



State of AI trust in 2026: Shifting to the agentic era

Findings from McKinsey's 2026 AI Trust Maturity Survey reveal progress in trust maturity, alongside persistent gaps in strategy, governance, and risk management.



By [Gabriel Morgan Asaftei](#)

Develops applications powered by artificial intelligence and machine learning for real estate, retail, and financial services companies



By [Roger Roberts](#)

Advises clients in a broad range of industries as they address their toughest technology challenges—and make the most of emerging opportunities



By [Abby Sticha](#)

Advises clients in the banking and technology, media, and telecommunications industries on AI and agentic AI strategy, solution development and building, and AI trust



By [Cécile Prinsen](#)

Coleads our McKinsey AI Trust service line in Europe, the Middle East, and Africa, and leads our work in data and technology risk with financial institutions

March 25, 2026 - AI adoption is accelerating rapidly, with organizations moving beyond experimentation toward scaled deployment of gen AI and, increasingly, agentic AI across core business functions. But as AI systems take on greater autonomy—making recommendations, triggering actions, and interacting with other systems—the consequences of failure grow materially. In this environment, AI trust and the responsible AI (RAI) practices that enable trust are no longer a tangential concern but a foundational requirement for realizing the full potential of the technology.

Trust underpins two critical outcomes. First, it enables organizations to realize value from AI investments by supporting sustained adoption and integration into core workflows. Second, it is essential if organizations are to successfully manage an expanding and evolving risk landscape. In the age of agentic AI, organizations can no longer concern themselves only with AI systems *saying* the wrong thing; they must also contend with systems *doing* the wrong thing, such as taking unintended actions, misusing tools, or operating beyond appropriate guardrails.

To understand how organizations are responding to this shift, McKinsey conducted the 2026 AI Trust Maturity Survey. Taken between December 2025 and January 2026, the survey gathered responses from approximately 500 organizations across industries and regions, with respondents who hold direct responsibility or expertise in AI governance, risk management, or AI investment decisions. Their responses were assessed using the McKinsey AI Trust Maturity Model, a framework based on five dimensions of RAI: strategy, risk management, data and technology, governance, and, new this year, agentic AI governance and controls, reflecting the growing importance of governing increasingly autonomous AI systems (Exhibit 1). We assess RAI maturity across four levels, from the development of foundational RAI practices to the implementation of a comprehensive and proactive program.

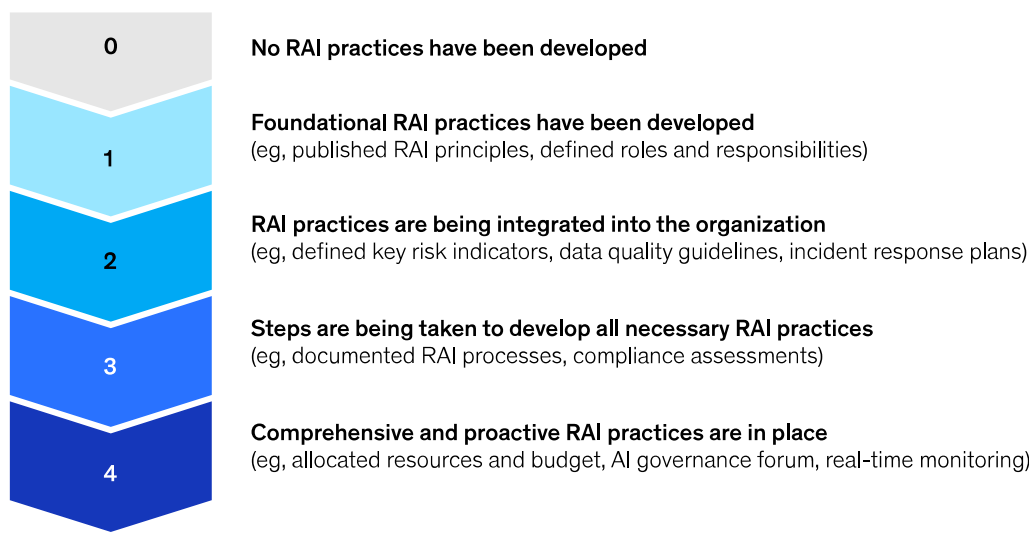
Exhibit 1

The AI Trust Maturity Model is a responsible AI framework that encompasses five dimensions.

Dimensions of responsible AI (RAI)

Strategy	Risk management	Data and technology	Governance	AI agent governance and controls
Defined vision and principles for AI or agentic AI governance aligned with client's values, strategic goals, and regulatory standards	Proactive monitoring and steps to reduce risks (incl model risk management, third-party solution auditing) through review and challenge (eg, red teaming, war gaming)	At-scale checks to instill confidence in data quality, which underlie AI solutions and ensure readiness of supporting AI governance tooling and development protocols	Talent, governance, and team structure in place to maintain AI governance across all client activities (incl processes to identify, assess, and mitigate AI risks); plan for AI-led shifts in skills, training programs, change management, and communication	Structures, oversight models, and technical guardrails to ensure that AI agents operate within defined boundaries (including agent registration and documentation, testing and evaluation of agent behavior and tool use, defined human-in-the-loop oversight, and real-time monitoring and controls)

Levels of RAI maturity



McKinsey & Company

At a glance: Key insights from the 2026 AI Trust Maturity Survey

The survey uncovers ten key insights about AI trust, which fall across three themes: the current state of AI trust, emerging risks and challenges, and how organizations are responding to close gaps and enable scale.

State of AI trust today

Despite improved maturity on the whole, significant variance still exists, but companies are increasingly recognizing the importance of RAI investment:

1. RAI maturity continues to improve, yet strategy, governance, and agentic AI controls lag behind, with only about 30 percent of organizations reaching a maturity level of three or higher in these dimensions.
2. RAI maturity varies by industry and region: Asia–Pacific leads globally, and technology, media, and telecommunications and financial services outperform other sectors.
3. Investment in RAI is strongly associated with higher RAI maturity and realized value.

Emerging risks and challenges

Increased adoption or consideration of adoption is leading to more risks and challenges as well as declined confidence in handling these risks:

4. Security and risk concerns are the top barrier to scaling agentic AI.
5. Inaccuracy and cybersecurity remain the most frequently cited AI risks as adoption expands.
6. Active mitigation lags behind risk awareness across nearly every AI risk category.
7. AI incident frequency remains stable, but confidence in organizational response has declined.

How enterprises and organizations are responding

As organizations implement RAI, they're facing barriers and differences in maturity while RAI is increasingly becoming a business enabler:

8. Knowledge and training gaps are the leading barrier to RAI implementation.
9. Organizations with explicit accountability for RAI achieve higher maturity scores than those without clear accountability.
10. AI trust is increasingly viewed as a business enabler rather than a compliance exercise.

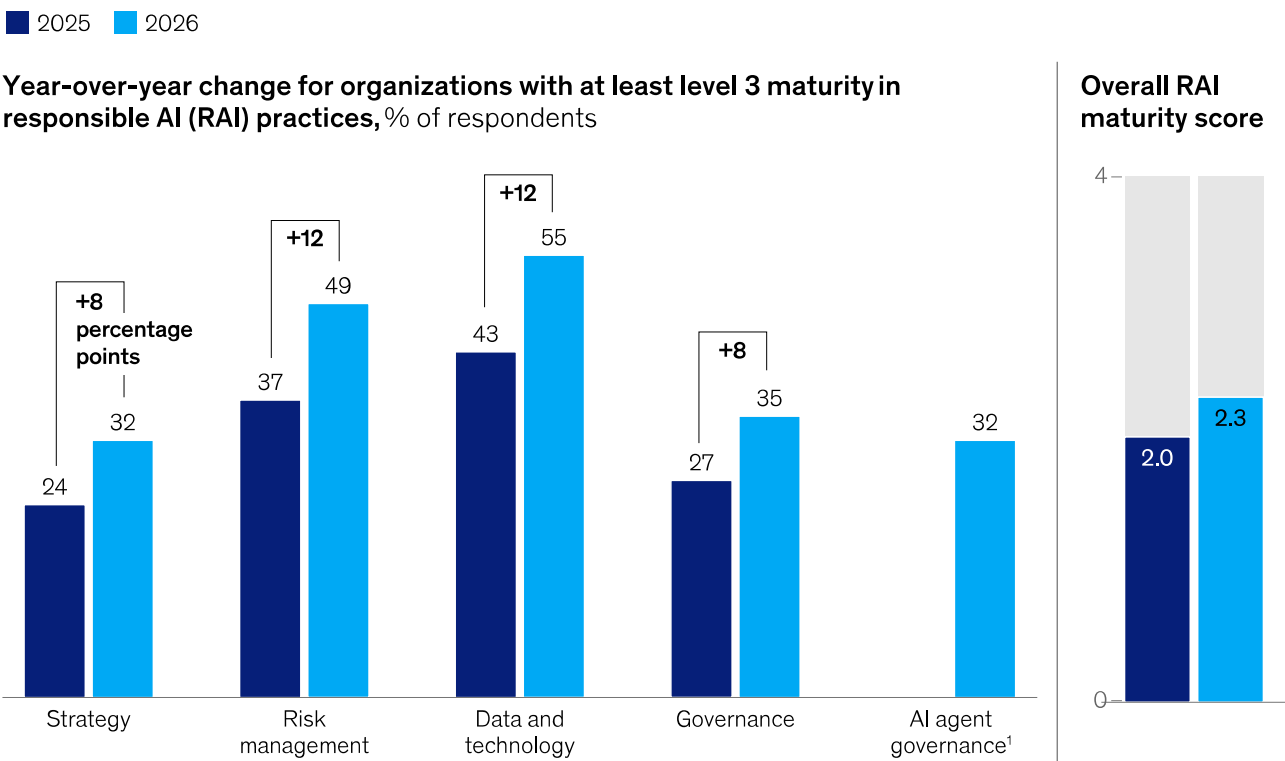
The state of AI trust today: Scrambling to keep up with the pace of change

The AI Trust Maturity Survey provides a detailed view into how organizations are approaching issues of trust and responsibility. Below, we examine the above ten insights in detail, looking at overall maturity levels, how progress differs across industries and regions, and how investment in RAI efforts relates to both trust capabilities and realized AI value.

Insight 1: RAI maturity continues to improve, yet strategy, governance, and agentic AI controls lag behind

The average RAI maturity score increased to 2.3 in 2026, up from [2.0 in 2025](#). However, only about one-third of organizations report maturity levels of three or higher in strategy, governance, and agentic AI governance (Exhibit 2). This imbalance suggests that while technical and risk management capabilities are advancing, organizational alignment and oversight structures are struggling to keep pace with the rapid expansion of AI use.

Responsible AI maturity is improving, but strategy, governance, and agentic AI governance and controls are lagging behind.



¹Section added in the 2026 AI Trust Maturity Survey (not included in 2025 survey).
Source: McKinsey AI Trust Maturity Survey, Dec 2024–Jan 2025 (n = 759); McKinsey AI Trust Maturity Survey, Dec 2025–Jan 2026 (n = 496)

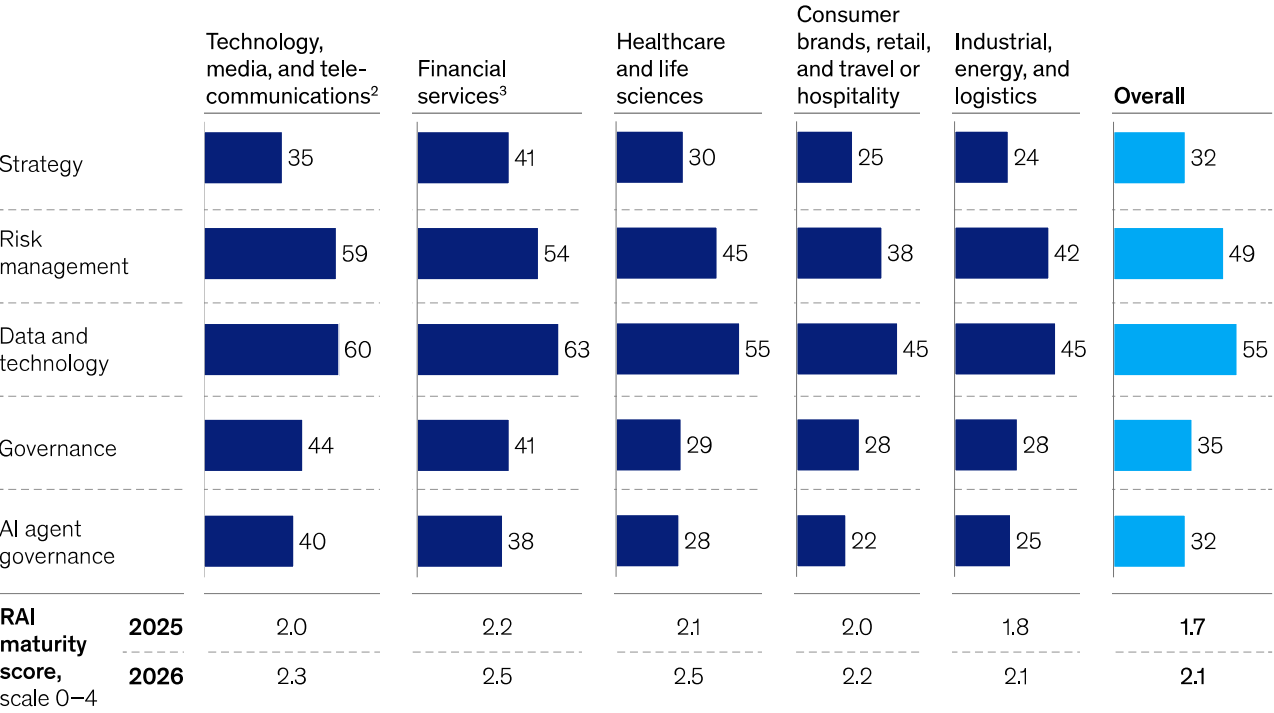
McKinsey & Company

Insight 2: RAI maturity varies by industry and region

Technology, media, and telecommunications and financial services continue to lead in RAI maturity, driven by stronger risk management and data foundations (Exhibit 3). Regionally, Asia–Pacific leads overall maturity, while governance and agentic AI controls lag behind data and technology across all regions, indicating a globally consistent governance gap (Exhibit 4).

Responsible AI maturity varies by industry, with technology, media, and telecommunications and telecommunications and financial services leading.

Share of respondents¹ with at least level 3 maturity in responsible AI (RAI) practices, %

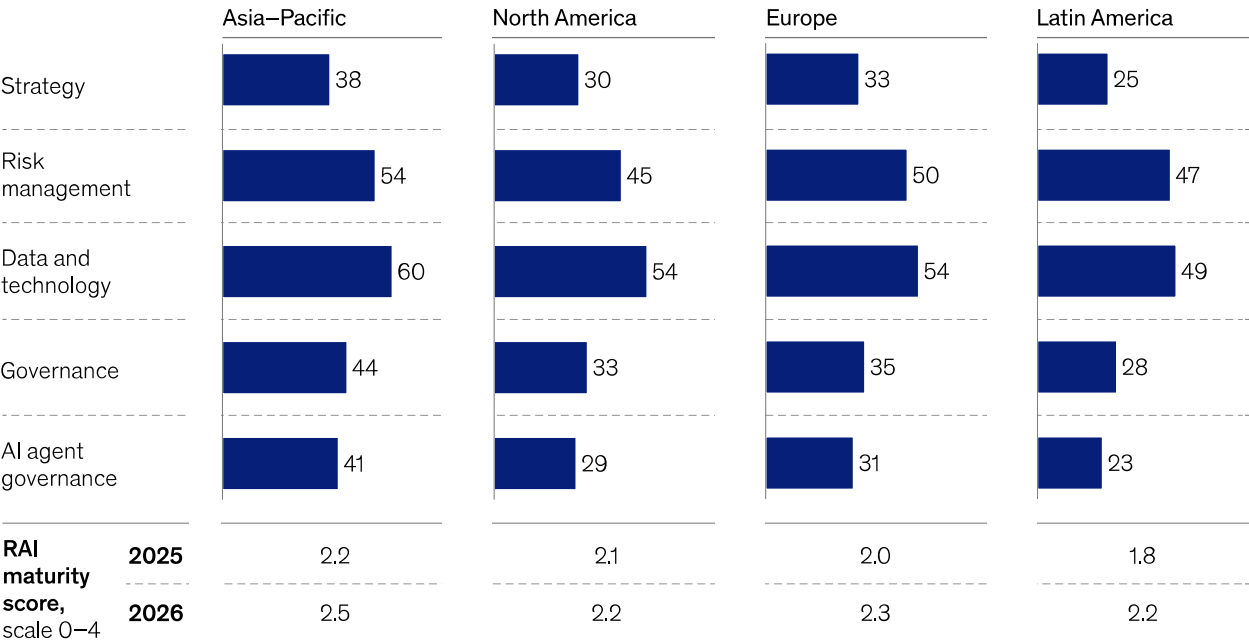


¹Technology, media, and telecommunications: n = 101; financial services: n = 146; healthcare and life sciences: n = 50; consumer brands, retail, and travel or hospitality: n = 100; industrial, energy, and logistics: n = 99. ²Includes semiconductors. ³Includes banking and payments and insurance and asset management. Source: McKinsey AI Trust Maturity Survey, Dec 2024–Jan 2025 (n = 759); McKinsey AI Trust Maturity Survey, Dec 2025–Jan 2026 (n = 496)

McKinsey & Company

Responsible AI maturity varies by region, with Asia–Pacific leading overall.

Share of respondents¹ with at least level 3 maturity in responsible AI (RAI) practices, %



¹Asia–Pacific: n = 95; Europe: n = 157; North America: n = 190; Latin America: n = 43.
Source: McKinsey AI Trust Maturity Survey, Dec 2024–Jan 2025 (n = 759); McKinsey AI Trust Maturity Survey, Dec 2025–Jan 2026 (n = 496)

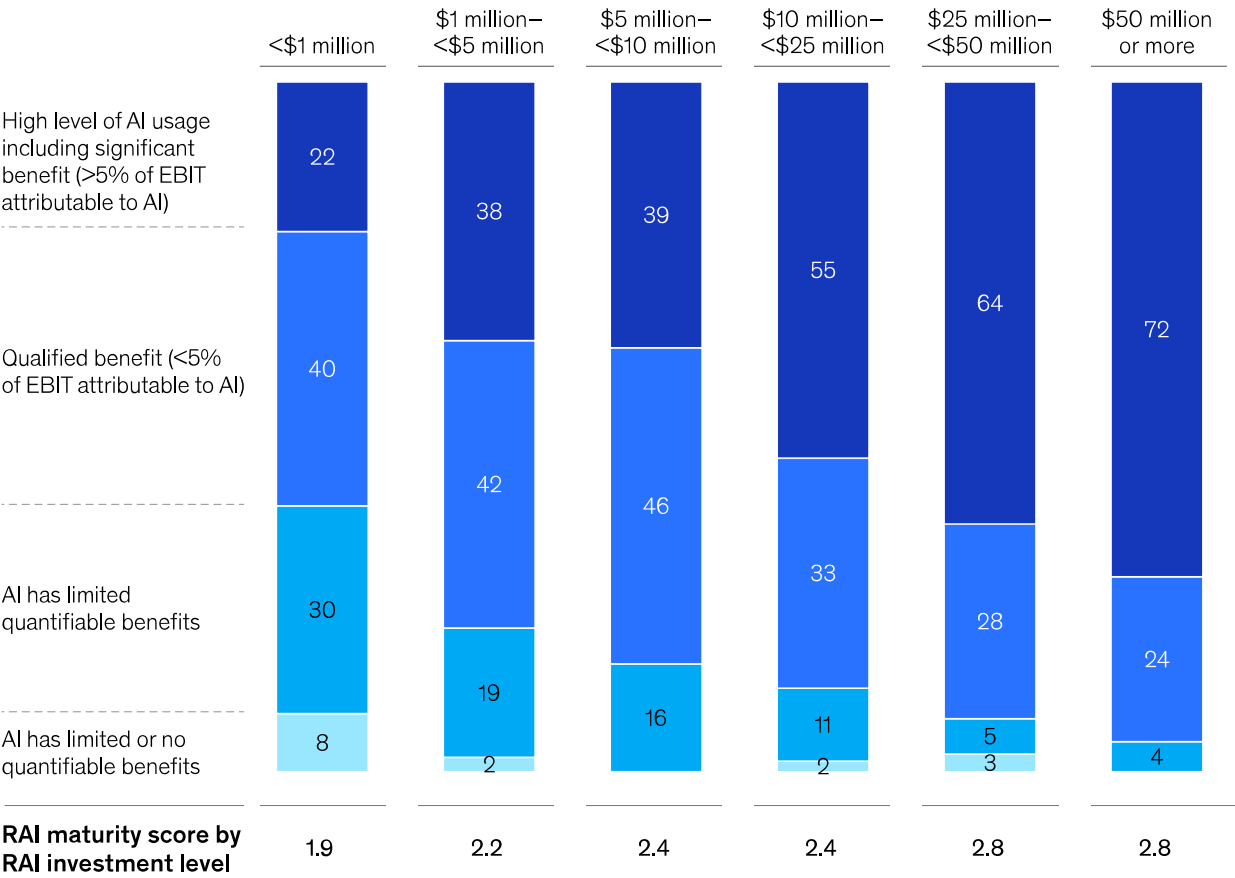
McKinsey & Company

Insight 3: Investment in RAI is strongly associated with higher RAI maturity and realized value

Organizations investing \$25 million or more into RAI initiatives report significantly higher maturity scores and are far more likely to realize material AI benefits, including EBIT impact above 5 percent (Exhibit 5). This relationship reinforces that RAI investment is not a tax on innovation but a key enabler of sustained value creation.

Investment in responsible AI is associated with greater realized AI value and higher responsible AI maturity.

AI benefits to the organization by level of investment in responsible AI (RAI),¹
% of respondents



Note: Figures may not sum to 100%, because of rounding.
¹Examples include hiring of RAI professionals; building or purchasing technical systems to comply with RAI principles; legal or professional services related to RAI; developing or integrating oversight, monitoring, and auditing tools; implementing workflow- or action-level safeguards and controls; conducting model evaluation, safety testing, or red-team exercises; upgrading security, access management, or data protection systems related to AI use; building and conducting training programs. "Unknown" RAI investment not shown (n = 21).
Source: McKinsey AI Trust Maturity Survey, Dec 2025–Jan 2026 (n = 496)

McKinsey & Company

Emerging risks: The key challenges organizations face today

The findings also explore the evolving risk landscape, including barriers to scaling agentic AI, the risks that organizations prioritize, and the disconnect between risk awareness, mitigation efforts, and incident preparedness.

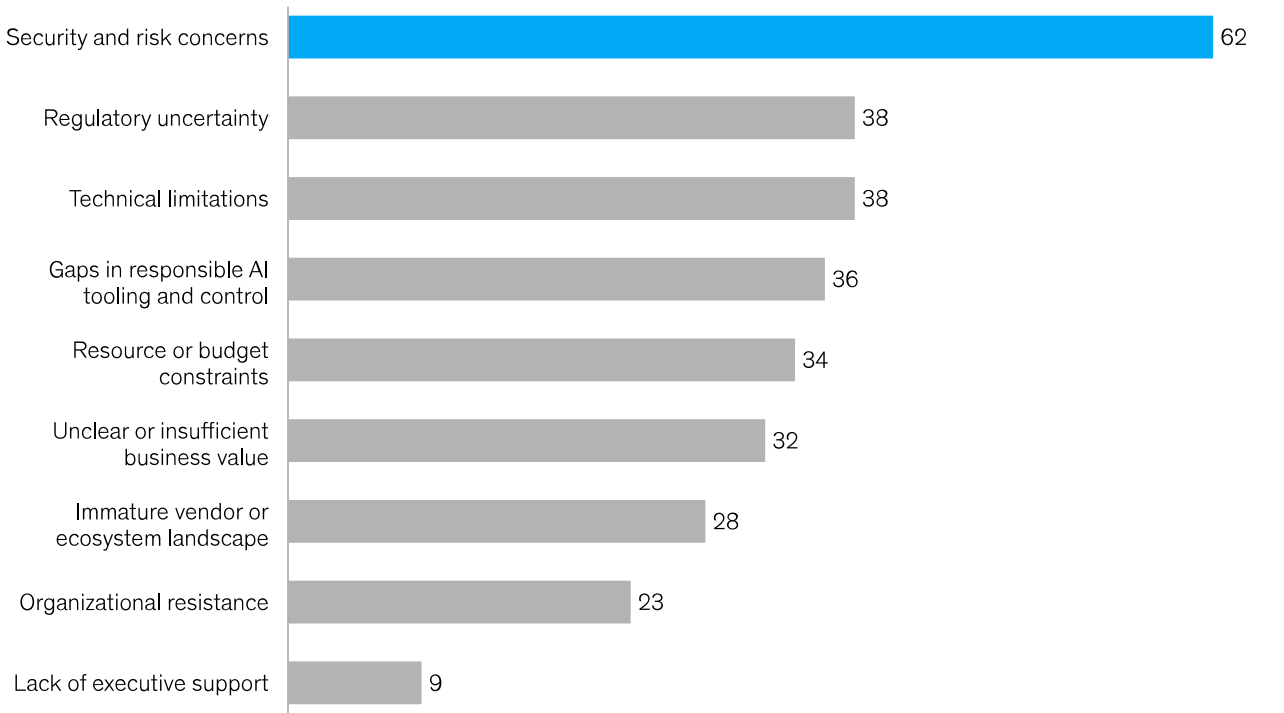
Insight 4: Security and risk concerns are the top barrier to scaling agentic AI

Nearly two-thirds of respondents cite security and risk concerns as the top barrier to fully scaling agentic AI, well ahead of regulatory uncertainty or technical limitations (Exhibit 6). This suggests that organizations are less constrained by experimentation capabilities and more by confidence in their ability to safely deploy autonomous systems at scale.

Exhibit 6

Security and risk concerns are the most frequently cited obstacle to fully scaling agentic AI.

Main obstacles preventing organizations from reaching fully scaled agentic AI,¹ % of respondents



¹"None" (1%) and "Other" (2%) answer shares are not shown. Respondents were asked to choose all that apply.
Source: McKinsey AI Trust Maturity Survey, Dec 2025–Jan 2026 (n = 496)

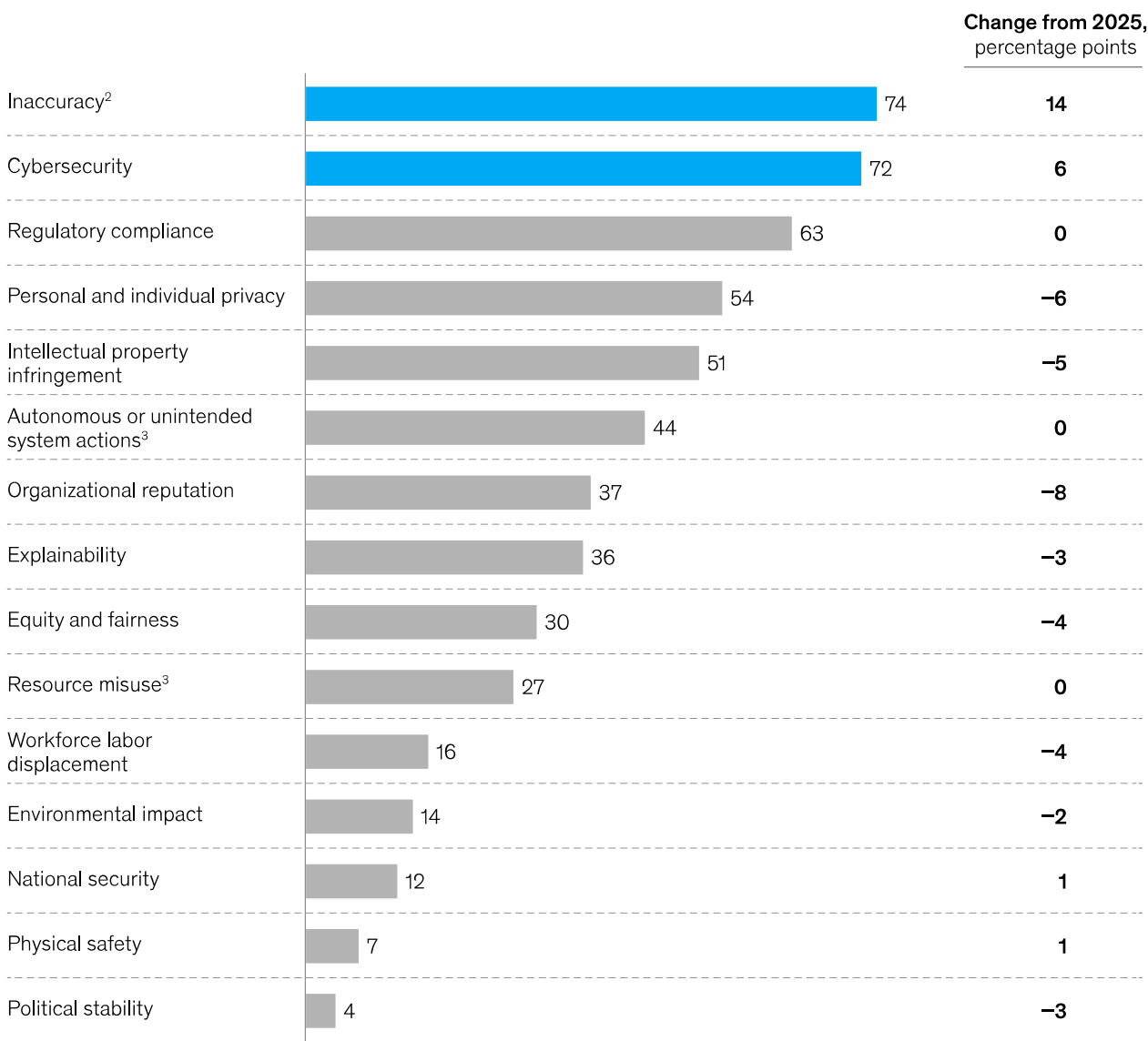
Insight 5: Inaccuracy and cybersecurity remain the most frequently cited AI risks as adoption expands

As AI adoption grows, 74 percent of respondents identify inaccuracy and 72 percent cite cybersecurity as highly relevant risks (Exhibit 7). These risks remain foundational concerns even as newer agentic risks emerge, highlighting that organizations must manage both traditional model risks and the expanded threat surface introduced by autonomy.

Exhibit 7

As AI adoption expands, inaccuracy and cybersecurity stand out as top-of-mind risks.

AI risks that organizations consider relevant and are working to mitigate,¹ % of respondents



¹"Other" (<1%) and "None" (<1%) answer shares are not shown. Respondents were asked to choose all that apply.
²Includes inconsistency and robustness failures.
³"Autonomous or unintended system actions" and "Resource misuse" were added in the 2026 AI Trust Maturity Survey (not included in 2025 survey).
Source: McKinsey AI Trust Maturity Survey, Dec 2024–Jan 2025 (n = 759); McKinsey AI Trust Maturity Survey, Dec 2025–Jan 2026 (n = 496)

McKinsey & Company

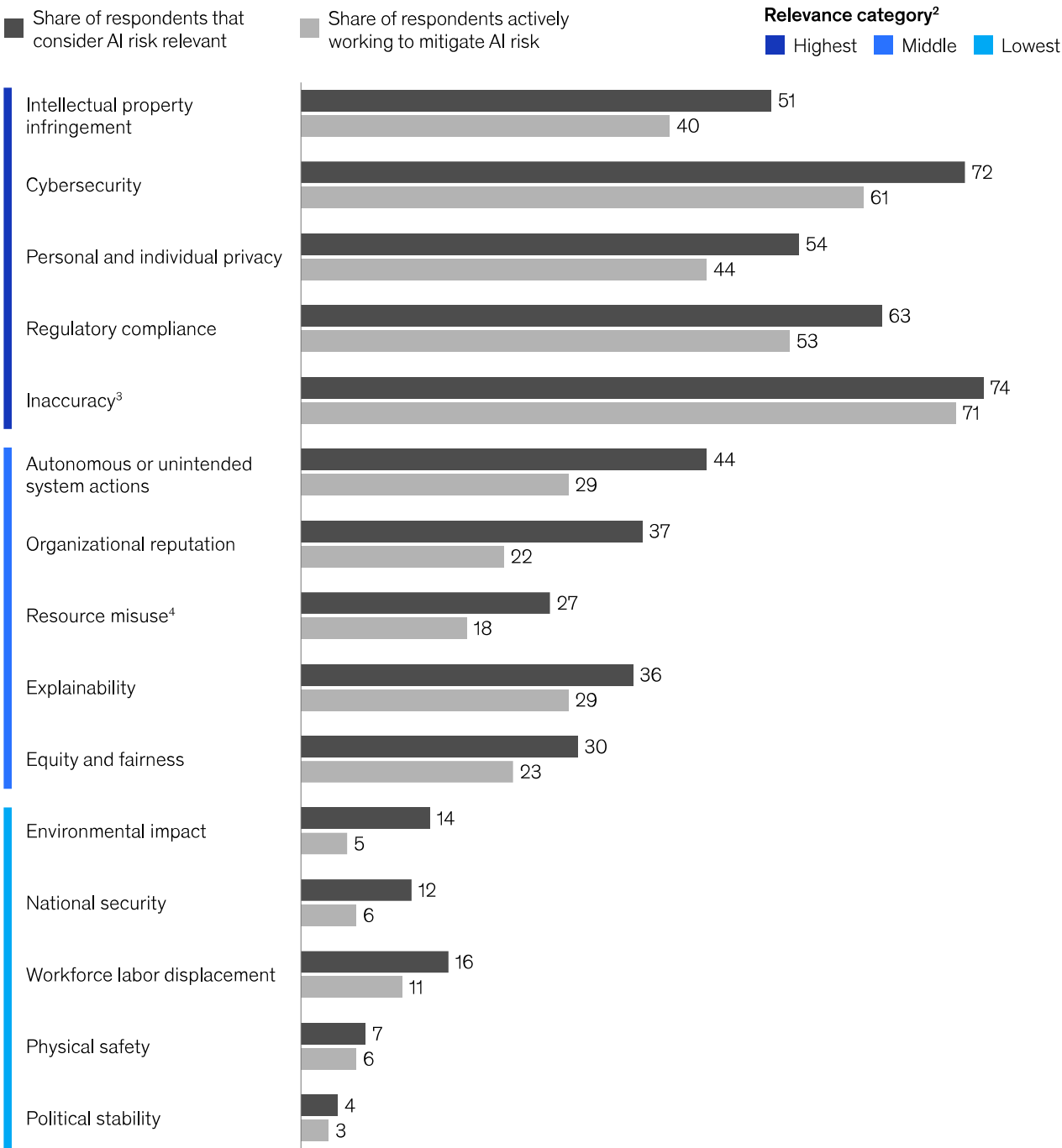
Insight 6: Active mitigation lags risk awareness across nearly every AI risk category

Across almost all risk types, respondents report a meaningful gap between the risks they consider relevant and those they are actively mitigating (Exhibit 8). This gap is especially pronounced for intellectual property infringement and personal privacy, suggesting that risk awareness is outpacing the implementation of controls, processes, and tooling needed to manage it effectively.

Exhibit 8

Active mitigation lags behind perceived relevance across nearly every risk category.

AI risk relevance and active work in mitigating risk,¹ % of respondents



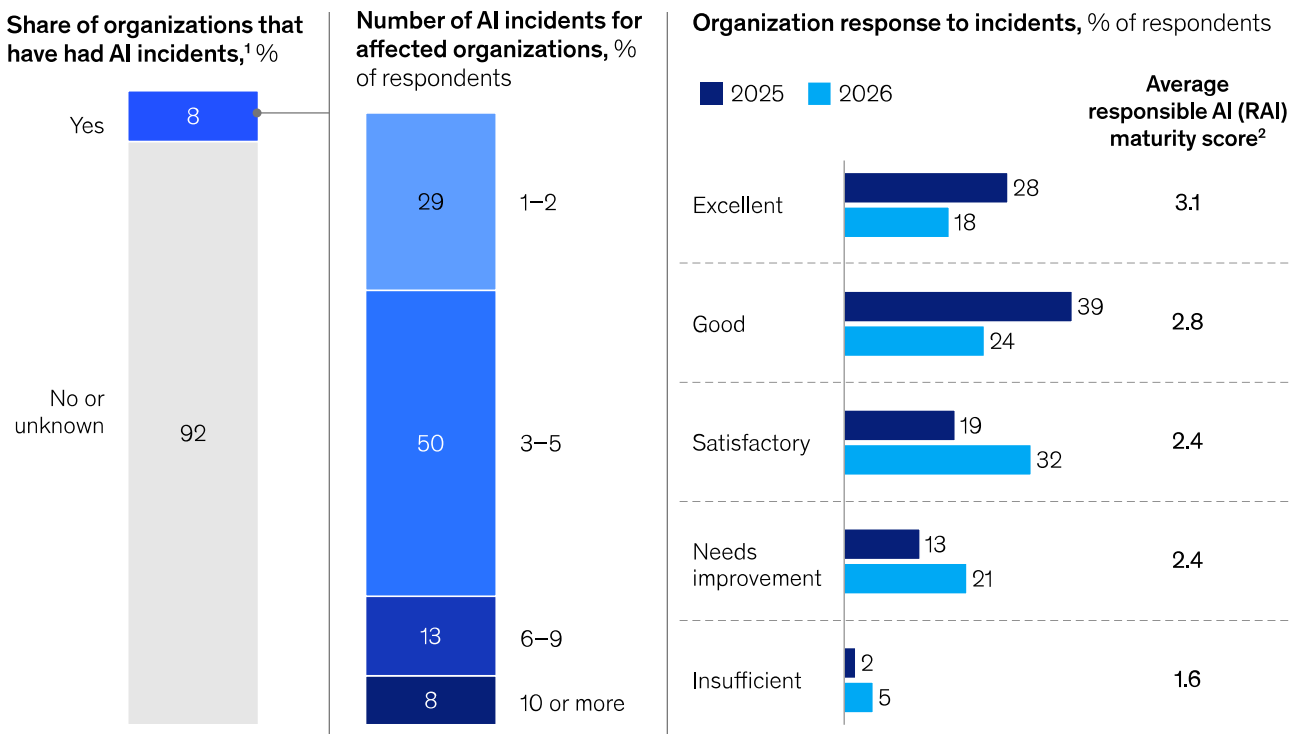
¹"Other" (1% or less) and "None" (1% or less) answer shares are not shown. Respondents were asked to choose all that apply. ²Highest-ranked risk by relevance is defined as the top 5 risks by share of respondents that consider AI risk relevant; mid-ranked risk by relevance is defined as the middle 5 risks by share of respondents that consider AI risk relevant; lowest-ranked risk by relevance is defined as the bottom 5 risks by share of respondents that consider AI risk relevant. ³Includes inconsistency and robustness failures. ⁴"Autonomous or unintended system actions" and "Resource misuse" were added in the 2026 AI Trust Maturity Survey (not included in 2025 survey).
Source: McKinsey AI Trust Maturity Survey, Dec 2025–Jan 2026 (n = 496)

Insight 7: AI incident frequency remains stable, but confidence in organizational response has declined

The share of organizations reporting AI-related incidents has remained steady at roughly 8 percent, but perceptions of incident response quality have deteriorated. Almost 60 percent of respondents who experienced incidents report satisfactory or negative views of their organization’s response, indicating that while incidents may not be increasing, preparedness and response capabilities are failing to keep pace with growing system complexity (Exhibit 9).

Exhibit 9

The number of AI incidents remained constant since 2025, but confidence in the response to such episodes has declined.



¹According to the OECD, an AI incident is an event, circumstance, or series of events in which the development, use, or malfunction of one or more AI systems directly or indirectly leads to any of the following harms: injury or harm to the health of a person or groups of people; disruption of the management and operation of critical infrastructure; violations of human rights or a breach of obligations under the applicable law intended to protect fundamental, labor, and intellectual property rights; and harm to property, communities, or the environment.
²Average maturity calculated by taking the average of levels across organizations.
Source: McKinsey AI Trust Maturity Survey, Dec 2024–Jan 2025 (n = 759); McKinsey AI Trust Maturity Survey, Dec 2025–Jan 2026 (n = 496)

How enterprises and organizations are responding to emerging risks and challenges

In response to emerging risks and challenges, organizations are working to strengthen the foundations of AI trust—closing capability gaps, clarifying accountability, and building the RAI capabilities needed to ensure trust accelerates innovation rather than constrain it.

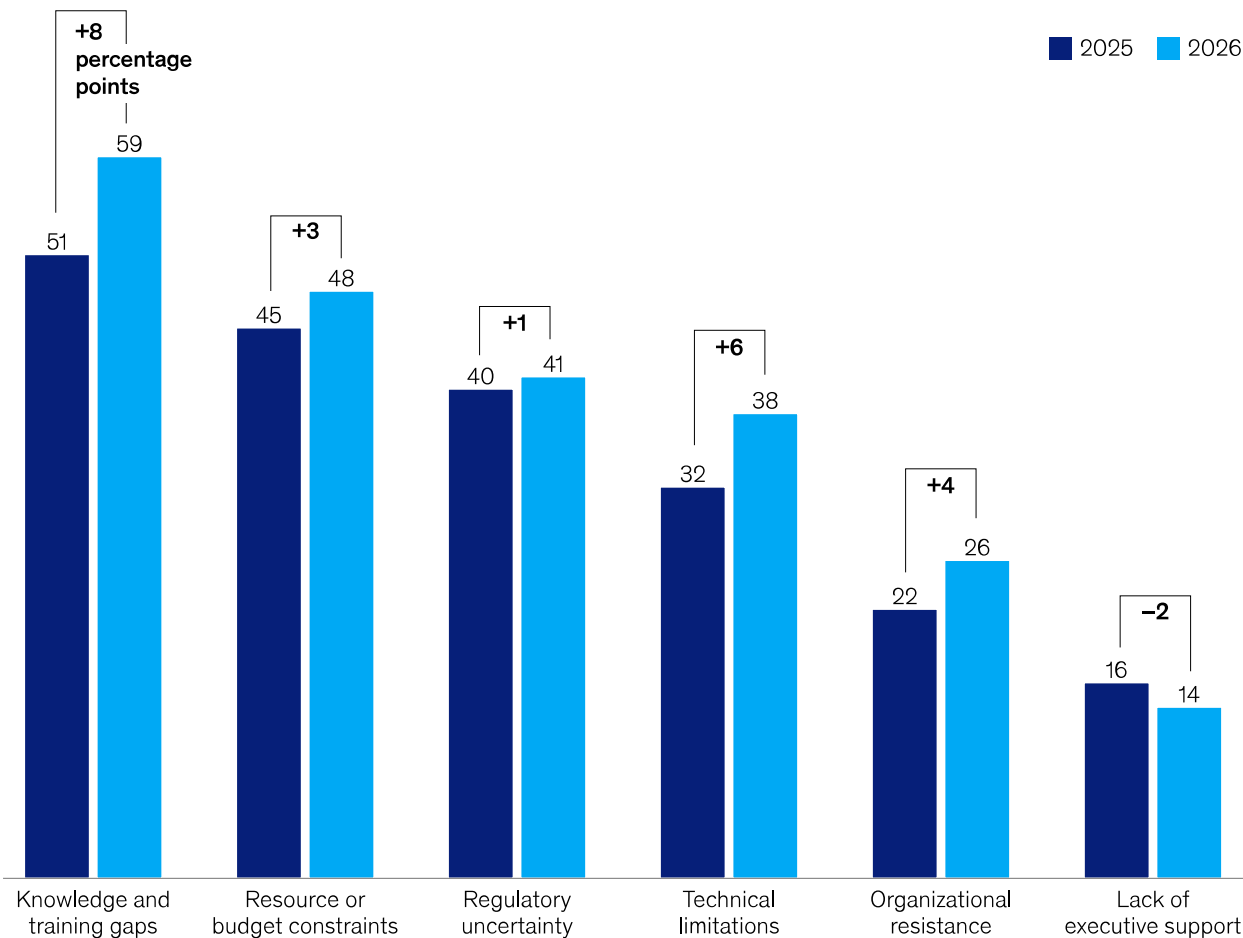
Insight 8: Knowledge and training gaps are the leading barrier to RAI implementation

Nearly 60 percent of respondents cite knowledge and training gaps as the primary barrier to implementing RAI practices, up from about 50 percent last year (Exhibit 10). While executive support has improved, the data suggests that organizations continue to struggle with building the skills, awareness, and operational muscle required to embed RAI consistently across teams.

Exhibit 10

Knowledge and training gaps and resource and budget constraints are the largest barriers to implementing responsible AI measures.

Main obstacles preventing organizations from implementing responsible AI measures,¹
% of respondents



¹"None" (2%) and "Unknown" (2%) answer shares not shown. Respondents were asked to choose all that apply.
Source: McKinsey AI Trust Maturity Survey, Dec 2024–Jan 2025 (n = 759); McKinsey AI Trust Maturity Survey, Dec 2025–Jan 2026 (n = 496)

McKinsey & Company

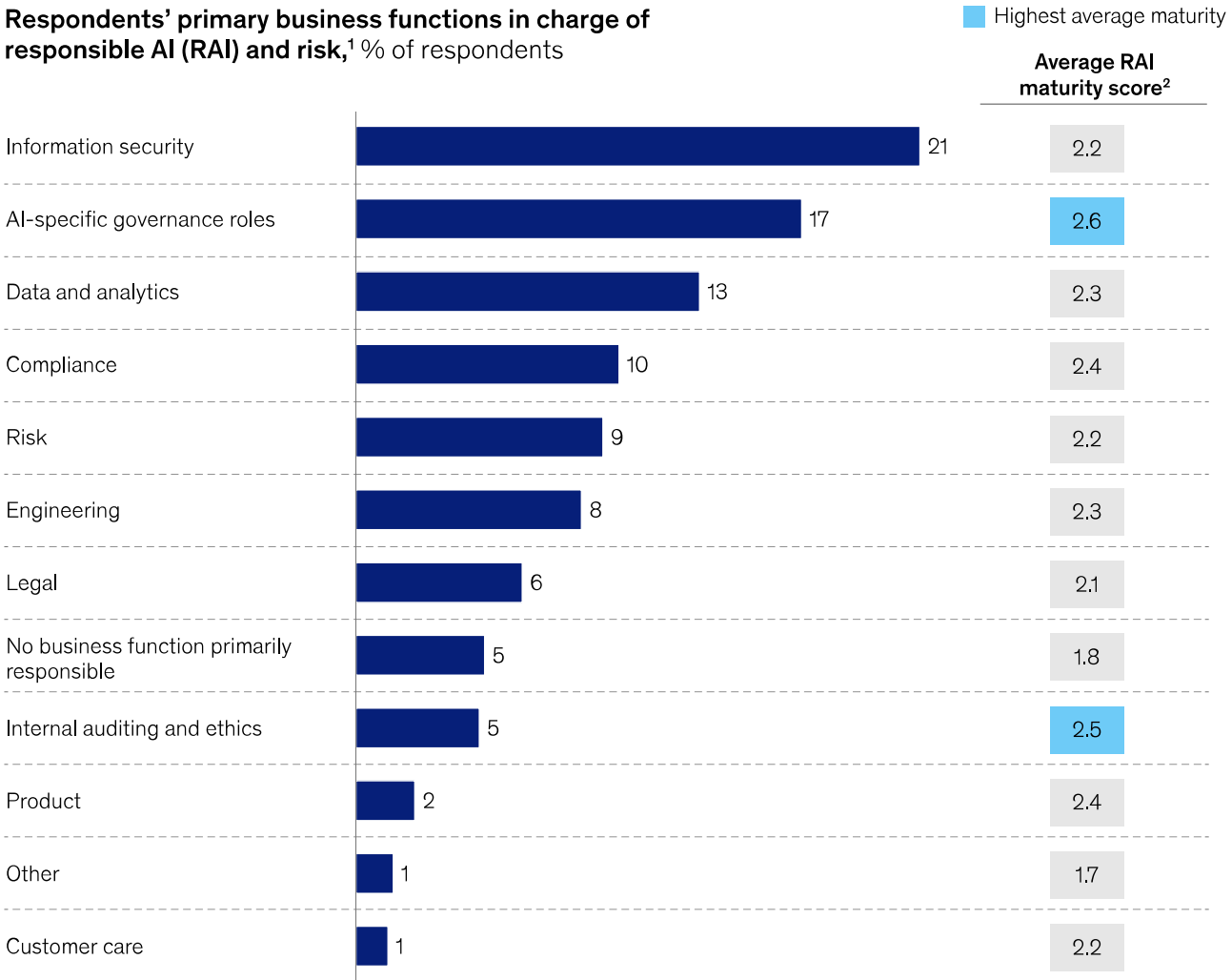
Insight 9: Organizations with explicit ownership for RAI have higher maturity than those without clear accountability

Organizations that assign clear ownership for RAI—particularly through AI-specific governance roles or internal audit and ethics teams—exhibit the highest average maturity levels, with an average score of 2.6. In contrast, organizations without a clearly accountable function lag behind materially

(scoring an average of just 1.8), reinforcing the importance of explicit ownership and decision rights (Exhibit 11).

Exhibit 11

Explicit accountability for responsible AI leads to a greater level of responsible AI maturity.



¹"Unknown" responses not included (n = 9).
²Average maturity calculated by taking the average of levels across organizations.
Source: McKinsey AI Trust Maturity Survey, Dec 2025–Jan 2026 (n = 496)

Insight 10: AI trust is increasingly viewed as a business enabler rather than a compliance exercise

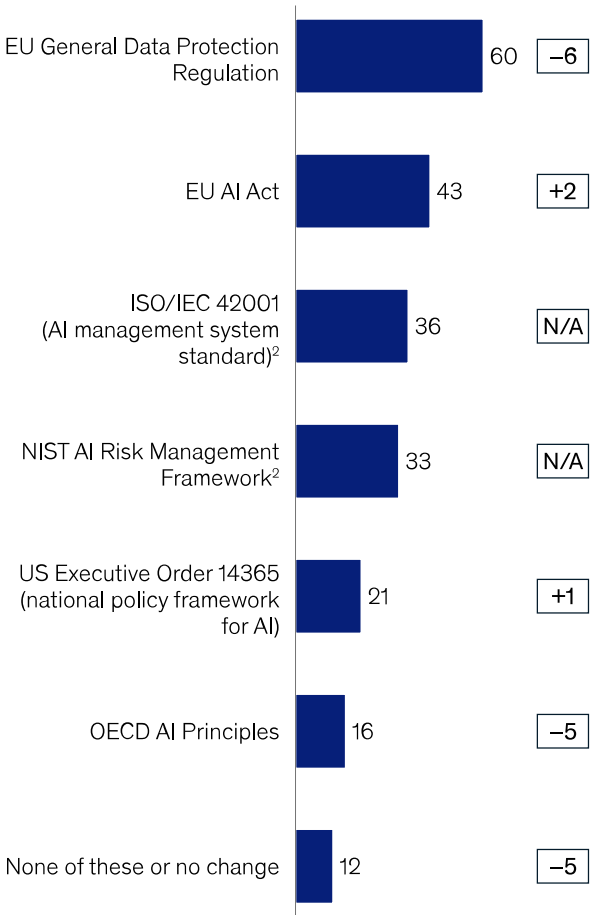
Respondents report improvements in business outcomes, operational efficiency, and customer trust more frequently than negative outcomes (Exhibit 12). At the same time, the perceived influence of some regulatory frameworks has declined, suggesting a shift from compliance-led motivation toward value- and performance-driven adoption of AI trust.

Exhibit 12

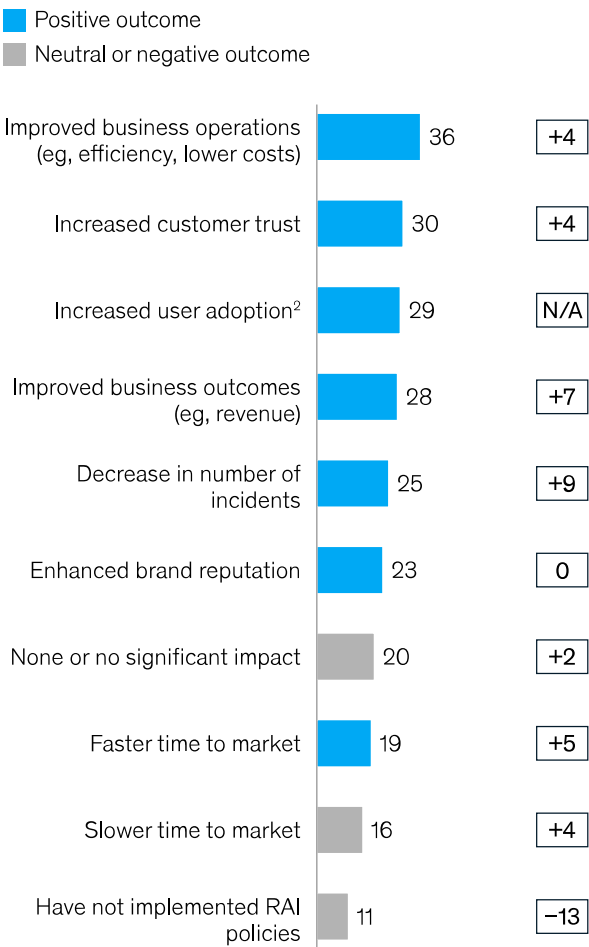
AI trust is increasingly viewed as a business enabler.

xx Change 2025–26, percentage points

Share of organizations that are influenced by specific AI regulations when making responsible AI (RAI) decisions,¹ % of respondents



Reported organizational results from implementing RAI policies, % of respondents



¹“Other” (4%) answer share not shown. Respondents were asked to choose all that apply.
²Added in the 2026 AI Trust Maturity Survey (not included in 2025 survey).
Source: McKinsey AI Trust Maturity Survey, Dec 2024–Jan 2025 (n = 759); McKinsey AI Trust Maturity Survey, Dec 2025–Jan 2026 (n = 496)

Looking ahead: Trust as the enabler of scale

As AI systems become more autonomous and embedded in critical workflows, gaps in governance and risk management will become increasingly costly. Organizations that fail to establish clear accountability, robust controls, and effective monitoring mechanisms risk slower adoption, higher incident impact, and diminished stakeholder trust.

Conversely, organizations that treat AI trust as a core business capability, rather than as a compliance requirement, are better positioned to scale AI adoption to its full potential. These capabilities are not something that you can buy and install but rather require a concerted combination of policies, processes, people, and technology to build the foundations needed for [agents, robots, and people to work together in new ways](#). Building a trustworthy innovation engine early will define those who capture long-term value from AI in the agentic era.

[Gabriel Morgan Asaftei](#) is a partner in McKinsey's New York office, where **Abby Sticha** is a consultant; [Roger Roberts](#) is a partner in the Bay Area office; and **Cécile Prinsen** is an associate partner in the London office.

Ushering in a new era of trusted AI

Artificial intelligence will compel organizations to elevate their compliance. Six levers hold the key.



By [Henning Soller](#)

Leads work on technology architecture and large-scale IT, data, and AI transformations across public- and private-sector institutions. Advises governments and enterprises on building scalable, sovereign-grade digital and AI platforms to enable resilient, future-ready ecosystems



By [Anselm Ohme](#)

Serves clients across industries on AI, tech regulation, and cybersecurity transformations



By [Fares Darwazeh](#)

An alumnus of McKinsey's Riyadh office



By [Thao Dürschlag](#)

Serves clients in the banking sector on AI transformations, focusing on governance and operating model

March 30, 2026 - Each week seems to bring another announcement about new AI tools with the potential to reshape how organizations function. This progress has been accompanied by the growing recognition of the importance of [trusted AI](#), defined by the OECD as corresponding solutions that are reliable, transparent, fair, resilient, and accountable.¹ The concept of trusted AI brings together technical robustness and security with ethical and social responsibility.

The [rapid advancement and widespread adoption of AI](#) have made [data privacy and AI regulation](#) critical topics on the global regulatory agenda. Yet the implementation of these guidelines and the regulatory landscape itself have become increasingly complex over the past several years.

Despite the higher regulatory burden, many organizations continue to depend on manual processes to implement regulatory and AI-related controls, adding layers and cost. As McKinsey's 2025 analysis of regulatory technology (RegTech) shows, financial institutions relying on manual compliance systems often fulfill only a fraction of their obligations, leaving them at higher risk of penalties and operational inefficiencies.

There is a high degree of difficulty in striking the right balance in compliance. For example, a US-based bank's legacy system met just 75 percent of requirements before the adoption of an [automated RegTech solution](#) streamlined data mapping, raising compliance to above 95 percent.

Given these competing challenges, companies should consider viewing compliance not as a cost but as an enabler of AI scale. This approach entails adopting a more streamlined, scalable, and digital approach to regulatory adherence. Once organizations overcome some common obstacles in the evolving regulatory environment, six levers drawn from best practices can help

organizations elevate their AI compliance efforts. Companies that get it right can not only unlock faster AI deployment but also reduce their risk.

Understanding the challenges

The widespread adoption of AI chatbots and voice bots, augmenting human behavior with human voice, the rise of personalized, tailored marketing, and rapid advancements in generative AI in new products and services offered to customers have prompted regulators to establish clear guardrails, with the aim of providing customers and employees with greater control over their data and preventing fraud. Regulations such as the General Data Protection Regulation (GDPR) and the European Union AI Act (as part of the EU Data Governance Act), as well as cybersecurity frameworks such as the Digital Operational Resilience Act (DORA), have emerged as key pillars. Similar initiatives have been introduced globally, including in Brazil, Thailand, and Türkiye, as well as by individual US states, such as the California Consumer Privacy Act (CCPA).

Despite the efforts of policymakers to provide clear guidance on compliance, the relationship between various regulations remains confusing. The result is that no universal standard exists to guide companies in simultaneously managing data privacy, verifying trusted AI, and promoting resilience. Conflicting requirements across jurisdictions further complicate compliance; even within regions such as Europe, interpretations of regulations such as the GDPR vary significantly between countries.

Organizations have tended to respond to these regulatory demands in an ad hoc manner. Often, impending compliance deadlines spur companies to reactively implement transparency measures and controls. For example, some organizations have deployed surveys and interviews to compile process repositories. This fragmented and inconsistent documentation undermines the organization's ability to maintain accurate, updated compliance records and controls as well as alignment across business units. Similarly, controls—such as those governing the use of personal data or adherence to the principles of the EU AI Act²—are frequently implemented as manual checks. Last, companies can

fail to assign responsibility for compliance, which can end up distributed across multiple departments, such as legal, IT, and compliance.

Adopting sustainable solutions for digital and AI compliance

To consistently achieve digital and AI compliance with emerging regulations, organizations should consider six levers.

1. Centralize compliance within existing teams

Compliance efforts should be incorporated into established processes and overseen by teams with a holistic view of the organization. For instance, data privacy teams could also take responsibility for compliance with the EU AI Act, or teams in charge of complying with ISO 9001 (a standard for quality management systems) could oversee the creation of process repositories. Centralizing oversight can eliminate duplicate and inconsistent documentation—for example, separate process representations for ISO 9001 and GDPR—thereby improving efficiency and reducing complexity. Such approaches are increasingly visible across industries, especially in banking, where we have seen several major institutions introduce centers of excellence that orchestrate data and AI responsibility among IT, business, and risk functions.

While execution necessarily spans legal, IT, data, and business functions, organizations should consider anchoring AI compliance accountability within the risk or compliance function, given its mandate for enterprise-wide control, regulatory interpretation, and independent oversight.

2. Digitalize controls through the use of legal tech tools

Many organizations are in the early stages of adopting [legal tech tools](#), with factors such as cultural resilience, data security concerns, and skills gaps impeding integration. Still, these tools hold significant potential. As one example, generative AI tools can automate policy adherence checks, flagging exceptions for further review. Tools that identify personally identifiable information (PII) can reduce the time and errors associated with manual processes. A comprehensive review of existing controls, with a focus on automating repetitive tasks such as data retention management, can not only streamline compliance efforts but also enhance accuracy and consistency.

3. Simplify overarching AI governance structures

Many organizations have overly complex governance setups, which hinder compliance efforts. This is especially true for multinational corporations with numerous subsidiaries, joint ventures, and regional entities. Simplifying policy frameworks across legal entities and consolidating oversight under a single chief compliance officer can dramatically reduce complexity and facilitate data transfer and compliance checks. Organizations should consider moving toward a federated model in which the risk or compliance function provides enterprise-level governance and challenge, while legal, IT, data, and business teams own implementation across the AI life cycle.

Regulators themselves are increasingly acknowledging the need for simplification and encouraging companies to streamline their approaches to compliance. For instance, the European Commission's Digital Omnibus initiative explicitly aims to reduce the administrative burden and overlapping requirements across digital, AI, and data laws, creating a more coherent and predictable regulatory landscape. Similarly, international policy bodies such as the G7 and

OECD have emphasized the importance of harmonizing AI and data governance frameworks globally to minimize fragmentation and ease cross-border compliance.

4. Design change management processes for all employees

To facilitate the integration of compliance into daily operations, organizations would need to prioritize change management for their workforce. Targeted training and awareness programs for employees can reinforce the importance of data privacy, ethical AI, and regulatory compliance in their daily roles. Tailored, role-specific guidance for functions such as IT, legal, and operations supports alignment with compliance requirements across the organization. For instance, we have seen a bank achieve a successful technological transformation by dedicating 34 percent of its investment to structured change management efforts.

In addition, fostering a culture of accountability and ethical decision-making, supported by a clear commitment and communication from leadership, helps weave these principles into the organizational fabric. Empowering employees to understand and embrace these responsibilities creates a workforce that actively supports compliance and resilience efforts.

5. Build resilient and ethical AI systems across the value chain

To ensure both resilience and ethical behavior, organizations must go beyond compliance controls to focus on the technical foundations of their AI systems. This lever includes designing technology architectures that are robust, scalable, and capable of withstanding disruptions such as cyberattacks, system failures, or regulatory changes.

Equally important is the development of ethical AI systems, which requires organizations to implement frameworks that promote transparency and fairness while mitigating bias. Such frameworks include the OECD AI Principles and the IEEE P7001 transparency standard. Critical components of this effort are regular audits of AI models, explainability mechanisms, and adherence to ethical guidelines throughout the AI life cycle. By addressing these technical aspects, organizations can build AI systems that both meet regulatory requirements and earn the trust of customers, employees, and other stakeholders.

6. Undertake continuous monitoring and improvement

Compliance is not a static goal but an ongoing process that requires continuous monitoring and improvement. Organizations can use AI tools that can proactively identify and address potential risks in real time. Establishing feedback loops with employees, customers, and other stakeholders provides valuable insights into the effectiveness of compliance practices and opportunities for refinement. Periodic reviews of policies, controls, and governance structures help them keep pace with evolving regulations and business needs to maintain operational resilience over time.

Next steps for implementation

To fully harness these levers, organizations could take the following steps:

- *Assess the current state.* Conduct a thorough benchmarking exercise (such as [McKinsey's AI Trust Maturity Model](#)) to compare the organization's current setup against industry best practices. A full evaluation of the organizational structure, policy frameworks, and control mechanisms can highlight gaps between an organization's AI impact aspirations and its AI trust and governance readiness.

- *Identify automation opportunities.* Collect and review all relevant controls, along with their current implementation methods, to flag opportunities for automation. Organizations should prioritize controls that are time-intensive, error-prone, or repetitive. Candidates for automation are processes that occur on a defined schedule (such as monthly reviews, quarterly audits, or continuous monitoring) rather than one-off checks.
 - *Develop a unified program.* Translate insights from the benchmarking and automation assessments into a comprehensive program, with a central oversight team in the lead. This program should focus on simplifying governance structures, digitalizing controls, and standardizing compliance efforts across the organization.
-

The era of trusted AI demands a fundamental shift in how organizations approach compliance. The six levers offer a path for organizations to transition from reactive, manual compliance processes to a proactive, digital-first approach that reflects the demands of the new regulatory landscape. These changes can also reduce the burden of compliance and position organizations to thrive in an increasingly regulated digital environment.

Henning Soller is a partner in McKinsey's Frankfurt office, **Anselm Ohme** is a consultant in the Berlin office, **Fares Darwazeh** is an alumnus of the Riyadh office, and **Thao Dürschlag** is an associate partner in the Munich office.

¹ "AI principles," OECD, accessed February 20, 2026.

² The act sets requirements for all companies that use, deploy, or integrate AI systems. The obligations center on governance, accountability, and oversight: ensuring the system's performance, addressing bias and unfairness, maintaining human oversight, and being able to demonstrate compliance.