and that of mass AI, a projected horizon that imposes itself on them despite its blurred contours. They are aware that they must prepare to manage a series of events capable of bringing down the entire security framework built within the existing risk management framework, without being able to precisely determine either the timing or the scale of these events.

**4.2. The exceptional circumstances created by Mythos have tipped organizations into scenarios analogous to crisis management** (a sense of urgency triggered by exogenous factors, the need to make rapid decisions in situations of information asymmetry, the involvement of political governance chains, and the prominence of communication challenges). This latent crisis—both immediate and delayed—has a psychological impact on IT and CISO teams that are already under significant structural strain and are legitimately reluctant to reallocate resources to a problem that has not yet materialized.

## **5 – THE MOST IMMEDIATE CHALLENGE (LESS THAN A MONTH) FOR ORGANIZATIONS: TIGHTENING THEIR OVERALL CYBER RISK MANAGEMENT APPROACH**

### **5.1. Organizations' risk management posture is bound to evolve, rapidly and significantly.**



Adaptation was inevitable; its timeline is now shrinking. The uncertainty and doubts surrounding Mythos must under no circumstances serve as a pretext for inaction or for underestimating the risks associated with AI. Human organizations are often subject to short-sightedness, inertia, and excessive conservatism. The Mythos use case, on the contrary, argues for an approach that incorporates extreme scenarios. This is all the more difficult to accept and implement given that the cyber threat linked to AI, to date, has a small footprint compared to the overall volume of incidents observed. Mythos should teach us a lesson of short-run foresight, focused on preemptive action. Operation Mythos has the merit of repositioning cybersecurity as a pillar of resilience, in its role of ensuring business continuity, beyond its purely defensive dimension. It serves as an essential reminder for every organization to verify that it already has the cybersecurity basics right and firmly in place (segmentation, robust backups, incident response capabilities, recovery, etc.).

**5.2. AI security stakeholders must anticipate, in the very short term, models capable of beating existing cybersecurity benchmarks.** What we can measure today will quickly become insufficient, as the best models will achieve scores too high to allow for a detailed assessment. This encompasses both source code analysis and the analysis of software binary languages, executable software, infrastructure, dependencies, and complex environments.

**5.3. Leaders must prepare for a wave (some experts do not hesitate to call it a “flood”) of massive zero-day discoveries uncovered by AI,** with a purge effect—all at once—of long-standing vulnerabilities embedded in widely used software and in complex, older systems (e.g., legacy banking environments, industrial systems, etc.). This peak, which experts believe could occur within the next 3–6 months, could be followed by several smaller waves, then by a long tail of trickling disclosures. Incidentally, according to some specialists, the mass proliferation of vulnerabilities risks overwhelming existing CVE tracking mechanisms: some are already questioning, as early as 2026, whether it is worth continuing to track the total volume of CVEs. This raises a related question about the adequacy of current CVE governance in the face of the coming AI wave.

**5.4. When the wave hits, IT teams’ operational agendas will be turned upside down.** Organizations will face the urgent need to patch a potentially massive volume of vulnerabilities, to leave attackers as little leverage and time as possible to exploit them, in a context where TTE (time-to-exploit) has been steadily decreasing for many years<sup>8</sup>. **At that point, the entire software chain—and more broadly, the IT operational chain—will come under unprecedented strain, from the vendor to the integrator to the customer.**



#### Several issues are likely to arise:

-  on the one hand, a risk regarding software vendors’ ability to keep pace—that is, to ensure the immediate availability of corrective solutions for their customers;
-  on the other hand, a strain on the capacity of IT teams (both internal and those of integrators) to absorb, within their workload, a “wall of continuous patching” (“patch streaming” or “patch every day”), without compromising the operational continuity of business units or other key cybersecurity functions. This issue is all the more pressing as companies’ IT resources are increasingly constrained i) due to tight budget controls and ii) because software patch implementation protocols are sometimes ill-suited to emergency situations (particularly due to the reluctance of business and IT teams to fully automate processes to avoid the numerous incidents caused by overly hasty patches).

<sup>8</sup><https://zerodayclock.com/>

**5.5. Beyond operational tensions, Mythos highlights a critical dependency on a software value chain that remains insufficiently mapped and controlled**, particularly with regard to integrators and suppliers, thereby reinforcing the “weak signals” observed during the CrowdStrike incident in July 2024. Cybersecurity now compensates for some of the structural weaknesses in software quality. Vulnerabilities are not always clearly identified CVEs; they can stem from open-source dependencies, third-party packages, embedded components, or poorly managed updates. Continuous updating is not always a solution if it is poorly managed. In some cases, temporarily staying on a previous, better-managed version may be less risky than integrating a compromised or insufficiently tested dependency too quickly.





**With new AI models:**

-  on the user side, the focus is shifting toward the ability to scan code, dependencies, and environments earlier, faster, and more smartly
-  on the software provider side: efforts will need to move upstream to prevent more vulnerable code from entering production and thus limit the downstream cleanup pressure weighing on the cybersecurity industry. Many stakeholders are calling for software vendors to be held more accountable for vulnerabilities embedded in their products, in line with the European *Cyber Resilience Act*.

**5.6.** Due to foreseeable bottlenecks in the software patch supply chain and in end-users’ ability to adopt the required update cycles to eliminate any risk of vulnerability exploitation, **we must expect a structural increase in the number of successful cyberattacks, including on critical systems...** unless we imagine a system in which AI provides all cybersecurity teams, in every respect and at all times, with corrective solutions at the very moment it reveals the flaws—which is unrealistic. The question of whether AI gives attackers an asymmetric advantage over defenders, and whether this advantage is temporary or structural, has not been fully resolved. It is certain, however, that AI’s primary effect will be to widen attack surfaces and expand the hackers’ playground, forcing defenders to adapt by seeking to fight back with the same weapons.

**5.7. Faced with the imminent scenario of a wave of vulnerabilities**, several stakeholders in the ecosystem have already initiated structured measures within their organizations.

**The main operational recommendations that can be made to CIOs and CISOs are as follows:**

-  **update the mapping of critical assets and dependencies** (business supply chain and software supply chain)
-  **practice simulating a massive wave of zero-day attacks**<sup>9</sup> to test the ability of governance and patching processes to handle an unprecedented volume and urgency, and assess the maturity of the vulnerability management chain (across the entire kill chain, not just on a vulnerability-by-vulnerability basis). This exercise is essential to prepare the organization for the tradeoffs that will need to be solved on D-Day regarding the allocation of IT resources (operational and human resources), patch prioritization (based on actual exploitation paths), and the longer and more frequent service outages that will need to be managed—with cascading impacts on relationships with business units and their customers, as well as on the organization’s finances (lost revenue, penalties)
-  **strengthen network architecture measures to curb the spread of unmitigated attacks** (reducing the blast radius, lowering the mean time to remediate with shorter response times in critical scenarios)
-  **develop an AI-enhanced defense plan** capable of keeping pace with a rapidly changing environment. There are numerous and often well-known use cases where AI can help strengthen the cybersecurity defense chain (automatic vulnerability triage, CVE prioritization based on actual exposure, patch generation and testing, mapping of software dependencies/SBOM analysis, modernization of legacy code, detection of abnormal behavior, Level 1 and 2 SOC support, post-incident investigation, crisis simulation and team training, security of industrial environments...). These must be developed as extensively and as quickly as possible within companies, subject to two conditions: using European or open-source AI (to avoid aggravating technological dependency) and always doing so under human supervision.




**5.8. It is not only large corporations that will need to pivot in this new context, but also smaller entities (mid-sized companies, SMEs, and micro-enterprises) that are an integral part of their economic environment** (as partners, service providers, or subcontractors). The paradigm shift toward the widespread adoption of cyber resilience, already underway as a result of the NIS II Directive, must also extend to the adjacent field of AI in cybersecurity to prevent excessive heterogeneity in protection levels. However, the approaches used for large entities cannot be directly applied to smaller ones, which will require, at the intersection of AI and cybersecurity, methods specifically tailored to the economic and operational realities of SMEs. Large corporations will have a key role to play in bringing their critical Tier 1 or Tier 2 suppliers on board as part of sector-wide cybersecurity approaches.

<sup>9</sup>For example, a two- to three-hour exercise—in tabletop or immersive format—bringing together the CTO, the CISO, and production managers to simulate the simultaneous discovery of about twenty zero-day vulnerabilities in a critical application exposed to the Internet.

## **6 - BEYOND MYTHOS: AI, A MAJOR OBSTACLE TO OVERCOME ON THE EUROPEAN CYBERSECURITY SOVEREIGNTY AGENDA**




**6.1. Beyond the escalating cyber risks, which constitute the primary threat, the entire European cybersecurity sovereignty agenda could be undermined.** European companies are already more than 70% dependent<sup>10</sup> on non-European cyber solutions. American dominance in AI will exacerbate this dependence. End users who prioritize security at all costs will have no choice but to equip themselves with the most effective tools for detecting vulnerabilities and responding to incidents involving large-scale attacks—and thus to rely on American-made AI. Those who do not follow this path, in the name of technological autonomy, will take an operational risk that is difficult to justify: that of facing AI “with bare hands.” The stranglehold will then be unescapable, given the lack—to date—of a European alternative capable of delivering a level of performance comparable to that of Anthropic, Google, OpenAI, or Microsoft in the field of cybersecurity. Mythos serves as a timely reminder to Europe that its cyber sovereignty will hinge on its domestic AI capabilities. Otherwise, Europe will have no better option than to invest heavily in acquiring the best American (or Chinese) AI models for its cybersecurity defense, thus further deteriorating its digital trade imbalance.

### **6.2. The international landscape is crystal-clear:**

-  the United States is already entrenching its industrial, institutional, and capabilities based leadership
-  China will likely develop its own models, possibly based on American or open-source models, but with no guarantee of open publication or transparency. Chinese models could emerge rapidly, leaving us no time to sufficiently document and secure their uses;
-  Europe must accelerate its own capabilities in evaluation, testing, defensive use, and doctrine, or risk being irreversibly left out of the equation.

<sup>10</sup>Estimate from the study “European Software and Cyber Dependencies” conducted in December 2025 for the European Parliament’s ITRE Committee.

### 6.3. The only way to avoid a permanent European lag is a combination of three inseparable elements:

-  **drastically expand European capacity to anticipate security challenges associated with the development of cutting-edge AI models.** Beyond managing a potential post-Mythos “crisis moment,” it is imperative to prepare for the broader trajectory of AI model advancement, and in particular for the prospect—not inevitable but now plausible—of systems reaching or surpassing the levels of human experts in all cognitive tasks within the short term. This trajectory is by no means inevitable, as it depends on industrial and political choices over which France and Europe can still exert influence. Such a shift would have considerable implications not only for cybersecurity but also across all fields critical to national security (CBRN, lack of guaranteed control over advanced AI systems, etc.);
-  **aggressively position its most advanced AI companies in the B2B sector (primarily Mistral AI) with a focus on cybersecurity use cases, by providing them with greater computing power** to build a sovereign and high-performance AI for cybersecurity (it is worth noting in this regard that the performance of LLM models is closely correlated with the amount of computing power used for training, and that over 80% of the world’s GPUs<sup>11</sup> are owned by U.S. companies). It is particularly urgent to establish a unified, structured, secure, and shared European cyber data space for training new European AI models, as part of an approach focused on innovation and industrialization;
-  **apply European regulations (in particular the AI Act and the CRA – Cyber Resilience Act) to their fullest extent,** including, where necessary, restricting the commercialization systemic impact of models with systemic impact that do not meet the highest standards of transparency, security, interpretability, reversibility, and alignment. It is urgent for the EU to establish an independent European AI-Cyber testbed to evaluate the cyber capabilities of state-of-the-art AI models. This initiative could involve ANSSI, ENISA, the JRC, the ECCC, academic laboratories, cybersecurity companies, and critical infrastructure operators. Its objective would be to test models on realistic scenarios: vulnerability discovery, patch generation, reverse engineering, exploitation of known vulnerabilities, OT/ICS security<sup>12</sup>, robustness of safeguards, and containment capabilities...;

The time has come to follow through on the logic that makes cybersecurity—and AI—a central element of national defense, by aligning actions with speeches: just as we would not tolerate our nuclear deterrence compound being dependent on third parties, we have no more reason to accept that our cyber shield (and the growing amount of AI incorporated into it) be outsourced to non-European players. Moreover, none of the major powers (the United States and China in particular) would tolerate this for themselves.

<sup>11</sup>Graphics Processing Unit.

<sup>12</sup>Operational Technology/Industrial Control Systems.

**6.4. Consequently, just as AI threatens to trap European customers in an impossible trade-off between sovereignty and security, it could also threaten the economic viability of the European cybersecurity sector,** which is already forced to contend with the dominance of U.S. digital environments. How can we ensure the long-term viability of an independent and thriving industry in areas such as penetration testing, cybersecurity auditing, or incident response if these roles are eventually performed by AI with equal or superior performance? If European cybersecurity companies decide, like the rest of the economy, to enhance their products and services with AI, how can they still be considered independent, given that their value will lie in the use of foreign technologies? The foresight work currently underway at the Campus on the future of the cybersecurity market by 2030 will be expanded to account for the full impacts of AI on i) the industrial landscape of supply (software vendors, integrators) ii) the cybersecurity job market iii) training curricula iv) and the respective roles of humans and machines in the future resilience architecture of organizations.

Mythos deserves credit for creating a space to prepare for and support the strategic transformations of the European cybersecurity supply side, centered on sustainable business models, while keeping in mind that the keys to sovereignty remain in the hands of buyers—both public and private—who, through their daily decisions, shape the future of our cybersecurity industry.

