CHEAT SHEET

TANIUM

Technology Primer

Top 10 risks & remedies for adopting Generative AI

Mitigating the risks of AI adoption requires a combination of robust cybersecurity practices, secure development methodologies, regular model audits and user education. For IT leaders, these are six areas that should command most of your attention.

Data Privacy and Protection

Risk: Multimodal generative AI often requires access to large datasets, potentially containing sensitive or personal information. Unauthorized access or data breaches can lead to privacy violations.

Remedy: Implement robust data encryption, access controls, and data anonymization techniques to protect sensitive data. Ensure compliance with relevant data privacy regulations.

Model Security

Risk: Multimodal AI models may be susceptible to attacks, including adversarial attacks or model inversion attacks, compromising their integrity.

Remedy: Regularly audit and evaluate AI models for bias. Implement bias-reduction techniques and promote fairness in AI development.

Adversarial Attacks

Risk: Malicious actors may attempt to manipulate multimodal AI systems by providing deceptive inputs, leading to malicious outputs.

Remedy: Implement robust input validation and anomaly detection to identify and prevent adversarial attacks. Train models to recognize and reject malicious inputs.

Supply Chain Risks

Risk: Multimodal AI models can be vulnerable to supply chain attacks, where attackers manipulate training data or models during development. LLM poisoning next level of risk.

Remedy: Verify the integrity of the AI supply chain, including data sources, model development, and deployment processes, to prevent malicious tampering.

Prompt Injection

Risk: Malicious actors may inject harmful or inappropriate prompts into generative AI systems, leading to the generation of harmful or undesirable content.

Remedy: Impletement strict input validation and filtering mechanisms to prevent prompt injection. Ensure that only authorized and validated prompts are accepted.

Prompt Jailbreaking or Unauthorized Control

Risk: Attackers may attempt to exploit vulnerabilities in generative AI models to gain unauthorized control or manipulate their behaviour.

Remedy: Continuously monitor and update AI models to address security flaws and vulnerabilities that could lead to jailbreaking or unauthorized access.

In the race for AI supremacy, basic data and IT hygiene matters. Tanium helps you manage this in the following ways:

Visibility & Control

Real-time data is required for next-gen AI to operate, so you will need cyber tools that can assess your asset lifecycle, deal with patch management and correlation of your CMDB. Learn more about Tanium here.

Continuous Safeguarding

Tanium is a platform that helps you manage multi-factor authentication, vulnerability assessment, supply chain management and data security. Discover more about Risk & Compliance here.

Detections & Countermeasures

Tanium can be used to test your response plans, carry out testing war games and provide table-top active hunting of cyber threats lurking undetected on the network. Find out more about Incident Response here.

Tanium delivers the industry's only true real-time cloud-based endpoint management and security offering. Its converged endpoint management (XEM) platform is real-time, seamless, and autonomous, allowing security-conscious organizations to break down silos and reduce complexity, cost, and risk. Securing more than 32M endpoints around the world, Tanium's customers include more than 40% of the Fortune 100, 7 of the top 10 U.S. retailers, 9 of the top 10 U.S. commercial banks, all 6 branches of the U.S. military, and MODs and DODs around the world. It also partners with the world's biggest technology companies, system integrators, and managed service providers to help customers realize the full potential of their IT investments. Tanium has been named to the Forbes Cloud 100 list for eight consecutive years and ranks on the Fortune 100 Best Companies to Work For. For more information on The Power of Certainty™, visit www.tanium.com and follow us on LinkedIn and X.