Deloitte.

Together makes progress

A foundation of confidence

Deloitte Trustworthy AI™ Solutions: **AI Risk Management & Governance**

The most recent Deloitte State of Generative AI in the Enterprise survey found that many organizations' approach to risk management and governance for their artificial intelligence (AI) initiatives was not as advanced as their AI use itself. That may suggest that in developing these controls, speed is important but scale may be just as vital a consideration.

It may seem counterintuitive, but it may be advisable in many cases not to govern too much too quickly. An over engineered risk and governance program may slow and bureaucratize the business just when it needs to be at its most agile, and AI programs may remain too long in proof of concept mode as a result.

AI risk management and governance should play a dual role: not only safeguarding the business, but also helping it unlock innovation and new value. With controls scaled to need, organizations can progress along their AI journeys with both energy and confidence.

Not all risks are created equal

Establishing and operating strong governance frameworks is part of the broader mandate of risk management, which uses technology and operational awareness to identify, assess, mitigate, and monitor AI related risks.

The risks from AI use can include bias, errors, and unintended consequences. When those risks occur, they can have legal, ethical, and reputational consequences for an organization. These failures can originate in the design of an AI system, in its implementation, or its use and may involve either deliberate or inadvertent missteps. To address this broad threat, AI risk management technology helps anticipate, perceive, and address risks at a pace and volume manual processes cannot match.

The risk approach also adjusts to direct governance efforts where they will do the most good. For example, a function such as in store navigation assistance is a low risk concern, while something like facial recognition for customer identification carries a high risk profile. The risk management system accounts for all identified risks, but not by devoting the same resources to each one. Instead, each organization can use a custom configured scoring system based on its unique risk tolerances. How many users does an issue affect? What regulations apply? Is there a potential impact on human health and safety? How sensitive is the data involved? These and other considerations can be used to "score" the different risks an organization may face, then calibrate the application of governance measures in a proportional way. The result is more agile and cost effective than a one size fits all approach would be.

Governance tuned to each threat level

One way to keep these safeguards at the right scale and pace is to align them with specific implementations, rather than applying them enterprise wide. Each effort to establish technology guardrails, compile documentation, establish permissions, or other steps can be carried out to suit a particular model or use case as they emerge. They can apply according to risk level, from low to moderate to high. This goes for AI in any form, whether it's rule based, machine learning, Generative AI, agentic AI or another variety.

- **Continually assess.** Inventory AI assets to evaluate what is currently in place, what capabilities are emerging, and what blind spots may exist.
- **Focus on speed to value.** Small, practical processes can foster buy in and build momentum.
- **Be risk based.** Apply governance controls at first to specific, identified risks, and fast track more general concerns.
- **Be nimble.** Flexibility and scalability allow the governance landscape to adapt as the risk landscape evolves.
- **Keep going.** From this agile start, use performance measurement to tune processes, technologies, and people for continuous improvement.

AI Risk Management & Governance in practice: Case example

A large materials management and remediation services company was ramping up its use of AI and needed a scaled approach to risk management and governance to accompany its first deployed applications. Deloitte helped the company up AI governance and risk management frameworks to facilitate real time monitoring of AI and Generative AI projects, models, and data sets across the organization.

Using Deloitte's AI Governance approach enabled on a central platform, the project set up a comprehensive governance framework for Generative AI implementation, established an AI Steering Committee, defined new roles and responsibilities, and facilitated team member sessions to finalize and activate those elements. The governance approach used interactive dashboards and was designed to expedite self service AI development with an online approval process. A newly established governance charter, governance operating model, and risk management framework provided the foundations for the new system.

The process gave the organization visibility into its AI asset risk for the first time and increased the chief information security officer's confidence level in reporting to the board on AI progress.

Tools that can bring this approach to life

Conducting an agile risk management process can reveal an organization's AI assets and the risk profile associated with them sometimes more comprehensively than before, or even for the first time. This can inform a more comprehensive view of areas where improvement is proceeding and which areas instead require attention. Using that insight, a governance process can apply the right approach as needed to balance caution with momentum.

To help achieve these aims, Deloitte offers AI Risk Management & Governance support as part of its overall suite of Trustworthy AI™ Solutions. This process uses proprietary accelerators and deeply researched intellectual property, along with the power of technology enablers like IBM WatsonX Governance, OneTrust, ServiceNow, and others. Moving step by step through the governance life cycle from intake, assessment, and mitigation through development, deployment, operation, and monitoring, Deloitte's solution mitigates AI risks by establishing strong enterprise policies, operating models, procedures, controls and mitigations, and governance frameworks. It helps promote compliance with evolving regulations and implements scalable risk management in ways that enhance operational trust.
As a result:

- Processes run faster at scale with streamlined decision making.
- Improved model performance and outputs lead to better outcomes and user satisfaction.
- Brand value and reputation enjoy protection against threats.
- Clarity and control give an organization the confidence to make bold AI investments.

Growing power applied with evolving controls

In this fast growing technology environment, enterprise trust is made up of twinned phenomena: human trust and Trustworthy AI. Understanding and controlling AI risks requires contributions from both disciplines.

Deploying AI without understanding its risks and building governance to manage them can be like driving a car without brakes. But over building those safeguards can be like driving a car with the parking brake on. The challenge is to find balance. Taking risk management and governance one step at a time, each in its full context, can help make that happen.

The Deloitte difference

At Deloitte, we believe trust is essential to scaling AI with confidence. It must operate on two levels: the system must be designed to perform reliably, and people must feel secure in using it.

That's why we've built an integrated platform for Trustworthy AI one that brings together machine level governance and human centered design. It's engineered to help organizations develop AI systems that are secure, transparent, explainable, and aligned with intended outcomes.

Backed by Deloitte's AI Institute, supported by global research, and informed by deep experience across industries, this platform helps organizations embed trust into AI development from day one transforming it from a reactive concern into a proactive capability for responsible growth and long term value.

Derek Snaidauf
Principal
Deloitte Transactions and Business Analytics LLP
dsnaidauf@deloitte.com