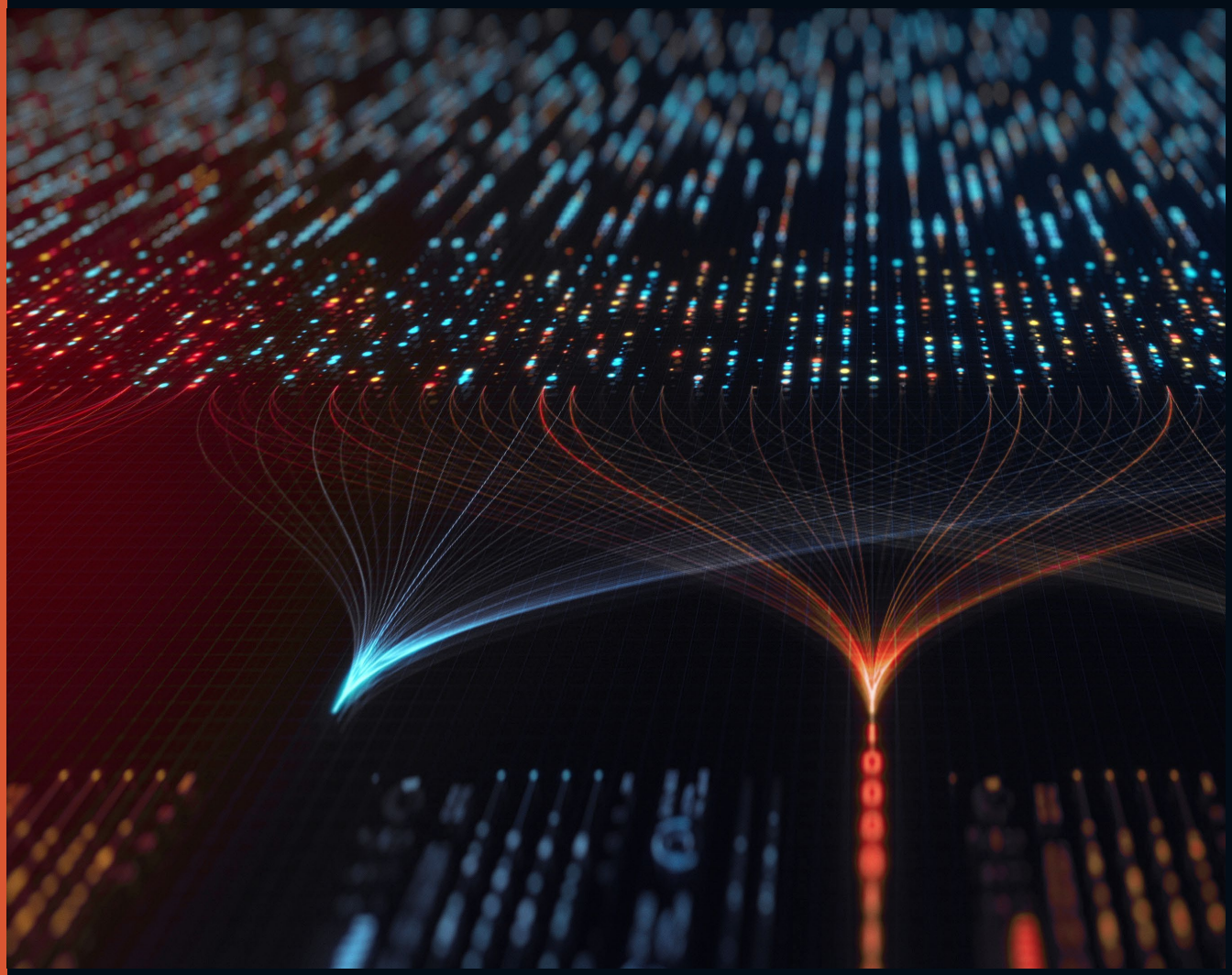
KOVRR

How to Build a Durable AI Governance Programme: *A 3-Pillar Framework*



www.kovrr.com

AI Governance Is Much More Than a Policy Challenge

Organisations are deploying AI faster than they are learning to govern it. Tools are getting provisioned, and vendors are being onboarded, while oversight structures lag somewhere behind, often by months. A recent survey found that 45% of employees have used AI tools for work without informing their manager, a figure that speaks less to bad intent than to how far ahead of governance most organisations' AI footprints already are.

A policy document gets written and approved, and leadership moves on, assuming the hard part is done. However, the distance between having a governance policy and enforcing one is where the majority of AI programmes are accumulating their greatest exposure, because AI governance is only effective when it is operationalised. The existence of written commitments means very little when teams are each managing their piece of the picture without a coherent view of the whole.

The result of this incoordination is often a set of, albeit well-intentioned, frameworks that remain misaligned, along with a broad set of AI systems and assets running across environments and business units that no stakeholder fully owns. Regulatory pressure is likewise mounting, with the EU AI Act alone carrying penalties that exceed those of the GDPR, and organisations that have not at least begun to address that underlying fragmentation will find the broader costs of unpreparedness difficult to absorb.

Substantive, scalable AI governance starts by building an in-depth understanding of what data AI systems are touching, who owns it, who approved its use, and whether the controls around it were implemented based on realistic usage patterns. Arriving there demands that data functions, AI oversight, and security operate from a shared model rather than parallel ones, with aligned policies and accountability structures that are enforced consistently. Organisations that invest in creating that foundation now are the ones positioned to govern AI defensibly long-term.

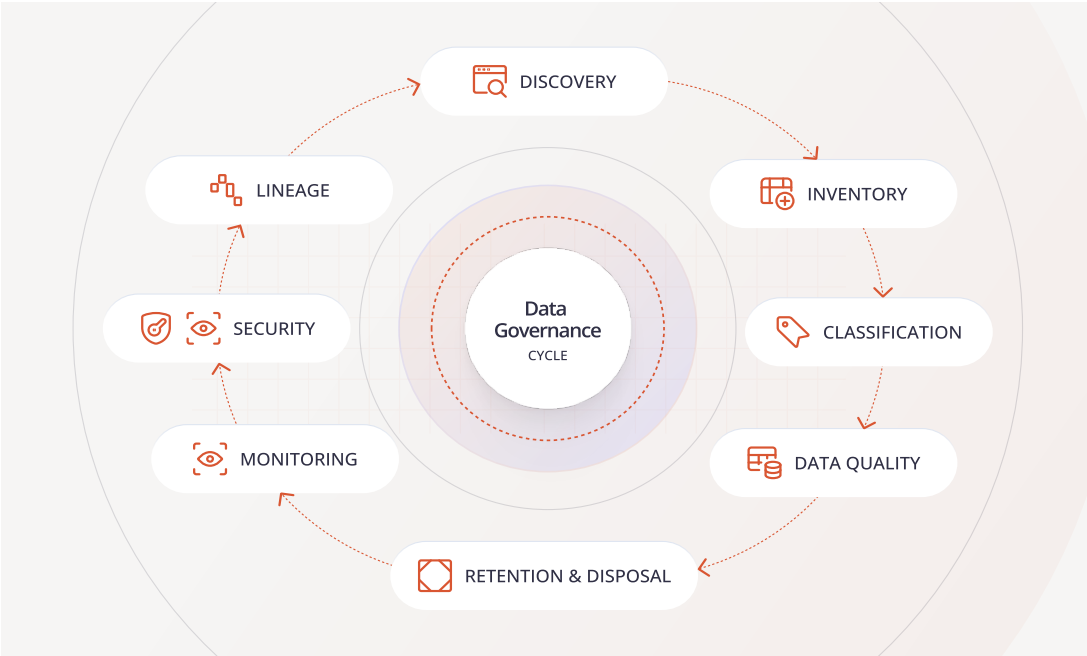
Pillar 1: Data Governance

> *"You can't govern AI without governing the data that feeds it."*

Data governance has been around for decades, but most organisations have yet to get it right, not because they have ignored it, but because the practice has long been treated as too complex to fully take on, and because most programmes were never designed for the environment that now exists. In recent years, that complexity has compounded, with data no longer living in a handful of known, stable repositories, but distributed across SaaS applications, cloud platforms, databases, email systems, file shares, and archiving solutions, moving continuously between them.

Consequently, most organisations face a level of exposure they heavily underestimate. Teams believe they have data governance in place, and on paper, they do, but it usually only applies to a single platform, leaving the rest of the environment with minor oversight. Data from a recent Kiteworks study confirms the scale of this blind spot, revealing that only 36% of organisations have visibility into where their data is processed, trained, or inferred. This means the vast majority of the modern enterprise data footprint is essentially operating in the dark.

In an AI context, the stakes of this incomplete coverage become impossible to ignore. Training, inference, and retrieval data all flow from an ecosystem that is only partially mapped. When that underlying data is poor quality or inadequately traced, the consequences manifest directly in model behaviour, resulting in drift, hallucinations, and outputs no one can explain. Mapping this foundational data is therefore the absolute prerequisite for any defensible AI governance programme.



The traditional data governance lifecycle, from discovery and classification through to retention and disposal.

Practical Starting Points for Data Governance

Discovery must precede everything. Long before anything can be inventoried, organisations need an automated means of finding where data actually lives, through tools like Microsoft Purview for Azure environments, Amazon Macie for AWS, and SaaS security posture management for SaaS applications.

Once surfaced, that data needs to be classified and labelled, and only then can it populate a dedicated inventory of which datasets are being used, by which systems, for which purposes, and under whose authorisation, mapped across the full environment rather than just familiar platforms. A coverage assessment then maps where genuine visibility exists and where blind spots remain, providing the basis for prioritising where to act first.

From there, data access controls need to be tied explicitly to AI use cases, ensuring that access to sensitive data is a deliberate decision rather than an inherited default. Data governance policy also needs to connect to whatever intake process exists for new AI tools, so that governance happens at the point of adoption rather than being retrofitted after the fact.

Where Data Governance Often Breaks Down

The most common point of failure is undocumented accountability. Organisations assume that because roles exist, responsibilities are understood. That assumption, however, is precisely where governance unravels. When the mandate of each function is never formally documented, teams operate on different interpretations of who owns what. Team A believes Team B is handling data quality, while Team B considers it someone else's remit entirely. The result is that critical governance activities either happen in duplicate or not at all.

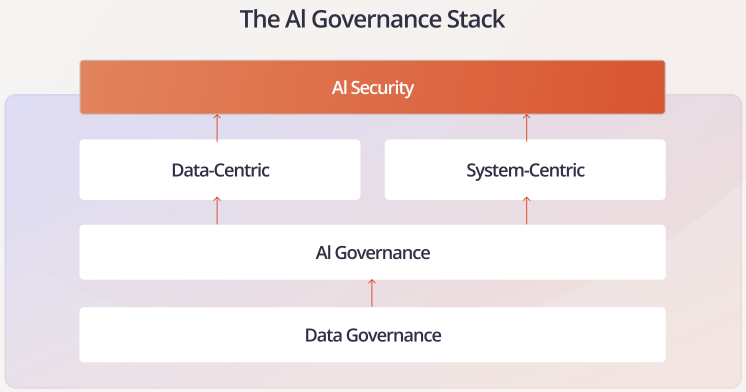
The deeper structural issue is that data governance has historically been treated as a data team responsibility. However, for AI, those governance processes become an enterprise-wide dependency, and redefining roles alone is insufficient without the documented mandates to back them up. Extending accountability to other domains requires not just redefining roles and decision rights, but committing them to paper and communicating them explicitly. Undocumented accountability is functionally the same as no accountability at all.

Pillar 2: AI Governance

> *"Policies don't govern AI, but processes do."*

With data governance in reasonable shape, the next challenge is governing the AI systems built on top of it. Some of those systems are built in-house, while others arrive as capabilities embedded within vendor and SaaS products, such as Microsoft Copilot, surfacing inside tools already in use without any deliberate decision to adopt them.

That breadth is part of why the prevailing instinct in the current market of treating AI governance as a policy exercise, in which a responsible use document gets drafted and the matter is considered addressed, falls short. Governance, however, requires operational embedding and will only take hold once it is incorporated into how AI systems are approved, deployed, and monitored across their lifecycle.



Data Governance

Traditional data security and lifecycle management

- Data classification
- Privacy
- Access controls
- Retention & deletion
- Data quality

AI Governance

Policies, oversight, and accountability for AI

- AI policies & standards
- Risk management
- Compliance & audit
- Human oversight
- Accountability

Data-Centric

Protect sensitive data used by AI models/tools

- Training data
- Prompt inputs
- RAG / retrieved data
- Model outputs
- Logs & telemetry

System-Centric

Protect AI systems, models, infrastructure and tools

- Models & model endpoints
- Infrastructure & platforms
- AI applications & agents
- Deployment environments
- Supporting systems

AI Security

Protect AI models and systems from AI-specific threats

- Prompt injection
- Model drift
- Hallucinations
- Adversarial attacks
- Model theft / abuse

AI governance builds on data governance and splits into two distinct dimensions. AI security spans both.

The Two Dimensions of AI Governance

A useful, yet frequently overlooked distinction in the AI governance space is the one between data-centric and system-centric. Understanding both and how they relate to one another is critical to building an AI governance programme that addresses the full scope of AI risk.

Data-centric AI governance concerns the data that AI systems are built on and operate from, such as training data, retrieval-augmented generation data, and inference data. Authorisation, classification, and lineage apply here, as well, just as they do in traditional data governance, extended specifically to AI use cases. Organisations that focus exclusively on the AI system without governing the data underneath it are addressing the symptom while leaving the cause unexamined.

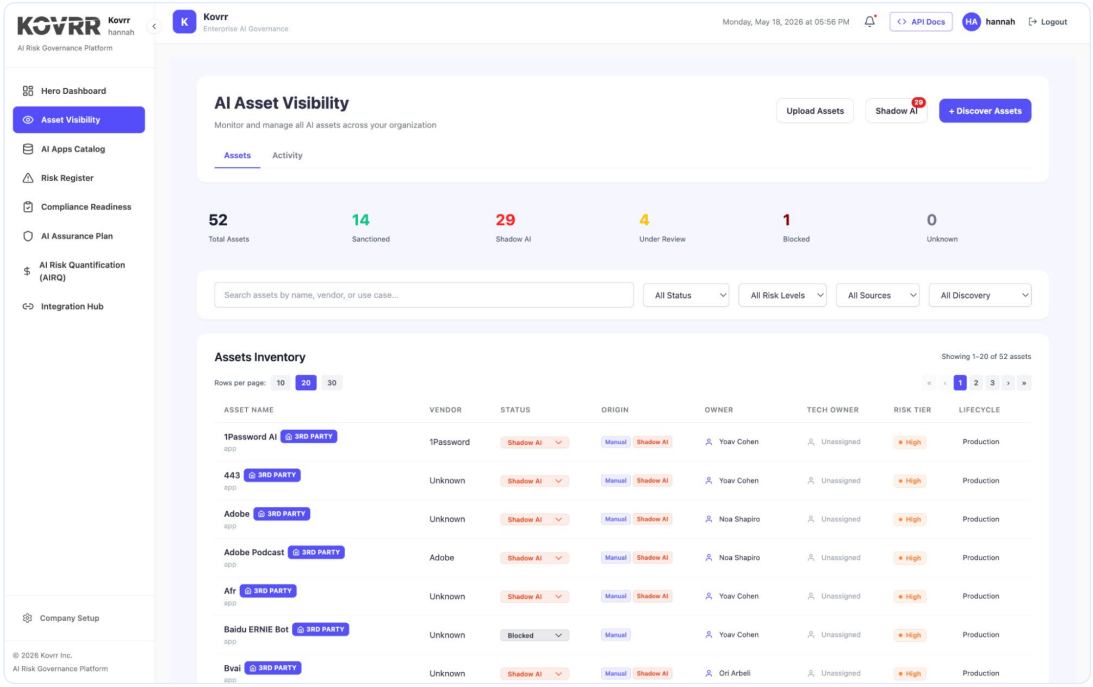
System-centric AI governance concerns the behaviour of the AI system itself. Even a system built on perfectly governed data can produce outputs that are unfair, unexplainable, or vulnerable to manipulation, which is why this dimension requires its own governance

sub-classification, one that defines what a model is permitted to do, monitors whether it stays within those boundaries, and adapts as the system and its operating context evolve.

Building an AI Inventory

Just as data must be discovered before it can be governed, AI systems must likewise be surfaced before they can be inventoried. That discovery is the precondition for everything that follows, since an organisation cannot govern what it has not yet found. Once surfaced, those systems populate an AI inventory, a living registry of every AI system in operation, the models powering them, the vendors behind those models, and the business functions they serve.

This registry is the foundational data-centric governance artifact for any AI programme, and it operates alongside the data inventory established earlier, the two together capturing both the systems in use and the data feeding them. Without that comprehensive AI asset visibility, there is no authoritative basis from which accountability can be assigned or risk systematically assessed.

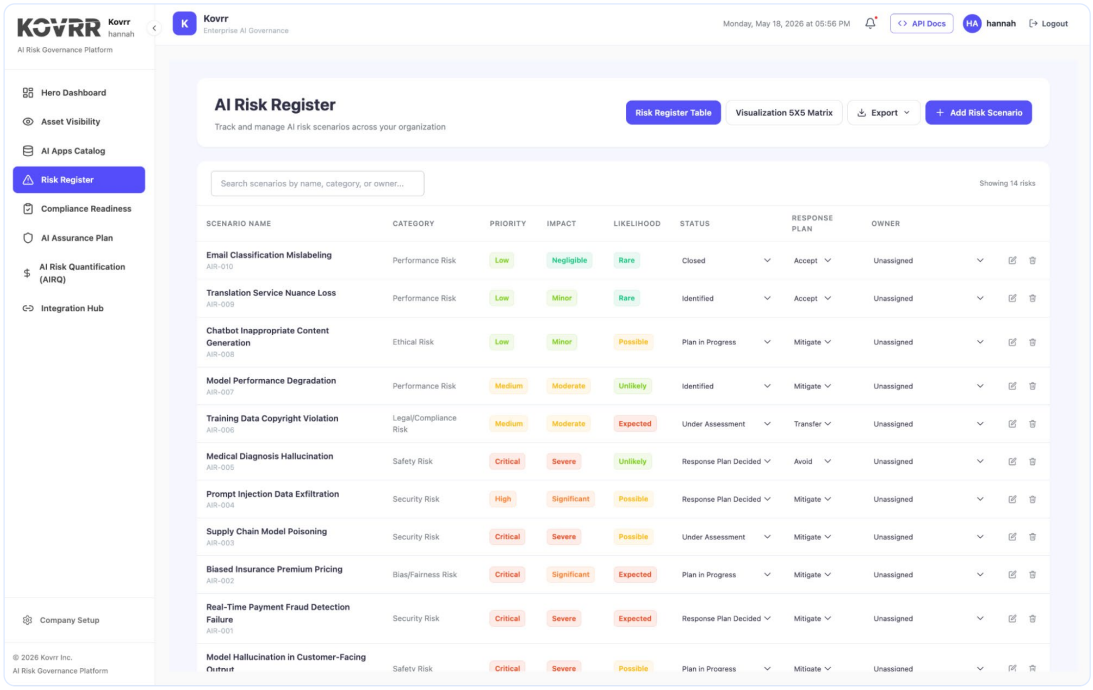


Kovrr's AI Asset Visibility capabilities surface every AI tool across the organisation, including sanctioned, shadow, and blocked.

A complete inventory needs to account for sanctioned AI, shadow AI, and systems that have been explicitly blocked. In most organisations, shadow AI is far more prevalent than leadership realises, and it represents the more consequential governance challenge. Surfacing shadow AI requires detection capability, but sustaining control over it requires an approval process functional enough that teams have no incentive to circumvent it.

Key AI Governance Components

While the AI and data inventories establish visibility, the artifacts built on top of it establish accountability. Three warrant prioritisation above all others, each operating across both the data-centric and system-centric dimensions. The first is the AI risk register, a living document that translates inventory entries into assessed risk, capturing the risk profile of each system and the residual exposure that remains after controls have been implemented. The register is the mechanism through which governance becomes ongoing rather than episodic.



Kovrr's AI Risk Register centralises AI risk and compliance scenarios in a structured, living register, linking each risk to affected assets, governance domains, and mitigation plans.

The second component is the use case intake and approval process, a structured review that evaluates data sensitivity, regulatory exposure, and control requirements. The third, accountability mapping, is the practice of formally assigning ownership over each AI system at both the business and technical levels, covering not only what data the system touches but also how it is expected to behave and who is responsible when it does not.

Aligning With AI Governance Frameworks and Regulations

Established frameworks offer a significant head start. The NIST AI RMF and ISO 42001, for example, each approach AI governance from a different angle, but share a common insistence that AI governance is a continuous management responsibility that evolves alongside the systems it covers. Alongside these governance frameworks, security-specific frameworks address the threat layer, with the OWASP LLM Top 10 cataloguing the most critical vulnerabilities in large language model applications and MITRE ATLAS mapping the tactics adversaries use against AI systems.

For organisations operating in or serving European markets, the EU AI Act carries mandatory obligations, including incident reporting requirements already in force for some system categories, with penalties that exceed those of the GDPR.

How GRC Teams Can Operationalise Ongoing Oversight

GRC teams already own the risk register and control framework, making them the natural home for AI governance operationalisation. Extending those structures to cover AI systems means integrating them into the enterprise risk register and building AI-specific criteria as part of a continuous assurance model, with real-time visibility into the AI estate rather than periodic point-in-time review.

That cadence needs to span both the data-centric and system-centric dimensions, ensuring the data feeding AI systems remains governed and traceable, and that model behaviour stays within defined boundaries as systems evolve.

Tools that span both sides of the AI governance equation, such as Kovrr's AI Asset Visibility and AI Risk Register tools, help designated teams maintain continuous oversight even as their AI footprint grows, combining automated discovery, real-time monitoring of the AI estate, and continuous risk scoring across both dimensions.

Where AI Governance Often Breaks Down

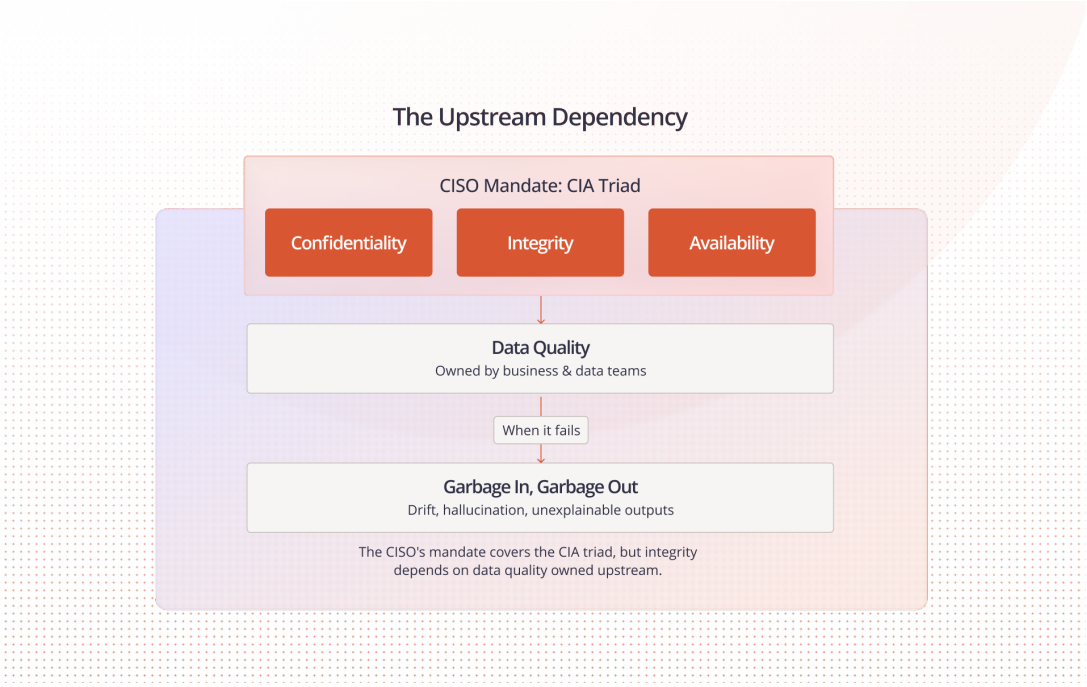
The most common misstep organisations make is treating AI governance as a one-time exercise. A risk assessment gets completed at deployment, and governance attention moves on, leaving the system without any structured oversight of how it evolves. Given that AI systems can change materially over time without anyone explicitly changing them, that posture creates compounding exposure.

The second detrimental pattern stakeholders fall into is conflating nominal ownership with functional ownership. Accountability mapping exercises frequently produce a list of names next to systems, but a name on a document carries no weight without the organisational mandate to act on it. AI governance programmes that obscure that distinction tend to discover its limits at the worst possible moment.

The third failure point is operational siloing between the two dimensions. An organisation can have mature data-centric governance and rigorous system-centric controls, each functioning well within its own lane, and still carry significant risk at the seam between them. When changes on the data side go uncommunicated to the teams governing the system built on top of it, behavioural guardrails calibrated against an earlier state continue operating as though nothing has changed.

Pillar 3: Security

> *"AI introduces new attack surfaces, and existing security frameworks need an AI-aware layer."*



The CISO's mandate covers the CIA triad, but integrity depends on data quality owned upstream.

Security spans both the data-centric and system-centric dimensions, warranting dedicated attention because the threat surface AI introduces is materially different from what most security programmes were built to address. The CISO's mandate covers confidentiality, integrity, and availability, but those three properties are only as strong as the data quality and lineage that sits upstream of them. When that foundation is weak, the security function is left accountable for outcomes it has no direct control over.

Why Traditional Cybersecurity Frameworks Fall Short

While traditional cybersecurity frameworks were built around threat models that have always been in flux, shaped by organisational changes, shifting attack trends, and the steady evolution of technology, AI accelerates that pace dramatically. Systems learn, adapt, and can shift behaviour without anyone explicitly changing the underlying code, making them difficult to secure using controls designed for static assets. The dynamic nature of AI systems means the attack surface moves with the model rather than the perimeter.

MITRE ATLAS documents how adversaries exploit this, operating at layers traditional security tooling was not built to monitor, manipulating models through their own inputs, and moving data in ways that leave no trace at the network perimeter.

Third-Party and Vendor AI Risk

Organisations now consume AI through a mix of SaaS platforms, API-based model providers, and systems built in-house, with the balance varying across companies. Some start with SaaS and move toward internal development; others have been building AI capabilities in-house from the start. What stays constant is that a significant portion of the AI attack surface sits with vendors not under the organisation's control.

A vendor can update the underlying model without any visible change to the product, and the organisation may have no mechanism to detect that the system it reviewed at onboarding is no longer the system it is running today.

Sensitive organisational data submitted through a third-party AI tool may be retained or used for training in ways the organisation never reviewed or approved. Contractual protections that adequately cover traditional software relationships often say nothing about model training data or what happens to proprietary inputs after they leave the organisation's environment. AI vendor risk assessments need to be redesigned to account for that complexity, with AI-specific criteria embedded in the process from the outset.

Where Security Often Breaks Down

The most common misstep is assuming that existing security controls automatically extend to AI environments. Security programmes that were rigorous and well-designed before AI entered the picture can develop significant blind spots simply because the threat model changed and the controls did not. The distance between what a security framework was built to cover and what an AI environment demands tends to remain invisible until something goes wrong.

The second detrimental pattern is vendor risk complacency. Organisations that hold their own AI systems to high security standards often apply a much lower bar to the third-party AI tools their employees are using daily. Given that a significant portion of enterprise AI exposure runs through SaaS and API-based tools, that asymmetry represents a material and frequently underestimated risk.

Building a Cohesive AI Security and Governance Programme

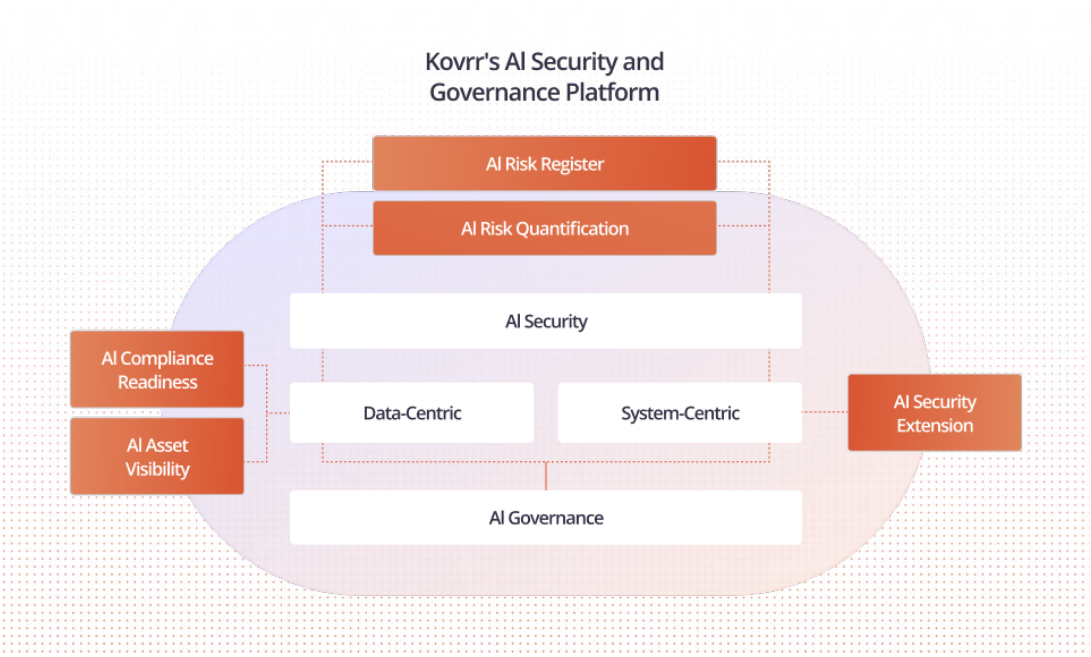
Data governance, AI governance, and security are often treated as three separate workstreams, each with its own owner, cadence, and reporting line. Unfortunately, that structure produces pockets of progress and systemic blind spots in equal measure. The handshakes between the three functions are where most governance programmes quietly unravel, and keeping those connections intact requires a single coordinating function with the mandate and authority to hold all three together.

Indeed, the instinct in most organisations is to distribute that responsibility across the CISO, CIO, and COO. However, shared ownership of a problem tends to produce the same outcome as no ownership at all. Friction accumulates, and accountability diffuses,

leading governance activities to fall into the space between functions. A nominated owner with a clearly documented mandate, backed by genuine organisational authority, is what prevents that.

Bringing the Dimensions Together

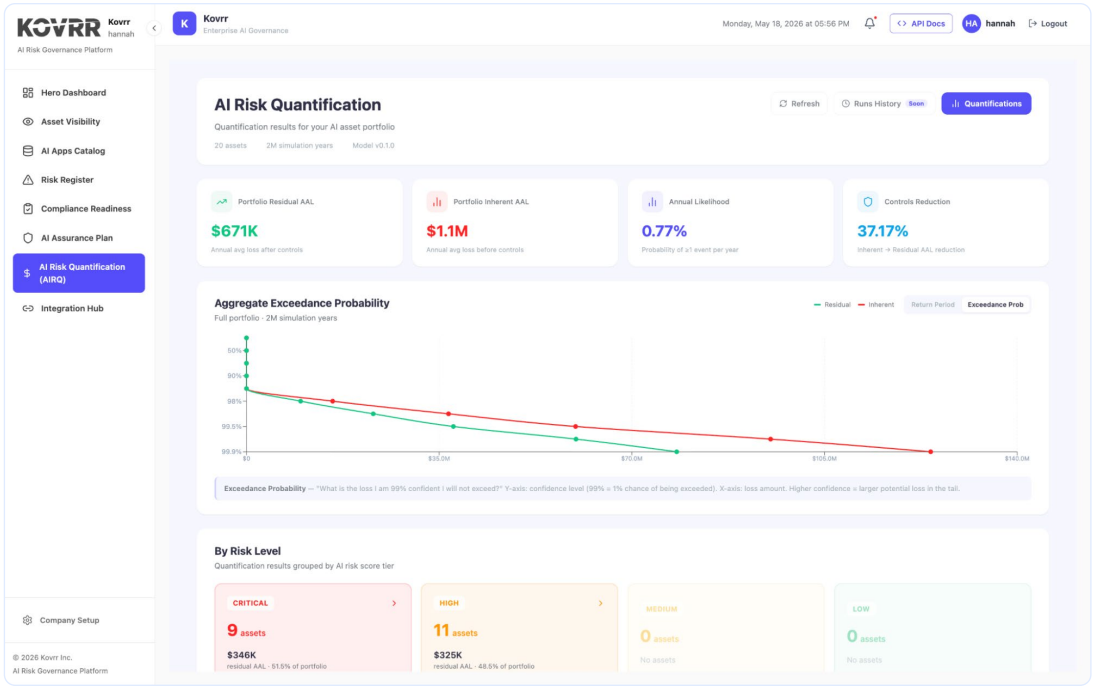
Mature data-centric governance and rigorous system-centric controls are necessary conditions for effective AI governance, but the organisations that mitigate AI risk effectively are the ones that treat the two dimensions as a single, coordinated programme rather than adjacent responsibilities that happen to share a domain. That connection requires visibility across both sides simultaneously, accountability structures that span the seam between them, and the ability to translate what is happening across the AI environment into terms that leadership can act on.



Built around that requirement, each capability within Kovrr's AI Security and Governance Platform surfaces a distinct layer of telemetry across an organisation's AI use cases. AI Asset Visibility and AI Compliance Readiness map what exists and how it aligns with governance frameworks. Kovrr Browser Protect, alongside upcoming agentic monitoring and control capabilities, feeds real-time browser and endpoint activity back into that visibility layer while also providing the controls to enforce policy at the point of interaction. Finally, the AI Risk Register and AI Risk Quantification translate the full picture into structured risk and financial exposure.

The value lies in how each capability contextualises the others. A shadow Copilot deployment surfaced by AI Asset Visibility only becomes governance when its risk scenarios appear in the AI Risk Register, its day-to-day usage is monitored through the browser layer, and policy actively controls interaction. That degree of cross-capability coordination is what governance at AI's scale and speed demands. The result is a governance programme that remains coherent as the organisation's AI footprint grows.

Making the Business Case



Kovrr's AI Risk Quantification engine translates AI exposure into financial terms, illuminating the data needed to prioritise governance investment against a defined risk appetite.

Getting organisational buy-in for AI security and governance investment requires speaking the language of business impact rather than compliance risk. Leadership teams respond to exposure they can measure, such as the financial consequence of an AI system failure, the cost of a regulatory penalty relative to the cost of prevention, or the specific areas where AI risk exceeds the organisation's stated appetite. Organisations that can articulate their AI risk in financial terms are significantly better positioned to secure the investment their governance programme requires.

Kovrr's AI Risk Quantification (AIRQ) software was built for precisely this critical business conversation. By modelling AI exposure against how specific models and vendors are deployed across the organisation, AIRQ reports risk at the level where decisions are made, tracks whether mitigation initiatives are actually reducing that exposure, and quantifies which findings carry the most financial weight.

The new generation of AI governance frameworks generates more findings than security teams can address at once. AIRQ gives security and GRC leaders the evidence base to focus on the work that delivers the greatest risk reduction, make the case to the board in tangible terms, and demonstrate that governance spending is proportional to the risk being managed.

From Checkbox Compliance to Defensible Governance

Governance maturity is a spectrum. Most organisations begin somewhere between having nothing and having a policy document, and progress is incremental. The meaningful distinction is between governance that exists to satisfy an audit and governance that

would hold up under genuine scrutiny. Defensible governance means being able to account for what happened when something goes wrong. Organisations that can answer with evidence rather than approximation have built something worth having.

Governance as a Competitive Advantage

AI governance is a programme that evolves alongside the organisation running it. The ones that invest in getting the foundations right now, before they are compelled to by an incident or a regulator, are the ones positioned to scale AI responsibly and defend their decisions when it matters. Those that wait will find themselves building under pressure, with considerably more exposure than they would have had otherwise.

The three pillars outlined here reflect the operational reality that practitioners are navigating every day. Progress does not require perfection. An organisation that has mapped its AI inventory, documented its accountability structure, and stress-tested its vendor relationships is in a materially stronger position than one that has not, regardless of how much further there is still to go. The work compounds, and getting started is the most important step.

To explore how Kovrr's AI Security and Governance Platform can help operationalise all three pillars as a single, connected programme, book a demo today.

ABOUT THE AUTHORS

YAKIR GOLAN
Co-founder and CEO, Kovrr

Yakir Golan is the CEO and Co-founder of Kovrr. He began his career in the Israeli intelligence forces and later built multidisciplinary expertise across software engineering, product development, and enterprise risk strategy. Over the past decade, he has worked closely with CISOs, risk executives, and boards to strengthen how organisations govern cyber and AI-related risk at scale. Yakir holds a BSc in Electrical Engineering from the Technion, Israel Institute of Technology, and an MBA from IE Business School in Madrid.

CHRISTIANE (CHRISI) BAETZ
CISSP, CISA

Christiane (Chrisi) Baetz, CISSP, CISA, is a CISO and board-level cyber risk advisor with over 17 years of experience helping organisations turn cyber risk into a driver of enterprise value. She has held senior security leadership roles across financial services, critical national infrastructure, and PE-backed businesses, leading transformation programmes ranging from £1m to £37m and building security operating models that hold up under regulatory and investor scrutiny. Her recent work spans IT and OT security, cloud transformation, and emerging AI security, with a particular focus on translating complex cyber risk into commercially relevant decisions at board level. She currently serves as a cyber security delivery partner at the Financial Conduct Authority.