Quick Guide: AI Attack Protection

Practical Steps to Reduce AI-Enabled Security Risk

By: Shawn Awan, SkySec Founder & Lead Assessor

What is AI?

Artificial Intelligence (AI) refers to systems that can analyze data, recognize patterns, and generate responses or actions that appear human-like.

Many organizations already use AI in email filtering, chatbots, analytics, and automation tools..

How is AI Weaponized?

Attackers use AI to make traditional attacks faster, more convincing, and harder to detect. AI doesn't create new threats—it amplifies existing ones.

Common weaponized uses include:

- *Writing highly realistic phishing emails and messages*
- *Generating fake voices, images, or videos (deepfakes)*
- *Automating password attacks and reconnaissance*
- *Mimicking employee communication styles*

Why this Matters? (Impact)

Most ransomware doesn't come from the malware itself – it's usually due to failing to prepare, unclear ownership, and false confidence in tools or reports.

Actions to Take Today (Checklist)

1. *Harden identity controls by enabling MFA & restricting privileged (admin) access*
2. *Strengthen verification procedures by requiring 2nd verification for payments & password resets*
3. *Train staff on AI-enabled social engineering – focus on realistic scenarios, not generic ones*
4. *Control AI usage internally by defining which AI tools are allowed and not allowed*
5. *Monitor for abnormal behavior such as Model Context Protocol and injection patterns*

***Bonus Tip:** Conduct internal AI pipelining & indirect prompt injection vulnerability testing!*

Common Mistakes (Reason for Failures)

- *Treating AI threats as futuristic instead of current*
- *Assuming AI attacks only affect large enterprises*
- *Allowing unrestricted AI tool use with sensitive data*
- *Relying solely on technology without process controls*

SkySec finds gaps focused on real-world attack paths, including AI-enabled threats.
Contact us www.skysecsecurity.com

SkySec