



# Bridging the AI Model Governance Gap

# Executive Summary

**Enterprise AI adoption is soaring, but gaps in model governance threaten to slow progress. A recent Anaconda survey of over 300 AI practitioners and decision-makers reveals both a strong awareness of AI supply chain risks and notable shortcomings in AI model governance practices.**

These challenges are emerging against a rapidly evolving market backdrop. Forrester predicts spending on off-the-shelf AI governance software will quadruple between 2024 and 2030, reaching \$15.8 billion and capturing 7% of overall AI software spend. The more value and risk tied to AI, the more urgently companies need governance and security compliance solutions that aren't just quick fixes.

Organizations are still grappling with foundational governance challenges against this backdrop of accelerated investment and rising expectations. The survey findings show where teams are gaining traction and where critical governance gaps remain.

- **Open-Source Security Concerns:** While 82% of organizations validate open-source Python packages for security, nearly 40% of respondents still frequently encounter security vulnerabilities in their AI projects. Over two-thirds report deployment delays due to security issues, underscoring tension between rapid innovation and risk management. As AI agents and applications become more prevalent, resolving this tension in a trusted way will be critical for all enterprises.
- **Model Lineage & Monitoring:** 83% track the origins of foundation models, and 81% maintain documentation for model dependencies. Yet 30% have no formal model drift monitoring in production, and only 62% of those tracking models use comprehensive documentation.
- **Toolchain Fragmentation:** Only 26% of enterprises have a highly unified AI development toolchain, with the rest using partially unified or fragmented tools.
- **Governance Priorities Misaligned:** Package security vulnerabilities are the most common AI development risk (39%), yet less than one in five respondents (19%) name security verification components as their top governance priority. Instead, priorities slightly skew toward regulatory compliance (25%) and model performance monitoring (23%).
- **Top Priorities:** When asked what would most help improve model governance, the number one answer (29%) was better-integrated tools combining development and security workflows. Better visibility into model components (23%) and team training (19%) were also highly ranked.

“

One of the biggest keys for AI model governance is building comprehensive evaluation frameworks that include both performance metrics and security considerations. This means establishing processes that assess not just model capabilities, but also potential vulnerabilities, adversarial exploits, and unintended behaviors. Given the evolving regulatory landscape and licensing uncertainties, organizations need governance systems that can adapt to changing compliance requirements while maintaining clear accountability structures. Creating robust benchmarking that ties specific safety considerations, to include guardrails that bound the permissions of deployed applications informed by a rich set of data about the AI assets being deployed is something I haven't seen enough attention on.”



**Greg Jennings**

VP of Engineering for AI, Anaconda

## METHODOLOGY

309 individuals participated in our online survey conducted in April 2025. Respondents consisted of IT, ML/AI, DevOps, and data governance professionals currently working with or making decisions about AI models within their organization. All responses are self-reported. Note: All percentages are rounded to the nearest whole percent. Due to rounding, some numbers may not equal 100.

# The Urgency of AI Model Governance

Today's AI systems are constructed atop vast, fast-moving supply chains of open-source code, cloud platforms, and third-party models, creating huge, dynamic ecosystems. AI is driving value in domains from customer service to fraud detection. However, this rapid innovation comes with new governance challenges. Poorly governed AI models don't just underperform—they can erode trust, obscure transparency, and expose businesses to compliance risks.

The survey data underscores a fundamental “governance gap.” Organizations know governance is essential, yet many lack the integrated processes to achieve it end-to-end. One respondent bluntly described their ideal AI platform: “It will be totally compliant, clear coding AI policies to prevent running into document issues and documentation, which we clearly lack,” pointing to internal roadblocks in achieving end-to-end governance.

To get to this ideal platform, we asked respondents about the biggest concern in AI model governance today. The most frequently cited issues were security, data privacy, and lack of transparency, alongside fears of bias and lack of accountability. Yet organizations are racing to innovate with AI and worry that stronger controls could slow them down.

Improving governance doesn't mean stifling innovation. Instead, it empowers teams with tools and practices that make doing the right thing—secure, compliant, well-monitored AI—as easy as doing the fast thing.

Striking the right balance between innovation and oversight is the central challenge of AI governance. It's not a choice between experimentation and responsible controls because both are necessary.

For many organizations, the tension between speed and security, flexibility, and accountability plays out most clearly in using open-source tools and components. These tools are foundational to modern AI development but introduce new risks. When paired with unified, transparent platforms, open-source tools are crucial to allow organizations to move quickly while maintaining the accountability that stakeholders and regulators expect.



# Security Compliance: Protecting the Open-Source AI Stack

## Widespread Awareness, Patchwork Execution

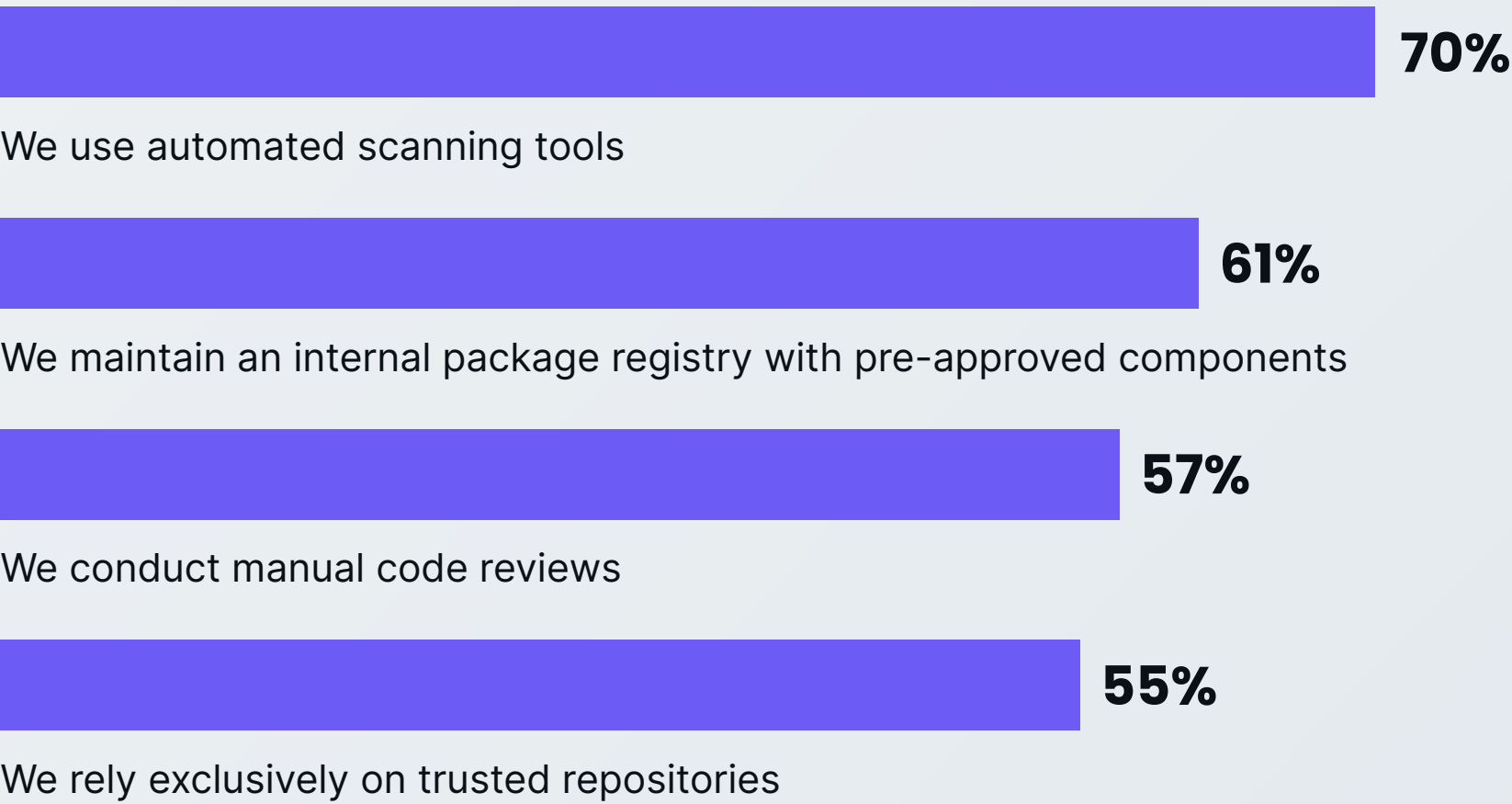
Open-source software is the lifeblood of AI development, but it also introduces supply chain risks to manage. AI practitioners show broad awareness of security and compliance needs. Over 82% of organizations have some process to validate Python packages and dependencies for security compliance, whether through automated vulnerability scanners, internal package registries, or manual code reviews. Most organizations combine approaches. For example, 70% use automated scanning tools, 61% maintain an internal vetted package repository, and 57% conduct manual reviews.



**Over 82%**  
of organizations have some process to validate Python packages and dependencies for security compliance, whether through automated vulnerability scanners, internal package registries, or manual code reviews.

### How does your team currently validate that the Python packages and dependencies used in your AI models meet security and compliance requirements?

Select all that apply.



SECURITY COMPLIANCE

Security in Practice Still Falls Short of Security in Policy

This layered defense is a positive sign, but security issues remain a frequent pain point. Nearly 40% of respondents said security vulnerabilities in open-source components (packages) are the most common risk in AI development, ranking above data leakage (28%) and environment inconsistencies (20%). Additionally, 67% of organizations experience delays in AI deployments due to security concerns, with 14% saying security significantly slows progress.






These delays often result from manual reviews, inconsistent tooling, and a lack of visibility into dependencies. There’s a clear cost to patching and vetting dependencies: Over 80% of teams spend more than 10% of their AI development time troubleshooting dependency conflicts or security issues, and more than 40% spend over a quarter of their time on these tasks.

Time spent resolving issues is a major drag on productivity. These findings point to a governance gap between policy and practice. Many enterprises think they are managing open-source risk—80% feel confident in their ability to remediate vulnerabilities. Yet the ongoing prevalence of incidents and delays suggests that the current tools are not keeping up.

67%

of organizations experience delays in AI deployments due to security concerns.

Most Common Risks in AI Development

-  Security vulnerabilities in packages
-  Data leakage or compliance issues
-  Inconsistent environments
-  Model drift or failure
-  Shadow IT (unauthorized tools)

“

If there’s no transparency, you really don’t have control in your AI model governance. At Anaconda, we’ve been confident that models will get smaller, and people will want to run them locally on their own data. We’re trying to make it easy for people to do that in a safe, efficient way without the risk of their data getting leaked.”



**Peter Wang**  
Chief AI and Innovation Officer  
and Co-Founder, Anaconda

SECURITY COMPLIANCE

# From a Reactive to Proactive Model of Governance

Manual reviews and disparate scanners can only scale so far, as AI projects pull in ever-expanding open-source libraries and pretrained models. Organizations must shift from reactive to proactive security governance. As AI projects become more prevalent and commingled with operational workflows, this will be more critical to track systematically.

When asked to pick one area to enhance in their AI supply chain, the top choice (27%) was implementing stronger security controls in development workflows. Likewise, 58% of organizations now track compliance metrics (e.g., adherence to regulations) as a KPI for their AI initiatives, indicating that leadership is paying attention to governance.

Best-in-Class Practices

-  Using trusted, curated package distributions to minimize known vulnerable components
-  Implementing continuous dependency monitoring for new vulnerabilities
-  Enforcing policy guardrails, such as restricting unapproved packages or licenses

27%

of respondents chose implementing stronger security controls in development workflows as the top area to enhance AI supply chain.

“

Organizations can strengthen governance without slowing development by centralizing package management and defining clear policies for how code is sourced, reviewed, and approved. Adopting a unified approach to tooling simplifies security tasks like vulnerability scanning, license tracking, and access control. These steps help create a more predictable, well-managed development environment, where innovation and oversight work in tandem.



**Greg Jennings**  
VP of Engineering for AI,  
Anaconda

# Model Lineage & Performance Monitoring: Ensuring Accountability

## Documenting AI Models is Common but Inconsistent

Beyond code security, governing AI model behavior over time is another critical pillar. Models are not static assets; they evolve (or degrade) with new data, and their provenance can be complex, especially with pretrained foundation models. The survey reveals high intent: 83% of organizations claim they track the origins and lineage of the foundation models they use.

Over 62% do this through comprehensive documentation of all model sources and versions, while others admit to only partial or ad-hoc documentation. Additionally, a substantial majority (81%) maintain documentation of model dependencies for regulatory compliance – for example, keeping records of packages, training datasets, model parameters, and performance metrics.

However, that number may be overinflated as companies have different views of “comprehensive” or experience levels with AI models. Even differing roles within the same organization may view documentation from individual lenses.

Emerging AI regulations and the need for auditability likely drive this emphasis on documentation, yet the consistency and depth of these practices vary. Nearly 1 in 5 organizations still have no formal documentation of model dependencies or lineage, which could prove problematic when an audit or incident occurs.

In open-ended responses, many professionals stressed “lack of transparency” as a chief concern, reflecting that stakeholders worry about black-box effects and unknowns in the model supply chain. This concern aligns with recent [McKinsey findings](#); 40% of respondents cited explainability as a key risk in adopting generative AI, yet only 17% are actively working to address it.



So many organizations regularly rely on tools like ChatGPT. You can take their documentation from the website and say, ‘This is what the model does’ and believe that’s comprehensive documentation. I would challenge users to be able to clearly define a benchmark or process for a model. The only standard is that there isn’t one, which is an area for improvement across the industry.”



**Peter Wang**

Chief AI and Innovation Officer  
and Co-Founder, Anaconda

## Nearly 1 in 5

organizations still have no formal documentation of model dependencies or lineage.

MODEL LINEAGE & PERFORMANCE MONITORING

Monitoring in Production is Gaining Traction, Though Gaps Remain

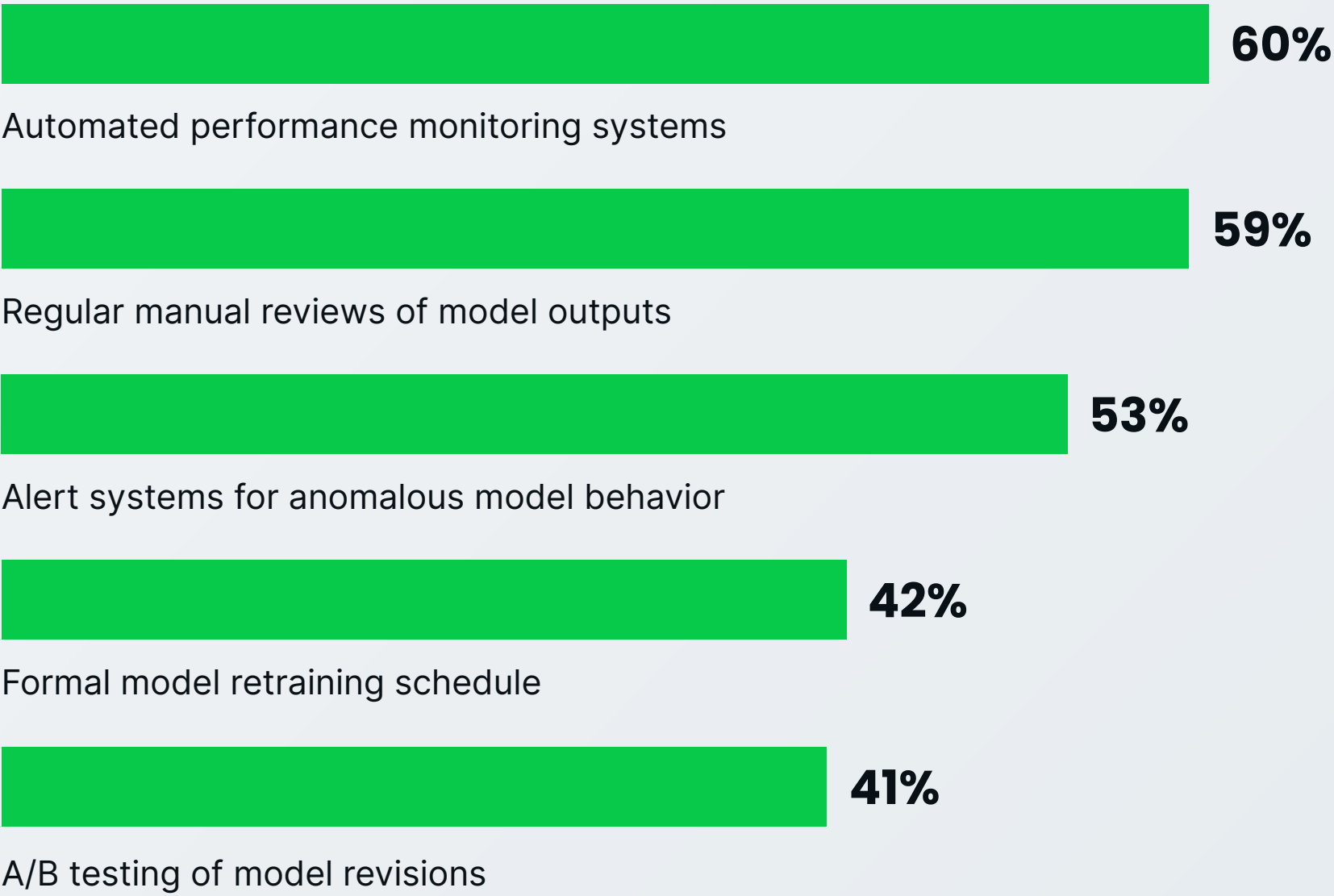
Monitoring models in production is another area with mixed maturity. 70% of teams report having mechanisms to monitor model performance and detect “model drift” (i.e., when a model’s accuracy degrades, or it behaves unexpectedly). The most common practices include automated performance monitoring (60%) and alerting systems for anomalies (53%), often coupled with periodic manual reviews of model outputs (59%). About 42% even schedule regular model retraining or employ A/B testing of model versions to catch degradation.


These are encouraging signs that enterprises recognize the importance of ongoing validation post-deployment. Still, 30% have no formal drift monitoring at all, leaving a significant blind spot.

Many people aren’t using the full suite of modern tools for monitoring. For example, roughly half do not have automated anomaly alerts or A/B tests in place.

As AI models move from pilot to production, this gap can lead to undetected performance issues, bias drift, or even outages if models make uncontrolled decisions. When asked which aspect of model governance will be most critical in the next year, 24% cited performance monitoring and drift detection (a close second to regulatory compliance at 25%).

Which of the following practices do you have in place to monitor for “model drift” or unexpected behaviors in deployed models?  
Select all that apply.





**30%**

of respondents stated they have no formal drift monitoring at all, leaving a sufficient blind spot.

MODEL LINEAGE & PERFORMANCE MONITORING

Weave Provenance Into the AI and ML Deployment Process

A unified approach to AI development makes it easier to embed documentation and monitoring throughout the model lifecycle. By capturing environment details, data lineage, and version metadata as part of everyday workflows, teams can ensure that provenance is recorded automatically, not as a side task. This approach reduces reliance on manual tracking or siloed knowledge and makes critical information accessible to both technical and business stakeholders.

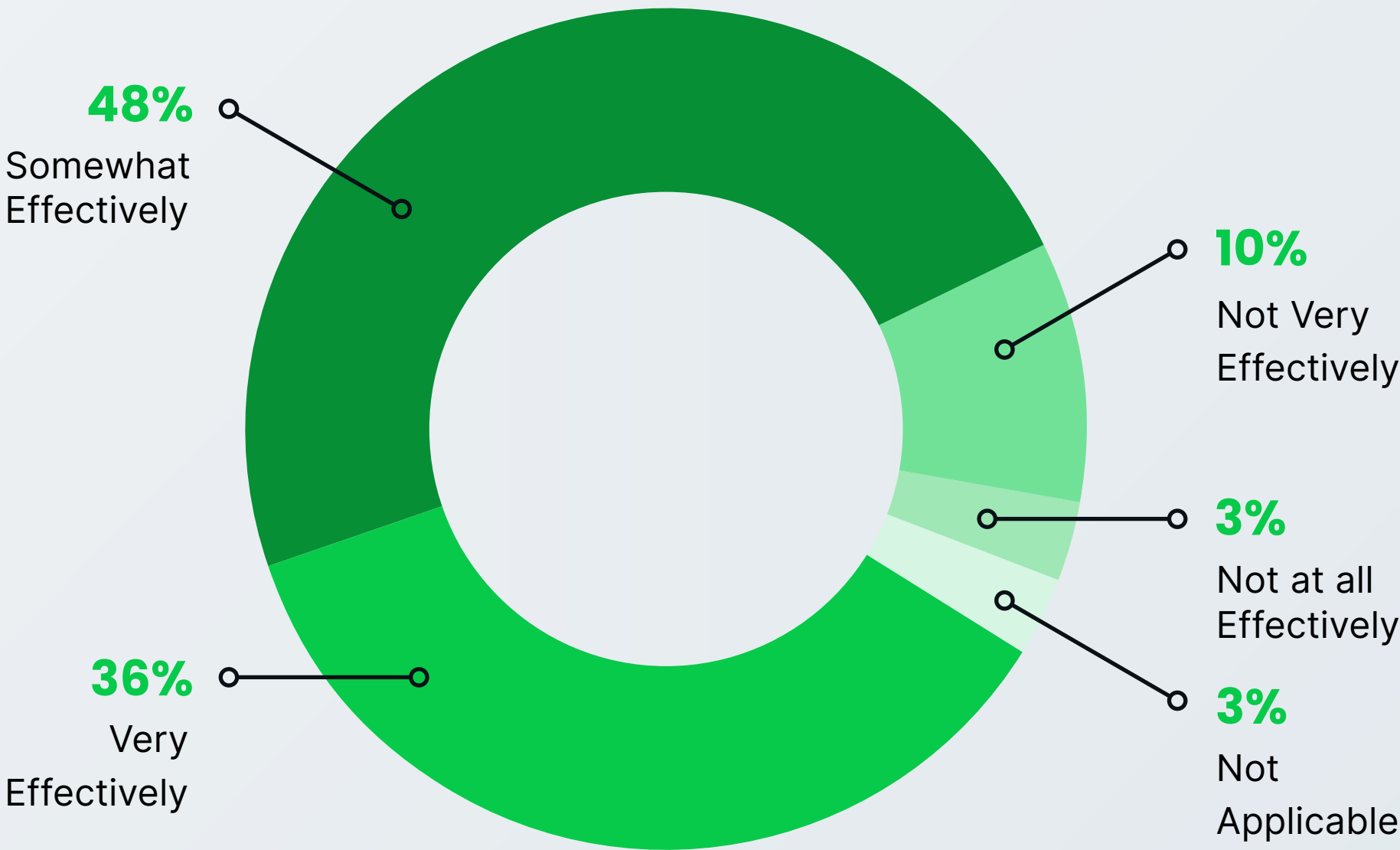
But access remains a major hurdle. Only 36% of organizations say their business stakeholders can very effectively self-serve information about a model's origin, components, and limitations. Most others said it's only "somewhat" effective or worse, often requiring significant effort to dig up details.



Only 36%

of organizations say their business stakeholders can very effectively self-serve information about a model's origin, components, and limitations.

How effectively can your business stakeholders currently access information about the origin, components, and limitations of AI models in production?



# Toolchain Integration: From Fragmentation to Unified Platforms

## Fragmentation Is a Barrier to Governance

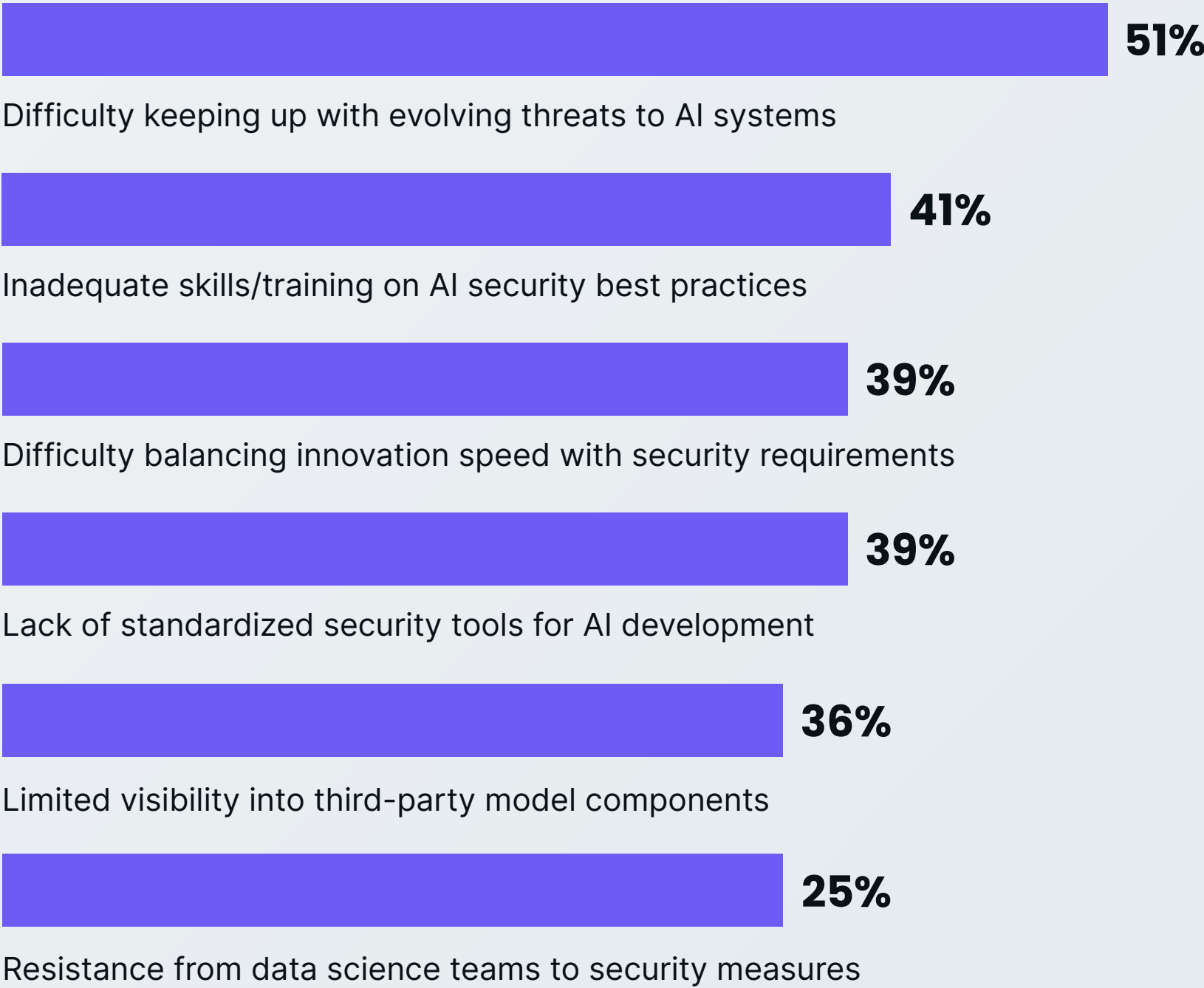
One clear theme in the survey is that toolchain fragmentation remains a challenge for many organizations. AI development involves a plethora of tools: notebooks and IDEs, data pipelines, ML frameworks, model registries, CI/CD, and more. Governance can fall through the cracks if these tools are integrated poorly. Only 25% of respondents described their AI toolchain as “highly unified” across the model lifecycle. The majority are still grappling with a mix of systems; 43% said their toolchain is “somewhat unified” (standardized in parts, but still multiple systems), and another 25% admitted it’s fragmented across teams. A small but significant 5% have “highly fragmented” toolchains with limited integration—essentially, each team using different, siloed tools.

This fragmentation directly undermines governance. It leads to inconsistent security controls and processes, duplicate efforts, and visibility gaps. For example, if data scientists download packages or build models in their own siloed environment, security teams might not know until deployment time, if at all. In the survey’s list of top operational challenges, “fragmented toolchain across teams” was noted by 17% of respondents, which likely factors into other challenges like keeping up with evolving threats (51%) or balancing speed and security (39%).



**Only 25%**  
of respondents described their AI toolchain as “highly unified” across the model lifecycle.

### What are your biggest operational challenges in implementing consistent security controls across your AI model development workflows?



## TOOLCHAIN INTEGRATION

# Unifying Tools and Teams to Improve Oversight

Encouragingly, enterprises recognize the value of unification. Standardizing tools and processes across teams was the #3 improvement priority for the coming year, noted by 21% of respondents. More strikingly, when asked what would most help improve their AI supply chain governance, the top answer (29% of respondents) was “integrated tools that combine development and security”. Better visibility into dependencies (23%) and more team training (19%) followed. These results clearly indicate a desire for shared platforms where ML practitioners and governance teams can collaborate more effectively.

These findings align with Forrester’s outlook that AI governance is consolidating around unified platform systems that balance model performance, latency, and cost while offering observability and explainability. Consolidation is not just about convenience; it’s a prerequisite for managing complexity and scale.



## Top Five Priorities for Improving AI Model Supply Chain Management This Year

- 1 Implementing strong security controls
- 2 Improving documentation and traceability
- 3 Standardizing tools and processes across teams
- 4 Enhancing monitoring capabilities
- 5 Training teams on governance best practices

## TOOLCHAIN INTEGRATION

# Integration Must Be Both Technical and Cultural

Unification is not only a technical task. It's also cultural. Nearly a quarter of survey respondents (25%) cited "resistance from data science teams to security measures" as a key challenge. Integrated platforms can help overcome this tension by embedding governance within familiar workflows rather than layering it on after the fact. The goal is to reduce friction by aligning tools, teams, and incentives.

# 25%

of respondents cited  
"resistance from data science  
teams to security measures"  
as a key challenge.



“

For organizations seeking to scale AI responsibly, there's a clear takeaway: Unify wherever possible. Consolidated environments for development, monitoring, and governance can simplify oversight, streamline collaboration, and reduce risk, helping ensure that innovation and compliance go hand in hand.



**Greg Jennings**

VP of Engineering for AI, Anaconda

# Governance Culture and Metrics: Aligning Technology with Leadership

## Measuring What Matters

Technology alone cannot close the governance gap. Organizations also need the right policies, skills, and leadership support. A positive sign is that 90% of organizations have begun measuring KPIs related to their AI model supply chain's health. Many organizations are actively communicating the value of governance to executive leadership in quantitative terms.

The most common metrics include compliance with regulatory requirements (58%), time to remediate vulnerabilities (53%), number of security incidents (51%), and percentage of models with complete documentation (43%). Additionally, reflecting business priorities, 61% measure productivity/efficiency gains from governance efforts, and 46% tie metrics to business outcomes like revenue or customer satisfaction.

## Communicating ROI

However, not everyone has cracked the ROI code. Nearly 1 in 5 respondents (19%) admitted they struggle to measure or communicate the ROI of AI governance investments effectively, and 5% do not attempt to measure it. For some, governance is still seen as a cost center or insurance policy rather than a direct contributor to value. Gaining leadership buy-in requires highlighting how good governance reduces risk exposure (e.g., fewer incidents, avoiding costly reworks) and accelerates delivery through more reliable pipelines. This connection is backed by external research: [McKinsey](#) found that CEO oversight of AI governance strongly correlates with higher self-reported bottom-line impact from generative AI initiatives. In other words, championing governance at the top will likely deliver better business results.

Framing governance improvements in terms of business risk reduction and value creation (for example, "our governance program prevented X potential breaches" or "enabled Y% faster compliance audits") will help sustain funding and buy-in at the C-suite level.

## Extending to New Frontiers

One notable cultural shift is the rise of generative AI in coding ("vibe coding"). Our survey asked how organizations govern using AI assistants in software development. Only 34% have formal policies and tools in place for AI-assisted coding.

The rest are either applying existing (perhaps outdated) frameworks (25%), developing new policies (21%), or have no specific governance (9%), with a small minority (4%) outright banning AI coding tools. There's a governance lag in adapting to new AI paradigms.

As generative AI becomes ubiquitous in coding, data generation, decision support, and more, organizations must extend their governance programs to cover these novel use cases.

That means updating ethical guidelines, data security policies, and quality checks to account for AI-generated content and code.

“

There are benchmarks that training organizations run in their own models. For example, there's HellaSwag, TruthfulQA, and WinoGrande, and they all ask the model a variety of questions to measure reasoning capabilities, application-specific queries, how well the model does against a battery of software development tasks, how it acts as an agent, and so on. But the problem is when people start testing to those metrics and training against the test set—in other words, they're cheating the system. That's why the metrics have to continually grow, evolve, scale, and incorporate upstream sources.



**Greg Jennings**

VP of Engineering for AI,  
Anaconda

# Closing the Governance Gap

The AI model governance gap is a solvable challenge. The survey data shows that enterprise leaders are aware of the critical issues: security vulnerabilities, compliance demands, model drift, tool fragmentation, and more. Many organizations have pieces of the solution in place, but few have unified it all into a seamless whole. The result is often a patchwork of controls that leaves weak spots—undermining trust and slowing down AI deployments. Closing this gap requires a strategic approach that combines people, processes, and technology. Enterprises should foster a culture where governance is viewed as an enabler of innovation rather than a hindrance. They should implement clear policies and invest in skills development, so teams understand why governance matters. Finally, they should equip those teams with integrated tools that make doing the right thing the path of least resistance.

AI governance isn't just a compliance box to check. It's a foundation for faster, safer innovation. Enterprises investing in culture, clarity, and consolidation will be best equipped to thrive in the era of generative AI.

Narrowing the governance gap is about bringing the same discipline to AI pipelines that we bring to traditional software engineering and IT operations, without losing the agility and creativity that AI innovation requires. It's a challenging balancing act—but entirely achievable with today's tools and a committed organizational approach.



# Build a Stronger, More Comprehensive Governance Approach



## Work Towards a Cohesive AI Governance Stack

AI projects span data pipelines, model development, deployment infrastructure, and compliance reviews. However, in many organizations, a different team using individual tools owns each piece. Form a cross-functional working group—bringing together data science, ML engineering, IT, and compliance—to map the end-to-end model lifecycle and align on a unified governance framework. Select as many interoperable tools as possible that allow for shared visibility and policy enforcement across stages.



## Secure the AI Supply Chain in Layers

Unlike traditional software, AI systems ingest volatile sources—open-source packages, pretrained models, public datasets, and external APIs. This increases risk and the likelihood of unexpected behavior. Establish layered controls:

- Use only trusted, internally mirrored packages and models.
- Require attestations or SBOMs (software bills of materials) for third-party code and models.
- Implement scanning at ingestion, integration, and deployment points to catch emerging risks.



## Treat Model Lineage as Critical Infrastructure

Non-determinism in AI means two training runs on the same data may—and often do—yield different outputs. That’s why tracking lineage isn’t optional—it’s foundational. Automate capture of training data sources, versioned code, parameter settings, and dependency graphs as part of your CI/CD pipeline. Store this metadata centrally to support reproducibility, audits, and model debugging.



## Monitor Deployed Models Like Live Systems

AI models drift, degrade, or behave unpredictably when exposed to new data. Unlike static applications, their performance must be observed continuously. Define meaningful performance metrics (not just accuracy, but fairness, latency, or confidence thresholds), and set automated alerts when these indicators fall out of range. Use shadow testing, canary releases, or A/B comparisons to monitor changes safely before they impact users.



## Operationalize Responsible AI Through Policies and Habits

Responsible AI isn’t a one-off training or a values statement—it’s embedded behavior. Establish formal approval checklists before models go live (e.g., bias testing, robustness evaluations, lineage sign-off). Use code review workflows that include security and ethical considerations. Encourage teams to document model limitations and intended use. These small habits compound to create a governance culture.



## Measure and Communicate Governance Outcomes

Don’t just check compliance boxes—measure how governance improves your AI delivery. Track cycle time from model prototyping to deployment, number of incidents tied to untracked changes, or percentage of models with complete metadata. Regularly report these metrics to leadership. Celebrate wins, such as improvements in incident response time or audit-readiness milestones, to build momentum.



Anaconda is built to advance AI with open source at scale, giving builders and organizations the confidence to increase productivity, and save time, spend and risk associated with open source. 95% of the Fortune 500 including Panasonic, AmTrust, Booz Allen Hamilton and over 50 million users rely on the value The Anaconda AI Platform delivers through a centralized approach to sourcing, securing, building, and deploying AI. With 21 billion downloads and growing, Anaconda has established itself as the gold standard for Python, data science, and AI and the enterprise-ready solution of choice for AI innovation. Anaconda is backed by world-class investors including Insight Partners.

Learn more at [anaconda.com](https://anaconda.com).

