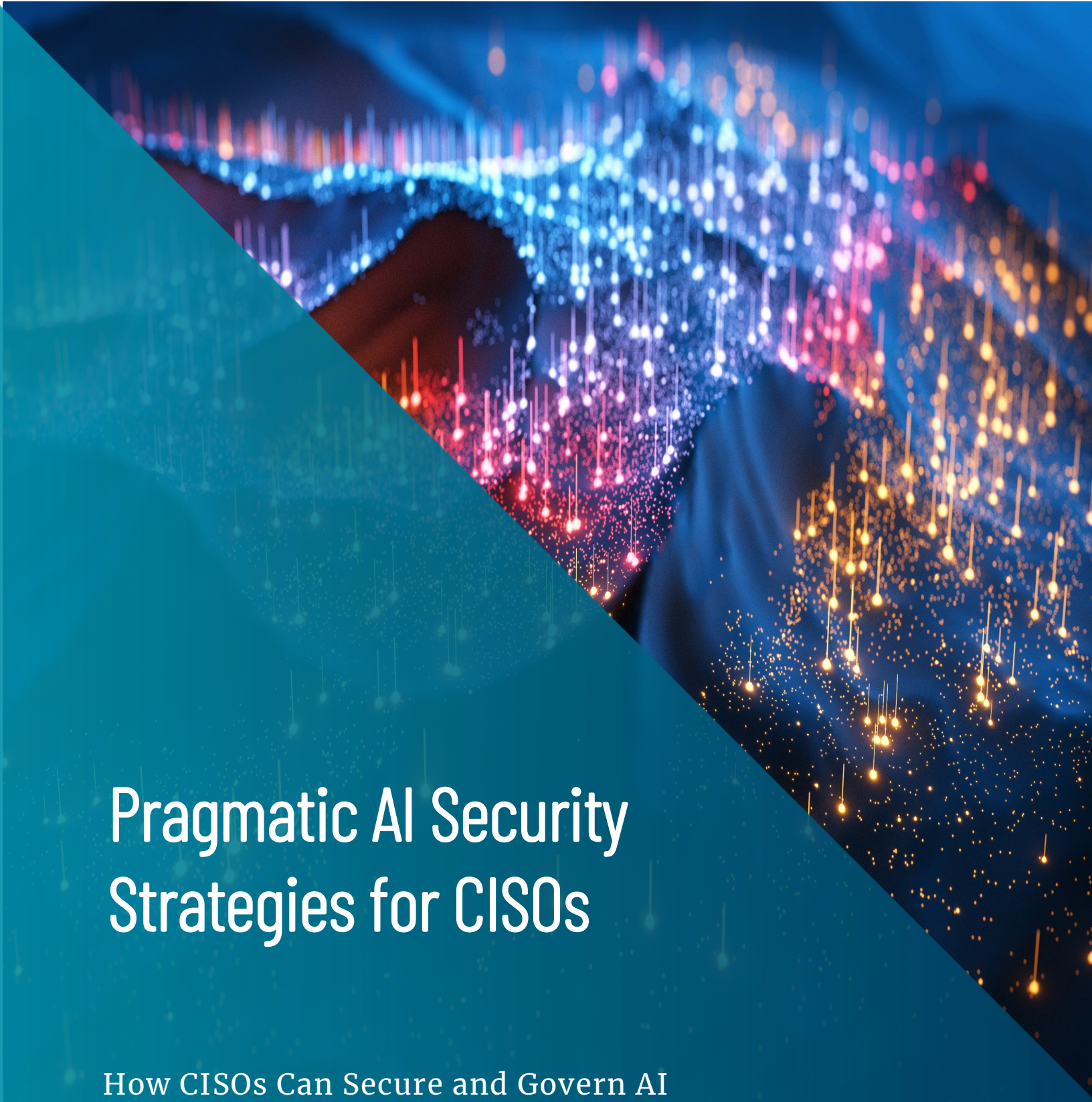

Pragmatic AI Security Strategies for CISOs

How CISOs Can Secure and Govern AI Without Slowing Business Operations

protiviti®
Global Business Consulting

Pragmatic AI Security Strategies for CISOs

How CISOs Can Secure and Govern AI Without Slowing Business Operations

Artificial intelligence (AI) is transforming how organizations work, compete, and serve customers. Many enterprises are moving quickly to implement AI in their business, eager to capture productivity gains and new capabilities.

Business leaders are pressing chief information officers (CIOs) to remove barriers so they can engage with new AI solutions and providers as fast as possible. The CIO, caught in the middle, is struggling to balance the needs of the business while ensuring that the organization is adequately protected.

For the chief information security officer (CISO), this is a familiar but intensified dilemma. Most CISOs understand the significant risks and challenges AI can bring to the organization, especially in areas like data security, privacy, model integrity, and novel attack vectors such as prompt injection or data poisoning.

In response, CISOs are racing to implement robust AI security standards and protocols to protect the organization. Yet those very procedures, if applied uniformly, can slow the organization and create overhead that business leaders don't want to absorb.

The predictable result: frustration, friction, and the temptation for teams to route around the rules, sometimes adopting "shadow AI" and thereby increasing risk.

The answer is not to slam on the brakes or to look the other way. The answer is pragmatism: a risk based model that enables speed where risk is low and applies deeper governance where risk is high.

61%

of CFOs and finance leaders rate data security and privacy as a high priority for the coming year

Why One-Size-Fits-All AI Controls Backfire

Traditional controls assume a relatively stable IT stack and a known vendor landscape. AI breaks both assumptions. AI systems:

- Touch sensitive data (training, tuning, prompts, outputs)
- Behave probabilistically, introducing explainability and fairness considerations
- Depend on complex supply chains (models, weights, datasets, vector stores, plug-ins)
- Change fast, making static assessments obsolete.

Imposing heavyweight reviews on every AI experiment slows learning and invites work-arounds. Conversely, green-lighting everything invites unacceptable exposure. A pragmatic approach threads the needle.

In Protiviti's 2025 Global Finance Trends Survey, AI usage in finance teams jumped from 34% in 2024 to 72% in 2025.

Establish Streamlined Fast Paths for Lower-Risk AI

It would be beneficial to CISOs to develop a **streamlined path to lower risk AI solutions** that business leaders can engage with quickly and easily. The key ingredients:

1. **Pre-vetted providers and platforms:** Focus on better known AI solutions with reputable companies that typically employ robust security protections, documented controls, and enterprise contracts (e.g., SSO, data residency options, SOC 2/ISO certifications).
2. **Low-risk use cases:** The fast path should be reserved for use cases that avoid high-stakes decisions and minimize the handling of sensitive information. Examples include summarizing public content, marketing ideation (without personally identifiable information, or PII) or efficiency helpers inside secured office suites.
3. **Guardrails for data:** If any sensitive information is used, it must run in a protected environment (e.g., private tenants, encryption in transit/at rest, strict access controls, data loss prevention, prompt/response redaction, and strong logging).

This approach preserves velocity while ensuring sensible protection and, crucially, it gives the business a legitimate, sanctioned path so it doesn't feel compelled to circumvent security.

Operationalizing Risk: AI Green/Yellow/Red Zones

A clear, easy-to-use zoning model helps everyone understand expectations and speeds decisions.

Green Zone — Fast-Path

- **Use cases:** Low-risk tasks; no human resources decisions, hiring, credit decisions, medical determinations, or other regulated/high-bias scenarios
- **Data profile:** Public or non-sensitive; no regulated data (e.g., protected health information, payment card industry data, special-category personal data)
- **Providers:** Reputable, approved vendors with enterprise controls and standard terms

Yellow Zone — Due Diligence Required

- **Use cases:** Moderate risk (e.g., customer support assistants, internal analytics that may touch confidential — but not highly regulated — data).
- **Data profile:** Confidential data permitted with guardrails (masking, role-based access, tenant isolation).
- **Providers:** Less-known vendors or new capabilities from known vendors; require security and privacy questionnaires, contract addenda and limited pilots.

Protiviti's Second Annual Generative AI Study found that the number of AI projects in production almost doubled in the last year.

Red Zone — High-Scrutiny

- **Use cases:** High risk or high-impact: HR solutions, hiring decisions, lending/insurance underwriting, safety-critical decisions, legal determinations, or any activity where bias and explainability are regulatory concerns
- **Data profile:** Sensitive or regulated data; any system making consequential decisions about people
- **Providers:** Unknown or unproven vendors; custom models that require model risk management

This simple taxonomy delivers clarity. It **enables** the organization to move quickly for AI use cases that do not create undue risk, while **preserving a robust governance structure** for higher-risk cases. It also creates a shared language between security and the business.

According to the 12th Annual Global Technology Audit Risks Survey by Protiviti and The Institute of Internal Auditors (IIA), 59% of IT audit leaders identify AI as a significant threat over the next two to three years.

Turning the Model Into a Repeatable Operating Mechanism

A pragmatic strategy works only if it's easy to follow. Consider these building blocks to make your zones real:

1. **AI Intake and Use Case Inventory**
   A short, self-service intake form that captures purpose, data types, users, and provider details. Auto-classify into green, yellow and red using simple rules. Maintain a living inventory to eliminate shadow AI.
2. **Acceptable-Use Policy and Quick-Start Playbooks**
   Provide concise do/don't guidance (e.g., "no secrets in prompts," "use only approved connectors"). Pair each zone with a one-page playbook: what teams must do, what security provides, and typical turnaround times.
3. **Prevetted Vendor Catalog**
   Keep a catalog of approved solutions with default configurations, security settings, data-retention defaults, and contract clauses. For yellow and red, include vendor due diligence templates and testing checklists.
4. **Protected Environments for Sensitive Data**
   When sensitive information is involved, require controlled runtimes: private large language model endpoints, encrypted vector stores, centrally managed keys, masking/tokenization pipelines, and strict egress controls.
5. **Human in the Loop for Consequential Decisions**
   For red-zone decisions, mandate human review and clear accountability. Document how overrides, appeals, and recourse work.

In Protiviti's 2025 Global Finance Trends Survey, only 41% of finance leaders expressed high confidence in navigating global challenges, underscoring the need for robust governance and risk management frameworks.

6. **Continuous Monitoring and Red-Teaming**
   Instrument usage: log prompts/responses (with privacy controls), track data flows, throttle abuse, and alert on anomalies. Schedule adversarial tests, bias audits, and periodic re-certifications for higher-risk systems.
7. **Training and Culture**
   Offer role-based training so product owners, developers, analysts, and leaders understand both the power and the limits of AI, as well as their responsibilities under each zone.

Metrics That Matter

To reinforce the partnership mindset, measure success in both **speed** and **safety**:

- **Time to approve** by zone (target: hours for green, days for yellow, defined service level agreement for red)
- **Control coverage** for high-risk systems (target: 100% for required controls)
- **Business value realized** (target: quantified efficiency or revenue outcomes tied to safe deployments)

When the business sees that security is accelerating the right things — and only slowing the wrong things — trust grows.

Aligning With Risk Tolerance and Regulation

A zoning model must reflect your organization's **risk appetite** and regulatory environment. For example:

- Highly regulated sectors (healthcare, financial services, public sector) may set a **higher bar** for red zone approvals and treat more use cases as yellow by default.
- Multinational companies need **jurisdiction-aware** defaults for data residency, cross-border transfers, and local AI and algorithmic accountability requirements.
- If you already use frameworks like NIST CSF (including the Govern function), NIST AI RMF, ISO/IEC 42001, or existing model risk management practices, map your zones and controls to those artifacts to avoid reinventing the wheel.

The CISO as Business Partner

Being seen as a partner to the business is essential to ensure that the business and security move forward together in this new territory of AI. The pragmatic approach — green, yellow and red zones fast paths for known providers and low-risk use cases, protected environments for sensitive data and heightened scrutiny where it counts — creates a **balanced, credible path**. It satisfies the desire to move quickly while also providing protection aligned to the organization's tolerance for risk.

About the author

Scott Laliberte (MBA, CISSP, CISA, CRISC, CISM)
Managing Director, Protiviti

Scott Laliberte is a managing director in Protiviti's Technology Consulting practice, serving as a trusted CISO advisor to boards and executive leadership teams. He specializes in helping organizations understand and manage cybersecurity risk in the context of business strategy, enabling innovation while safeguarding enterprise value.

With deep technical expertise and extensive business knowledge, Scott advises clients on leveraging emerging technologies and advanced methodologies to drive growth without compromising security. His unique ability to translate complex cyber threats into actionable business insights positions him as a strategic partner for organizations navigating digital transformation.

Previously, Scott led Protiviti's Global Cybersecurity practice, its Emerging Technology practice, and its attack and penetration labs, where he guided initiatives in technology innovation, cyber strategy, and advanced security architectures. His experience spans advising boards on cyber risk governance and helping enterprises build resilient security programs aligned with regulatory and operational priorities.
