DARKTRACE

The AI Maturity Model for Cybersecurity

A guide to the evolution of the security organization in the Era of AI

darktrace.com

Contents

Executive Summary

As artificial intelligence becomes foundational to modern cybersecurity, security leaders are increasingly challenged to distinguish meaningful innovation from market hype.

The critical question is no longer if AI should be adopted, but **how to measure its effectiveness** and strategically mature its use to deliver real security outcomes.

AI fundamentally allows software to perform more complex tasks than traditional automation, which is already in wide use. To get the most effective outcomes from security resources, organizations will need to take advantage of what AI can reliably do for them and the speed and scale and continuous operation that it can offer.

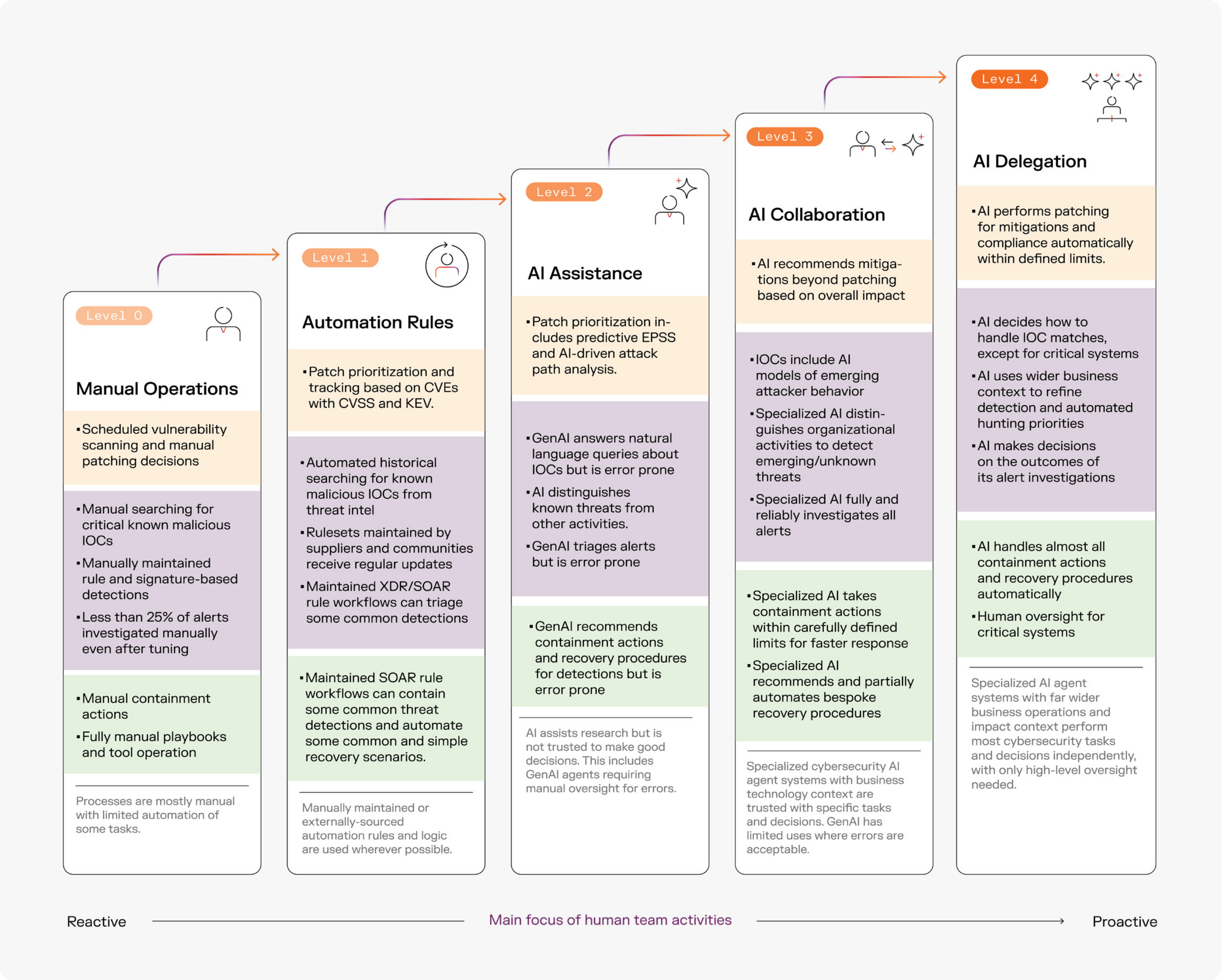
This guide introduces the **AI Maturity Model for Cybersecurity** — a structured framework designed to help CISOs and security leaders evaluate and advance their organization's AI capabilities across core operational areas. By assessing AI integration through progressive levels of trust and autonomy, the model enables a clear, actionable understanding of where your security operations stand today and how to evolve them for tomorrow.

The model provides a practical roadmap for aligning AI adoption with operational needs, identifying gaps, and prioritizing investments. It supports a results-driven approach to AI, one that enhances detection, response, and resilience across the security lifecycle. In an era where threats are growing in complexity and speed, the AI Maturity Model empowers security teams to move beyond experimentation and toward operational excellence, ensuring AI is not just present but impactful.

[1] Agentic GenAI refers to generative AI systems, typically built on large language models (LLMs), that produce new content, like text or code, based on training data. While useful in some scenarios, these systems can also "hallucinate" or generate inaccurate information, especially in high-stakes environments like cybersecurity. In this model, we differentiate Agentic GenAI from other types of AI (e.g., unsupervised machine learning) due to its specific risks and oversight requirements.

[2] While there is little research yet in the cybersecurity domain, more mature implementations of the same technology in other domains have shown GenAI to be untrustworthy in critical applications. Currently we have seen high error rates – with false information presented.

AI across security operations



Level 0

Manual Operations

- Scheduled vulnerability scanning and manual patching decisions

- Manual searching for critical known malicious IOCs
- Manually maintained rule and signature-based detections
- Less than 25% of alerts investigated manually even after tuning

- Manual containment actions
- Fully manual playbooks and tool operation

Processes are mostly manual with limited automation of some tasks.

Level 1

Automation Rules

- Patch prioritization and tracking based on CVEs with CVSS and KEV.

- Automated historical searching for known malicious IOCs from threat intel
- Rulesets maintained by suppliers and communities receive regular updates
- Maintained XDR/SOAR rule workflows can triage some common detections

- Maintained SOAR rule workflows can contain some common threat detections and automate some common and simple recovery scenarios.

Manually maintained or externally-sourced automation rules and logic are used wherever possible.

Level 2

AI Assistance

- Patch prioritization includes predictive EPSS and AI-driven attack path analysis.

- GenAI answers natural language queries about IOCs but is error prone
- AI distinguishes known threats from other activities.
- GenAI triages alerts but is error prone

- GenAI recommends containment actions and recovery procedures for detections but is error prone

AI assists research but is not trusted to make good decisions. This includes GenAI agents requiring manual oversight for errors.

Level 3

AI Collaboration

- AI recommends mitigations beyond patching based on overall impact

- IOCs include AI models of emerging attacker behavior
- Specialized AI distinguishes organizational activities to detect emerging/unknown threats
- Specialized AI fully and reliably investigates all alerts

- Specialized AI takes containment actions within carefully defined limits for faster response
- Specialized AI recommends and partially automates bespoke recovery procedures

Specialized cybersecurity AI agent systems with business technology context are trusted with specific tasks and decisions. GenAI has limited uses where errors are acceptable.

Level 4

AI Delegation

- AI performs patching for mitigations and compliance automatically within defined limits.

- AI decides how to handle IOC matches, except for critical systems
- AI uses wider business context to refine detection and automated hunting priorities
- AI makes decisions on the outcomes of its alert investigations

- AI handles almost all containment actions and recovery procedures automatically
- Human oversight for critical systems

Specialized AI agent systems with far wider business operations and impact context perform most cybersecurity tasks and decisions independently, with only high-level oversight needed.

Key takeaways from the maturity model

SOC fatigue is real, and AI can help: Most teams still struggle with alert volume, investigation delays, and reactive processes. AI adoption is inconsistent and often siloed. When integrated well, AI can make a meaningful difference in making security teams more effective

GenAI is error prone, requiring strong human oversight: While there is a lot of hype around agentic GenAI systems[1], teams will need to account for inaccuracy and hallucination in agentic GenAI systems.[2]

AI's real value lies in progression: The biggest gains don't come from isolated use cases, but from integrating AI across the lifecycle, from preparation through detection to containment and recovery.

Trust and oversight are key initially but evolves in later levels: Early-stage adoption keeps humans fully in control. By L3 and L4, AI systems act independently within defined bounds, freeing humans for strategic oversight.

People's roles shift meaningfully: As AI matures, analyst roles consolidate and elevate from labor intensive task execution to high-value decision-making, focusing on critical, high business impact activities, improving processes and AI governance.

Outcome, not hype, defines maturity: AI maturity isn't about tech presence, it's about measurable impact on risk reduction, response time, and operational resilience.

Why Darktrace?

Darktrace has been at the forefront of cybersecurity innovation since 2013, with AI embedded in its mission from the very beginning.

Long before AI became a buzzword, we were building, testing, and operationalizing it in live environments to solve real-world security challenges. Today, the cybersecurity community is experiencing a surge of interest—and uncertainty—around AI. Many organizations are asking critical questions: What types of AI are truly effective? How much trust should be placed in automation? And how can teams separate real innovation from hype? To help answer these questions, we developed a maturity model that enables security teams to benchmark their AI adoption and cut through the noise.

At Darktrace, we've learned that no single AI technique can solve cybersecurity on its own. That's why our Self-Learning AI uses a multi-layered AI approach, strategically integrating a diverse set of techniques both sequentially and hierarchically.

This layered architecture allows us to deliver proactive, adaptive defense tailored to each organization's unique environment. Some of the core techniques we use include:

Unsupervised machine learning
to understand what's normal within each unique environment

Bayesian probabilistic modeling
to update risk assessments in real time as new data is observed

Clustering algorithms
to map peer groups and detect subtle anomalies

Ensemble methods and meta-classification
to refine detection and reduce false positives over time

■ AI at work: Darktrace by the numbers

90 million

■ autonomous investigations

conducted in 2024

3 million

■ incidents

escalated for human validation

500,000

■ confirmed critical incidents

43 million

■ hours

of manual analyst work saved

Up to 30

■ full-time analysts' output

delivered by Cyber AI Analyst™

10,000+

■ global security teams supported

90%

■ reduction

in false positives in customer environments

Importantly, our Self-Learning AI doesn't rely on static threat signatures or external threat feeds. Instead, it continuously learns from live data within each organization, flagging deviations in real time.

This enables the detection of novel threats such as insider attacks, unknown vulnerabilities, lateral movement, and misuse of legitimate tools, and ensures outputs are always relevant, interpretable, and tailored to each organization's evolving digital landscape.

This multi-layered AI approach also powers every layer of the Darktrace ActiveAI Security Platform, from dynamic threat detection and deep investigation to autonomous response and predictive assurance.

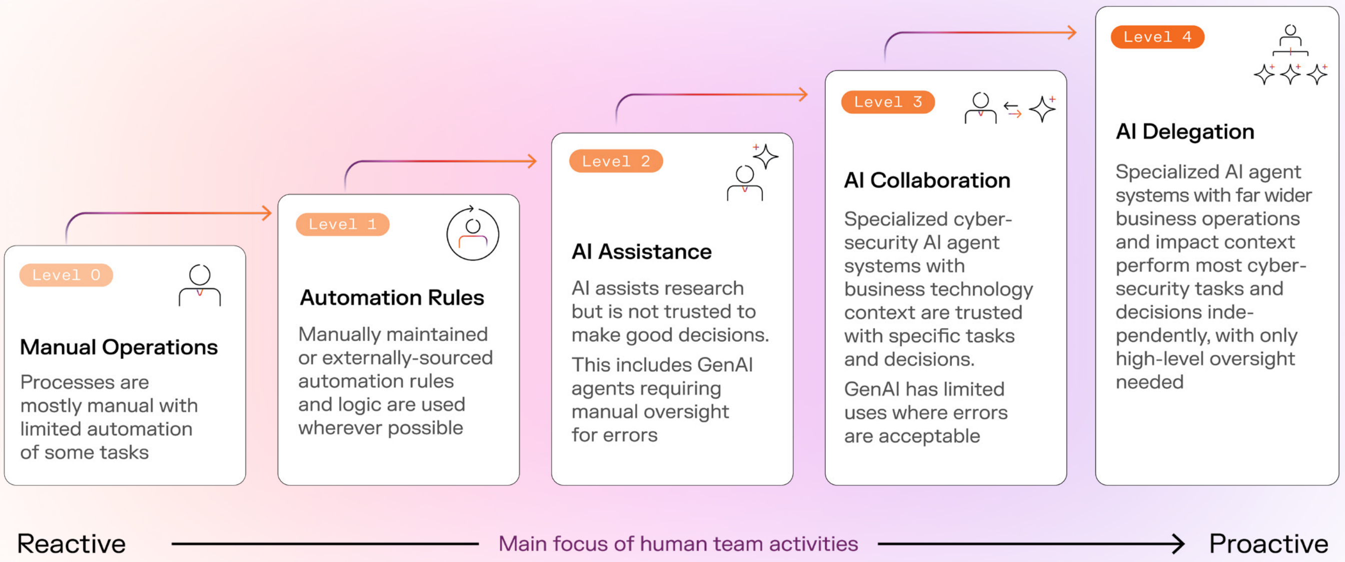
In 2024 alone, our AI conducted over 90 million autonomous investigations, helping nearly 10,000 security teams scale their capabilities, without increasing headcount.

Interested in learning about the different types of AI in cybersecurity?

Download our AI Arsenal Whitepaper to explore how AI models can be applied to cybersecurity, and how Darktrace's Self-Learning AI layers multiple techniques to deliver proactive, resilient threat defense.

Learn more ↗

Understanding outcomes, roles, and technology



Level 0

Manual Operations

Processes are mostly manual with limited automation of some tasks

Level 1

Automation Rules

Manually maintained or externally-sourced automation rules and logic are used wherever possible

Level 2

AI Assistance

AI assists research but is not trusted to make good decisions.

This includes GenAI agents requiring manual oversight for errors

Level 3

AI Collaboration

Specialized cyber-security AI agent systems with business technology context are trusted with specific tasks and decisions.

GenAI has limited uses where errors are acceptable

Level 4

AI Delegation

Specialized AI agent systems with far wider business operations and impact context perform most cyber-security tasks and decisions independently, with only high-level oversight needed

Reactive → Main focus of human team activities → Proactive

As AI capabilities evolve, their impact on cybersecurity **cannot be measured by technology alone**.

Meaningful transformation occurs when outcomes improve, staff responsibilities shift, and toolsets mature together. This model evaluates AI maturity across three dimensions: cybersecurity outcomes, the evolving role of people, and the underlying technology. By focusing on these interrelated aspects, security leaders can better assess where they stand and what progress looks like in real operational terms.

Each level of maturity represents a step-change in efficiency, accuracy, and strategic value, offering a clear path from reactive security to AI-driven resilience.

Level 0

Manual Operations

People's role: Analysts perform all tasks manually, this includes triage, investigation, patching, and reporting.

Technology: Basic tools with no automation. High effort, low scalability.

Cybersecurity outcome: Reactive and labor-intensive. Most alerts go uninvestigated. Vulnerability management and compliance are inconsistent.

Level 1

Automation Rules

People's role: Analysts write and maintain rules. SOAR handles repetitive tasks but requires constant tuning.

Technology: Signature-based detection, scheduled scans, and predefined workflows.

Cybersecurity outcome: Some efficiency gains through rule-based automation (e.g., SOAR, XDR). Still limited by human bandwidth.

Level 2

AI Assistance

People's role: LLMs assist with research, summarization and recommendations. Analysts role matures to higher value task, detection engineering load is reduced but Agentic GenAI is error prone requiring deep oversight.

Technology: AI assists research but is not trusted to make good decisions. This includes agentic GenAI requiring manual oversight for errors..

Cybersecurity outcome: AI supports triage and compliance, but errors and oversight limit trust. Detection accuracy improves.

Level 3

AI Collaboration

People's role: Analysts oversee AI, focus on high-risk decisions, and refine detection strategies.

Technology: Specialized cybersecurity agentic AI systems embedding business technology context are trusted with specific tasks and decisions. GenAI has limited use cases, where errors are acceptable.

Cybersecurity outcome: AI performs full investigations, recommends containment and recovery actions, and prioritizes risks based on attack paths.

Level 4

AI Delegation

People's role: Human teams provide strategic oversight and intervene only in critical scenarios.

Technology: Specialized AI agent systems with far wider business operations and impact context perform most cybersecurity tasks and decisions independently, with only high-level oversight needed.

Cybersecurity outcome: AI autonomously handles most security tasks, detection, triage, containment, and recovery, at machine speed.

Understanding capabilities across security operations

As organizations advance through the AI maturity model, core capabilities within security operations will evolve and shift from fully manual processes to autonomous decision-making. The chart below illustrates an organization's evolution, with each row representing a foundational capability within security operations, such as alert investigation, threat detection, or incident response.

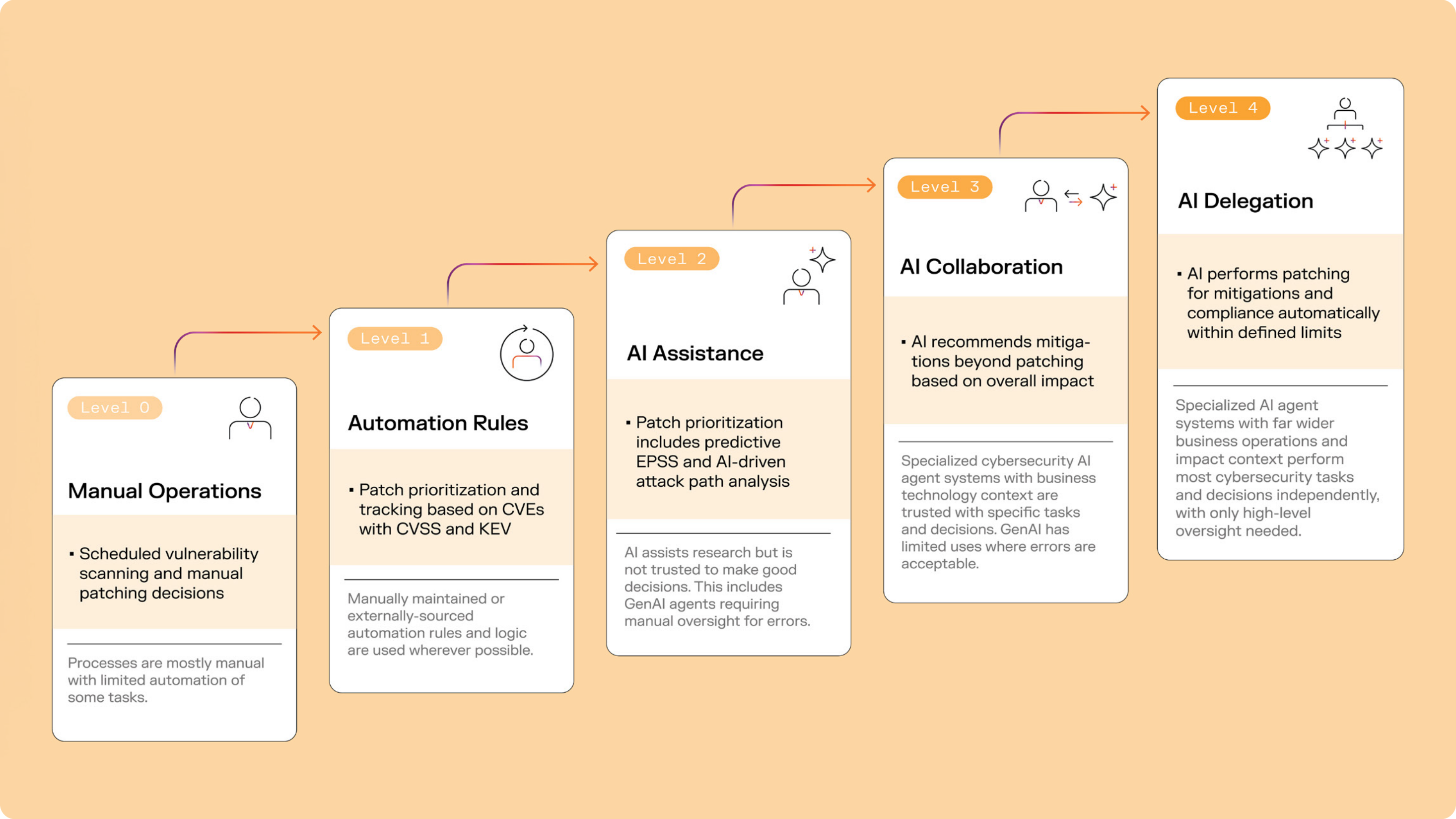
At each maturity level (L0 to L4), you'll see how AI changes the way work is performed, the quality of outcomes, and the role of human analysts.

The model provides a side-by-side view of the incremental impact AI has on efficiency, accuracy, and operational resilience, helping you benchmark current capabilities and identify realistic next steps for transformation.

Capability	L0 – Manual	L1 – Automation Rules	L2 – AI Assistance	L3 – AI Collaboration	L4 – AI Delegation
Alert Investigation	Manual triage; <25% alerts investigated	SOAR handles common triage; rest manual	LLMs assist triage; error-prone	AI fully investigates ALL alerts	AI autonomously investigates and escalates only critical cases
Threat Detection	Signature-based rules; low accuracy	Technique-based rules; limited scope	AI improves accuracy; still rule-bound	AI detects unknown threats via org behavior analysis	AI continuously adapts detections with minimal human input
Containment & Recovery	Manual playbooks and actions	SOAR automates common responses	LLMs recommend actions; oversight needed	AI executes containment and tailoredrecovery within limits	AI handles most containment and tailored recovery autonomously
Risk Management	Scheduled vulnerability scanning and manual patching decisions.	Patch prioritization and tracking based on CVEs with CVSS and KEV.	Prioritization includes predictive EPSS and AI-driven attack path analysis.	AI recommends mitigations beyond patching based on overall impact.	AI performs mitigations automatically within defined limits
Threat Hunting	Manual IOC searches; high effort	Automated IOC matching	AI finds similar threats; lowers skill floor	AI hunts emerging threats; expands reach	AI minimizes threat risk with minimal human time
People's Role	Fully hands-on; all tasks manual	Rule creation and SOAR maintenance	Oversight of AI suggestions	Strategic oversight and tuning of AI	High-level governance and critical decision-making only

Evolution of Risk Management

From periodic patching to continuous, prioritized mitigation



Level 0

Manual Operations

- Scheduled vulnerability scanning and manual patching decisions

Processes are mostly manual with limited automation of some tasks.

Level 1

Automation Rules

- Patch prioritization and tracking based on CVEs with CVSS and KEV

Manually maintained or externally-sourced automation rules and logic are used wherever possible.

Level 2

AI Assistance

- Patch prioritization includes predictive EPSS and AI-driven attack path analysis

AI assists research but is not trusted to make good decisions. This includes GenAI agents requiring manual oversight for errors.

Level 3

AI Collaboration

- AI recommends mitigations beyond patching based on overall impact

Specialized cybersecurity AI agent systems with business technology context are trusted with specific tasks and decisions. GenAI has limited uses where errors are acceptable.

Level 4

AI Delegation

- AI performs patching for mitigations and compliance automatically within defined limits

Specialized AI agent systems with far wider business operations and impact context perform most cybersecurity tasks and decisions independently, with only high-level oversight needed.

Risk Management contains Exposure Management, Vulnerability Management and Compliance scope of activities

Why risk management needs to evolve

Traditional approaches to risk management often lack the agility to keep up with today's evolving threats. Many programs still rely on fixed schedules and severity scores that don't reflect an organization's real-time exposure or operational priorities. As the attack surface grows and vulnerabilities are weaponized faster, organizations must move beyond patching by score and toward more context-aware methods.

The question isn't just how to apply AI, but how to ensure it informs the right decisions, at the right time, with clear accountability.

Transformation trends

- **From CVE lists to real-world exposure:** AI evaluates how vulnerabilities interact across systems, prioritizing based on real-world exploitability
- **From point in time patching to autonomous mitigation:** AI not only recommends but executes mitigations, reducing time-to-remediation
- **From static compliance to continuous assurance:** AI ensures compliance is maintained in real time, not just at audit checkpoints

Key takeaways

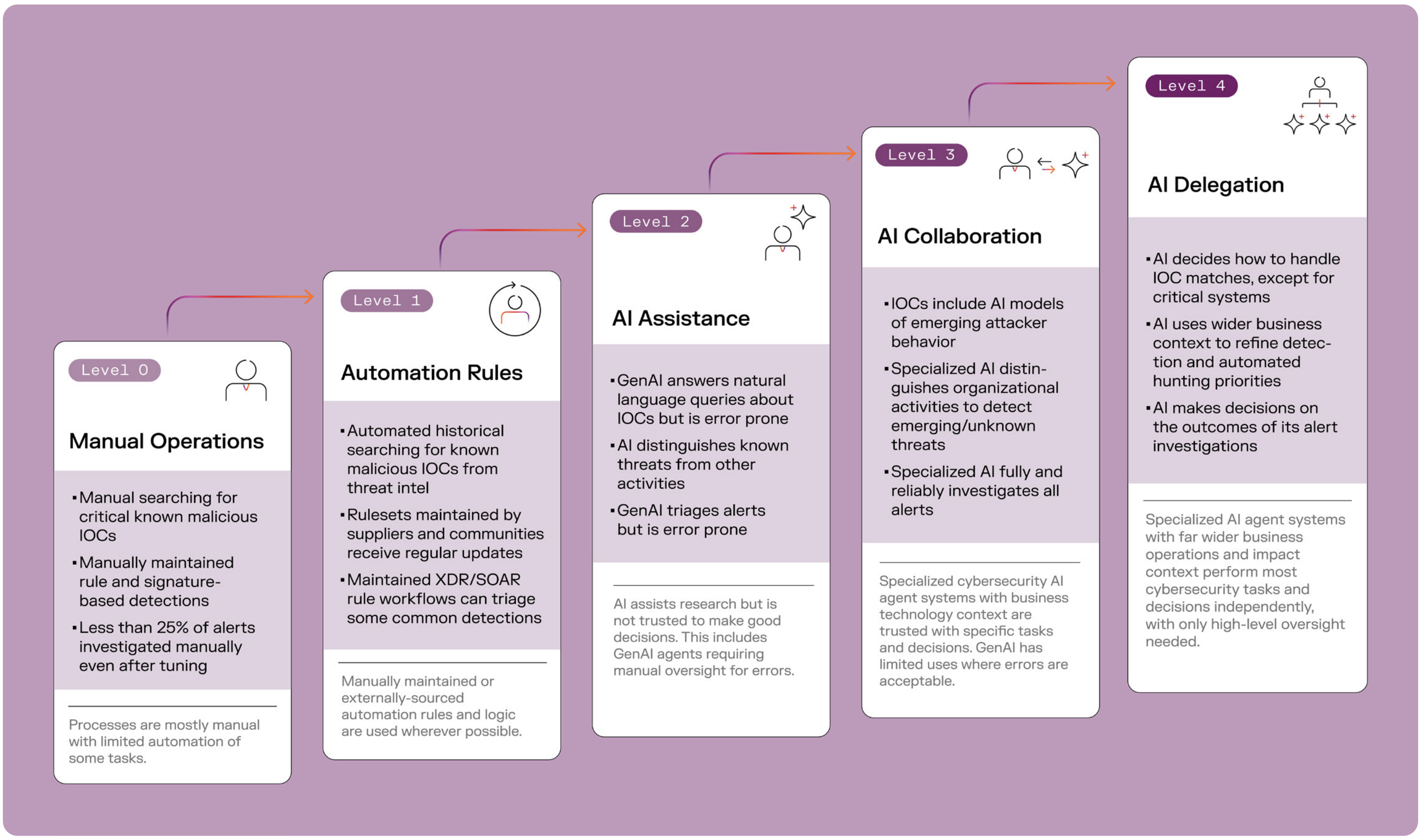
- **Reduce** manual workload and patching delays
- **Prioritize** what truly matters based on business risk
- **Identify** risk before it is exploited
- **Manage** risk across complex environments with minimal human effort

Risk management outcomes across maturity levels

Capability	L0 Manual Operations	L1 Automation Rules	L2 AI Assistance	L3 AI Collaboration	L4 AI Delegation
Overall Outcome	Highest known vulnerabilities are patched periodically where possible	Highest known vulnerabilities and those with active exploits are patched as soon as possible	Highest known vulnerabilities and those most likely to be exploited are patched when possible	Highest risks are mitigated using many methods, including where unknown exploits or misuse of legitimate services could do the most damage	Internal exposure risks are minimized while taking minimal people time
Vulnerability Prioritization	Manual CVE lists	CVSS, KEV-based scoring	Adds predictive EPSS insights and attack path analysis	Adds unknown exploits or misuse of legitimate services could do the most damage	Business-contextual, autonomous prioritization
Mitigation Actions	Manual patching	Scheduled patching	AI suggests actions	AI recommends and sequences mitigations	AI executes mitigations within limits
Exposure Visibility	Known assets only	Known assets + scheduled scans	AI discovers more assets	AI maps full exposure graph on business context	AI maintains real-time exposure awareness
Compliance	Point-in-time audits	More frequent checks	AI suggests improvements	AI recommends impactful changes	AI enforces compliance automatically
People's Role	Fully manual	Manual patch installation based on suggested CVE prioritization	Manual patch installation based on suggested CVE prioritization	Manual mitigation actions including patch installation and many more hardening and configuration changes	People oversee the AI and make high cost and risk decisions, notably about critical systems and services

Evolution of the SOC

Evolving threat intelligence, hunting, detection and alert investigation



Level 0

Manual Operations

- Manual searching for critical known malicious IOCs
- Manually maintained rule and signature-based detections
- Less than 25% of alerts investigated manually even after tuning

Processes are mostly manual with limited automation of some tasks.

Level 1

Automation Rules

- Automated historical searching for known malicious IOCs from threat intel
- Rulesets maintained by suppliers and communities receive regular updates
- Maintained XDR/SOAR rule workflows can triage some common detections

Manually maintained or externally-sourced automation rules and logic are used wherever possible.

Level 2

AI Assistance

- GenAI answers natural language queries about IOCs but is error prone
- AI distinguishes known threats from other activities
- GenAI triages alerts but is error prone

AI assists research but is not trusted to make good decisions. This includes GenAI agents requiring manual oversight for errors.

Level 3

AI Collaboration

- IOCs include AI models of emerging attacker behavior
- Specialized AI distinguishes organizational activities to detect emerging/unknown threats
- Specialized AI fully and reliably investigates all alerts

Specialized cybersecurity AI agent systems with business technology context are trusted with specific tasks and decisions. GenAI has limited uses where errors are acceptable.

Level 4

AI Delegation

- AI decides how to handle IOC matches, except for critical systems
- AI uses wider business context to refine detection and automated hunting priorities
- AI makes decisions on the outcomes of its alert investigations

Specialized AI agent systems with far wider business operations and impact context perform most cybersecurity tasks and decisions independently, with only high-level oversight needed.

SOC includes Threat Intelligence and Hunting, Threat Detection and alerting, Alert Investigation scope of activities

Security Operations Centers (SOCs) sit at the heart of cybersecurity defense, but today's demands are stretching them to the limit.

Teams face a constant barrage of alerts, often without the tools or time to determine which ones matter most. Even experienced analysts find it difficult to connect fragmented signals across tools and data sources. As infrastructure grows more complex and attacks more subtle, traditional workflows, built for linear processes, struggle to keep up with the speed and scale of modern threats. A new model is needed, one where AI augments investigation, prioritization, and response at scale.

By automating repetitive tasks, enhancing detection accuracy, and enabling proactive threat hunting, AI empowers SOCs to evolve from reactive, resource-intensive operations into intelligent, adaptive, and scalable defense systems.

01

Threat Intelligence & Hunting

Threat intelligence and hunting are critical for identifying both known and unknown threats. Traditionally, this has been a manual, high-effort process focused on known indicators of compromise (IOCs).

AI enables a shift toward proactive, behavior-based threat hunting that scales with the organization.

Transformation trends

- **Shift** from IOC-based to behavior-based detection
- Use of AI models to **identify** emerging threats before confirmation
- **Integration** of threat intelligence across multiple data sources
- **Reduction** in skill barriers for effective hunting

Key takeaways

- AI **expands** threat coverage beyond known IOCs
- **Reduces** manual effort and increases speed
- **Enables** earlier detection of novel threats
- **Elevates** analysts to strategic oversight roles

Key characteristics by maturity level

Maturity Level	Key Characteristics
L0 – Manual Operations	Critical known malicious IOCs are searched for manually; High effort for unknowns
L1 – Automation Rules	Automated historical searching for known malicious IOCs from threat intel
L2 – AI Assistance	GenAI answers natural language queries including data from integrations, but is error prone.
L3 – AI Collaboration	IOCs include AI models of emerging attacker behavior
L4 – AI Delegation	AI decides how to handle IOC matches, except for critical systems

Outcomes and people roles

Maturity Level	Outcomes	People Roles
L0	Known threats checked manually; high effort	L3 analysts perform all threat hunting manually
L1	Reduced effort for known threats	Analysts use tools to automate IOC lookups
L2	Broader coverage; GenAI lowers investigation skill floor at the expense of introducing errors.	L3 Analysts validate AI findings
L3	Likely threats emerging elswehere are checked for, even before they are confirmed as malicious.	L3 SOC analyst role merges with L2 as unknown threat detection and hunting grow closer
L4	Risks from threats and likely threats are minimised while taking minimal people time	Combined SOC analyst role oversees AI and acts on high importance situations

02

Threat Detection & Alerting

Detection and alerting are foundational to SOC operations, but traditional systems often generate excessive false positives and miss stealthy attacks.

AI enhances detection by learning organizational behavior and adapting to new threats in real time.

Transformation trends

- From static rules to **adaptive**, AI-driven detection
- **Context-aware** alerting based on organizational behavior
- **Reduced** false positives and alert fatigue
- **Continuous** improvement without manual tuning

Key takeaways

- AI **improves** detection accuracy and reduces noise
- AI **distinguishes** organizational activities to detect emerging/unknown threats. Detection engineering is mainly left for AI systems and overseen or intervened in high importance situations
- **Supports** scalable, resilient detection strategies

Key characteristics by maturity level

Maturity Level	Key Characteristics
L0 – Manual Operations	Manually created rule and signature-based detections.
L1 – Automation Rules	Tools make standard rulesets available, and further rules are applied by SIEM/XDR across multiple data sources.
L2 – AI Assistance	AI distinguishes known threats from other activities.
L3 – AI Collaboration	AI distinguishes organizational activities to detect emerging/unknown threats.
L4 – AI Delegation	AI continuously refines detections autonomously with broader organization context

Outcomes and people roles

Maturity Level	Outcomes	People Roles
L0	High false positives; low fidelity	Detection engineering adds and tunes new rules against alert rates and effectiveness
L1	Slightly improved accuracy	Detection engineering adds and tunes technique-based rules against alert rates and effectiveness
L2	Improved accuracy; still rate-limited	Detection engineering maintenance load is reduced as some rules are replaced by more accurate AI classifications
L3	Detects unknown threats; less maintenance	Detection engineering maintenance load is reduced as AI can distinguish the organisation's activities from others
L4	Broad detection with minimal effort	Detection engineering is mainly left to AI systems and overseen or intervened in high importance situations

03

Alert Investigation

Alert investigation is often the most time-consuming part of SOC operations. With limited resources, many alerts go uninvestigated, and potential threats often get missed. AI helps reduce this burden by triaging alerts, enriching them with context, and investigating routine cases.

This enables broader alert coverage, faster prioritization, and quicker escalation of truly risky activity.

Transformation trends

- AI systems **handle** all initial triage and investigation of alerts, reducing time to insight.
- Fewer alerts are missed or dismissed without review, thanks to **automated** analysis. Reducing overall risk
- AI **integrates** data from multiple sources, such as endpoint, network, and identity tools, for more context-rich decisions.
- Analysts are freed up to focus on unknown threats, complex investigations, and **refining** AI workflows

Key takeaways

- AI **enables** broader and faster alert investigation, reducing triage backlogs.
- Automation **lowers** the risk of missing critical alerts or mishandling early indicators.
- SOC analysts spend more time on **high-value** investigations and less on repetitive reviews.
- Investigation workflows continuously **improve** as AI learns from analyst actions and outcomes.

Outcomes across maturity levels

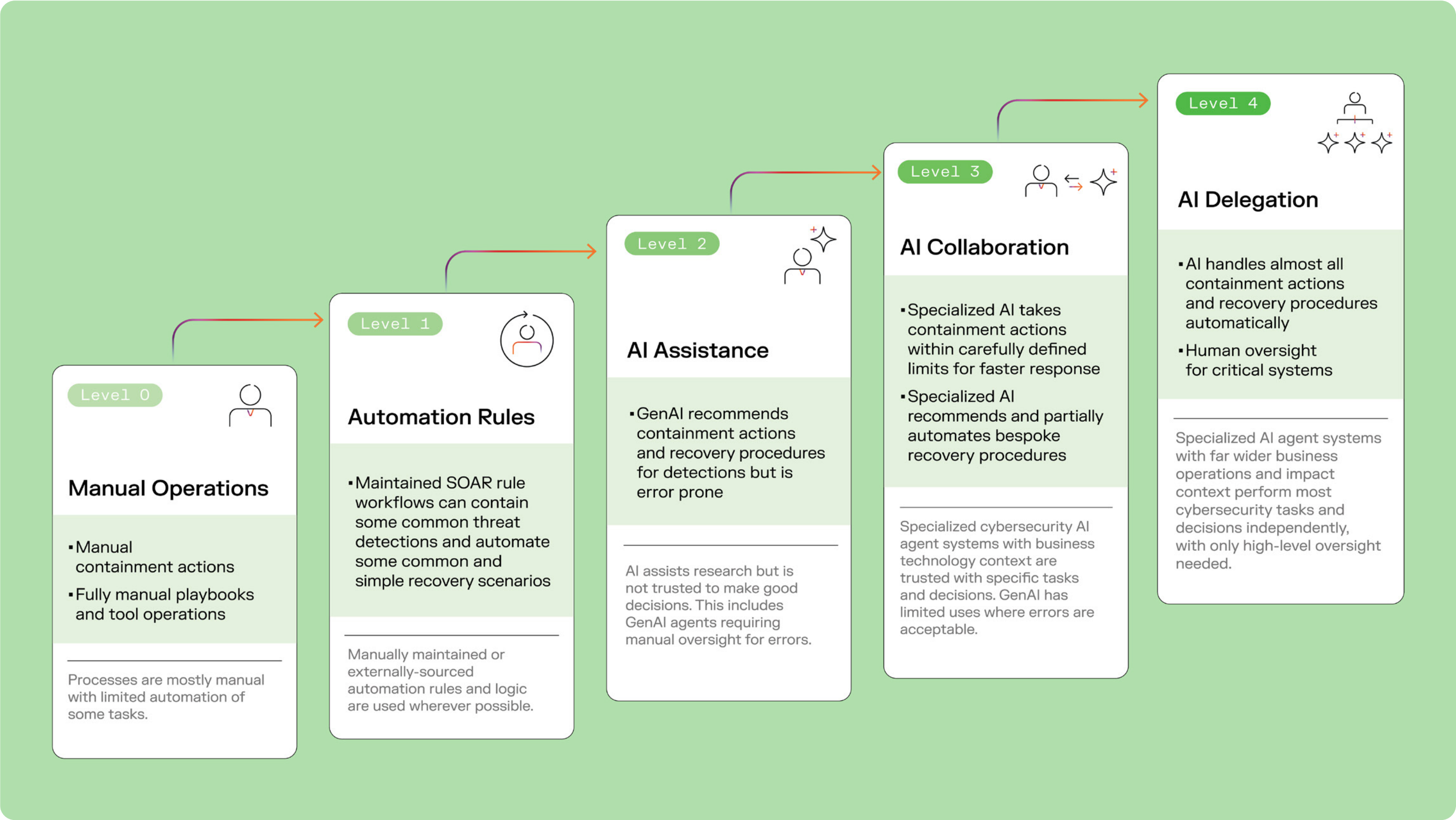
Maturity Level	Key Characteristics
L0 – Manual Operations	Manual triage; <25% alerts investigated
L1 – Automation Rules	SOAR handles common triage; rest manual
L2 – AI Assistance	GenAI assists triage; error-prone
L3 – AI Collaboration	Specialized AI fully investigates alerts and is not error prone.
L4 – AI Delegation	AI makes decisions on outcome of alert investigations.

Outcomes and people roles

Maturity Level	Outcomes	People Roles
L0	<25% of alerts investigated	L1/L2 analysts handle all investigations manually
L1	Slightly improved triage, most troublesome rules are handled by SOAR	L1 SOC analyst role effort is partially converted to writing and maintaining SOAR rules
L2	Detections are all triaged by GenAI, but these introduce errors of their own	L1 SOC analysts convert to L2 but have to deal with GenAI errors in triage
L3	Detections are fully investigated by an AI Analyst	L2/L3 roles merge to follows up on AI investigation, and hunting
L4	Machine-speed triage and resolution	Analysts oversee AI and handle critical cases

Evolution of Containment & Recovery

From static playbooks to adaptive, AI-driven response



Level 0

Manual Operations

- Manual containment actions
- Fully manual playbooks and tool operations

Processes are mostly manual with limited automation of some tasks.

Level 1

Automation Rules

- Maintained SOAR rule workflows can contain some common threat detections and automate some common and simple recovery scenarios

Manually maintained or externally-sourced automation rules and logic are used wherever possible.

Level 2

AI Assistance

- GenAI recommends containment actions and recovery procedures for detections but is error prone

AI assists research but is not trusted to make good decisions. This includes GenAI agents requiring manual oversight for errors.

Level 3

AI Collaboration

- Specialized AI takes containment actions within carefully defined limits for faster response
- Specialized AI recommends and partially automates bespoke recovery procedures

Specialized cybersecurity AI agent systems with business technology context are trusted with specific tasks and decisions. GenAI has limited uses where errors are acceptable.

Level 4

AI Delegation

- AI handles almost all containment actions and recovery procedures automatically
- Human oversight for critical systems

Specialized AI agent systems with far wider business operations and impact context perform most cybersecurity tasks and decisions independently, with only high-level oversight needed.

Containment & Recovery includes functions of initial containment, remediation and recovery scope of activities

When a threat is detected, the speed of containment and recovery often determines whether the impact is minor or business-altering. Traditionally, these actions rely on predefined playbooks and point in time steps, which can **delay response, allow threats to spread, and drive-up costs.**

AI changes this by speeding up decisions, recommending recovery paths, and, at higher maturity levels, taking containment and remediation actions on its own.

Early implementations might **automate common use cases** through a SOAR, while more advanced approaches use agentic AI to understand context and act within predefined risk boundaries.

Transformation trends

- Organizations **evolve** from static playbooks to dynamic response strategies based on the specifics of each incident
- Common incidents are addressed autonomously by AI, **freeing** analysts to focus on novel or high-risk threats
- Specialized AI systems can contain threats, isolate systems, and trigger recovery workflows within safe parameters **defined** by human teams
- AI systems **improve** over time by learning from past incident outcomes and analyst decisions

Key takeaways

- AI significantly **reduces** time-to-containment and recovery, helping prevent lateral movement and minimizing impact on business continuity
- AI **ensures** standardized, repeatable responses across incidents while creating bespoke responses for unique incidents
- Security teams spend less time on repetitive tasks and more time **shaping** long-term defense strategies
- AI systems **evolve** with use, refining responses based on past outcomes and analyst feedback
- Incident response becomes **scalable** with the organization without requiring proportional headcount increases

Outcomes across maturity levels

Capability	L0 Manual Operations	L1 Automation Rules	L2 AI Assistance	L3 AI Collaboration	L4 AI Delegation
Initial Containment	None. Manual Containment	Maintained SOAR rule workflows can contain some common detections	GenAI recommends containment actions but is error prone.	Specialized AI takes containment actions within carefully defined limits; faster response.	AI handles almost all containment actions automatically; human oversight for critical systems
Remediation and Recovery	Fully manual playbooks and tool operation	SOAR rule workflows can automate some common and simple recovery scenarios.	GenAI recommends recovery procedures for detections but is error prone.	Specialized AI recommends bespoke recovery procedures and partially automates	AI performs recovery procedures automatically within carefully defined limits
People Roles	Maintaining a limited number of playbooks and checking tool functionality	SOAR requires full time maintenance to handle common use cases, playbooks are maintained in SOAR with similar efforts	GenAI offers assistance in recovering from incidents, although they are error prone and require oversight	Specialized agentic AI offers assistance in recovering from incidents	People oversee the business-aware agentic AI and make only high cost and risk decisions

Conclusion

Reclaiming time and resources

As organizations progress through the AI maturity model, one of the most significant benefits is the **reduction in time and effort** spent on repetitive, high-volume, and low-value tasks. This graph below highlights how the primary focus areas of security teams shift as organizations progress through the AI Maturity Model for Cybersecurity.

Rather than thinking about AI adoption as a single-step transformation, security leaders can use this view to understand which operational responsibilities evolve at each maturity level, from manual triage and patching to cross-functional improvement and oversight of AI-driven decisions. It provides a clear, role-based perspective on how AI gradually takes over specific tasks, freeing up human expertise for higher-level strategy and continuous improvement. **This helps security leaders align AI investments with tangible operational outcomes and workforce planning.**

Top time constraints	L0 – Manual Operations	L1 – Automation Rules	L2 – AI Assistance	L3 – AI Collaboration	L4 – AI Delegation
#1	Alert investigation (triage)	Alert investigation (triage)	Alert investigation (investigation)	Alert investigation (investigation)	AI oversight and key decisions (all)
#2	Vulnerability management (patching)	Vulnerability management (patching)	Internal exposure management (patching)	Internal exposure management (mitigations)	Continuous improvement (all)
#3	Threat detection and alerting (tuning)	SOAR maintenance (triage/containment)	SOAR maintenance (triage/containment)	Continuous improvement (all)	Continuous improvement of AI systems not focused on one task (all)

What this tells us

At lower maturity levels, security operations teams are often weighed down by labor intensive triage, patching, and rule tuning. These are repetitive and reactive tasks that demand significant time and resources. As organizations advance into the stages of AI Assistance and Collaboration, much of this operational burden begins to shift. AI systems take on triage and detection duties, enabling human analysts to focus more on investigation and mitigation.

At the highest level, AI Delegation, the role of the human team evolves further. Teams are now focused on strategic oversight, governance, and the continuous refinement of AI-driven processes to ensure resilience and effectiveness.

Strategic implications on investing in the RIGHT AI for CISOs

Free up analyst time: AI reduces the burden of triage and detection, allowing analysts to focus on higher-value activities like threat hunting and response strategy.

Shift from execution to oversight: As AI matures, the role of the security team evolves from doing the work to ensuring the work is done right.

Invest in continuous improvement: At higher maturity levels, the most valuable use of human time is in refining AI models, improving data quality, and aligning security with business risk.

Build for scale: AI enables SOCs to scale without proportional increases in headcount, making security operations more sustainable and resilient.

Self-assessment checklist for existing tools

Use this self-assessment rubric to evaluate where your current tools and processes align within the maturity model framework.

The rubric below serves as a guide to help you determine the level of maturity your existing solutions have achieved. As you work through this exercise, use the accompanying self-assessment sheet to document your tools and insights for each stage of the evaluation.

		L0	L1	L2	L3	L4
Category	Assessment Criteria	Manual Operations	Automation Rules	AI Assistance	AI Collaboration	AI Delegation
Risk Management (includes Exposure, Vuln Mgt)	Are risks prioritized based on exploitability and business impact?	Scheduled vulnerability scanning and manual patching decisions.	Patch prioritization and tracking based on CVEs with CVSS and KEV.	Prioritization includes predictive EPSS and AI-driven attack path analysis.	AI recommends mitigations beyond patching based on overall impact.	AI performs patching for mitigations automatically within defined limits.
Threat Hunting	Is hunting proactive and AI-assisted?	Critical known malicious IOCs are searched for manually	Automated historical searching for known malicious IOCs from threat intel	GenAI answers natural language queries including data from integrations, but is error prone.	IOCs include AI models of emerging attacker behaviour	AI decides how to handle IOC matches, except for critical systems
	Are unknown threats regularly investigated?	Rarely	Occasionally	Prompted by AI	AI finds unknown threats based on Organization behavior	Continuously by AI
Threat Detection	Are detections adaptive and context-aware?	Static rules	Technique-based rules (XDR/SOAR)	AI distinguishes known threats from other activities.	AI distinguishes organizational activities to detect emerging/unknown threats.	AI uses wider business context to refine detection and automated hunting priorities
	Is Detection Engineering automated?	Detection engineering load is fully manual and siloed	Standard and community rulesets are used, along with cross-domain rules	Detection engineering maintenance load is reduced as some rules are replaced by more accurate AI classifications	Significant reduction as AI can distinguish the organization's activities from others	Detection engineering is mainly left to AI systems and overseen or intervened in high importance situations
Alert Investigation	Is triage automated or AI-assisted?	Manual triage	SOAR workflows triage common alerts	GenAI workflows triage alerts but are error prone	Specialized AI fully investigates alerts and is not error prone.	AI makes decisions on the outcomes of alert investigations.
	Are investigations enriched with context?					
Initial Containment	Can containment occur autonomously?	No	SOAR workflows for known cases	GenAI recommends containment actions for detections but is error prone.	AI executes actions within limits	AI contains threats autonomously
	Is containment speed sufficient?	Always long delays	Short delays for known threats	Short delays for common scenarios	Near real-time even for novel scenarios	Near real-time and aware of wider business context when taking decisions
Recovery	Are recovery actions adaptive and AI-driven?	Manual playbooks	SOAR workflows for common cases	GenAI recommends recovery procedures for detections but is error prone.	Specialized AI recommends bespoke recovery	AI executes recovery autonomously within carefully defined limits

How does your security stack match up?

Add your tools and evaluate their maturity level

Category	Assessment Criteria	L0 Manual Operations	L1 Automation Rules	L2 AI Assistance	L3 AI Collaboration	L4 AI Delegation
Risk Management (includes Exposure, Vuln Mgt)	Are risks prioritized based on exploitability and business impact?					
Threat Hunting	Is hunting proactive and AI-assisted?					
	Are unknown threats regularly investigated?					
Threat Detection	Are detections adaptive and context-aware?					
	Is Detection Engineering automated?					
Alert Investigation	Is triage automated or AI-assisted?					
	Are investigations enriched with context?					
Initial Containment	Can containment occur autonomously?					
	Is containment speed sufficient?					
Recovery	Are recovery actions adaptive and AI-driven?					

■ **About Darktrace**

Darktrace is a global leader in AI cybersecurity that keeps organizations ahead of the changing threat landscape every day. Founded in 2013 in Cambridge, UK, Darktrace provides the essential cybersecurity platform to protect organizations from unknown threats using AI that learns from each business in real-time. Darktrace's platform and services are supported by 2,700+ employees who protect nearly 10,000 customers globally. To learn more, visit www.darktrace.com.

North America: +1 (415) 229 9100 Europe: +44 (0) 1223 394 100 Asia-Pacific: +65 6804 5010 Latin America: +55 11 4949 7696

darktrace.com | info@darktrace.com