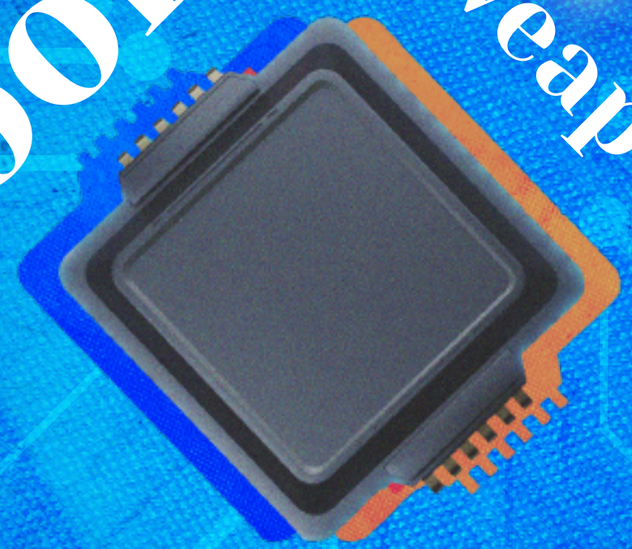
AI & CYBERSECURITY

tool or weapon



AI can predict attacks and malware before they actually occur based on training data from past attacks. But, AI can also generate false positives, or be manipulated by attackers, which means human oversight is a must.

AI can automate the identification of vulnerabilities in IT systems and generate malware and malicious code that can be adapted to different environments without requiring much skill on the part of the attackers.

Increasingly, IT systems are deployed with phishing-detection or prediction mechanisms based on AI analysis. However, AI detection is not perfect, so it is important to always check the recipient, time and subject of the email.

Highly personalised and convincing phishing messages can be created using generative AI. In addition, AI systems can autonomously maintain interactions with victims over time until the attack is successful. This can allow attackers to carry out multiple attacks in parallel.

AI can predict and automate responses to certain types of threats, such as isolating a compromised system or blocking malicious IP addresses. This can reduce response times and minimise damage.

AI systems can also be compromised or exploited by attackers to gain unauthorised access to data or other systems. As AI-powered services become more accessible to the public, without taking security into account, this risk will become more prevalent.

AI can scale to handle large volumes of data and security events, predict attacks, and potential information leaks, which is particularly useful for organisations with extensive and complex IT environments.

Company employees might resort to external/public AI services that are not secure, leading to potential sensitive information and personal data breaches. Additionally, malicious links inside AI responses may lead to cybersecurity breaches.

EUROPEAN CYBER SECURITY MONTH

EDPS

edps.europa.eu