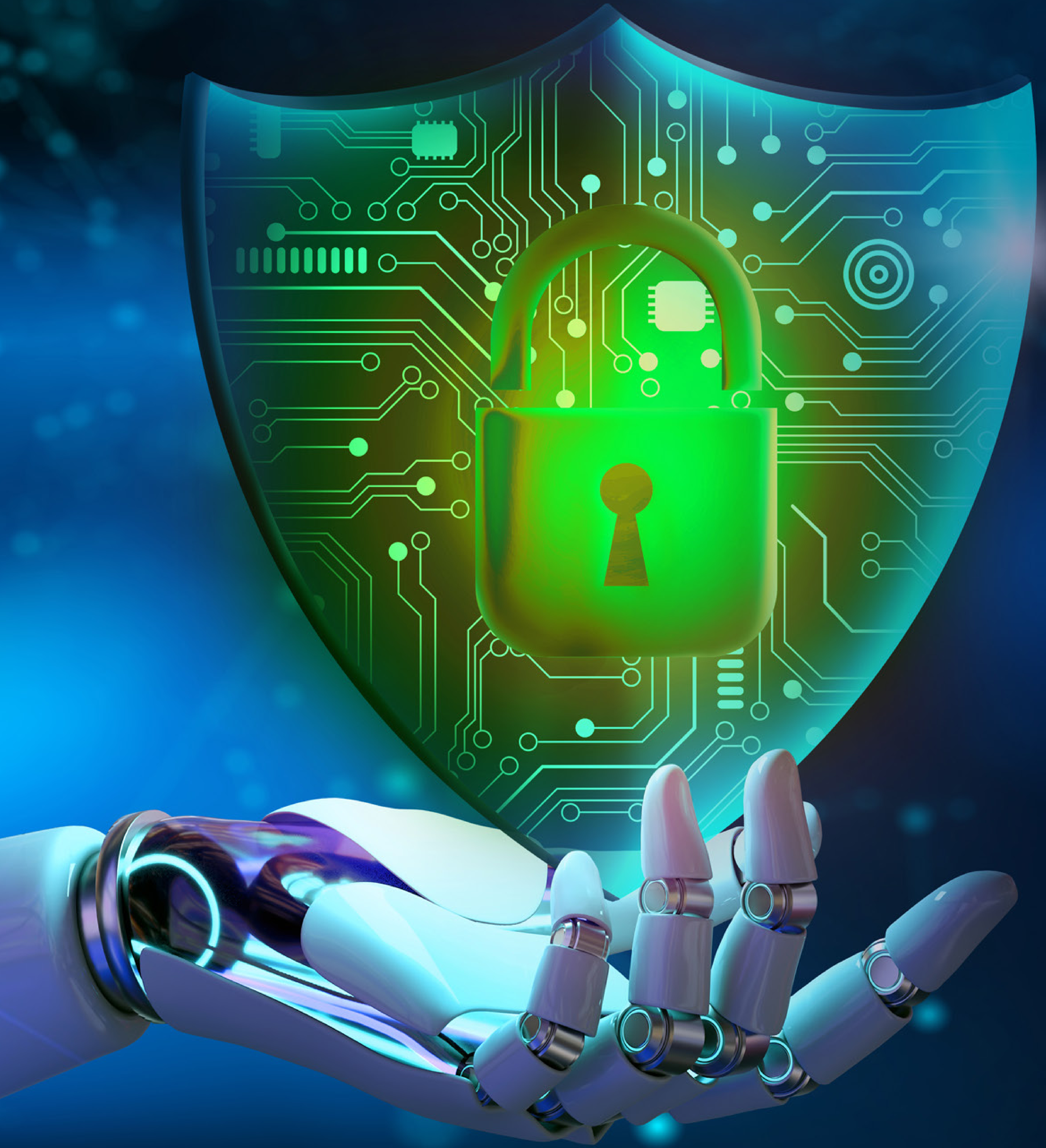
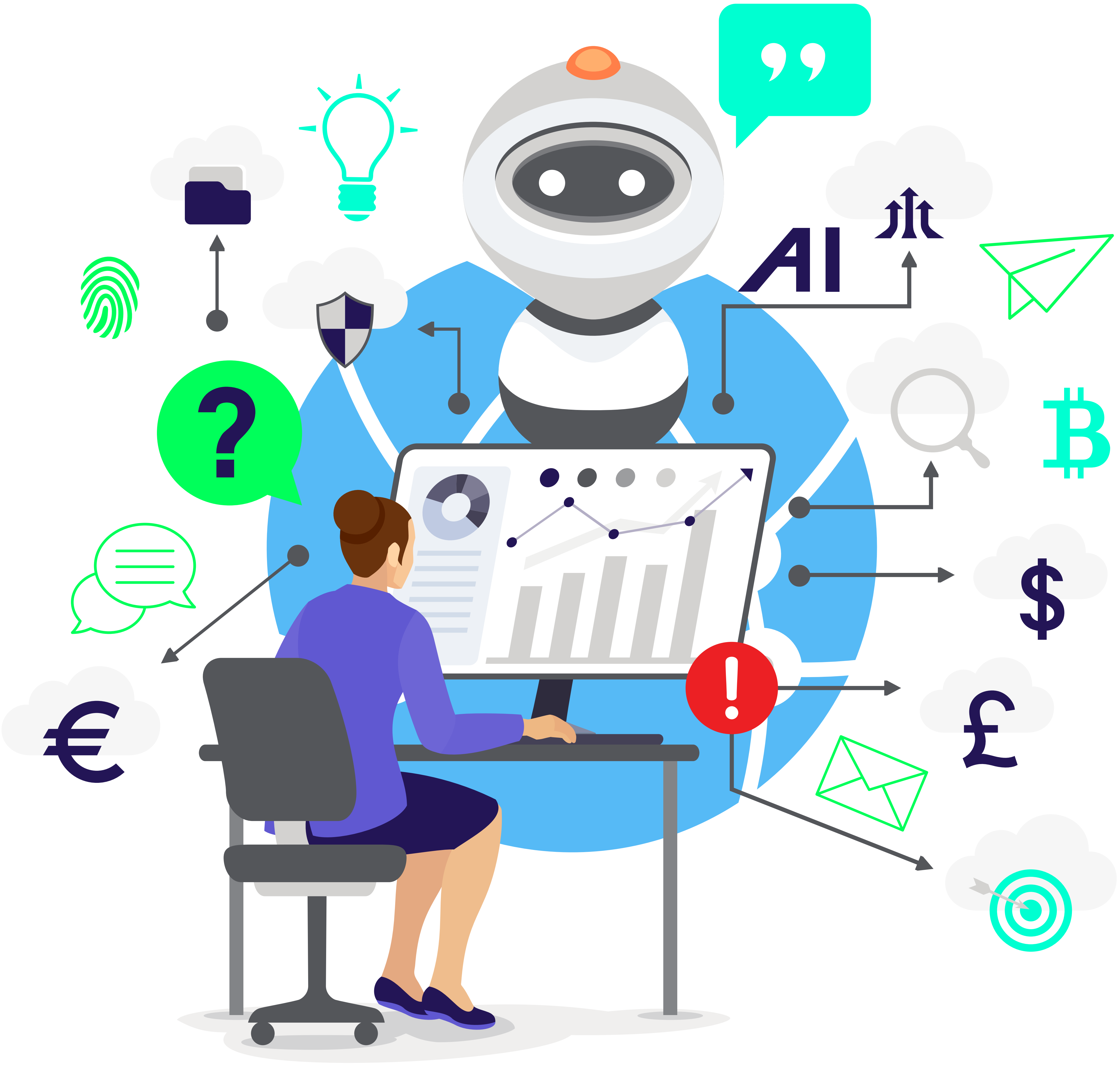
Delinea



The High Stakes of Securing AI

A security leader's guide

The stakes of securing AI



Organizations are ramping up AI use. In early 2024, 65% of organizations surveyed by McKinsey regularly used Generative AI, double the percentage from 10 months before.

The stakes are high. AI models can execute tasks, interact with systems, and even create and manage credentials, making them a major risk if compromised. The rise of Agentic AI – AI systems capable of making independent decisions – is increasing the use of machine identities and expanding the attack surface at many organizations.

If an attacker can successfully attack a machine identity, they can successfully attack AI.

"AI poisoning" and other adversarial attacks

With the right access, these powerful tools can be weaponized. Attackers can leverage AI agents to mix harmful or deceptive data into datasets used to train a machine learning model. If poisoned, AI agents could be manipulated to bypass security controls, escalate privileges, or exfiltrate data.

For example, recently discovered DeepSeek vulnerabilities highlight how AI systems can be exploited if security isn't prioritized from the start. Researchers found flaws in DeepSeek that allowed sensitive data leakage and manipulation of AI outputs. This reinforces why you must treat AI identities with the same level of security as human users, ensuring rigorous access controls and continuous monitoring.

Auditors are asking: Do you know what is going on inside your company and do you have a plan?

Machine identities, including AI agents, are hot button areas for compliance auditors and cyber insurance providers.

Regulatory frameworks such as NIST, ISO, and GDPR are increasingly including guidance on securing AI like the NIST AI risk management framework. To increase security and resilience, they encourage accountability and transparent use of AI based on valid and reliable data sources.

At the core, they want to know whether you, as a security leader, understand the use of AI in your organization and whether you are taking steps to protect it.

To prepare for a future AI regulatory audit, it's a good idea to:

- Evaluate data including collection, storage, and processing methods, ensuring compliance with data privacy and security regulations such as GDPR, CCPA, PCI, HIPAA, and others.
- Create and review documentation about your AI usage and security practices to make sure you understand and can communicate about your data sources, model architectures, and decision processes.
- Conduct red team attack exercises that highlight unmanaged, vulnerable machine identities, including AI.

Security leaders must influence others to secure AI

Delinea's Identity Security Report found that 28% of businesses consider securing machine identities, such as AI, a top priority. Considering the rapid rise of AI and the associated risk, we believe businesses should be more worried.

While business leaders are focused on developing policies based on ethics, privacy, and compliance requirements for AI (see the Securing the Use of Generative AI Technologies Delinea whitepaper for more information), security leaders must figure out how to implement security controls and supporting technology.

Doing so requires influencing others in your organization – developers, Linux admins, and system architects, etc. – to follow best practices, even if it means changing their processes. As you likely don't have direct influence over their work, it's essential to build their awareness of the problem and create security solutions that are easy for them to adopt without hindering productivity.

Key terms you need to know

Organizations are comprised of **different identity types**:

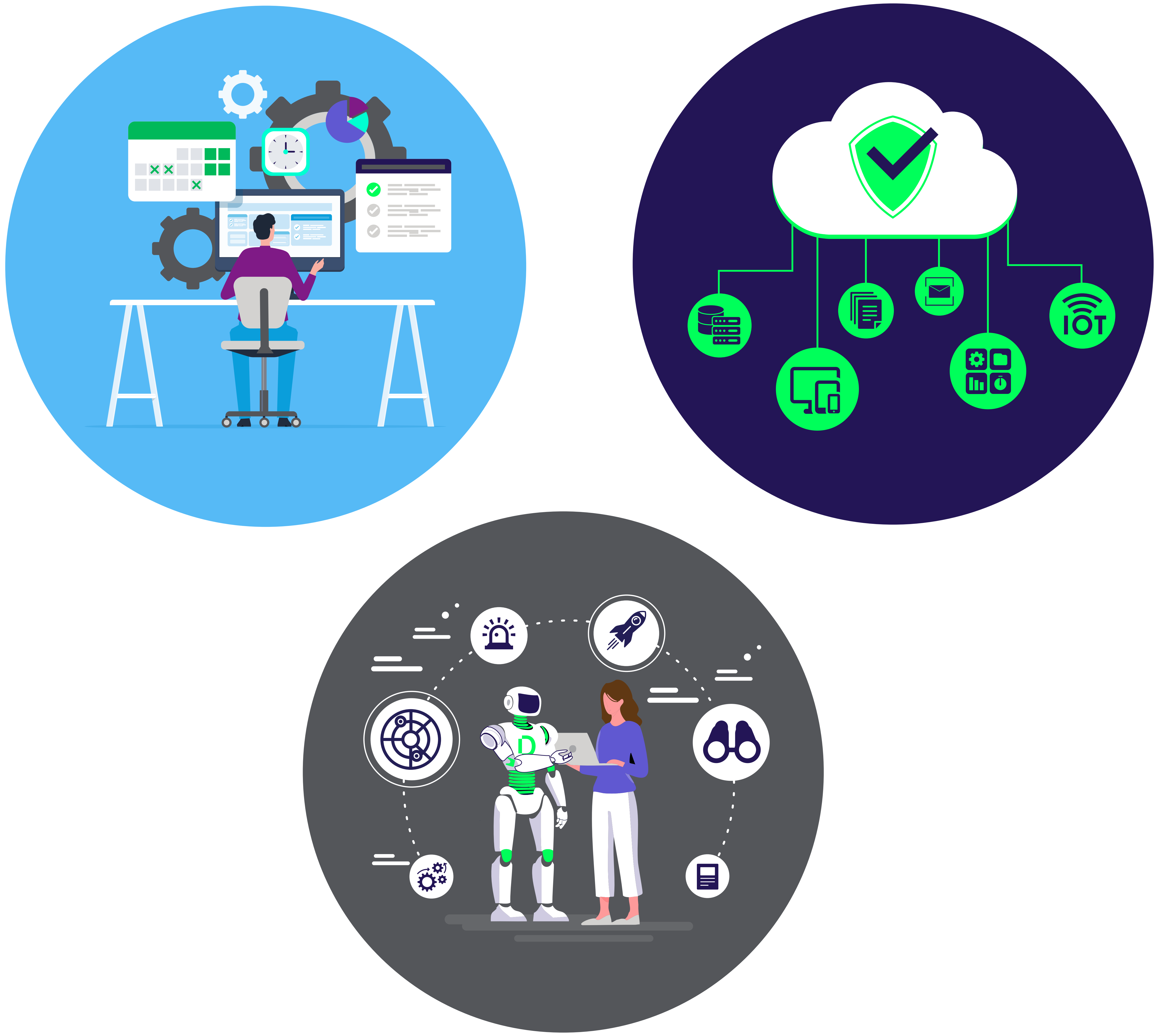
- **Human identities** such as IT admins, application owners, workforce users, and developers.
- **Machine identities** such as service accounts, local accounts, managed identities, and AI identities. AI identities are a type of machine identity that leverages other machine identities to interact with systems and perform tasks autonomously.

Machine identities are used to interact with systems, exchange data, and perform tasks autonomously, via:

- AI agents
- APIs
- Services
- Applications
- Containers
- Automations
- DevOps

Credentials are the keys that allow identities to authenticate and unlock system access.

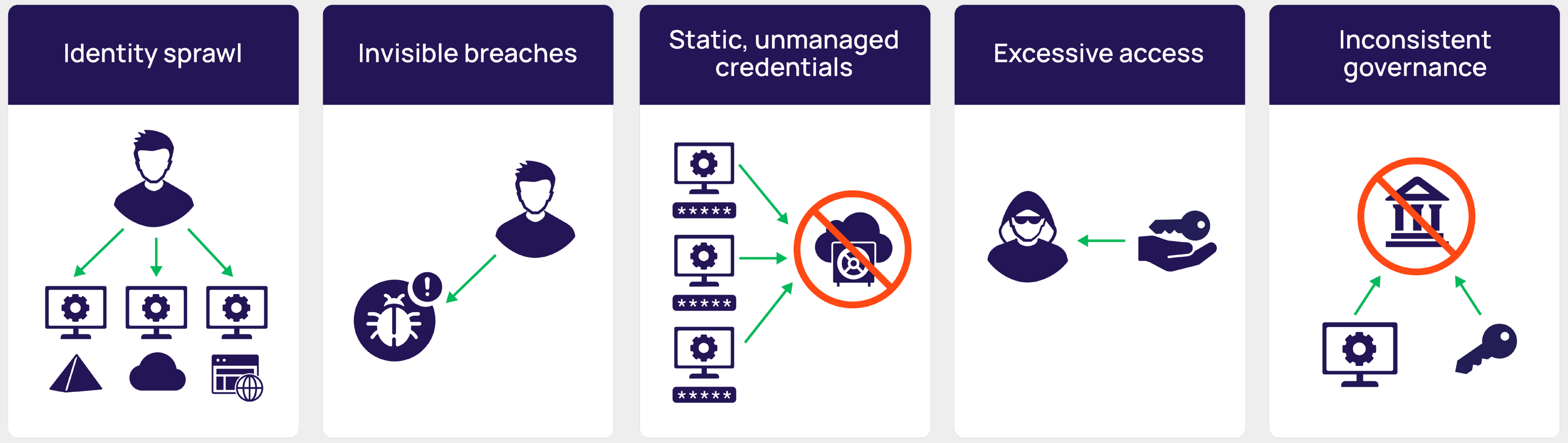
- Humans typically log into systems via credentials such as usernames and passwords.
- Machine identities leverage credentials such as SSH keys, API keys, certificates, and OAuth tokens.
- All these credentials can be called "secrets."



Challenges to securing AI

Securing AI is easier said than done for many organizations. Several issues stand in the way.

Common challenges encountered with securing AI



Source: Delinea

Identity sprawl

AI identities are everywhere across on-prem, cloud, and SaaS environments. This obscures your visibility and makes human oversight challenging. You aren't sure: *What does AI have access to? Who is really in control?*

Invisible breaches

Unknown identities create invisible breaches without clear avenues for detection and remediation. You wonder: *Can you trust AI? What is the line of reasoning? Can it be corrupted or manipulated?*

Before the advent of AI, we asked this question about code. However, you could inspect the code and ensure it was doing the right thing. If an AI's LLM has 30 billion plus parameters, you can never check everything.

Unlike malware injections that can be identified by checking code, AI poisoning is much more difficult to detect and remove. In fact, it's virtually impossible to discover AI poisoning in your LLM if data has already been diluted. It's like mixing sugar into water and then trying to get the sugar out.

That's why it's so important to block poisoning from happening in the first place.

Static, unmanaged credentials

If they aren't centrally managed in a vault, static credentials can't be provisioned, deprovisioned, or rotated regularly or immediately in the event of a breach. Any person or system in possession of a valid credential can potentially gain access to your environment.

Ultimately, the goal is to move from static to dynamic, ephemeral credentials where possible. Until then, managing static credentials securely in a vault with appropriate governance is critical.

Excessive access

Overly permissive identity provisioning creates unnecessary risk. Users often don't know what access AI identities need. As a result, they give AI identities access to more data and business systems than they need to accomplish their goals.

Inconsistent governance created "Shadow AI"

Oftentimes, AI identities have short lifespans and require frequent provisioning and deprovisioning. Without proper human oversight, AI identities can be governed incorrectly and easily become orphaned as part of your organization's "Shadow AI."

Step-by-step approach to securing AI

Let's take a look at how to address these challenges to secure AI.

Discovery and inventory

As a first step, discover and inventory all AI identities, secrets, accounts, and credentials throughout your entire infrastructure, across on-prem, cloud, and SaaS environments.

Where might you find AI? Data models and training data may be stored and operating in:

- Applications
- Databases
- Code repositories

Identity posture and threat analysis

Continually monitor and audit machine identity and AI access to detect unauthorized and unexpected access and remediate threats.

Protected credentials

Vault and secure machine identities and AI identities, just as you would for human identities. Use an encrypted, centrally managed vault to manage static secrets through automated rotation and/or generate dynamic, ephemeral credentials. Your vault should provide an immutable audit trail for compliance.

Privileged secure access

Authenticate and authorize all machine-to-machine communications with dynamic authorization. Access controls based on predefined roles and permissions adapt to the current context, ensuring privileged access for AI identities remain secure.

Zero standing privilege

Ensure consistent least privilege and just-in-time access to limit the rights of AI agents throughout your environment. In particular, authenticate and authorize all machine-to-machine connections for granular access and set up controls to prevent unauthorized or lateral access to servers.

Identity governance

At every point in the lifecycle from provisioning to deprovisioning, move from ad-hoc to contextual decisions that right-size identity access. Ultimately, shift from manual work like access reviews to automated lifecycle management that spins identities up and down, so that the window of opportunity attackers can take advantage of shrinks.

You can accomplish all of these steps with Delinea →

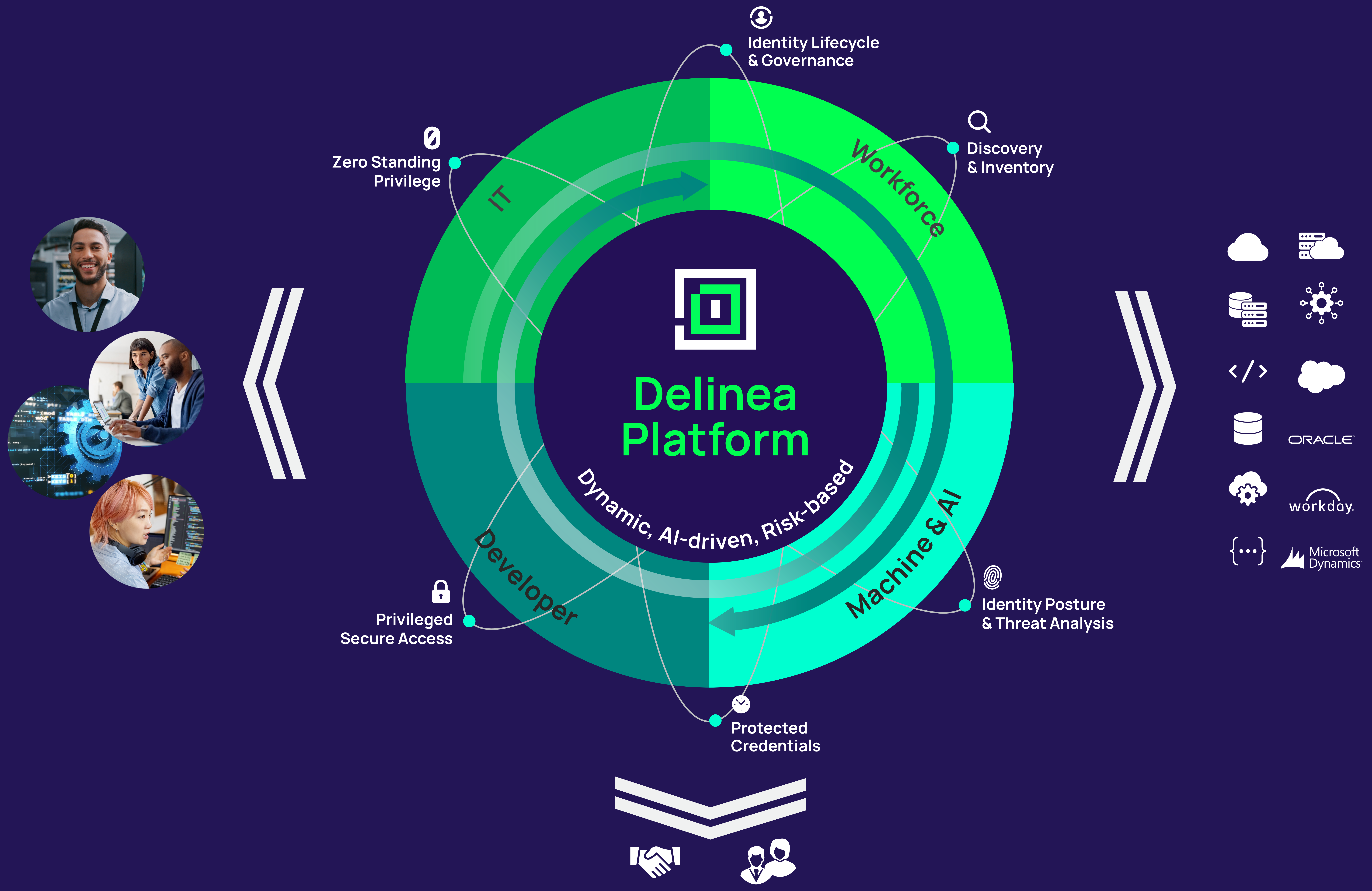
The Delinea Identity Security Platform

Securing AI isn't a siloed, standalone program. It's one use case for identity security. To be most efficient and effective, it should be a part of your overall identity security strategy.

Luckily, securing AI follows the same principles you're already using to manage and secure human identities and machine identities. And if you're already working toward securing these in an integrated solution, you're well on your way to managing AI use cases too.

Ultimately, your organization needs a solution that addresses the security best practices described in this eBook, a solution that is managed by a central team in a connected ecosystem with centralized access and credential management, integrated monitoring and reporting, and context-based detection and remediation. This approach gives you end-to-end visibility and enables comprehensive auditing and reporting for compliance.

See the Delinea Platform in action.



Delinea Platform

Dynamic, AI-driven, Risk-based

IT

Workforce

Machine & AI

Developer

Identity Lifecycle & Governance

Discovery & Inventory

Identity Posture & Threat Analysis

Protected Credentials

Privileged Secure Access

Zero Standing Privilege

Delinea

Securing identities at every interaction

Delinea is a pioneer in securing human and machine identities through intelligent, centralized authorization, empowering organizations to seamlessly govern their interactions across the modern enterprise. Leveraging AI-powered intelligence, Delinea's leading cloud-native Identity Security Platform applies context throughout the entire identity lifecycle across cloud and traditional infrastructure, data, SaaS applications, and AI. It is the only platform that enables you to discover all identities — including workforce, IT administrator, developers, and machines — assign appropriate access levels, detect irregularities, and respond to threats in real-time. With deployment in weeks, not months, 90% fewer resources to manage than the nearest competitor, and a 99.995% uptime, the Delinea Platform delivers robust security and operational efficiency without complexity. Learn more about Delinea on **Delinea.com**, **LinkedIn**, **X**, and **YouTube**.