



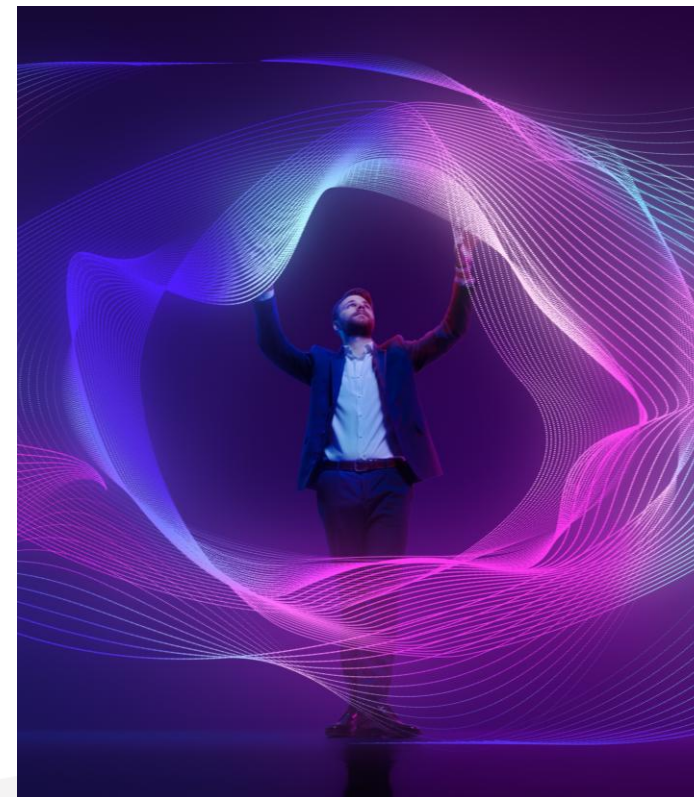
EUROPEAN UNION AGENCY
FOR CYBERSECURITY

Cyber on steroids, game changer, or light at the end of the tunnel?

Nuno Carvalho, ENISA

5 Prediction for 2027 (or earlier)

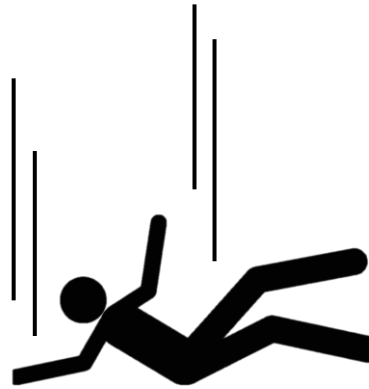
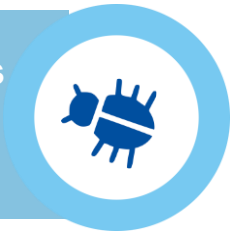
1. The current frontier models **will look dumb**
2. An **open source / open weight** equivalent of MYTHOS 5 will be **available**
3. **All vulnerability finding & patching** will be **done by AI**
4. **95%-99%** of the **code** will be **written by models**
5. The first examples of **superhuman attacks** (on cryptography, physics, exploits)



The opportunity and the risk

OPPORTUNITY

Finding Vulnerabilities
Creating Patches:
Faster & Cheaper



RISK

Patching Vulnerabilities:
Slow & Human in the Loop



Vulnerabilities
are now found
very fast ...
**But they aren't
patched at the
same speed.**

Ai accelerates vulnerability pressure



01 · OPEN-SOURCE EXPOSURE

Projects and SBOMs

Open-source ecosystems face an upgraded threat as discoverability and attack reach expand.



02 · LOWER BARRIERS

AI amplifies attackers

AI lowers the barrier to entry for malicious actors and gives them disproportionate capabilities.

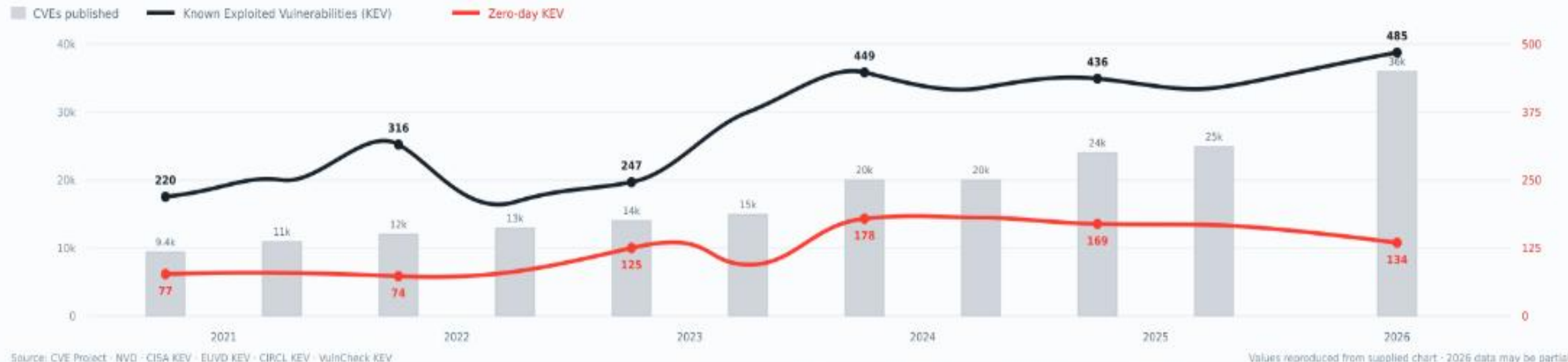


03 · DEFENSIVE PRESSURE

Defenders cannot catch up

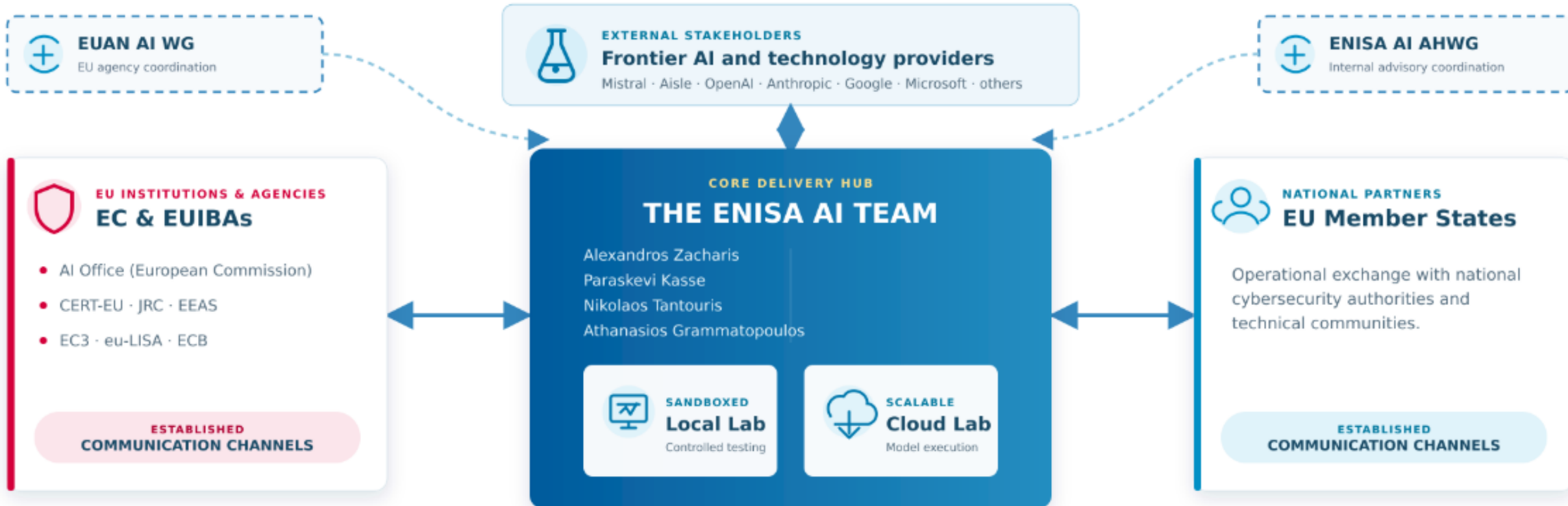
Traditional capacity cannot keep pace with the new rate of vulnerability discovery.

Published vulnerabilities and known exploitation signals



We are not alone!

STAKEHOLDER AND DELIVERY NETWORK



The role of ENISA under the EU AI Action Plan

Pillar 1



Making frontier AI safe, accessible & deployable

Establish sovereign evaluation infrastructure via access to frontier & open models.

Pillar 2



Preparing the EU's cyber ecosystem

Guidance, recommendations, advisories and more practical advice.

Pillar 3



Scaling European AI capabilities for cyber

Building capacity both in technology and human capital.

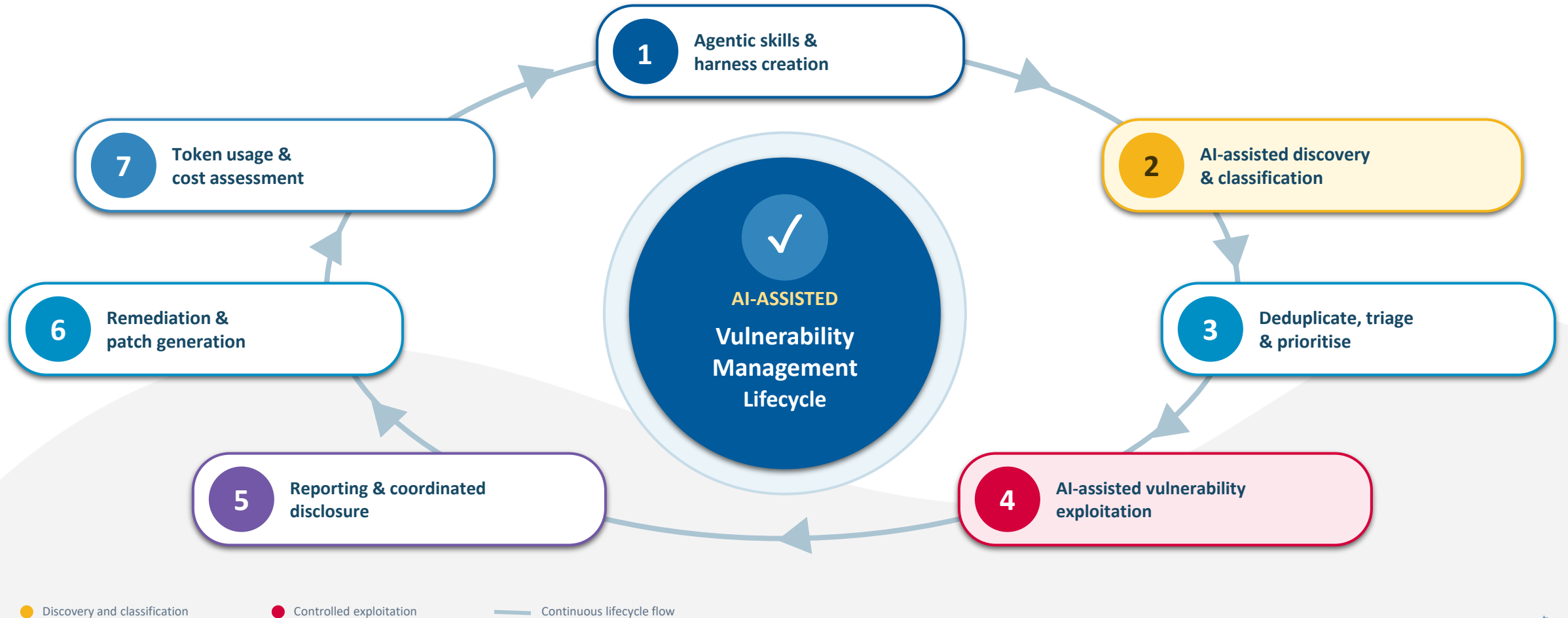
- Help create a **Contingency Plan** under the European Blueprint along with EC
- Build together with Joint Research Centre (JRC) a **secure testing platform for AI** in cybersecurity use cases, including "cyber ranges" starting with the first cases in Q4 2026) and expanding them in the next 2 years

- **Issue guidance, recommendations, advisories and best practices** on protecting against AI-powered threats and secure integration of AI into cyber operations
- Help **modernise vulnerability management** practices/tools for the AI era
- Launch a **pilot Critical Open Source Resilience Campaign** to accelerate patching using AI (Q4 2026)

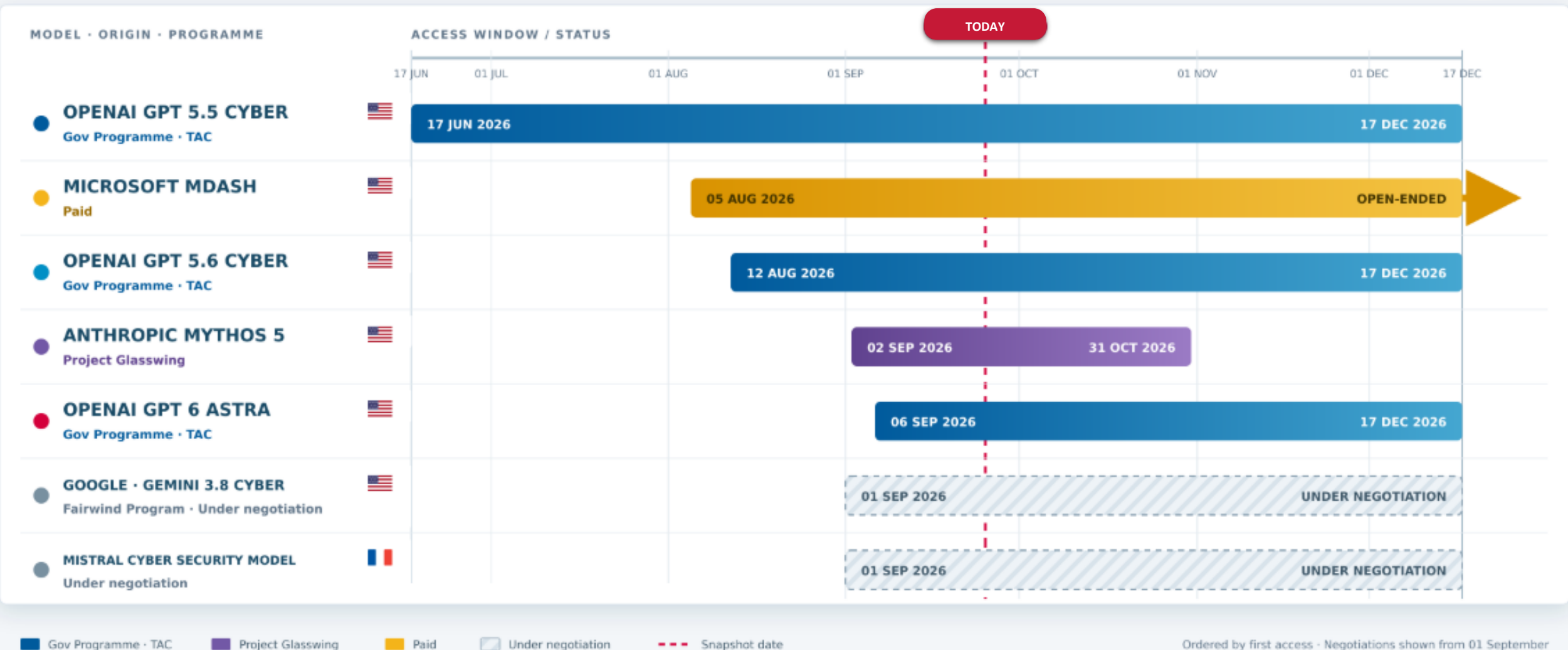
- Contribute to the launch an **EU Grand Challenge** on AI-assisted vulnerability remediation (Q4 2026)

ENISA's Cases Studies & Experiments

Seven-step AI-assisted vulnerability management testing cycle



Our Access to Frontier AI Models



Observations on AI Powered Vulnerability Management



DOING IT RIGHT

Safe execution + complete context + human validation = actionable security findings



01 · CONTROL THE ENVIRONMENT

Secure operation

- Sandbox AI harness deployment
- Apply least-privilege access
- Use trusted skills and plugins
- Control cloud data transmission
- Design for supervision fatigue

SAFER TESTS · FEWER ACCIDENTS



02 · SUPPLY THE RIGHT CONTEXT

Context is key

- Define project scope and purpose
- Explain user roles and trust levels
- Describe the deployment model
- Include external security controls
- State accepted and out-of-scope risks

MORE TRUE POSITIVES · GREATER ACCURACY



03 · VALIDATE AND SHAPE THE OUTPUT

Human in the loop

- Remove duplicates and out-of-context results
- Filter, reduce and prioritise high-volume output
- Validate relevance, severity and impact
- Tailor reports to the target audience
- Improve output while reducing human overhead

FASTER RESOLUTION · HAPPIER MAINTAINERS

Our Delivery Roadmap

DELIVERY TIMELINE

SEPTEMBER



OCTOBER



NOVEMBER



DECEMBER



ADVISORY

KEY ACTION 4

**Technical Advisory on
AI-Assisted Software
Development**

SEPTEMBER 2026

JOINT ADVISORY

KEY ACTION 4

**Joint Technical Advisory
on the Use of Frontier AI
for Vulnerability
Discovery**

OCTOBER 2026



REPORT

KEY ACTION 4

**Current State of Agentic AI
in Cybersecurity Operations
& Skills**

ENISA GITHUB REPOSITORY

ADVISORY

KEY ACTION 4

**Technical Advisory:
Defending Against
AI-Powered Threats**

NOVEMBER 2026

CAMPAIGN

KEY ACTION 6

**Pilot Critical Open Source
Resilience Campaign**

DECEMBER 2026

PLATFORM

KEY ACTION 3

**Secure Testing Platform
for AI in Cybersecurity
Use Cases**

INCLUDING CYBER RANGES
DECEMBER 2026



Thank you for your attention

ENISA

European Union Agency
for Cybersecurity

Athens Office

Agamemnonos 14
Chalandri 15231, Attiki, Greece

Brussels Office

Rue de la Loi 107
1049 Brussels, Belgium

ai@enisa.europa.eu

Our Lab & Harness

SAFE EXPERIMENTATION LAB – FOUR DESIGN QUESTIONS

AI MODELS

Which models are appropriate? Open or closed weight, general-purpose or cyber-specialised, public or restricted access.

MODEL EXECUTION

Where will the model run? Local, corporate or provider cloud; what data can leave the environment; what guardrails and resource limits apply.

AI CLIENT

What can the agent access? Files, shell, tools, APIs or external services; what privileges are necessary; how are usage and costs monitored.

WORKFLOW

How is work made repeatable? Controlled prompts, approved skills/plugins, standard inputs, evidence capture and reporting templates.

