TLP: CLEAR

EUROPEAN UNION AGENCY
FOR CYBERSECURITY

ENISA's view on Cybersecurity in the Frontier AI Era

JULY 2026

ABOUT ENISA

The European Union Agency for Cybersecurity, ENISA, is the Union's agency dedicated to achieving a high common level of cybersecurity across Europe. Established in 2004 and strengthened by the EU Cybersecurity Act, the European Union Agency for Cybersecurity contributes to EU cyber policy, enhances the trustworthiness of ICT products, services and processes with cybersecurity certification schemes, cooperates with Member States and EU bodies, and helps Europe prepare for the cyber challenges of tomorrow. Through knowledge sharing, capacity building and awareness raising, the Agency works together with its key stakeholders to strengthen trust in the connected economy, to boost resilience of the Union's infrastructure, and, ultimately, to keep Europe's society and citizens digitally secure. More information about ENISA and its work can be found here: www.enisa.europa.eu.

CONTACT

For contacting the authors please use info@enisa.europa.eu

For media enquiries about this paper, please use press@enisa.europa.eu.

AUTHORS

ENISA

ACKNOWLEDGEMENTS

ENISA would like to acknowledge the valuable insights received from the EU CSIRTs Network, EU-CyCLONe, the ENISA Advisory Group, the ENISA Cyber Partnership Programme, industry representatives, the open-source community and academia.

LEGAL NOTICE

COPYRIGHT NOTICE

English PDF Web TP-01-26-015-EN-N 978-92-9204-801-3 2314-9434 10.2824/9549088

VERSION HISTORY

DATE	VERSION	MODIFICATION
4 May 2026	0	Internal draft
8 May 2026	1.1	Shared with EU-CyCLONe and CSIRTs Network
29 May 2026	1.2	Feedback integrated from CSIRTs Network and EU-CyCLONe
11 June 2026	1.3	Further integration of comments from EU-CyCLONe
17 June 2026	2	Proofread and graphic design Discussion and feedback from ENISA Management Board
18-19 June 2026	3	Minor edits Shared under embargo with the Management Board, ENISA NLO, AG, CSIRTs Network, EU-CyCLONe and CPP.

Table of contents

Executive summary

Frontier AI models are challenging traditional security paradigms by compressing the vulnerability management lifecycle and attack chain, from discovery to exploitation. The frontier model landscape is evolving rapidly; it is expected that open-weight models may reach a similar level of capability within 9 to 12 months and that existing models when coupled with skilled security experts can yield comparative results[1]. While ENISA acknowledges the potential benefits of these technologies to improve security, this note aims to focus on some of the immediate and mid-term cybersecurity challenges.

This note provides national competent authorities in Member States and EU policymakers, defenders, and service providers with an initial set of recommendations to support them in their respective roles towards developing the necessary operational capabilities to face machine-speed threats. The recommendations are not an all-inclusive checklist. ENISA aims to further refine and expand these recommendations in close cooperation with Member States and EUIBAs and will align these to upcoming European Commission Action Plan.

To identify the appropriate response to the emergence of AI and its impact on cybersecurity, ENISA held a series of engagements where our stakeholders raised the following:

- there is the likelihood that attackers will have access to exploits before fixes are released (so called *negative time-to-exploit*[2]);
- AI amplifies challenges related to legacy systems and products that will soon reach or have reached their end-of-life and end-of-support;
- due to an expected increase in patch release frequency, patching may lead to an increase in service disruptions;
- open-source require a strategy to prevent maintainers to be overloaded with vulnerability reports;
- SMEs, part of the backbone of the EU economy, may require additional support, in particular in terms of guidance and access to the latest models;

1 https://aisle.com/blog/ai-cybersecurity-after-mythos-the-jagged-frontier (accessed 18 June 2026)

2 https://suzulabs.com/suzu-labs-blog/mean-time-to-exploit-has-gone-negative.-security-strategy-has-to-change (accessed 18 June 2026)

- cybersecurity should be positioned as a strategic use case for European investment in AI, as a need exists for the EU-based organisations to have access to and develop their own AI models,
- security fundamentals matter more than ever in the age of AI;
- resources need to be shifted from discovery to risk-based prioritisation of vulnerabilities through higher-speed triage, remediation and risk reduction;
- defensive AI tooling need to be integrated into the software development lifecycle to support secure by design practices;
- human-gated AI workflows need be integrated across incident response and threat modelling, by upskilling and reskilling the cybersecurity workforce;
- architectural solutions must use an assume-breached mindset, while acknowledging that zero trust approaches will require a deep transformational process;
- Cybersecurity as Code: means machine-speed threats need to be addressed with machine-speed defences (Vulnerability Management as Code, Incident Response as Code, Security by Design as Code, Security Architecture as Code);
- AI-driven defensive capabilities that can detect, correlate and respond to threats at machine speed need to be deployed, and
- the Cyber Resilience Act's Single Reporting Platform must be leveraged, once it is functional, to address the challenges posed by new developments.

Impact on Vulnerability & Patch Management

Over the past few years, ENISA has urged the cybersecurity community to adapt its vulnerability management practices in response to advancing AI capabilities. We have noted the steadily narrowing window between vulnerability discovery and exploitation: this period has gone from years, to months, and now (potentially) to hours or even minutes. Industry research underscores the urgency. Attackers can weaponise new vulnerabilities within 15 minutes of disclosure, and the median time from initial access to data exfiltration has been compressed to 72 minutes[3], meaning that security operation centres (SOCs) must further adopt automation, and explore new strategies to reduce their Mean Time to Detect (MTTD), Mean Time to Respond (MTTR), and Mean Time to Contain (MTTC) to levels commensurate with the evolving threat landscape.

Historically, the Window of Exposure (WoE) was a manageable gap; the inherent friction of human-led research and the manual development of exploits provided defenders with a 'grace period'. However, data synthesised by researchers shows that the delta between discovery and weaponisation is approaching zero[4].

ENISA's current capacity, as well as that of manufacturers, national CSIRTs, and national coordinated disclosure processes, may need further reinforcement or reconsideration to manage a large wave of newly discovered vulnerabilities.

The NCSC-NL noted[5] that the speed of autonomous agents eliminates the defender's advantage of "private discovery". Once a vulnerability is found by an AI agent, the transition to a weaponised exploit is no longer a human-led engineering project but a machine-speed computationale task. The message to organisations should be clear:

3 https://securitybrief.com.au/story/ai-fuelled-cyber-attacks-now-steal-data-in-72-minutes (accessed 18 June 2026)
4 https://zerodayclock.com/ (accessed 18 June 2026)
5 https://www.ncsc.nl/nieuws/anthropics-frontiermodel-mythos-vraagt-om-directe-actie (accessed 18 June 2026)

> *"Do not treat this as a "next trend", but as a structural shift in the pace of attack and defence[6]".*

Also, as CERT-EU noted[7]:

> *"What makes the latest generation of models particularly dangerous is not just the volume of vulnerabilities they find. It is their ability to chain findings across multiple steps, reason about application logic, and produce exploitation paths that previously required deep specialist knowledge. This is what turns a list of individual flaws into a working attack."*

This capability extends beyond vulnerability chaining. Frontier AI models can creatively reason about application logic, credentials, configurations, and API access patterns, going beyond static code analysis to understand how applications work, and can orchestrate entire autonomous attack campaigns from reconnaissance through lateral movement to exfiltration, not merely discover individual vulnerabilities.

As mentioned by NCSC IE:

> *"Now is the time to review asset inventories, prioritise patching discipline, and assess exposure to unsupported components".*

Europe's particular context also needs to be acknowledged as noted by Belgium's CCB:

> *"The EU, as an innovative block with a high density of SMEs and a deeply interconnected supply chain, is particularly sensitive to the disruptions brought by Frontier AI".*[8]

To navigate these changes, the cybersecurity community must prepare for seven structural challenges derived from current ENISA's assessments:

- **The Velocity Asymmetry and the Authority Gap**. In the new environment, the primary latency is no longer technical, but procedural. The "Authority Gap" refers to the inability of human Change Advisory Boards (CABs) to authorise interventions in just the few minutes available to counter autonomous exploits. Plainly, it means organisations are too slow to authorise a meaningful antidote. The necessity of autonomous patching, where the risk of an automated update breaking a system or causing significant downtime is statistically weighed against the near-certainty of an AI-driven breach will be challenging. Particularly in the short-term, larger organisations will likely be able to establish resilient patching practices quicker, while smaller ones will face challenges due to a lack of resources, security focus, and the necessary skills.
- **Economic Devaluation of Discovery**: Vulnerability discovery is becoming increasingly industrialised, which lowers the marginal cost of finding flaws and raises the volume of disclosures faster than many organisations can remediate them. As an example, an industry representative shared with ENISA that their organisation used to receive a few hundred reports of possible vulnerabilities a year, of

6 Original (NL/machine translated): *Behandel dit niet als "volgende trend", maar als structurele verschuiving in het tempo van aanvallen én verdediging.*

7 https://cert.europa.eu/blog/ai-vulnerability-discovery-defenders-must-adapt (accessed 18 June 2026)

8 Centre for Cybersecurity Belgium (CCB) – Food for thought: The Cybersecurity Revolution in the Age of Frontier AI

which only one hundred would traditionally receive a CVE ID. However, the organisation went from about ~80 CVEs in Q1 2025 to close to 500 in Q1 2026, and then the number jumped to about 500 per day when they used Frontier AI-enabled tools. ENISA's analysis of vulnerability disclosure[9] notes that disclosure economics depend on incentives, coordination, and the costs borne by both researchers and defenders, making rapid triage and remediation a more critical bottleneck than discovery alone. Vulnerability prioritisation mechanisms like the Exploit Prediction Scoring System (EPSS) developed by FIRST, and the Vulnerability Exploitability eXchange (VEX) can directly assist organisations in focusing scare resources.

- **Technical Debt as an Existential Risk**. Considering legacy systems today, AI-assisted analysis and reverse engineering can accelerate code understanding and vulnerability discovery, but this does not make these systems automatically 'undefendable'. Further complexity is added by the already existing backlog of vulnerabilities and technical debt. System hardening, network segmentation, security monitoring, and staged modernisation can still reduce risk. In addition, these transformational needs may push small and medium organisations towards broader cloud adoption as entities look at whether the risk, operational constraints, and maintenance burden justify continued operation versus retirement or replacement.
- **The Verification Bottleneck and the 'Truth' Crisis**. While Advanced AI models may significantly reduce false positives, the sheer volume of 'true' reports creates a verification bottleneck. As pointed out in its studies, ENISA notes that the quality of vulnerability information is often insufficient for rapid action. Technical teams now face a challenge of scale, that is how to audit and validate AI-generated patches so that they do not introduce new vulnerabilities into the codebase. An additional burden for technical teams should be expected, as low-risk vulnerabilities can no longer be deemed harmless, as they can be chained to create working exploits.
- **N-Day Weaponisation**. Public vulnerability disclosure and patch releases can still provide attackers with useful information, because patch diffs[10] and binary changes have historically enabled reverse engineering and exploit development. ENISA's vulnerability management guidance recognises that the exploitation window can begin before patch deployment and that exposure often persists while organisations test, coordinate, and schedule updates. Such strategies must, however, be pursued within the constraints set by some of the mission critical systems deployed by Operators of Essential Services. In the short term, such organisations will need to invest in proactive security measures in the absence of a patch and/or methods for faster patch consumption.
- **The Secure SDLC revolution**. There is an opportunity to shift the focus from traditional vulnerability and patch management towards the stronger enforcement of Secure Software Development Life Cycle (SSDLC) practices, including more comprehensive assessments of potential vulnerabilities throughout the development process. This needs to happen across all phases of product development in order to minimise the introduction of vulnerabilities from the outset. At the same time, integrators and end users should be enabled to detect vulnerabilities as early as possible, particularly those arising from combinations of configurations and system integrations. Regarding vulnerability detection and

9 ENISA - Economics of vulnerability disclosure (accessed 18 June 2026)
10 Patch Diffing | CVE North Stars (accessed 18 June 2026)

remediation, the objective should be to move from a model based primarily on oversight towards one based on the integration of security directly into development and operational processes at company level. This also requires a clear focus on increasing the accessibility and adoption of relevant security technologies and practices, and the reskilling and upskilling of the existing workforce.

- **AI-Generated Vulnerability Reports**. By early 2026, AI-generated vulnerability reports had begun to overwhelm parts of the open-source disclosure pipeline.A global vulnerability disclosure coordination and bug bounty platform paused new Internet Bug Bounty submissions[11] after a surge of mixed-quality AI-assisted reports, while curl shut down its CVD programme in January 2026 because maintainers could not sustainably absorb the reporting volume and noise. The underlying problem was not just volume, but signal dilution meaning that human reviewers were spending increasing amounts of time separating real issues from repetitive, superficial, or poorly validated submissions. That shift turns vulnerability reporting from a discovery channel into a triage bottleneck, where the limiting factor becomes reviewer capacity rather than the supply of findings. However, according to recent ENISA discussions with industry, the last few months have seen an important increase in AI report quality, therefore there is a large amount of 'signal' currently being reported.
- **The Risk of 'Inside-Out' Attacks**: Distinct from the vulnerability-exploitation paradigm above, adversaries are increasingly compromising the AI and software supply chain to land inside infrastructure directly, for example by arriving through a trusted software update or a compromised open-source dependency. In these scenarios, the attacker never needs to breach perimeter defences as they are inside the environment and only need to move laterally to access and exfiltrate data. This threat model demands detection and response strategies that go beyond perimeter-oriented defences and assume that adversaries may already be present within trusted environments.

11 https://www.darkreading.com/application-security/ai-led-remediation-crisis-prompts-hackerone-pause-bug-bounties (accessed 18 June 2026)

Security fundamentals matter more than ever

Security fundamentals will not change, but new AI models are compressing the entire attack lifecycle, from reconnaissance to lateral movement, in ways that stress-test these fundamentals, forcing defenders, manufacturers and service providers to accelerate their cybersecurity initiatives.

1. Vulnerability management

To defend against lightning-fast automated cyberattacks, critical service providers may be forced to use autonomous patching, as organisational approval is simply too slow. However, this creates a new risk: automated updates could accidentally break systems and cause unwanted downtime. Furthermore, verifying these AI-generated patches may become a major bottleneck. System managers may face challenges to build tools that can quickly check and approve these fixes without accidentally introducing new bugs into the code.

2. Incident response

For defenders, incident response becomes a race against AI-orchestrated campaigns, with SOCs needing to validate intrusions against live telemetry potentially within hours or minutes. Research[12] indicates that in 75% of breaches, logging existed that should have flagged anomalous behaviour but signals were fragmented across different tools and not acted upon. That gap was manageable when attacks moved at human speed but at AI speed, it will become untenable.

Service providers, including newly expanded providers under the scope of NIS2, must harden operations against AI-scale threats such as adaptive phishing, model poisoning, and supply-chain pivots. Risk assessments should blend logging and monitoring requirements with runtime AI guards, access controls, and threat intelligence sharing, to enable real-time detection and response.

ENISA's Technical Implementation Guidance on NIS2 Risk Management[13] reinforces this by operationalising documented risk frameworks, incident handling policies, supply-chain security, threat hunting capabilities, processes now under pressure by the speed and autonomy of AI.

12 Palo Alto Networks, 2025 Unit 42 Global Incident Response Report (accessed 18 June 2026)
13 ENISA – Technical Implementation Guidance (accessed 18 June 2026)

A parallel constraint is the ability of national authorities to absorb and act on incident data at scale. Mandatory incident reporting, combined with a broader scope of regulated entities, will sharply increase reporting volumes, just as AI accelerates vulnerabilities, exploitation, and breaches. This creates a structural risk of simultaneous surges in incidents exceeding national capacity. Proactive scanning of critical infrastructure by National CSIRTs becomes even more relevant.

In such scenarios, key questions include how to prioritise incidents when they all meet regulatory thresholds, which CSIRTs structures can operate without prioritisation, and how sectoral CSIRTs can be supported when national capacity is saturated. Governments should act and integrate industry into its crisis planning, establishing surge capacity models, defining clear prioritisation and escalation frameworks, and stress-testing multi-incident scenarios.

3. Product Security

As the frontier AI capabilities will enable software developers, owners and manufacturers to secure their code in the long run, the immediate problem is the security of legacy products and in particular open-source components which form critical part of the technology stack in use. There is also the risk that Union's small manufacturers will not get access to software/product testing capabilities enabled by new AI models due to costs.

Manufacturers must adjust their product lifecycles for compliance with the Cyber Resilience Act (CRA). Developers of AI models should train their models to be structurally constrained (stronger than a guardrail), and ensure that such models are suggesting code that is secure by design/default. ENISA's proposal for machine-processable security attestations and implementation of SBOM strategies can operationalise the CRA security requirements; embed threat modelling, least privilege, and lifecycle attestations from the outset.

4. Talent and human capital

A cross-cutting challenge will be to ensure that foreseen solutions will keep the human in the middle. While AI can help build system fasters, or defend networks better, humans will still need to understand the environments they assume responsibility for. Cybersecurity risks related to AI will require raising the level of awareness, understanding and skills, across all levels of organisations, substantially faster.

The way ahead: building AI-resilient defences

Frontier AI models demand a fundamental shift if defenders are to achieve operational parity with, or stay ahead of attackers through structured frameworks and more adaptive practices. The recommendations below set a clear **direction** for what organisations should aim for through European coordination, national enforcement and defender operations. They are **neither exhaustive nor mutually exclusive**.

The CRA's **Single Reporting Platform** (SRP) and EU Vulnerability Database (EUVD) can help support the operationalisation of some of these recommendations. For national authorities and defenders, the development of new frameworks and the reinforcement of existing platforms can function as force multipliers. For example, correlated EUVD data can help identify supply-chain hotspots, systemic weaknesses and vulnerability trends accelerated by AI-assisted discovery. CRA-related reporting may help validate vendor hardening claims against real-world exploitability and incident patterns. When combined, over time, NIS2 incident reporting and other situational awareness reports, can help close the gap between vulnerability disclosure to coordinated response.

Recommendations

- **European Level:**
  - At the European level, the existing legal frameworks, including NIS2, CRA and the EU AI Act[14] should be leveraged, to ensure that systemic risks stemming from the most advanced AI models are assessed and mitigated. To this end, it may be useful to establish EU-wide state-of-the-art benchmarks for the security evaluation of advanced AI models, including standardised testing against cyber ranges, exploitability metrics, and simulations of chained attacks, to set a consistent bar for assessing their capabilities.
  - Leverage and build on existing initiatives at European and Member State level that consider the use of future AI models for cybersecurity.
  - European institutions, national authorities, CSIRTs, regulators and operators of essential services hold cybersecurity data that could be highly valuable for training, evaluating and fine-tuning defensive AI systems. This includes incident reports, vulnerability disclosures, telemetry, malware samples, abuse reports and sectoral threat intelligence. Subject to strong safeguards, anonymisation, access controls and clear legal bases, such datasets could become a European strategic asset. They could support trusted European cyber-AI models and serve as leverage in procurement negotiations.
  - In accordance with the AI Act and the Code of Practice for General Purpose AI, adequate risk mitigations for AI models with advanced cyber capabilities need to be identified.
- **National Authorities:**
  - Run AI-powered threat hunting operations and publish anonymised datasets from Frontier AI simulations.
  - Require critical infrastructure operators to attest zero-trust baselines, resilient incident response capabilities during annual audits, and clear escalation paths for AI-accelerated incidents.
  - Direct government agencies and critical infrastructure operators to reduce ineffective security tooling and complement their cybersecurity stack with AI-enabled platforms capable of ingesting, correlating, and acting on threat data.
  - Develop common frameworks for the deployment of AI-enabled security tools that prioritise human oversight, auditability, and human-gated triage for decision making.
  - Develop evaluation capacities for products including AI-functionalities, in order to ensure and validate security levels to face cybersecurity threats
  - Proactively scan critical infrastructure components that fall under the responsibilities of national authorities.

14 Article 55 of the EU AI Act requires providers placing general-purpose AI models with systemic risk on the EU market to assess and mitigate those risks, regardless of where they are based. Through the GPAI Code of Practice, providers commit to identifying systemic risks, including large-scale cyberattack risks, applying safety and cybersecurity measures across the model lifecycle, and reporting relevant information and evaluation results to the European AI Office. These rules have applied since 2 August 2025 and will be enforced from 2 August 2026, with the Office empowered to investigate compliance, require mitigation measures, impose fines of up to 3% of global annual turnover and, in extreme cases, restrict, withdraw, or recall models from the EU market.

- **Defenders:**
  - Treat every environment as potentially already compromised and extend endpoint protection across all environments. Every organisation should conduct an analysis of its exposure to frontier AI and deploy real-time, ML-based prevention and detection on all on-premises and cloud hosts, including securing agentic endpoints and enterprise browsers, which represent critical new attack surfaces as AI agents increasingly operate autonomously on endpoints and within browsers.
  - Embed structured threat modelling early and continuously, leveraging advanced AI models with precise context to uncover combinations of weaknesses and run continuous simulations against live architectures, effectively turning AI into an always-on support capability for red teaming and defensive validation.
  - Build more dynamic incident response pipelines, using AI-assisted triage for real-time telemetry validation, prioritisation, and blast-radius containment, while keeping human review firmly in the loop to reliably hit 24-hour notifications.
  - Transform security operations into a near real-time function, targeting single-digit-minute mean-time-to-detect (MTTD) and mean-time-to-respond (MTTR) metrics.
  - Accelerate zero-trust segmentation, behavioural baselining, and more evasion-resistant detection layers that assume attackers will adapt and evolve in near real-time.

TP-01-26-015-EN-N

ABOUT ENISA

The European Union Agency for Cybersecurity, ENISA, is the Union's agency dedicated to achieving a high common level of cybersecurity across Europe. Established in 2004 and strengthened by the EU Cybersecurity Act, the European Union Agency for Cybersecurity contributes to EU cyber policy, enhances the trustworthiness of ICT products, services and processes with cybersecurity certification schemes, cooperates with Member States and EU bodies, and helps Europe prepare for the cyber challenges of tomorrow. Through knowledge sharing, capacity building and awareness raising, the Agency works together with its key stakeholders to strengthen trust in the connected economy, to boost resilience of the Union's infrastructure, and, ultimately, to keep Europe's society and citizens digitally secure. More information about ENISA and its work can be found here: www.enisa.europa.eu.

ISBN 978-92-9204-801-3

ENISA
European Union Agency for Cybersecurity

Athens Office
Agamemnonos 14
Chalandri 15231, Attiki, Greece

Brussels Office
Rue de la Loi 107
1049 Brussels, Belgium

enisa.europa.eu

Publications Office
of the European Union