



Darktrace / SECURE AI

Secure how AI is used across your business



THE CHALLENGE

AI Is Expanding Faster Than Security Can Keep Up

AI is now embedded across everyday business workflows—from enterprise tools to custom-built agents. But as adoption accelerates, most organizations lack clear visibility into how AI is actually being used.

This creates a new class of risk. Sensitive data can be exposed through prompts, AI agents can act beyond intended permissions, and unsanctioned tools can emerge without oversight. Traditional security approaches—built on rules and known threats—simply cannot keep up with the dynamic, context-driven nature of AI systems.

Why Darktrace?

Securing AI requires more than visibility; it requires a deep understanding of behavior, context, and intent across users, systems, and AI.

Darktrace brings a level of maturity and behavioural understanding that is uniquely suited to the complexity of AI environments. Achieved with Self-Learning AI that has been proven across more than 10,000 organizations, Darktrace learns the normal patterns of each business to interpret context, uncover subtle intent, and detect meaningful deviations, without relying on predefined rules or known attack signatures.

Just as importantly, Darktrace is proven in dynamic, cross-domain environments, where risk doesn't live in a single tool or control plane. We're built to see behavior as it moves across systems, which is exactly what securing AI requires.

By bringing this maturity to AI, Darktrace enables organizations to move forward with confidence, whether they are deploying GenAI tools, building agents, or uncovering previously unseen activity across their environment.

Visit darktrace.com for a demo

THE SOLUTION

Understand, Secure, and Control AI Usage

Darktrace / SECURE AI is designed to bring clarity and control to this new environment.

Rather than relying on predefined rules, it analyzes how prompts, agents, and AI systems behave within the context of your organization. By understanding intent across users and systems, it gives security teams the ability to see where risk is emerging—and act on it in real time.

The result is continuous oversight of AI usage across the business, even as adoption evolves and expands.

What It Delivers

Darktrace / SECURE AI gives organizations a unified view of how AI is being used—making it possible to understand risk across prompts, agents, development environments, and unsanctioned activity.

It enables teams to:

- **Monitor the prompts driving GenAI agents and assistants:** Gain real-time visibility into prompts, sessions, and responses to understand intent, detect conversational prompt attacks, and identify risky deviations.
- **Identify and secure AI agents in real time:** Discover AI agents across suppliers and environments, understand their access and interactions, and audit behavior across the digital estate.
- **Evaluate AI risk during development and configuration:** Gain visibility into AI agent identities and build activity to surface misconfigurations, over-permissioning, and anomalous development events.
- **Discover and control Shadow AI:** Identify unsanctioned or unexpected AI activity, distinguish misuse from approved usage, and apply policy to reduce data exposure and risk.