# Securing AI

## From the AI Act to the Next Wave of European Regulation and Beyond

Michaela Klopstra
Brussels, 09.06.2026

# Why This Matters Now

# The AI Security Inflection Point

- AI adoption accelerating across all sectors
- New attack surface + systemic risk
- Shift: AI → regulated technology

# AI-specific threat landscape

# AI-native threats (examples)

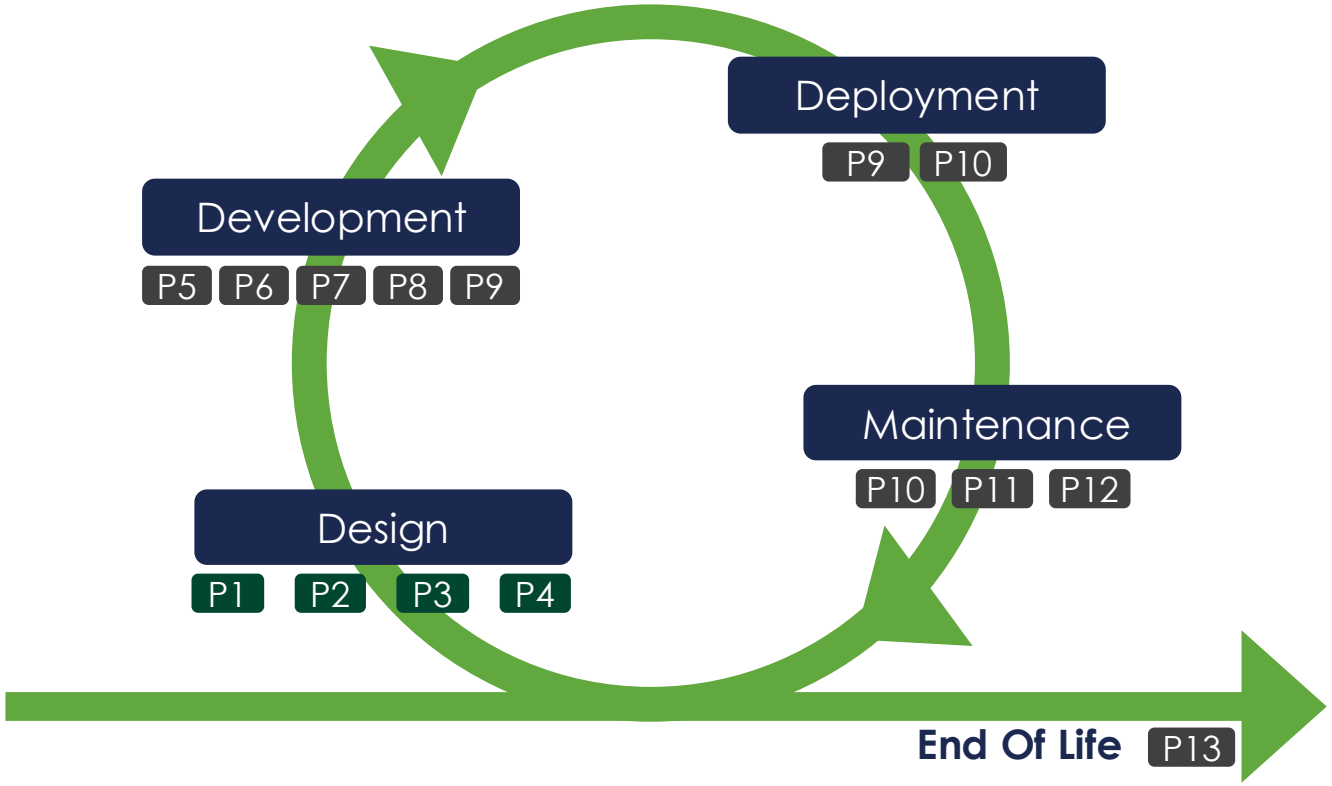
- Prompt injection & adversarial inputs
- Data poisoning (training-time attacks)
- Model extraction / exfiltration
- Output manipulation & hallucination exploitation
- Privacy leakage / inference attacks

# Supply chain risks (examples)

- Vulnerabilities in training data, pre-trained models, dependencies
- Third-party components massively increase attack surface
- Attacks propagate across ecosystems (e.g. compromised model/artifacts)

# Securing the AI lifecycle

# Security must be end-to-end




| | |
|---|---|
| P1 | Raise awareness of AI security threats and risks |
| P2 | Design your AI system for security as well as functionality and performance |
| P3 | Evaluate the threats and manage the risks to your AI system |
| P4 | Enable human responsibility for AI systems |
| P5 | Identify, track and protect your assets |
| P6 | Secure your infrastructure |
| P7 | Secure your supply chain |
| P8 | Document your data, models and prompts |
| P9 | Conduct appropriate testing and evaluation |
| P10 | Communication and processes associated with End-users and Affected Entities |
| P11 | Maintain regular security updates, patches and mitigations |
| P12 | Monitor your system's behaviour |
| P13 | Ensure proper data and model disposal |

Source: ETSI TR 104 128 „Securing Artificial Intelligence (SAI); Guide to Cyber Security for AI Models and Systems“

# The AI Act

# Key concepts

- Risk-based framework: obligations scale with risk

4 risk categories:

- Unacceptable → prohibited
- High-risk → strict regulation
- Limited → transparency
- Minimal → no obligations

- Extraterritorial scope: applies if AI affects EU users

# Obligations

Prohibited AI (ban)

- Social scoring, harmful manipulation, exploitation of vulnerable groups

High-risk AI (core obligations)

- Risk management system (lifecycle-based)
- Data quality & bias control
- Technical documentation & logging
- Transparency & human oversight
- Accuracy, robustness, cybersecurity
- Conformity assessment + EU registration
- Post-market monitoring & incident reporting

Limited-risk AI

- Disclosure: user must know they interact with AI
- Labelling (e.g., deepfakes, chatbots)

GPAI models

- Documentation, training data summary, copyright compliance
- Additional testing & risk controls for systemic models

# The Next Wave of European Regulation

# To consider

- AI Act Omnibus
  - clarifies existing AI Act requirements
  - extends compliance deadlines for high-risk AI systems (HRAIS)
  - introduces new rules on AI-generated intimate content
- Guidelines and Code of Practice (transparency requirements)
- Cyber Resilience Act
- Cybersecurity Act 2
- EU Cloud & AI Development Act
  - Security, sovereignty & resilience of AI/cloud infrastructure
- AI Liability Directive?

# Beyond

# What to expect

- Continuous, real-time governance of AI system
- Continuous monitoring obligations (not just post-market)
- Runtime compliance / “AI observability”
- Automated audit & reporting
- Integrated trust framework
- Infrastructure security + geopolitical risk management

# What organizations should do now

# Act

- Inventory and classify AI systems
- Introduce AI risk management framework
- Integrate AI into Cybersecurity governance and Supply chain security
- Secure the full AI lifecycle
- Prepare for:
  - Documentation
  - Testing
  - Continuous monitoring

# Thank You