Trellix

From Shadow AI to Strategic Adoption

Creating an AI-aware culture that drives innovation while securing critical data

Table of Contents

Introduction 3

Understanding AI use and data risk today 4

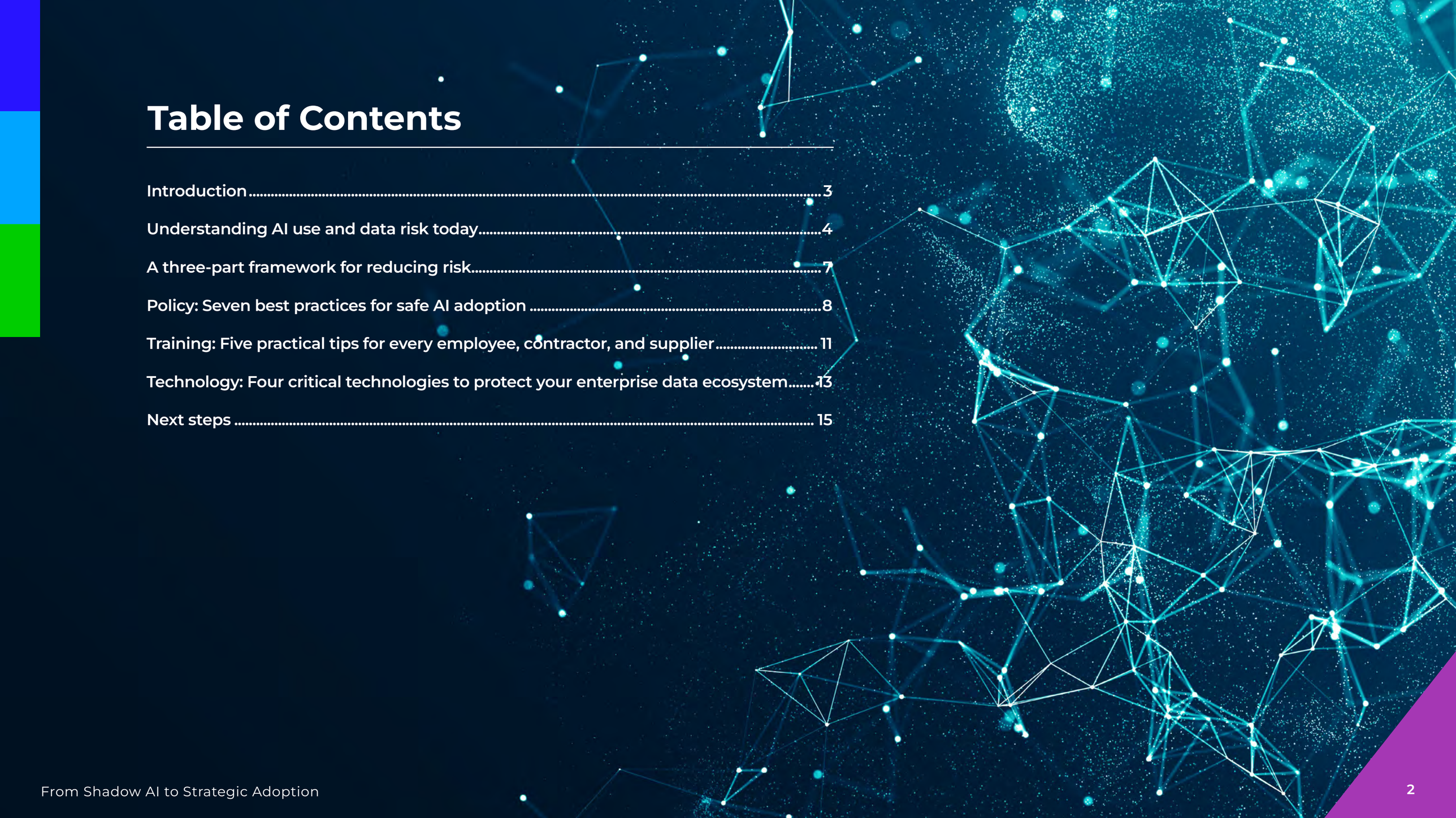
A three-part framework for reducing risk 7

Policy: Seven best practices for safe AI adoption 8

Training: Five practical tips for every employee, contractor, and supplier 11

Technology: Four critical technologies to protect your enterprise data ecosystem 13

Next steps 15



Introduction

AI has been around for decades, steadily advancing in capability up to the relatively recent integration of AI/ML into smart devices such as Alexa and Siri. While those advances were meaningful, none felt like the fast-moving freight train that has been Generative AI (GenAI) since the commercial launch of ChatGPT in late 2022.

The past few years have seen an explosion of GenAI embedded into personal and work activities. Organizations that started with a “block-first” or “block all AI” mentality are now embracing the efficiency gains of AI, which means security teams need to become business partners and technology accelerators, ensuring the safe and effective incorporation of AI across the environment. For data security teams, this presents an even more unique challenge, as the data needed to make GenAI work and the outputs from AI tools may contain sensitive information that requires diligent protection.

The industry has reached a tipping point at which business objectives, user activity, new (and potentially untested) technologies, and regulatory frameworks intersect.

- CISOs and other executive stakeholders are already at their breaking point supporting complex regulatory frameworks, with additional AI legislation planned globally that will further stress their resources.
- Every day, enterprise technologies are updated with built-in AI capabilities that must be tracked and understood.
- Employees, contractors, and suppliers are sharing more data that could be sensitive with a diverse universe of AI tools
- Even unsophisticated bad actors can create effective attacks using AI tactics

This ebook is for cybersecurity and IT leaders, data security professionals, governance and compliance officers, and anyone concerned with reducing risk to their organization’s sensitive data while securely adopting AI. In it, you’ll find a framework and best practices for securely adopting AI for innovation while protecting your critical data against Shadow AI and other AI-enabled threats.

Key eras in AI advancements

AI and machine learning (ML) have existed for 60-plus years.

1950s-1960s: Foundational research

The earliest developments of AI and ML focused on rule-based systems and basic algorithms for data analysis. In 1950, mathematician Alan Turing published “Computing Machinery and Intelligence” and proposed the so-called Turing Test for machine intelligence.

1990s: Complex calculations and pattern recognition

In 1997, IBM’s Deep Blue beat international champion Garry Kasparov at chess. Many consider that an important turning point that brought AI into the public’s view.

2010s: Smart devices

For over a decade, we’ve been interacting with smart devices such as Amazon’s Alexa, Apple’s Siri, and more.

2020s: Generative AI

Based on large language models (LLMs), Generative AI (or GenAI) began to produce new content based on training data. GenAI has been widely adopted since the introduction of OpenAI’s ChatGPT in late 2022, which brought GenAI into the mainstream and made it a popular culture fascination.

Understand AI use and data risk today

AI has become rapidly embedded in enterprises, but security teams have not kept pace with the potential exposure of sensitive data.

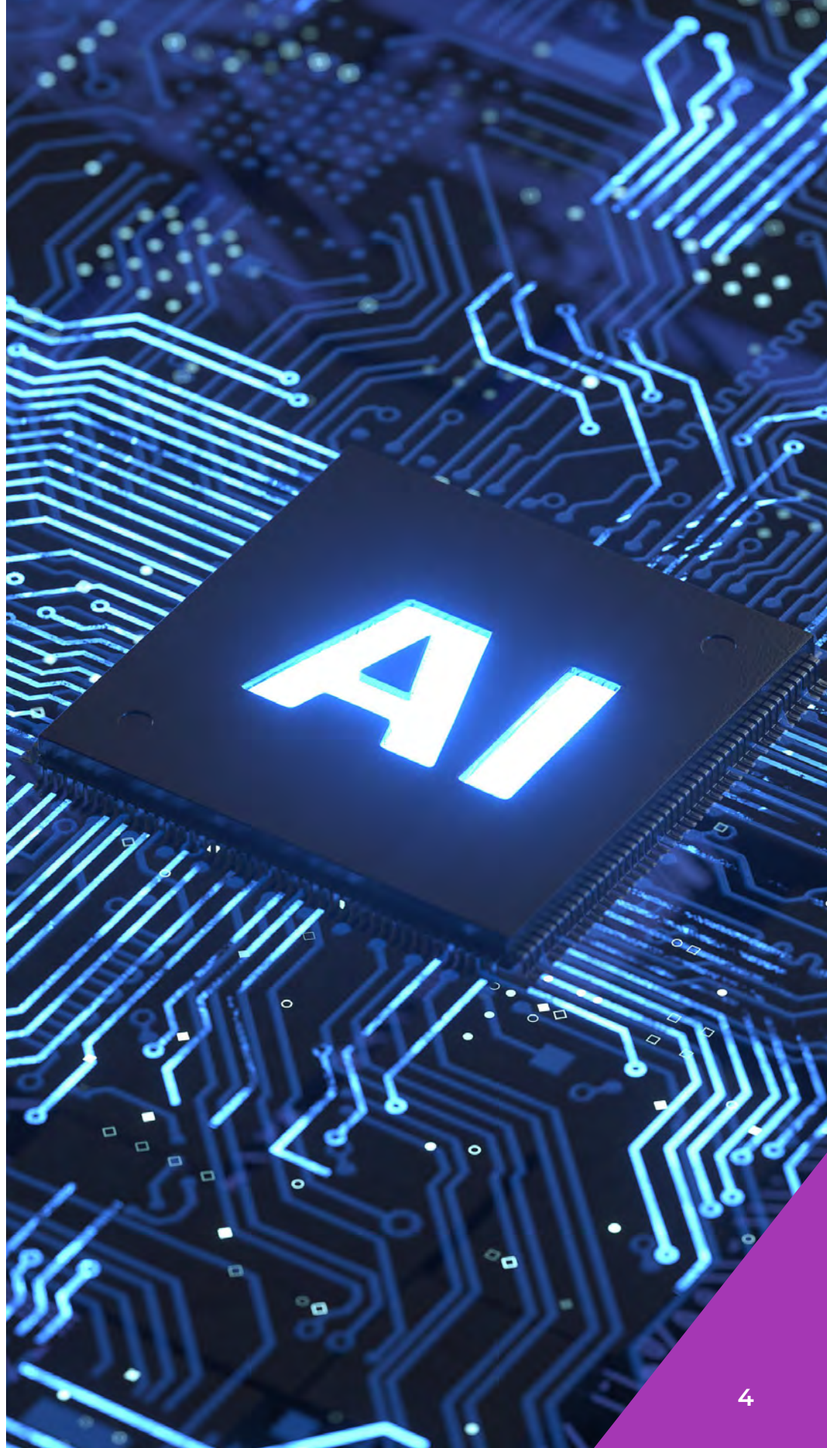
What's driving AI adoption

Enterprises, as well as individual users, are increasingly turning to AI for a multitude of reasons, including improved efficiency, increased processing power, and added resource capacity.

Efficiency: AI helps automate repetitive, cumbersome, and complex processes, freeing up time and resources for more strategic work. It also standardizes and optimizes workflows, which is crucial for scalability in today's fast-paced environments.

Processing: The advent of GenAI has improved capabilities to analyze vast amounts of data, and generate content. It's also becoming invaluable for improving and debugging code and summarizing data to provide meaningful analytics.

Resourcing: And as we transition to more advanced AI, it's transforming how organizations interact with their customers, clients, and partners. AI is being employed to provide customer service, answer medical questions, offer legal advice, and deliver creative content.



Common data risks with enterprise AI use

Security professionals face a daunting number of entry points for GenAI. Data can be entered, processed, and generated through a vast array of interfaces, making comprehensive data protection a significant challenge. Consider the following risks:

Unauthorized access: An employee, contractor, or supplier attempts to share unauthorized sensitive database information with an AI application or AI agent.

Model training risks: Sensitive data (PII, PHI, IP, etc.) can be ingested by models from data sources. Data can be absorbed through an insecure API or other interface.

Prompt injection attacks: A threat actor attempts to compromise critical information through a sophisticated AI-automated SQL injection attack.

Shadow AI: An employee uses an unsanctioned AI tool, evading AI acceptable use policies.

5 critical AI data exposure points

Enterprise data is increasingly exposed to AI. Today's organizations face five critical AI data exposure points:

Productivity suites: Native AI integration within email, documents, presentations, spreadsheets, etc., exposes data to AI through productivity tools such as Microsoft Copilot or Google Gemini.

Embedded SaaS: Built-in AI agents with access to confidential data are increasingly common in customer relationship management (CRM), HR information systems (HRIS), and enterprise resource platforms (ERPs).

URL-based or third-party: AI tools are used by employees to analyze data, attend company meetings, or review proprietary content like designs, product plans, or code (GPTs, meeting agents).

Specialized tools: AI applications that support business functions like AI coding assistants, contract lifecycle management, or financial reporting may access sensitive data.

Vendor/supplier AI: Enterprise data is exposed to AI applications and tools used by third parties with access to sensitive data. Organizations increasingly lack visibility into embedded AI tools across their supplier and third-party environments.

Compliance and regulatory expectations

For CISOs, other IT leaders, and risk officers, AI adds a new dimension to already significant regulatory compliance requirements. In a survey of global CISOs, 98% reported concern about the pace of regulatory change in cybersecurity, with 79% saying they believe the pace of change is not sustainable.[1]

New regulations may require organizations to provide an audit trail of data used by AI systems, ensure data privacy for users interacting with AI, and track sensitive data, along with demonstrating robust data governance and reporting programs. Security incidents involving shadow AI have added an average of $670,000 to the cost of a data breach—and that cost is only expected to grow.[2]

As AI use expands, the ability to conduct forensic investigations, provide audit trails, and produce reports for compliance will become critical.

5 key questions to assess your compliance capabilities

- Have you assessed your compliance requirements for data governance?
- Are you able to track sensitive data shared with AI and provide an audit trail?
- Are you able to track data exfiltration to AI tools?
- Can you conduct a forensic investigation if there's a data breach?
- Can you produce reporting for both regulators and executives on your ability to comply with industry and legal regulations?

[1] Trellix, The Mind of the CISO: CISO Crossroads
[2] IBM, Cost of a Data Breach Report, 2025

European Union Artificial Intelligence Act

Enacted in 2024, the EU AI Act is a comprehensive legal framework designed to foster trustworthy AI in Europe. Key points of the Act promote transparency, safety, and a focus on human-centric AI.

The Act classifies AI according to its risk and prohibits some types of AI systems. The obligation is on providers of AI systems to meet the Act's requirements, which vary according to risk tier. The EU AI Act requires technical documentation describing how data was collected, cleaned, and labeled for every AI model deployed. Higher-risk systems must meet robust data governance requirements as well as documentation programs to be in compliance.

Read more about the EU AI Act

A three-part framework for reducing risk

Secure AI adoption can be accelerated through a holistic approach that prioritizes protecting sensitive data. Technology solutions can help mitigate potential data loss, but an approach that encompasses organizational policy, user training, and real-time technology integration will offer stronger protection and enhance data security posture.

The Policy, Training, and Technology framework addresses these emerging challenges. We offer practical tips for cross-functional leaders to implement safe AI adoption practices that protect critical data and to build an "AI-aware culture" where employees serve as the first line of defense.

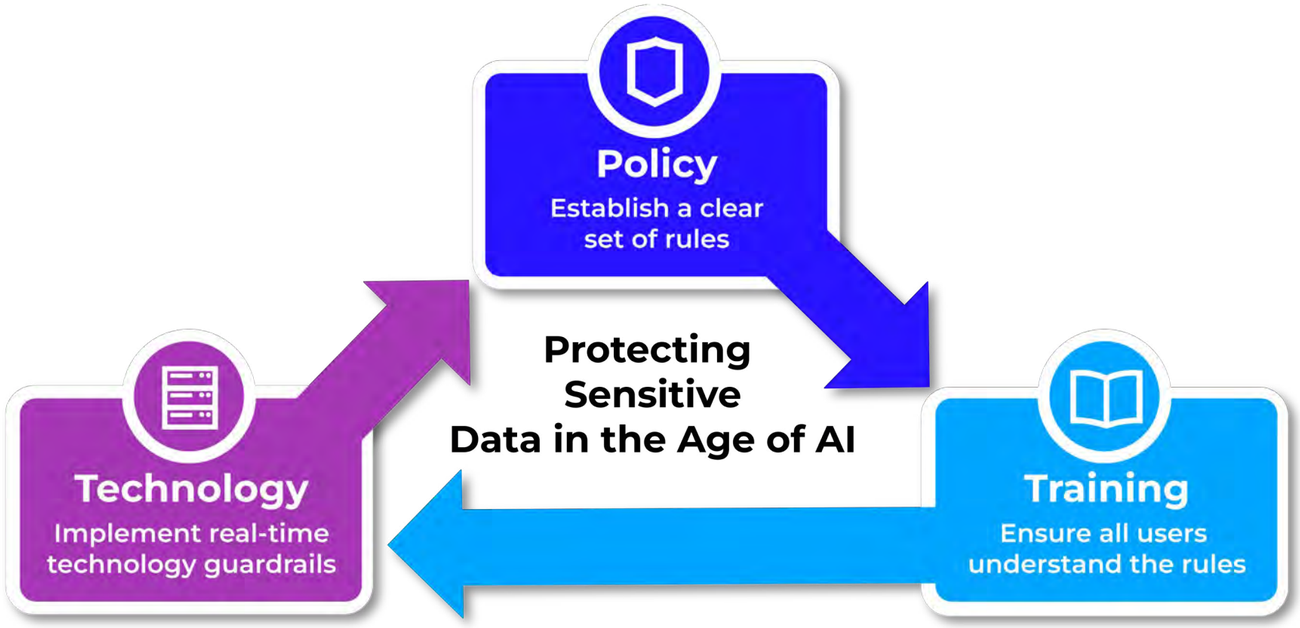


Figure 1. A three-part framework of policy, training, and technology accelerates secure AI adoption.

- **Policy:** Establish clear, actionable guidance on AI acceptable use, sensitive data handling, and expectations on compliance with global regulations. Policies should define boundaries while designating authorized use cases and tools.
- **Training:** Build an "AI-Aware Culture" organization-wide, ensuring all personnel (employees, contractors, third parties) understand acceptable use. Given the rapid pace of AI evolution, training should be an ongoing process with regular updates as technology changes.
- **Technology:** Implement technology to monitor sensitive data shared with AI, identify potential data exfiltration, stop advanced attacks, and create effective guardrails based on policy. Technology should alert users instantly upon policy violations, providing "in-the-moment" reinforcement and guidance.

Policy: Seven pillars of a modern AI acceptable use policy

The bedrock of protecting sensitive data against the rapid expansion of AI is to create clear and practical acceptable use policies (AUP). It's not sufficient to tack a few words about AI onto an existing AUP for general data handling. To be effective, your policy must move beyond "don't do this" and provide clear, actionable frameworks across seven critical domains impacted by the growth of AI. As technology changes, prepare to update it regularly.

1 Data classification and the hierarchy of risk

The cornerstone of data protection in the age of AI is a sophisticated understanding of data classification. Employees must be trained to recognize that large language models (LLMs) are essentially data sponges. Your policy should establish what is Public Data or safe for general prompts; Internal-Only Data, which may only be used with enterprise-grade, privacy-protected tools; and Restricted/Highly Sensitive Data, such as trade secrets or personally identifiable information (PII), that is strictly forbidden from entering any AI prompt. By defining these boundaries, you empower staff to use AI without inadvertently turning a confidential memo into public training data.

2 Securing intellectual property and ownership

One of the most complex legal frontiers in AI is the question of Intellectual Property (IP). An effective AUP must explicitly state that any work product, code, or strategic insight generated with the assistance of AI remains the company's exclusive property. Furthermore, it must address the "training leak" risk. Without specific configurations, many consumer-grade AI tools default to using user inputs to train future iterations of their models. Your policy should mandate the use of "opt-out" settings or enterprise licenses to ensure that your company's unique innovations don't become the "intelligence" sold to your competitors in the next model update.

3 Establishing ethical guardrails

AI-driven efficiency should never come at the expense of corporate ethics. Your policy must address the inherent risks of algorithmic bias and the “hallucination” of facts. It is vital to codify a “human-in-the-loop” mandate, requiring that no AI-generated output—be it a line of code, a legal summary, or an HR evaluation—is ever published or acted upon without rigorous human review. Transparency is equally critical; the policy should define when and how employees must disclose the use of AI in their deliverables to maintain trust with stakeholders and clients.

4 Navigating the compliance landscape

As global regulations like the GDPR and the EU AI Act evolve, your AUP acts as your first line of defense against regulatory non-compliance. Data protection in AI is particularly tricky regarding the “right to be forgotten”; once sensitive data is ingested into a model’s weights, it is nearly impossible to extract. Your policy should align AI usage with existing privacy frameworks, ensuring that no data subject to HIPAA, CCPA, or other regional mandates is processed through unvetted AI systems. This section should also outline the necessity of maintaining audit trails for high-stakes AI interactions to satisfy future regulatory inquiries.

5 Curating the approved AI toolkit

Shadow AI—the use of unauthorized, personal AI accounts for work tasks—is a primary vector for data leaks. To combat this, your policy should provide an “allow list” of approved tools, applications, and browser extensions that have passed a formal security review. By providing employees with a set of powerful, sanctioned tools, you reduce the temptation to use unvetted “free” versions that often have lower security standards. This section should also explicitly ban AI browser extensions that passively scrape screen data, which can inadvertently capture sensitive credentials or private communications.

6 Setting expectations for the supply chain

Your data is only as secure as the weakest link in your supply chain. Modern AUPs must extend their reach to third-party vendors and suppliers. It is no longer enough to secure your own house; you must require vendors to disclose if and how they are using AI to process your organization's data. This includes adding specific "AI riders" to service level agreements (SLAs) that prohibit your data from being used for model training by third-party providers. Clear expectations ensure that your data protection standards are upheld even when the data leaves your immediate control.

7 Incident response and the culture of reporting

Finally, a policy is only as good as its enforcement and the response to its breach. Because AI interactions happen at the speed of thought, accidental disclosures are inevitable. Your policy should foster a "no-blame" reporting culture, where employees are encouraged to immediately notify IT or Security teams if they suspect they have pasted sensitive data into a public AI. Rapid response can often mitigate the damage through prompt deletion requests or account resets. This section should also detail the offboarding procedures for employees to ensure that access to proprietary AI environments is revoked as strictly as access to any other sensitive internal system.

Training: Five practical tips for every employee, contractor, and supplier

AI is constantly transforming enterprise activities, and the only way to succeed in protecting sensitive data is to work together. Organizations that win this battle will make security a team sport. To do that, every person touching enterprise data has to know and understand how to handle data properly.

1

Don’t share sensitive information

The most effective way to protect data is to ensure it never reaches an AI’s memory in the first place. The primary directive for all users is simple: Do not share sensitive or confidential information with AI if you can avoid it. Whether it is a proprietary software script, a draft of a merger agreement, or internal financial projections, users should treat every prompt as a potential public disclosure. If a task can be achieved using generic descriptions or hypothetical scenarios rather than real-world secrets, that must be the default approach.

2

Use approved tools

Innovation often leads to Shadow AI, where well-meaning employees use personal accounts or unvetted plugins to move faster. However, your policy must mandate the use of only company-approved, secured AI tools and managed browsers. These enterprise-grade platforms are specifically negotiated to include data-out clauses, ensuring your inputs are encrypted and—crucially—excluded from the provider’s global training sets. By staying within these sanctioned environments, users benefit from the machine’s intelligence without sacrificing the organization’s digital sovereignty.

3 Sanitize the data before use

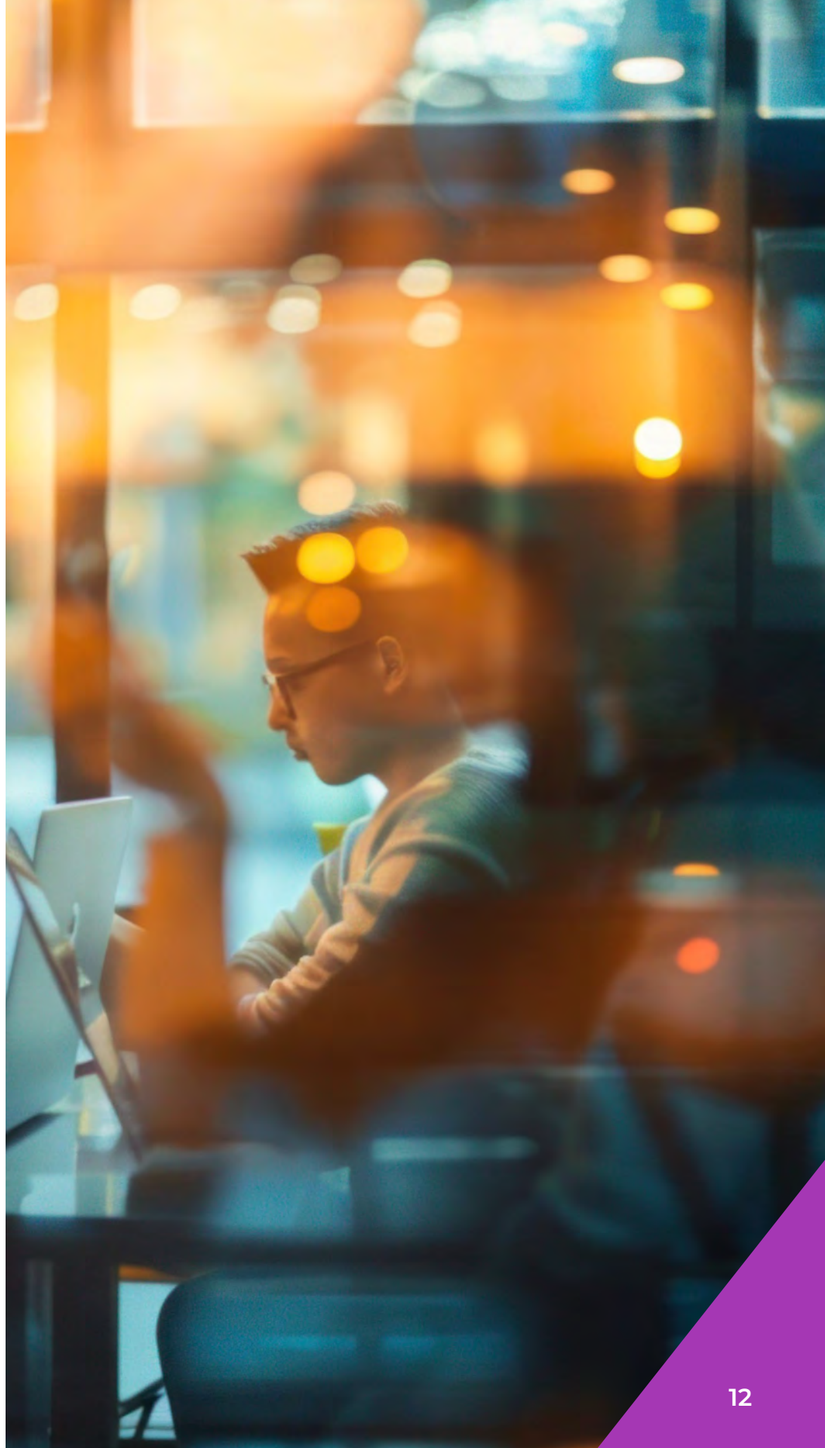
Before any dataset or document touches an AI interface, it must undergo a rigorous sanitization process. Users are responsible for anonymizing or de-identifying data before use with AI. This involves stripping away names, specific dates, or unique identifiers and replacing them with generic placeholders (e.g., changing “Mary Jones” to “Customer A”). By feeding the AI “clean” data, the organization extracts the necessary analytical value without ever exposing the underlying identities or specific secrets that constitute a compliance risk.

4 Review privacy settings and terms

In the rush to adopt new productivity-boosting tools, it is dangerously easy to click “Accept” on terms of service without a second thought. Every user has a proactive responsibility to review privacy settings and terms of use before accepting or integrating a new AI application into their workflow. Many “free” tools recoup their costs by claiming ownership of user inputs. Training should empower employees to look for specific red flags in terms of use that indicate their data might be repurposed for model training or third-party marketing.

5 Stay informed and share knowledge

Because AI technology evolves weekly, the security landscape is constantly shifting. A culture of protection requires every individual to stay informed and share what they know with colleagues. If a team member discovers a safer prompting technique or identifies a privacy flaw in a popular browser extension, that insight must be socialized immediately. This collaborative approach transforms a workforce from a group of individuals into a unified front, ensuring that as AI capabilities grow, the collective “immune system” of the organization grows with them.



Technology: Four critical technologies to protect your enterprise data ecosystem

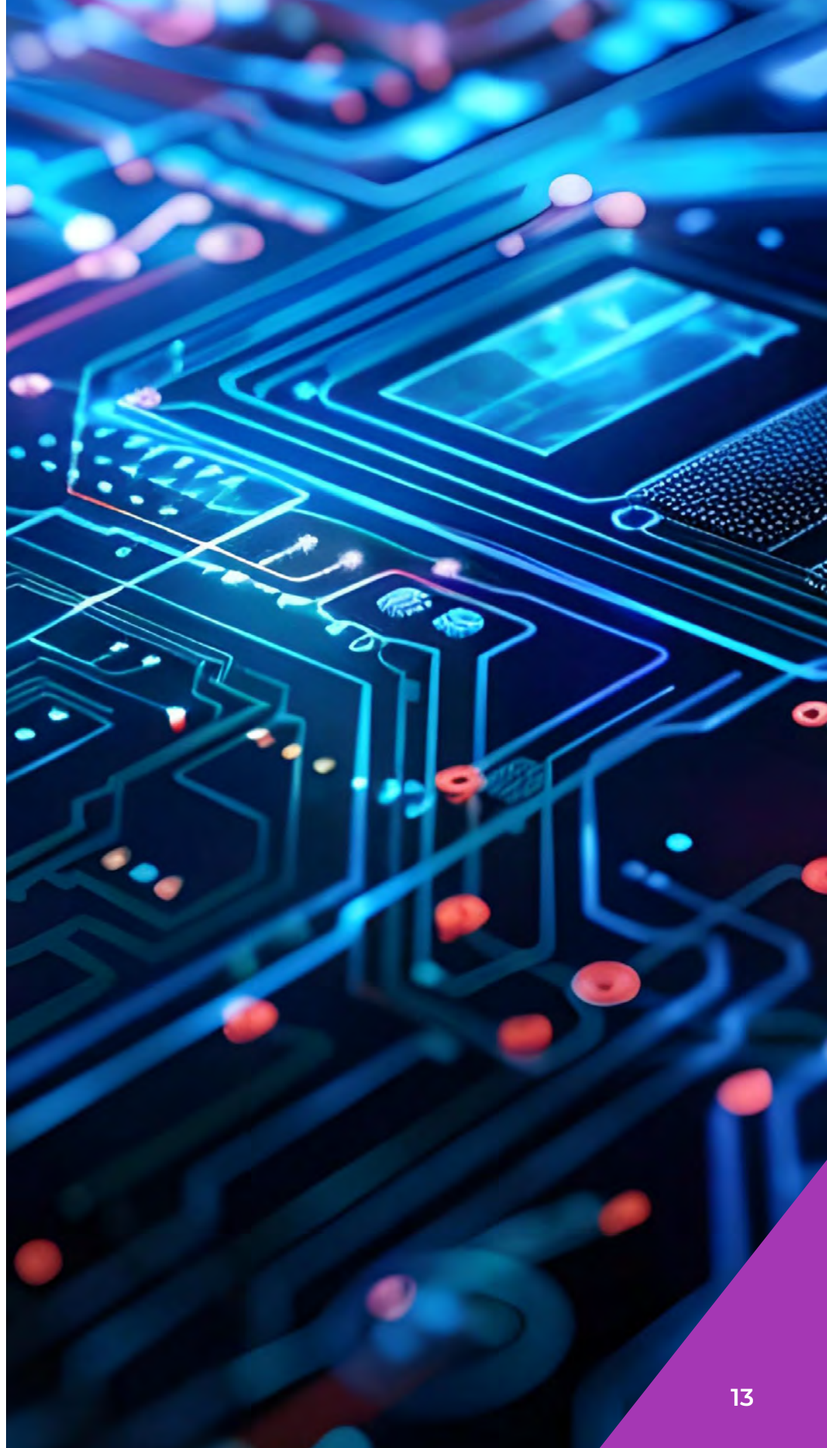
As AI adoption accelerates, the sheer volume and velocity of data movement can overwhelm traditional security measures. To maintain control, organizations must leverage a sophisticated suite of technology solutions—including data loss prevention (DLP), database security, and data encryption. However, in the age of generative AI, these tools are no longer “set and forget” systems; they must be transformed into proactive, intelligent guardians of the enterprise data ecosystem.

The first line of defense remains robust DLP. Modern DLP must be tuned to recognize the specific “DNA” of your company’s intellectual property and confidential information. Complementing this is data encryption, which serves as the ultimate failsafe; even if a data packet is intercepted or erroneously uploaded to a public cloud, high-level encryption ensures that the information remains unreadable and useless to unauthorized parties. Database security safeguards some of the most important data in any organization by limiting access and enforcing security controls.

1

Stopping database compromise

Perhaps the most critical modern data security frontier is database security. As organizations deploy AI agents to query internal data, the risk of privilege escalations and prompt injection attacks is skyrocketing. Modern database security tools act as gatekeepers, preventing AI agents from accessing unauthorized data by enforcing strict, identity-based access controls. If an AI agent attempts to query a table it isn’t cleared for, the security layer can instantly terminate the session before a single row of data is exfiltrated.



Furthermore, a patched database is a secure database. Tools can automate the "hygiene" of your data environment, ensuring your databases stay patched and up-to-date against the latest vulnerabilities that AI-driven malware might seek to exploit. Legacy databases, where vendors no longer offer patches, require the extra security provided by database security solutions as AI-based threats expand.

2 Real-time monitoring and blocking

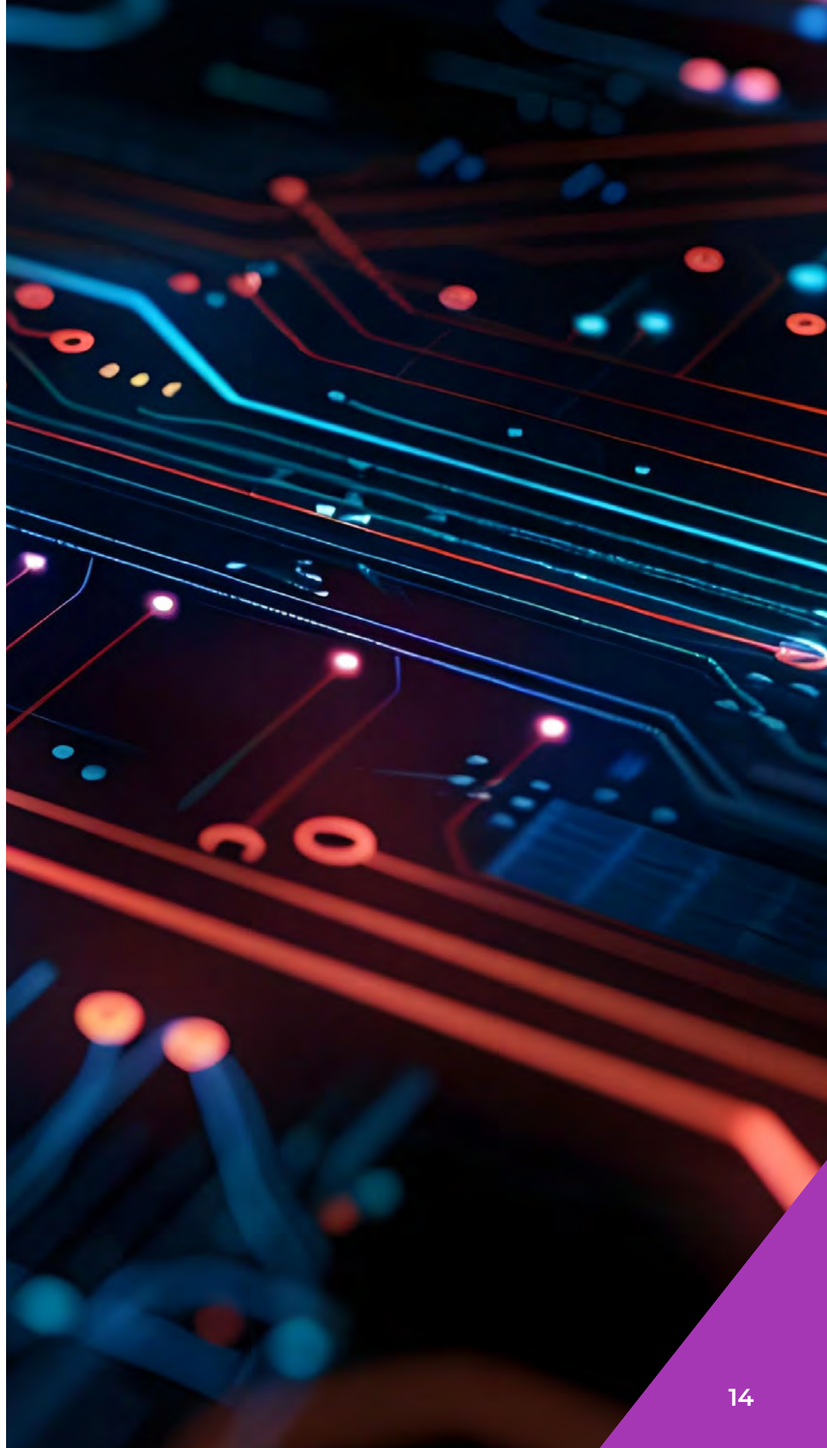
For data security tools to be successful, they must operate at the speed of the threat. Organizations require real-time monitoring and automated blocking of data exfiltration. Unlike traditional file transfers, an AI data leak can happen in a single copy-paste action into a browser. Your security stack must be capable of intercepting that action instantly, recognizing sensitive strings before they leave the endpoint, and blocking the transmission before the AI model can ingest the information.

3 Instant and customizable user coaching

Technology is most effective when it bridges the gap between security and education. Advanced data loss prevention tools now offer instant and customizable user coaching. Rather than simply blocking an action with a generic "Access Denied" message, these systems trigger a pop-up notification the moment a violation is attempted. This "teachable moment" explains why the action was blocked and directs the employee toward a sanctioned, secure alternative.

4 Granular visibility and compliance reporting

Finally, in a regulated world, security programs and technology aren't enough—you must be able to prove your actions. It is vital to ensure your technology tools offer strong compliance reporting capabilities. As auditors begin to scrutinize AI workflows, you need highly granular reporting that provides visibility into how employees are using both sanctioned and unauthorized AI tools.



Next steps

The age of artificial intelligence is not a distant milestone on a roadmap; it is the current reality of how we work today. As we have explored throughout this book, the rapid expansion of AI does not necessitate a total abandonment of traditional security principles or a wholesale attempt to block it all. Instead, it demands an evolution to modern practices. Protecting sensitive data in this new era is a multi-dimensional challenge that requires a seamless integration of clear policy, empowered, highly trained people, and proactive technology solutions.

Trellix can help you bridge the gap between the risks of rapidly advancing AI technology and the benefits of secure adoption, empowering your workforce to innovate with confidence.

To better understand your current level of exposure to the risks of rapidly advancing AI, take our online AI Data Risk Posture Assessment.

Explore more with these resources:

Web Page: Protecting Sensitive Data in the Age of AI

Solution Brief: Protecting Sensitive Data in the Age of AI with Trellix® Data Security

Data Sheet: Trellix Data Loss Prevention AI Data Risk Dashboard

Trellix is a global cybersecurity company delivering intelligence-led cyber resilience for security-conscious organizations at any stage of their journey. Transforming over 30 years of threat intelligence into high-fidelity detections and automating AI-driven detection and response across cloud, on-premises, air-gapped, and operational technology environments, Trellix helps customers adapt to the constantly evolving threat landscape.

More at *www.trellix.com*
Follow Trellix on *LinkedIn* and *X*.

About the Author

Laurie Robb
Director, Product Marketing
Trellix Data Security

Laurie has more than 25 years of experience in marketing communications across a variety of industries including cybersecurity, SaaS, technology management, and healthcare. At Trellix, she leads Product Marketing for Data Security, translating complex technical concepts into clear, compelling stories that fuel engagement and growth. She frequently writes about and presents on data security industry topics, delivering high-impact messaging and content that bridges the gap between technical and business audiences. She is passionate about simplifying complexity, helping cybersecurity companies stand out, and understanding customer challenges. Laurie is also a proud member of the leadership team for Women in Cybersecurity Northeast Ohio affiliate.