# Top Strategic Technology Trends for 2024: AI Trust, Risk and Security Management

Published 16 October 2023 - ID G00796422 - 10 min read

By Analyst(s): Avivah Litan, Jeremy D'Hoinne, Bart Willemsen

Initiatives: Artificial Intelligence; Evolve Technology and Process Capabilities to Support D&A

Generative AI has opened the floodgates for new AI initiatives, making the need to implement robust AI trust, risk and security management capabilities even more urgent. This research describes why IT leaders must build guardrails to keep AI efforts under control and enable AI's potential.

**Additional Perspectives**

- Summary Translation: Top Strategic Technology Trends for 2024: AI Trust, Risk and Security Management
  (24 November 2023)

## Overview

### Opportunities

- Business and application stakeholders have become more aware of the challenges of AI and supportive of dedicated AI security and risk management projects, but aren't prepared to postpone business AI initiatives. However, such initiatives still require a robust set of control measures for successful and durable deployment, thereby creating a schism in the organization.
- Organizations that limit AI security controls to narrow vectors, such as large language model prompts, will fail to adequately manage risk. AI trust, risk and security management (AI TRiSM) helps to consistently address ModelOps, AI-specific security and risk concerns, and mitigate against unintended outcomes.
- Organizations that have already implemented some components of AI TRiSM technology move AI models and applications into production more quickly and reliably.

## Recommendations

IT leaders responsible for artificial intelligence:

- Make the use of AI safer by enhancing and upscaling your application security and risk management programs to cover new AI attack and compromise surfaces when delivering and supporting AI-based solutions.
- Keep up with the increasing maturity of available controls to design, train and operate AI models and applications. Include data protection measures in the design of models and applications, and in training data when applicable. Continuously monitor models and applications in production for inaccuracies, drift and unintended outcomes.
- Get ahead of compliance issues by deploying AI TRiSM principles across business units to manage technical and organizational requirements. Do so because regulations and accepted industry wisdom are much slower and potentially less demanding than business imperatives.

# Strategic Planning Assumption

By 2026, enterprises that apply TRiSM controls to AI applications will consume at least 50% less inaccurate or illegitimate information that leads to faulty decision making.

# What You Need to Know

*This research is part of Gartner's Top Strategic Technology Trends for 2024.*

**Download the Executive Guide to AI Trust, Risk and Security Management.**

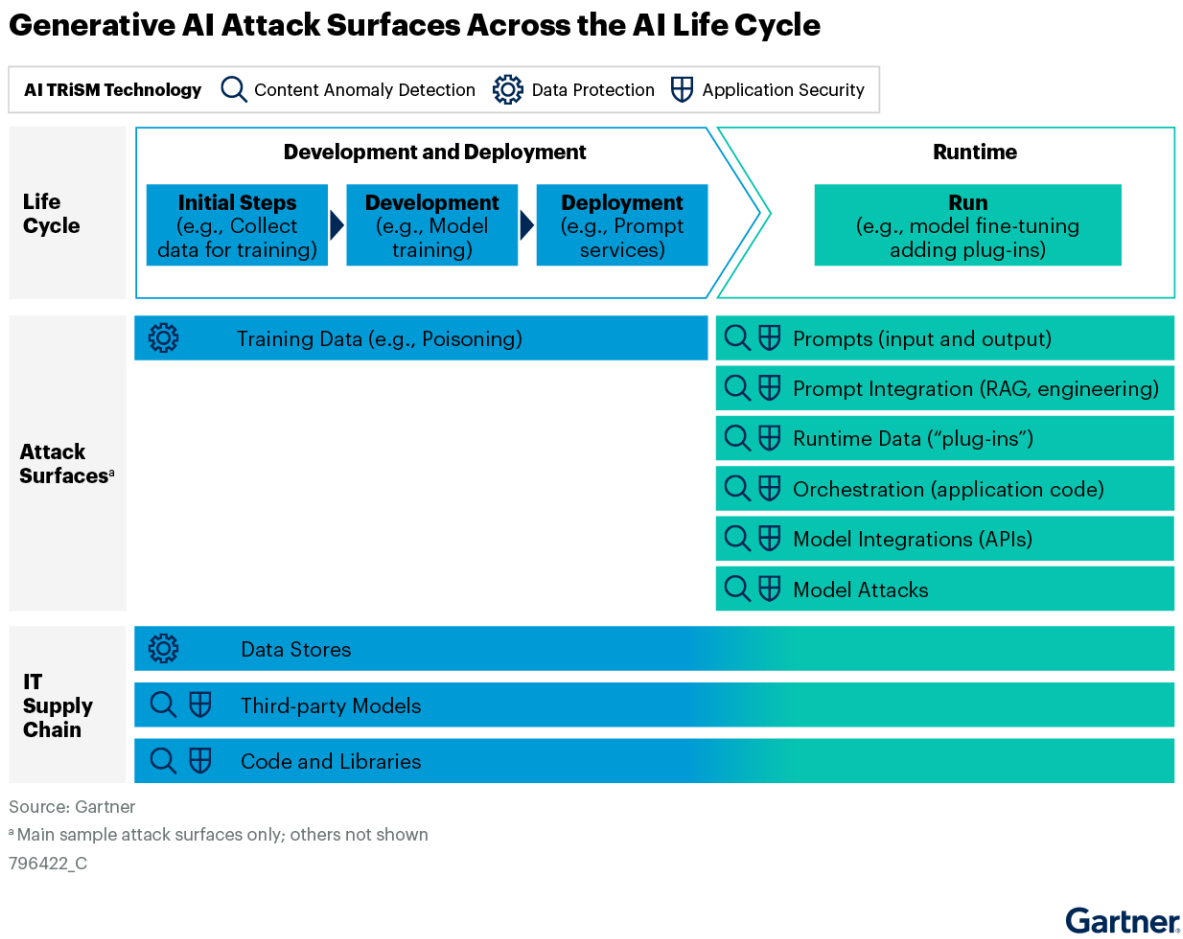
Though organizations develop, buy or deploy business solutions with AI incorporated as if they were traditional applications, they aren't. Developing AI applications in this way exposes the organization to additional risk. Organizations' growing dependence on AI increases the risk and impact of misperforming AI models and applications that use them. The consequences can be serious.

AI models and applications deployed in production should be subject to protection mechanisms. Using such mechanisms ensures acceptable use — based on predetermined intentions — that support sustained value generation.

In a Gartner poll conducted after the release of ChatGPT, 45% of organizations expected their investment in AI to increase. [1] When faced with the realization that more AI applications will go to production in a few months because of business imperatives, IT leaders often turn to security technologies purported to be relevant for AI. Unfortunately, these quick workarounds aren't sufficient to protect the new attack surfaces introduced by the use of AI models.

Good AI trust, risk and security management (AI TRiSM) also requires organizations to manage the entire AI development life cycle and complement existing application security testing and runtime controls (see Figure 1 for examples of generative AI attack and compromise surfaces and life cycle). Applicable life cycle steps will depend on the deployment option chosen. For example, training data is applicable only when enterprises train or fine-tune a model.

## Figure 1: Generative AI Attack Surfaces Across the AI Life Cycle

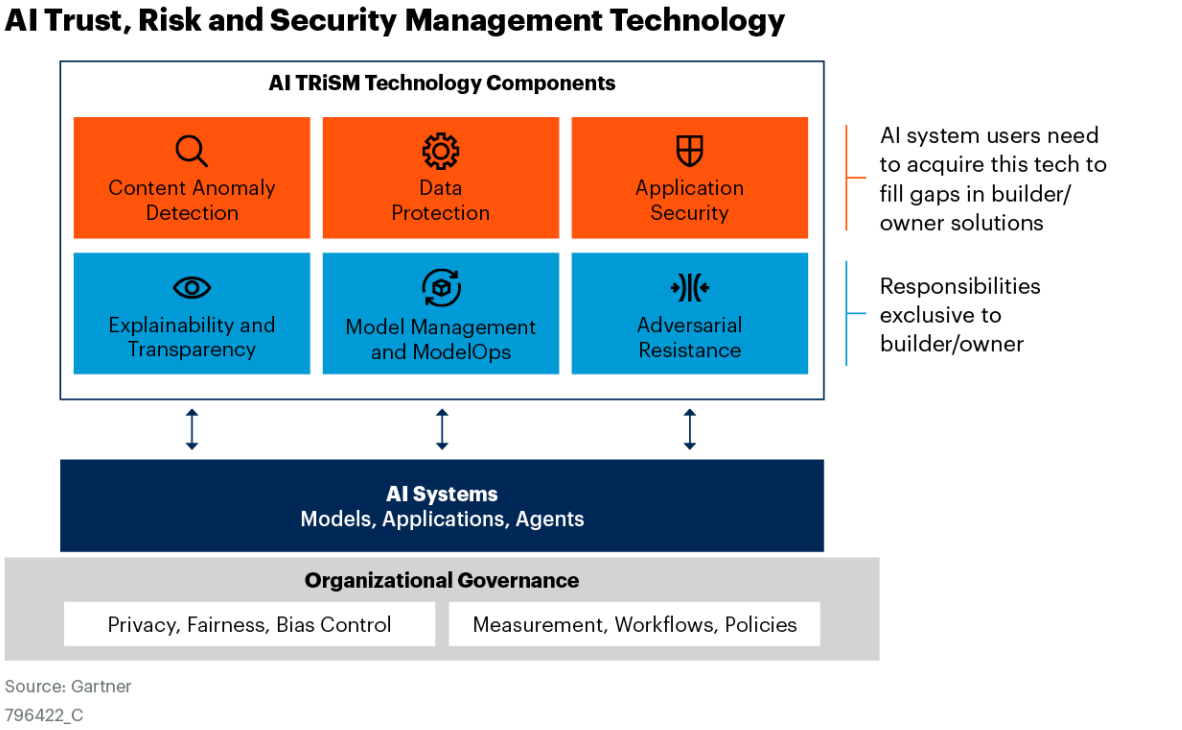



Source: Gartner
[a] Main sample attack surfaces only; others not shown
796422_C

Gartner

AI TRiSM (see Figure 2) is a framework of risk and security controls and trust enablers. It helps organizations govern and manage AI models and applications throughout their life cycles — and achieve business goals. It will also help organizations comply with upcoming regulations, such as the EU's AI Act. [2]

## Figure 2: AI Trust, Risk and Security Management




IT and business leaders must lead a cross-discipline team to manage AI trust, risk and security. The team must start by inventorying AI projects and datasets to track potential privacy, bias and copyright issues. It must establish acceptable use policies and implement ongoing monitoring controls for data and model drift, unintended outcomes, and overall model and application governance. An AI TRiSM program must also protect AI models, applications and datasets from adversarial activities — including adversarial inputs, prompts, completions and data compromises.

Organizations that initiate and sustain AI TRiSM efforts will achieve better business results through improved accuracy and performance of AI model and application outputs and outcomes.

## Profile: AI Trust, Risk and Security Management

### Description

AI trust, risk and security management (AI TRiSM) ensures AI model and application governance, trustworthiness, fairness, reliability, robustness, efficacy, security and data protection. AI TRiSM includes solutions and techniques for model monitoring, interpretability and explainability, data and content anomaly detection, AI data protection, model management and operations, adversarial attack resistance and AI-specific application security.

Some functions are solely the responsibility of AI model and application builders or owners, while the responsibility for other functions is shared across users and builders or owners.

### Why Trending

The democratization of AI has made the need for AI TRiSM even more urgent and clear. For example, ChatGPT has transformed how many organizations compete and do work. The risks associated with hosted generative AI applications are significant and quickly evolving. Without guardrails, any type of AI model can rapidly generate compounding negative effects that spin out of control, overshadowing any positive performance and gains from AI.

Those who use third-party AI services (i.e., search and chat) guided by TRiSM capabilities will benefit from more reliable information for decision making. Contrarily, those who use AI services *not* controlled by TRiSM capabilities will experience substantial growth in inaccurate information used in decision making.

Organizations must expand their AI TRISM practices to mitigate external AI risks because they can't directly control those risks' origins. Organizations that use AI models managed with TRiSM can enhance bias control in decisions while increasing fairness in AI-driven applications.

Ensuring that AI implementations are compliant, fair and ethical requires constant monitoring of AI models and applications. Risk management tools are vital to identify and eliminate drift and uncontrolled biases from training data and algorithms.

AI model explainability is also important, when possible, and must be constantly tested through model observations. This ensures that original explanations and interpretations of AI models remain active during model operations. If they don't, corrective actions must be taken.

Using AI TRiSM will enable organizations to remain competitive, with transparent AI readily understood and managed by business leaders. Such models and applications will deliver consistent improvement in intended outcomes and results, increasing internal and external acceptance. Organizations that actively use AI TRiSM controls move more of their AI projects into production and achieve more business value than organizations that don't actively manage these risks.

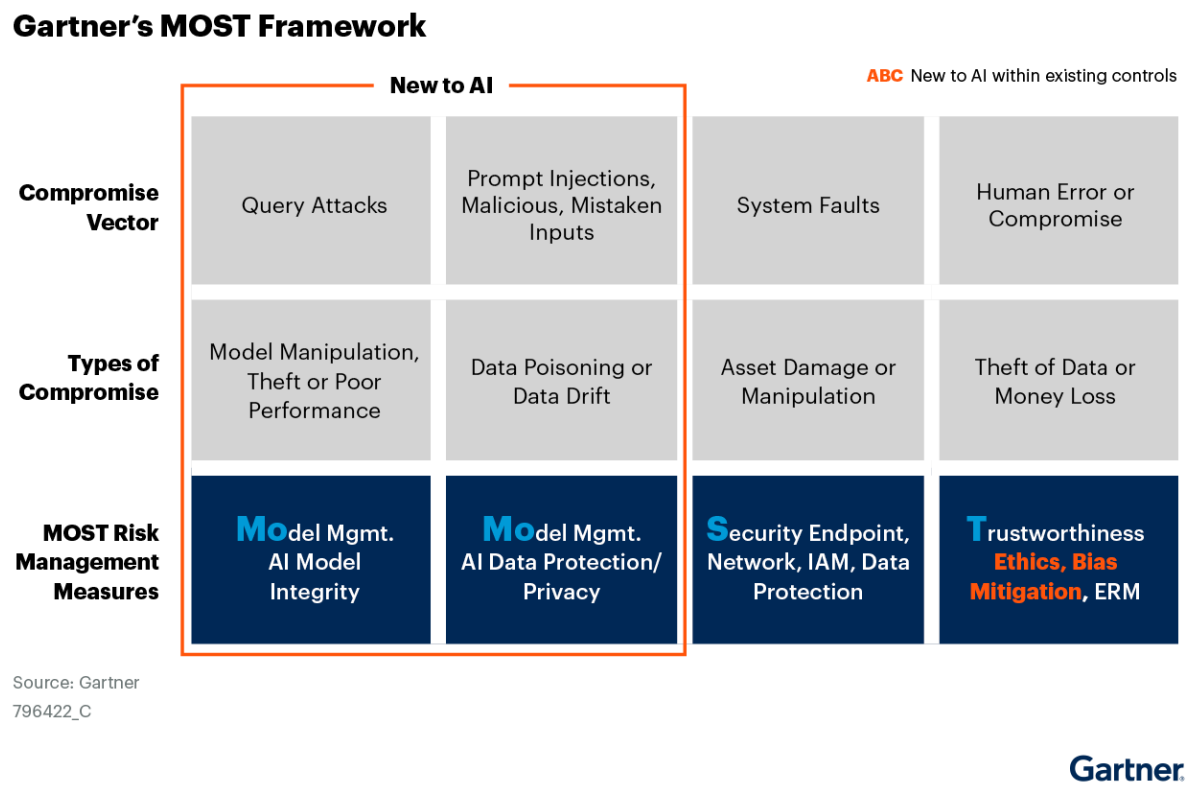
### Implications

Apart from providing operational benefits in the form of reliable and durable technology deployments, AI TRiSM also helps meet AI regulations proliferate worldwide and mandate auditable practices that ensure trust, transparency and human protection. IT leaders must implement AI TRiSM capabilities to ensure model and application reliability, trustworthiness, security and data protection.

However, many IT leaders fail to ensure that AI TRiSM controls are properly implemented and followed by their teams during the design stages of AI-based application delivery. Addressing them later often fails to provide the robust controls needed to enable organizations to adequately trust and rely on the AI model. This makes it more challenging, creates inefficiencies and opens the process to potential risks.

Gartner's MOST framework (see Figure 3) shows existing compromises and mitigation measures and those that are new to AI. IT leaders should familiarize themselves with these new forms of compromise, and the mitigating measures that AI TRiSM provides, so they can properly protect AI models and applications.

Figure 3: Gartner's MOST Framework




Organizations that use hosted models, such as large language models like GPT4, can't control model training and management, which belongs to model owners. They can, however, actively work with solutions that:

- Mitigate input/output risks using data and content anomaly detection
- Protect against data protection and privacy risks using data governance solutions that prevent sensitive data from leaving the enterprise
- Reduce application security risks using cybersecurity solutions that detect and prevent attacks

AI TRiSM requires participants from different units — such as AI, security, compliance and operations — to work together to implement new AI TRiSM measures. This can be challenging in some organizations because of competing priorities.

### Successful AI TRiSM Deployments

Examples of successful AI TRiSM capability deployments include those at:

- **Fidelity Investments:** Fidelity deployed hundreds of AI models through a model operations framework. It followed detailed control steps to consistently monitor deployments for potential problems such as drift, enabling Fidelity to respond to them before the need for escalation. As a result, Fidelity increased model-to-production speed by 100% and reduced the time to find and resolve issues by 80% (cutting it from weeks to hours). [3]

- **Harver:** Harver implemented TruEra software to comply with new legislation on bias reduction in algorithms used for hiring. Harver implemented TruEra for ongoing model performance testing, bias/fairness controls and root cause analysis to rapidly address any issues found. [4] The AI-enabled solution is annually evaluated and audited for compliance by the Babel Group.

### Actions

- Set up an organizational task force or dedicated unit to manage your AI TRiSM efforts.

- Work across your organization to manage best-of-breed toolsets as part of a comprehensive AI TRiSM program.

- Define acceptable use policies. Establish systems to methodically record and approve user applications and document uses.

- Monitor usage continuously against stated objectives, and adjust usage parameters as appropriate on an ongoing basis.

## About Gartner's Top Strategic Technology Trends for 2024

This trend is one of our Top Strategic Technology Trends for 2024. These are the trends we consider most relevant and impactful. Our trends fall into three main themes:

- **Protect your investment.** Preserve your investments and secure the benefits from past and future strategic technology decisions to make them durable.

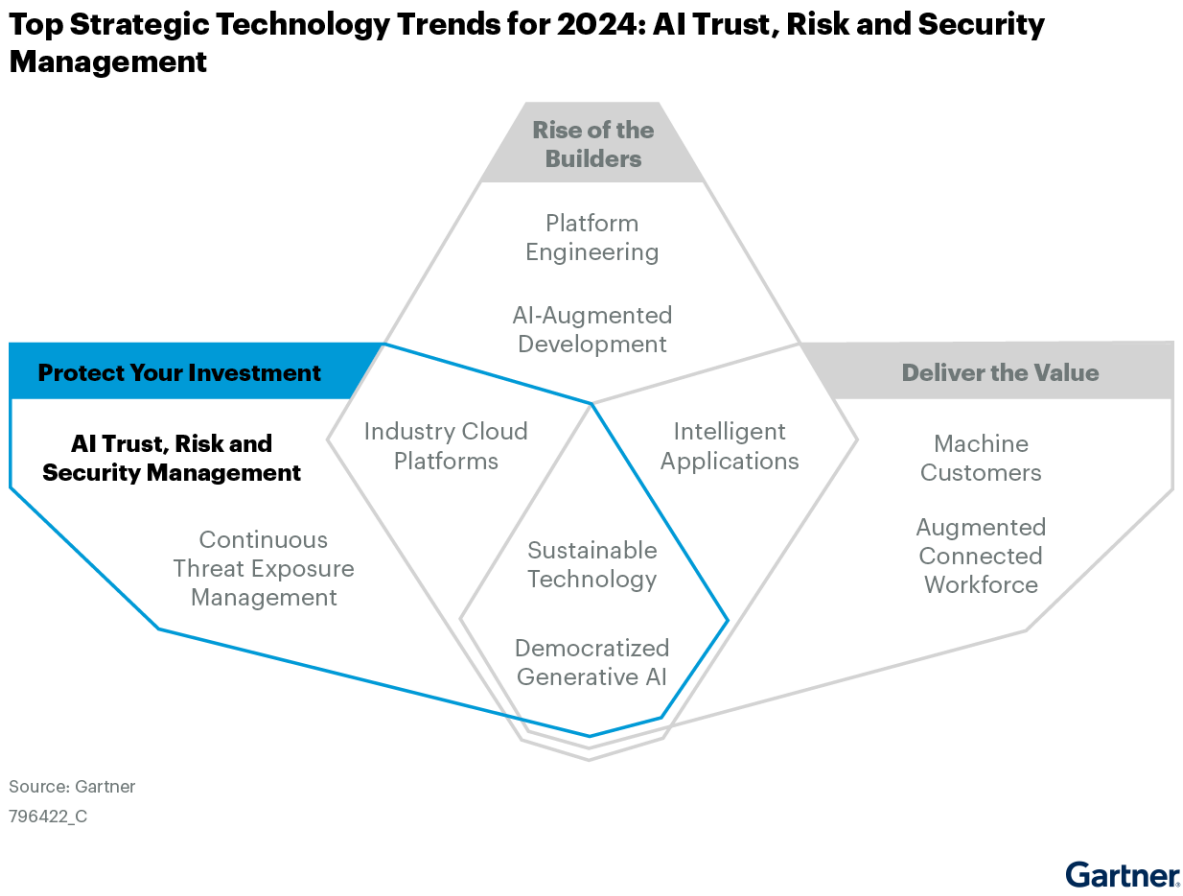
- **Rise of the builders.** Unleash creative powers by using the appropriate technology for the appropriate functions.

- **Deliver the value.** Refine and accelerate value optimization, built on top of durable operational excellence.

These technology trends don't exist in isolation — they interconnect (see Figure 4) and several fall into more than one theme. The trends' potential importance for your organization differs by organizational maturity, but also by industry, business needs and previously devised strategic plans.

Work with other executives to evaluate the impacts and benefits of our trends. This will enable you to determine which single trends — or strategic combination — will have the most significant impact on your organization, and the ecosystem in which it operates. Examine the trends' potential relative to your organization's specific situation, factor them into your strategic planning for the next few years, and adjust your business models and operations appropriately.

Figure 4. Top Strategic Technology Trends for 2024: AI Trust, Risk and Security Management



## Evidence

[1] On 30 March 2023 and 21 April 2023, we conducted a poll as part of our webinar, Beyond the Hype: Enterprise Impact of ChatGPT and Generative AI. We asked the webinar attendees about their plans for AI investment. The results were:

- Forty-six percent planned no change to investment
- Forty-five percent planned to increase investment
- One percent planned to temporarily suspend investment
- Zero percent planned to decrease investment
- Eight percent not applicable

[2] EU AI Act: First Regulation on Artificial Intelligence, European Parliament.

[3] Case Study: AI Model Operations at Scale (Fidelity)

[4] Harver Implements TruEra AI Quality Solution to Assess Algorithmic Fairness, Help Comply with New York Bias Audit Regulations in Automated Hiring, TruEra.

## Document Revision History

Top Strategic Technology Trends for 2023: AI Trust, Risk and Security Management - 17 October 2022

---

## Recommended by the Authors

Some documents may not be available as part of your current Gartner subscription.

Microsoft Azure OpenAI Service vs. OpenAI: Comparing GenAI Trust, Risk and Security

What Executives Need to Do to Support the Responsible Use of AI

Quick Answer: How to Make Microsoft 365 Copilot Enterprise-Ready From a Security and Risk Perspective

4 Ways Generative AI Will Impact CISOs and Their Teams

Market Guide for AI Trust, Risk and Security Management

---