Scrut Automation

EBOOK

AI in the trenches: Insights from seven security leaders redefining modern risk

Merritt Baer

Nicholas Muy

Aayush Ghosh

Srikanth Chavali

Sanket Naik

Ian Amit

Kevin Qiu

Scrut Automation

Featured contributors

Aayush Ghosh Chaudhary

Aayush Ghosh Choudhury is Co-founder and CEO at Scrut Automation. After advising Fortune 500 clients on digital transformation at McKinsey & Company and co-founding a procurement SaaS startup, he set out to reinvent the way organizations manage risk. Aayush serves on the Forbes Technology Council, regularly speaks at industry events like SecuFest and Black Hat, and is a recognized thought leader in AI-driven GRC and security automation.

Nick Muy

Nick Muy is CISO at Scrut Automation, where he leads cybersecurity efforts across compliance and AI risk management. With over a decade of experience in AI-powered threat modeling and enterprise security strategy, he has secured Fortune 500 environments and shaped national cyber policy at the U.S. Department of Homeland Security.

Merritt Baer

Merritt Baer is CISO at Reco and a former Deputy CISO at AWS, where she advised Fortune 100 companies and helped close $2 billion in cloud deals. She has held security roles in all three U.S. government branches and is a trusted advisor to AI and security startups. A Harvard Law graduate, she is a noted speaker, writer, and advocate for inclusion in cybersecurity.

Iftach Ian Amit

Iftach Ian Amit is the Co-founder and CEO of Gomboc AI, and a 25-year cybersecurity veteran. He has led security at Amazon, Cimpress, Rapid7, and more and was a 2025 SC Awards “Security Person of the Year” finalist. Known for advancing cloud infrastructure defense, he also mentors the next generation through IANS, DEFCON, and BSides Las Vegas.

Sanket Naik

Sanket Naik is Founder and CEO at Palosade, an AI-driven platform helping enterprises automate security operations at scale. Previously SVP of Engineering at Coupa, he built its global cloud and cybersecurity org from the ground up through IPO. He’s held roles at HP and Qualys and advises early-stage startups.

Srikanth Chavali

Srikanth Chavali is Co-founder and CPO at Kitecyber, where he leads cybersecurity R&D to combat first-order threats. With deep expertise in network architecture and TCP/IP security, he has built a team of researchers solving real-world vulnerabilities at scale. He brings over a decade of experience driving innovation in enterprise cyber defense.

Kevin Qiu

Kevin Qiu is a security practitioner with 10+ years of experience across industries. He has helped companies build category-defining trust centers to streamline security reviews. He advises several security startups and mentors professionals entering the field, drawing on his deep knowledge of cloud security, enterprise risk, and trust-led growth.

Scrut Automation

Table of contents

AI is no longer optional.
AI is inevitable.

"The key lies in finding the middle ground: adopting AI responsibly, supported by strong guardrails and governance frameworks."

Aayush Ghosh Choudhury
Co-founder and CEO, Scrut Automation

AI is here to stay—not just as a trend but as a long-term shift in how we live and do business. It is no longer a question of whether AI should be used but rather how it should be implemented responsibly, securely, and effectively.

As AI systems integrate into developer tooling, automation stacks, and production workflows, new failure modes are surfacing—from prompt injection and context leakage to agent overreach and control bypass.

Governance must evolve to cover model lifecycle integrity, dependency scrutiny, and real-time inference monitoring. Without visibility into inputs, outputs, and downstream actions, AI systems execute unaudited decisions at scale. Safe adoption depends on treating outputs as untrusted by default—embedding containment, validation, and rollback directly into the pipelines AI is accelerating.

In this eBook, we bring together perspectives from seasoned security and technology leaders who are already executing AI responsibly. They share how AI when implemented thoughtfully, can transcend traditional business barriers, reduce risk, and unlock new revenue opportunities.

"AI doesn't replace best practices—it demands them."

Kevin Qiu
Security Leader, Shiftsmart

01 The AI advantage — Business at the speed of thought

> “Execution speed must be matched by assurance. Enterprises embracing AI should also invest in explainability infrastructure and continuous validation, especially when AI outputs touch production systems.”
>
> 
>
> **Ian Amit**
> Co-founder and CEO, Gomboc AI

The enterprise has changed forever

AI heralds a foundational shift in how companies operate. The old boundaries between departments, roles, and responsibilities are dissolving, giving way to more fluid, cross-functional systems. According to McKinsey’s State of AI report, 78% of organizations have adopted some form of AI—a number that continues to rise. IDC projects global AI spending will reach $500 billion by 2027. In this rapidly evolving landscape, companies must move beyond experimentation toward execution—with governance and oversight built in. Decision velocity is no longer a luxury. It’s a competitive advantage.

As Ian Amit, Co-founder and CEO aptly notes, AI decisions aren’t just fast—they’re high-impact. And when those decisions touch customer experiences, financial operations, or security postures, trust and traceability become non-negotiable. They aren’t afterthoughts but prerequisites for innovation.

AI is also reshaping the foundations of enterprise strategy, forcing leaders to rethink what true differentiation looks like in an AI-native business.

Strategic planning evolves into real-time, data-driven decisioning powered by continuous feedback loops and probabilistic modeling. Talent acquisition prioritizes model literacy and AI integration skills. Value creation shifts from incremental efficiency to scalable innovation driven by AI-first architectures.

The AI Maturity Curve – Moving from task automation to enterprise reinvention

Figure 1. Stages of AI integration across the enterprise

How AI touches every business function

Product: User behavior insights driving rapid experimentation; generative tools assisting with UI/UX ideation.

Customer Support: Chatbots and virtual agents delivering 24/7 multilingual support; real-time escalation of high-priority issues.

Finance: AI-assisted forecasting, fraud detection, and automatic reconciliation of complex transactions.

Marketing: Personalized campaigns generated on demand; automated A/B testing; content scaled globally in minutes.

Sales: Lead qualification and prioritization done through predictive scoring; AI copilots drafting emails and scripts.

HR: Intelligent resume screening, sentiment analysis in exit interviews, and attrition risk models.

> "AI isn't just a new technology, it's new enterprise DNA. Just as the internet enabled any business to be born global, AI turns every idea into instant execution. If it can be thought of, it can be done. That's how Palosade was born, and we bring that same DNA to help organizations weave AI into their innovations and operating fabric without the fear or risk of malpractice."

Sanket Naik
Founder and CEO, Palosade.

02 The other edge of the sword — AI-driven threats on the rise

> “The impulse with AI is to try and solve every hypothetical risk, which is impossible. A pragmatic approach is essential.”
>
> 
> **Nick Muy**
> CISO, Scrut Automation

Tackling the dual nature of AI – From innovation to exploitation

MITRE’s “AI Trust Gap” notes that while AI can be transformative, it has the potential for misuse. In congressional testimony prepared by MITRE experts, they warn that adversaries are already weaponizing AI to “create polymorphic malware that evades traditional defenses,” underscoring that many existing detection tools simply aren’t built to keep up with code that mutates on the fly.

When AI turns malicious

AI is supercharging cyberattacks. Threat actors are now leveraging generative models and automation to scale deception, precision, and volume:

- **Deepfakes** are being used to impersonate executives in social engineering campaigns. In 2023, a UK energy firm lost over $250,000 when a fraudster used voice cloning to imitate its CEO.
- **LLM-powered phishing** creates emails indistinguishable from internal communications. A 2025 CrowdStrike Global Threat Report highlighted that phishing emails written by AI achieved a 54% CTR, whereas those crafted manually by humans only saw a 12% rate.
- **Polymorphic malware** evolves dynamically, rewriting itself in real-time to evade signature-based detection systems. In recent red-team tests, AI-written malware bypassed 80%-95% of traditional AV systems.

Signature-based security is falling behind modern threats

Traditional cybersecurity defenses rely on static indicators and known attack signatures. In a world of AI-powered threats, these defenses are rapidly becoming obsolete.

Figure 2. Why traditional defenses fail against modern threats

In analyses of encrypted command-and-control traffic (often used by AI-enhanced malware), behavior-based systems detected 82% of sessions, while signature-based tools saw only 18%.

Security teams can no longer rely on historical data alone. The new standard must be real-time, behavior-based detection and automated, context-aware responses.

Enter the autonomous adversary

The emergence of autonomous, self-directed threats marks a new era in cyber risk. We are no longer dealing solely with human-led attacks but with threats that think and act independently.

- **Agentic malware** can navigate networks and make decisions without human input
- **AI-driven reconnaissance tools** mine public and internal data to construct hyper-personalized attack paths
- **Prompt injection attacks** manipulate enterprise LLMs to leak information or execute unauthorized actions

These AI threats can identify vulnerabilities, bypass controls, and exfiltrate data with minimal human coordination, introducing a new threat category: **autonomous compromise.**

03 From outside to inside — the threats we let in

> "Within 24 hours of DeepSeek's launch, one financial services firm already had 500 employees using it, despite strict AI policies. Shadow AI moves faster than most controls, and traditional tools can't keep up with the speed of this adoption."
>
> 
>
> **Merritt Baer**
> CISO, Reco

Innovation without oversight has led to the rise of shadow AI

As AI adoption accelerates across every business unit, governance struggles to keep up. Employees are turning to powerful AI tools—often freely available—without a formal security review or oversight.

According to Gartner's 2025 Cybersecurity Innovations in AI Risk Management survey, 41% of employees were already using unauthorized "shadow AI" tools in 2022, and the company projects this will grow to 75% by 2027. These unvetted tools often introduce risks such as:

- ⚠ **Freemium LLMs** used to process internal memos and confidential data
- ⚠ **Marketing teams** are adopting AI content platforms without compliance oversight
- ⚠ **Developers** embedding AI APIs without validating licensing, data handling, or auditability

Why third-party AI risks are the new vendor blind spot

Even sanctioned platforms bring new risk vectors. A 2023 Cisco survey found that **60% of IT and security teams** say they can't see the specific prompts or requests made by employees using generative AI tools.

Common risks include:

- **Opaque decision-making:** Vendors may not provide insight into model logic or outcomes
- **Unpredictable updates:** AI models are often retrained silently, changing behavior without warning
- **Unknown training data:** Proprietary or biased data sources could create compliance or ethical issues
- **No audit trail:** Most models lack sufficient logging for incident reconstruction or forensics

When AI goes quietly astray – model drift and silent failure

AI systems degrade over time—a phenomenon known as model drift. This can be caused by shifts in user behavior, system data, or real-world context.

A multi-institution study (MIT, Harvard, University of Monterrey, and Cambridge) found that 91% of machine-learning models exhibit measurable performance decay if left in production without ongoing monitoring or retraining. Risks include:

- **Silent drift:** The model becomes less accurate while appearing functional
- **Performance bias:** Once-balanced outputs may skew based on outdated assumptions
- **Compliance failures:** A once-audited model may no longer meet fairness or privacy standards

The governance imperative

Unchecked AI is not simply a performance risk; it's a governance crisis. A well-intentioned LLM assistant can become a data exfiltration vector. A helpful third-party model can introduce biased outcomes or legal liability. Organizations must track AI use, vet vendors and models, implement monitoring tools, and ensure transparency and explainability across all AI initiatives.

> **"Unauthorized AI tools create new vectors for threats, with shadow AI usage typically exceeding officially sanctioned deployments by a factor of ten."**

Srikanth Chavali
Co-Founder & CPO, Kitecyber

04 Guardrails, not handcuffs — the new role of GRC in AI security

"With AI governance becoming a major focus point for security teams, consider using AI itself to quickly identify usage of new AI tools in the design review process. Existing teams are typically understaffed and overworked. Take advantage of AI to more quickly identify potential risks in the design review process before it's too late."

Kevin Qiu
Security Leader, Shiftsmart

Why GRC must evolve from reactive to proactive

Governance, Risk, and Compliance (GRC) functions were designed for systems that followed the rules. But AI learns, evolves, and sometimes behaves unpredictably. As enterprises race to adopt AI, they need a governance model that doesn't just mitigate risk but anticipates it.

Legacy controls focus on access, logging, and manual oversight. But AI introduces a new layer of complexity, including non-deterministic behavior, opaque "black box" decision-making, and dynamic model updates that often occur without notice.

The new mandate for GRC teams

GRC leaders must now partner closely with data science, engineering, IT, and legal teams to build AI-specific oversight. This includes:

- **Model documentation:** Covering purpose, training data, and known limitations.
- **Bias auditing:** Systematic testing for fairness and discrimination.
- **Explainability reviews:** Can the output be justified and understood?
- **Lifecycle tracking:** Monitoring for drift, updates, and anomalies.

> "The challenge for CISOs is to balance support for AI innovation with comprehensive risk management."

Merritt Baer
CISO, Reco

Smart AI doesn't always mean the right AI.

Even the most advanced models can misfire when they lack context—things like business logic, human nuance, or operational intent that give meaning to complexity.
Real-world misfires make this clear:

- **A security model** flags an engineer's 3 a.m. code push as anomalous, missing the fact that this is standard behavior during release week.
- **A compliance chatbot** auto-denies access to sensitive data, unaware the request came from a DPO during an audit week—adding friction, not security.
- **A vendor risk scoring engine** downgrades a partner due to lack of ISO 27001, without recognizing they're under a FedRAMP program—triggering unnecessary internal reviews.

In each case, the model wasn't wrong—it just wasn't grounded in the right context.
Pattern detection is not the same as understanding. For AI to make reliable decisions, it needs more than data—it needs environmental, procedural, and regulatory context.
This is where GRC becomes a strategic enabler. By embedding policy, ownership, and exception logic into AI systems, GRC leaders give AI the "ground truth" it needs to minimize false positives, reduce alert fatigue, and make smarter, safer decisions.

Embedding GRC into every stage of the AI lifecycle

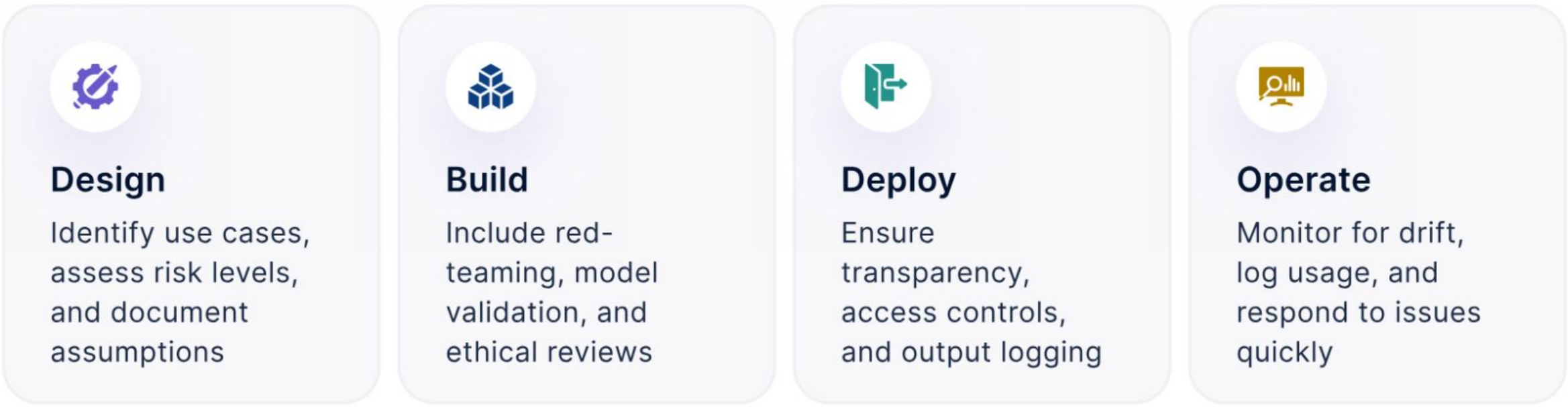

Figure 3. Integrating GRC in AI workflows

By integrating context, not just control, GRC enables smarter decision-making, faster detection of anomalies, and better alignment between AI systems and the humans they serve.

Navigating the regulatory maze

With AI systems touching sensitive data and decision-making workflows, embedding them into GRC frameworks is no longer optional. The global regulatory landscape around AI is evolving rapidly.

Meanwhile, industry-specific regulations like HIPAA, GDPR, and ISO 27001 are being interpreted to apply to AI systems, especially where data privacy, safety, or equity are impacted.

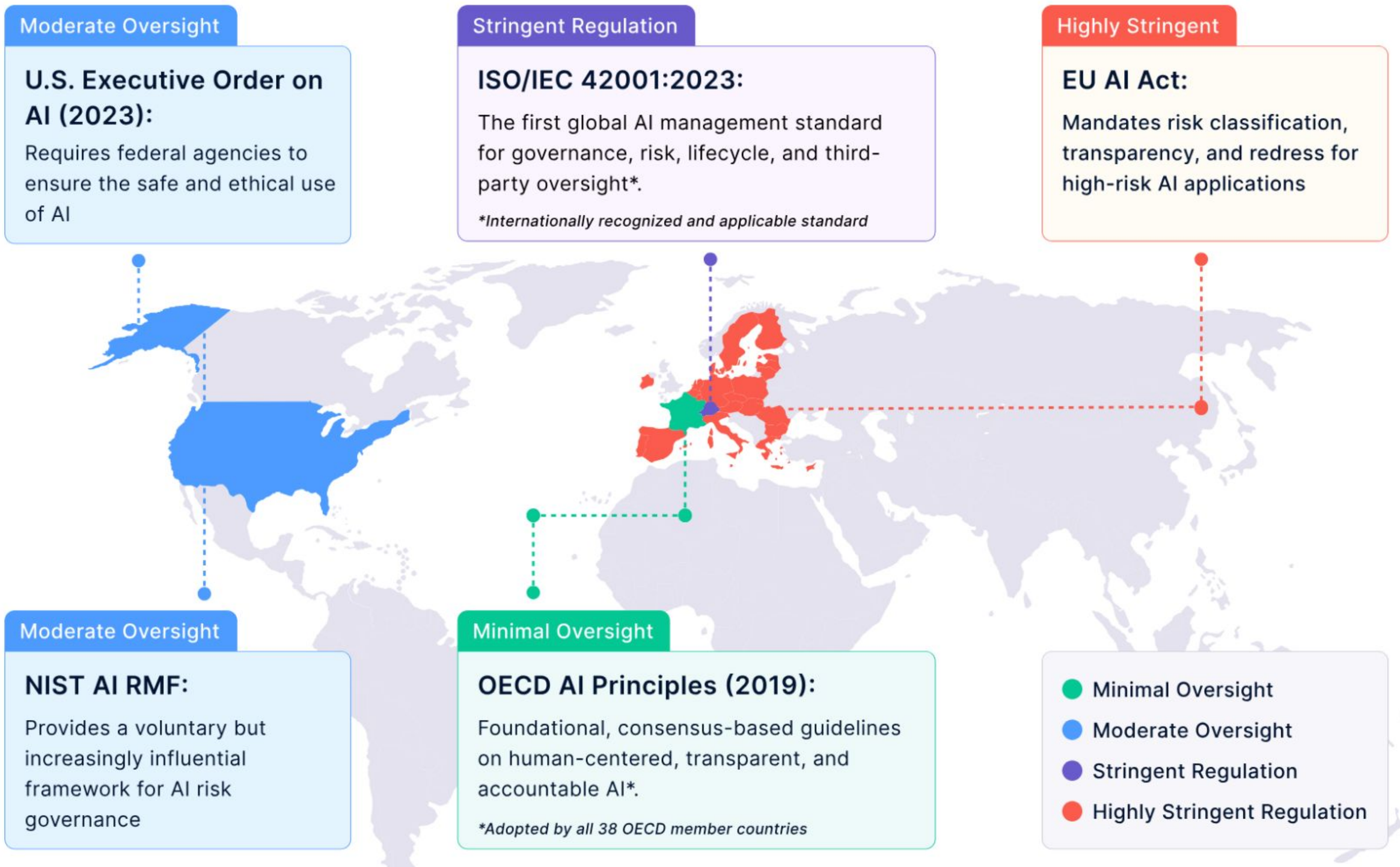

Figure 4. Popular AI governance frameworks

Related resources

- ISO/IEC 42001 Readiness Checklist
- Implement NIST AI RMF
- Navigate the EU AI Act

05 Redefining GRC with human + AI collaboration

> “AI is a brilliant assistant, but it still needs a boss. Without context, safeguards, and accountability, AI can shift from being an asset to a liability. The real power lies in the partnership: machines to scale, humans to steer. Palosade makes that collaboration seamless, governed, and safe, countering risks like non-determinism, hallucinations, and the absence of audit trails.”

Sanket Naik
Founder and CEO, Palosade.

The collaboration paradox

Over-automation introduces risk. Over-supervision throttles impact. Everyone agrees that human judgment is essential. But if every action requires human approval, is AI adding value?

The future of cybersecurity and governance isn’t human or machine — it’s both. AI brings unmatched speed and scale; humans bring context, judgment, and accountability.

But getting that partnership right requires redefining roles, controls, and trust boundaries. To resolve this paradox, organizations need a structured way to evaluate what AI can handle independently and what must remain under human control.

> “Organizations must develop comprehensive AI governance frameworks while simultaneously building internal capabilities for AI risk assessment and management.”

Srikanth Chavali
Co-Founder & CPO, Kitecyber

The tiered model for human + AI collaboration

A practical framework to define boundaries between autonomy and oversight. This model is dynamic, as AI proves itself and as governance strengthens, tasks can migrate between tiers.

TIER 1: Fully automated	TIER 2: Human-in-the-loop	TIER 3: Human-led, AI-informed
DESCRIPTION: AI handles low-risk, high-volume tasks with clear rules and minimal variation.	DESCRIPTION: AI suggests actions; humans review and approve before anything is executed.	DESCRIPTION: Humans lead complex decisions with AI providing insights and risk context.
EXAMPLE USE CASE: Spam filtering, anomaly alerts, automated log parsing in real time.	EXAMPLE USE CASE: Vendor triage, policy checks, access revokes, triage incident alerts.	EXAMPLE USE CASE: GRC audits, breach plans, legal review, AI model approval processes.

Figure 5. Three-tier model of AI-human decision-making

Building trust through proof, not promises

These safeguards allow companies to automate without abdicating accountability:

- **Accuracy benchmarks:** Evaluate AI performance on historical and live data. Set thresholds and test regularly.
- **Auditability:** AI systems should provide logs, lineage, and versioning. This ensures forensic traceability and enables root-cause analysis when something goes wrong.
- **Fail-safes:** Every AI system should include rollback options, alerts, and layered decision points to catch anomalies before damage spreads.
- **Human override:** Clearly defined escalation paths and manual controls must be embedded. In critical scenarios, human operators need both visibility and veto power.

Laying the groundwork for new skills, teams, and culture

As AI assumes more operational responsibility, security and GRC teams must evolve — not only in tools but in mindset.

- **Data fluency:** Teams must understand data sources, input integrity, feature engineering, and signal interpretation.
- **AI literacy:** Teams must know how models are trained, tuned, evaluated, and monitored. This includes familiarity with concepts like drift, hallucination, prompt injection, and model versioning.
- **Ethical reasoning:** Security professionals must assess fairness, inclusion, and the downstream effects of automated decisions.

> **"AI agents offer hope: they can identify patterns, contextualize alerts, and streamline investigations, allowing humans to focus on what actually matters rather than chasing false positives."**

Merritt Baer
CISO, Reco

Forward-looking organizations are retooling their org charts to reflect hybrid teams — blending analysts, ML engineers, GRC leaders, and business owners. In this hybrid model, AI becomes a collaborator — not a black box, not a silver bullet.

The next frontier is not just collaboration — it's co-evolution. Security teams must grow alongside their AI counterparts, shaping and being shaped by each new generation of tools.

This means evolving from static controls and audit checklists to dynamic oversight, continuous discovery, and cross-functional collaboration. AI fluency, ethical reasoning, and contextual governance aren't niche skills—they're becoming core to the DNA of every forward-looking security organization.

Want to see how top CISOs are operationalizing trust in AI? Watch the full "From Black Box to Boardroom" webinar on-demand to learn how leaders from ClickHouse, Bright Security, and Scrut are tackling shadow AI, audit readiness, customer trust, and AI governance in real-world settings.

06 The buying center shift — why security is back in control

"As AI permeates every role and workflow, security and governance must do the opposite, converge. Buyers are now the frontline of trust, shifting from tactical acquisition to strategic enforcement. At Palosade, we help organizations embed security, oversight, and alignment at the point of adoption."

Sanket Naik
Founder and CEO, Palosade.

For over a decade, the rise of SaaS has empowered business units to buy the tools they want, often with little IT or security oversight. Marketing bought CRMs, sales bought analytics platforms, and HR adopted AI-based screening tools.

But with the rise of AI, that purchasing autonomy is being reined in. Security and compliance teams are reasserting control, not to slow innovation but to protect the enterprise.

The evolution of the buying center

STAGE 1:

Centralized IT (Pre-cloud era)

All technology procurement flowed through IT. This ensured technical vetting but often slowed down business agility.

STAGE 2:

Functional SaaS adoption (Cloud boom)

Business units gained the power to procure their own tools. This accelerated productivity but fragmented security and compliance.

STAGE 3:

AI risk re-centralization (Now)

AI's impact on privacy, risks, and decisions is driving CISOs to regain control. Centralized oversight is now extremely critical.

Figure 6. The evolution of procurement decision-making

> "The days of black-box AI tools being approved without scrutiny are over. Procurement now demands proof. How does it work? Who's liable? Teams want model cards, audit logs, and clear proof of policy alignment."

Ian Amit
Co-founder and CEO, Gomboc AI

In fact, 78% of procurement leaders believe AI will disrupt the profession within just 3–5 years, underlining the urgent need for procurement functions to establish rigorous evaluation and governance processes for AI-powered solutions.

Implications for vendor evaluation and policy

As buying patterns evolve, evaluation criteria are also changing. Security teams are now incorporating AI-specific checks into the vendor lifecycle, asking targeted questions such as:

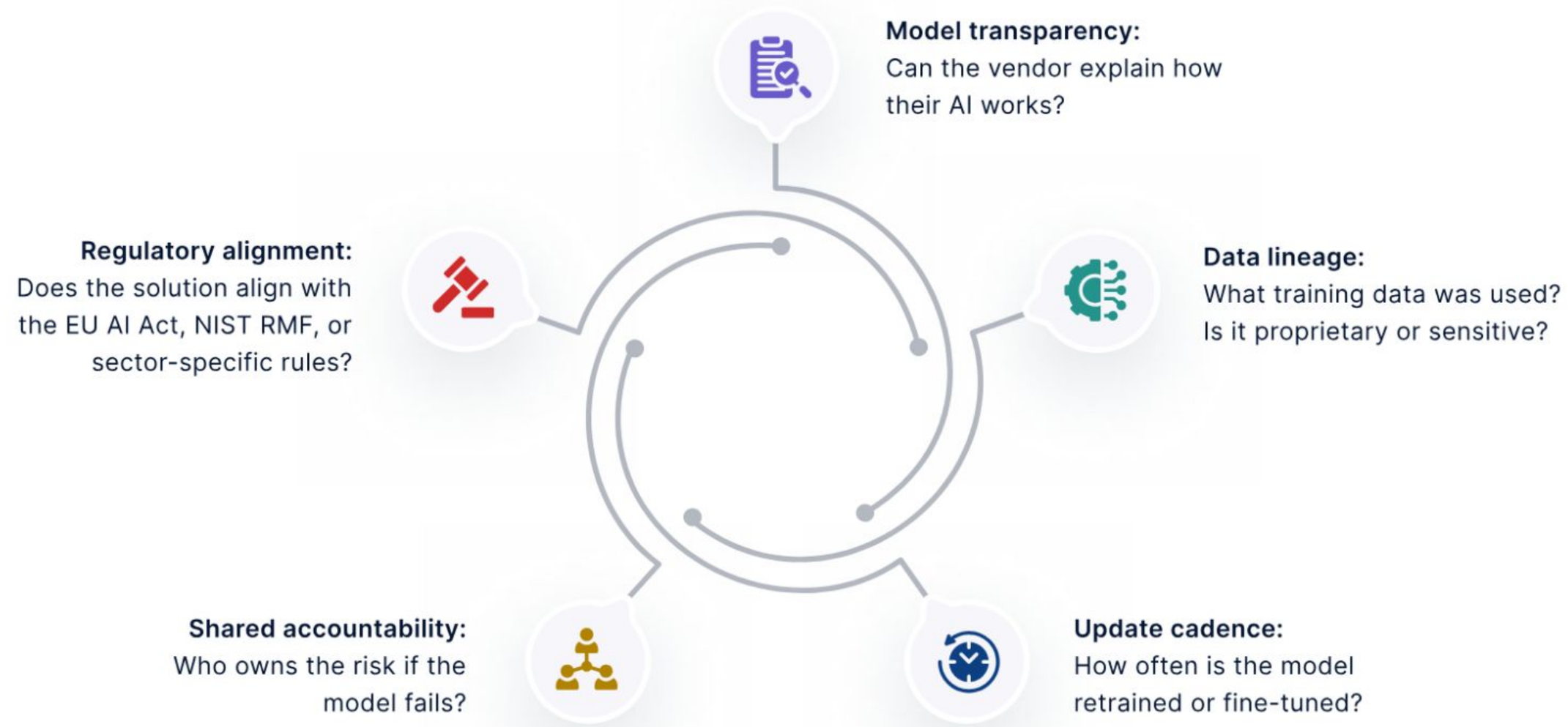

Figure 7. Embedding AI-focused checks for vendor evaluation

If AI is making decisions, security must shape the criteria by which those decisions and the tools behind them are evaluated. Leading organizations are formalizing this shift—rewriting procurement playbooks to include AI-specific risk assessments, demanding model documentation during RFPs, and requiring sign-off from both functional and security leaders.

Final roadmap — from reactive to ready

> "Apply the 80/20 rule: identify the 20% of AI use cases that carry 80% of the potential business risk, whether that's protecting sensitive training data or ensuring a customer-facing model doesn't create reputational damage. Focus your resources there first. Don't let the pursuit of perfect security for every AI tool paralyze your ability to manage the risks that truly matter."

Nick Muy
CISO, Scrut Automation

To prepare for an AI-defined future, organizations must:

Figure 8. Building an AI-Ready governance strategy

Every contributor to this ebook echoed the same truth in different ways. Visibility, accountability, and adaptability are now non-negotiable. AI doesn't wait for annual audits or security sign-offs. It evolves in production, learns from live data, and makes decisions that affect everything from compliance to brand trust.

Whether you're a CISO rethinking your procurement oversight, a GRC leader embedding explainability into policy, or an engineer balancing automation with accountability, the principles are the same.