Deloitte.

Securely adopting and scaling AI

Organizations are being tasked with leveraging artificial intelligence (AI) to achieve aggressive goals around operational efficiency, customer engagement, revenue, and profit. At the same time, AI introduces new security risks that can derail organizational goals if not properly managed.

Cyber for AI emphasizes the need for security to be integrated throughout the AI journey—from intake to secure development and testing, supported by governance and foundational cyber capabilities to manage AI risks—facilitating alignment with business growth objectives to leverage AI's transformative potential.

AI introduces new threats—and challenges executives to act

As AI evolves and adoption rapidly expands, it creates different security challenges, making it harder for business leaders to design strategies that effectively counter emerging threats.

Concerns for executives	
Shadow AI	**Model jailbreak**
Supply chain attacks	**Copyright violations**
Data poisoning	**AI model access**
Agent memory poisoning	**Bias and discrimination**
Sensitive data exfiltration	**Insecure agent**

...and many more

Common AI threats
What if unsecured AI turns a strategic investment into a financial disaster?
What if an AI misstep undermines our risk management efforts?
What if an AI vulnerability derails our digital transformation?
What if a single cyberattack jeopardizes the organization?

Deloitte blueprint to securing AI

Organizations need to act now and set the pace for long-term success—securing AI transformation requires a concerted effort across the organization and throughout the AI life cycle beginning with understanding how you consume AI, the different threats that shape your landscape, controls and guardrails required, and how these can be implemented **to derive sustainable value from AI amid rapid AI adoption**.

Understand your AI consumption patterns and use cases

Cybersecurity controls differ based on how you consume AI:

- **Internal AI:** Custom AI app developed or hosted in environment (e.g., internal AI chatbot)
- **External AI:** E.g., AI-powered productivity tools

Analyze the AI threats your organization could be exposed to

AI introduces new and unfamiliar attack surfaces, while intensifying existing attack surfaces with threats to the:

- **Application**
- **Data**
- **Model**
- **Infrastructure**
- **Life cycle**

Identify AI security controls and guardrails you need to manage AI threats

Mitigate AI risks by identifying essential security controls and guardrails, leveraging leading industry frameworks such as:

- **NIST AI RMF[1]**
- **ISO[2] 42001**
- **OWASP[3] LLM[4]**
- **OWASP ML[5]**
- **EU[6] AI Act**

Enforce identified security controls across organizational functions and domains

Enforcement of AI security and governance requires a concerted effort across the organization, **starting with leadership guidance** and continuing through to **implementation at the technology stack across cybersecurity domains** such as **identity and access management (IAM), data security, and cloud security**.

AI governance

Enhance the **governance of AI**

AI SSDLC[7]

Embed security leading practices throughout the **AI life cycle**

Secure AI tech ecosystem

Bolster foundational **cybersecurity tooling and configurations**

1. National Institute of Standards and Technology Artificial Intelligence Risk Management Framework. 2. International Organization for Standardization.
3. Open Web Application Security Project. 4. Large Language Model. 5. Machine Learning. 6. European Union. 7. AI Secure Software Development Life Cycle.

Connect to accelerate

Contact our leaders to dive deeper into the blueprint and reimagine what's possible for your organization.

Mark Nicholson
Principal
Cyber AI GTM Leader
Deloitte & Touche LLP
manicholson@deloitte.com

Naresh Persaud
Principal
AI Transform Leader
Deloitte & Touche LLP
napersaud@deloitte.com

Chris Knackstedt
Managing Director
Cyber for AI Leader
Deloitte & Touche LLP
cknackstedt@deloitte.com

Mehdi Houdaigui
Principal
Cyber Strategy & Transformation Leader
Deloitte & Touche LLP
mhoudaigui@deloitte.com

Nirmal Arava
Senior Manager
Cyber for AI Leader
Deloitte & Touche LLP
narava@deloitte.com

Steve Ruzzini
Senior Manager
Cyber AI GTM Activation Lead
Deloitte & Touche LLP
sruzzini@deloitte.com

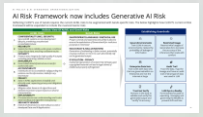
Organizations need to secure AI with strategic guardrails across the three layers of AI governance, AI SSDLC, and the secure AI technology ecosystem

Building upon the threats and security controls identified, securing AI requires a concerted effort across the organization, beginning with clear leadership direction and extending through each layer of implementation, from policy development to integration within the technology stack.

AI governance

- Establish and integrate **AI security and governance structures** (with cross-functional stakeholders and established accountability) into the larger AI enablement efforts.
- Set up **overarching AI security and risk management artifacts and processes** aligned with AI-specific risks and threats so that security is embedded into, and not outpaced by, AI enablement.

AI acceptable use policy, generative AI policy, and so on

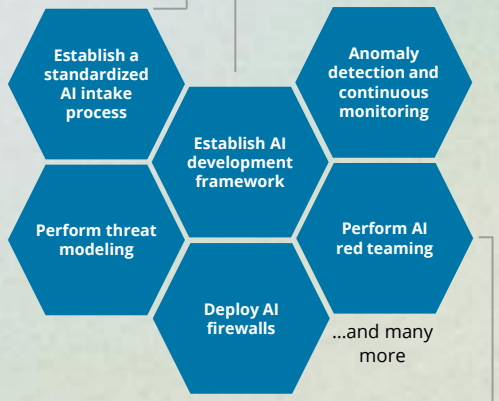


AI SSDLC

- An SSDLC for AI systems requires enhanced controls that address data quality, AI red teaming, model governance, and ongoing risk monitoring, extending well beyond standard software development practices. Security requirements, guardrails, and security review stage-gates are required to **promote and monitor secure AI development and deployment**.

Effectively govern and manage risks around proposed AI solutions

Facilitate secure and scalable deployment of AI models and applications



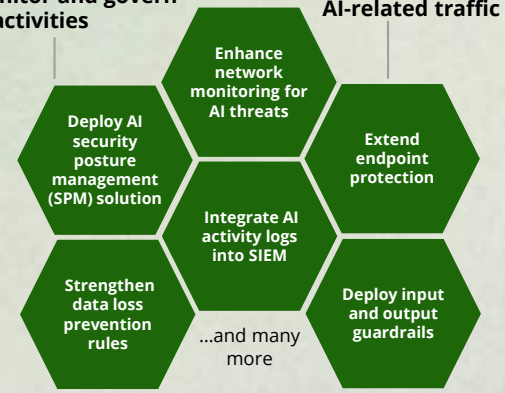
Simulate adversarial attacks and stress-test AI systems to facilitate broad security

Secure AI tech ecosystem

- Securing AI requires organizations to both **strengthen existing security tools and deploy new solutions tailored for AI risk management**.
- Integrate **broad protections across IAM, data, cloud, and each cybersecurity domain** to help strengthen your organization's defenses.

Detect shadow AI; monitor and govern AI activities

Detect anomalous AI-related traffic



Block sensitive data from being shared with external AI chatbots