White Paper, sponsored by Microsoft | December 2024

FROM RISK TO REWARD:

The Business Case for **Responsible AI**

Ritu Jyoti
Group Vice President/General Manager,
Worldwide Artificial Intelligence, Automation,
Data and Analytics Research Practice, IDC

Dave Schubmehl
Research Vice President,
Conversational Artificial Intelligence and
Intelligent Knowledge Discovery, IDC

Table of Contents

CLICK ANY HEADING TO NAVIGATE DIRECTLY TO THAT PAGE.

Executive Summary . . . 3

Introduction . . . 4

Key Findings from the Survey . . . 7

Responsible AI Tooling . . . 11

Additional Insights from the Study . . . 18

AI Adoption . . . 18

Important Use Cases . . . 20

Advice and Recommendations . . . 22

Conclusion . . . 28

Definitions . . . 30

Generative AI . . . 30

Responsible AI . . . 31

Responsible AI Attributes . . . 31

Appendix 1: Supplemental Data . . . 33

About the IDC Analysts . . . 37

Message from the Sponsor . . . 38

Executive Summary

Every organization needs to be responsible at the core in the AI era as it helps the organization accelerate realization of the benefits of AI. A responsible-at-the-core organization has the following foundational elements:

- **Core values and governance:** It defines and articulates responsible AI (RAI) mission and principles, supported by the C-suite, while establishing a clear governance structure across the organization that builds confidence and trust in AI technologies.
- **Risk management and compliance:** It strengthens compliance with stated principles and current laws and regulations while monitoring future ones and develops policies to mitigate risk and operationalize those policies through a risk management framework with regular reporting and monitoring.
- **Technologies:** It uses tools and techniques to support principles such as fairness, explainability, robustness, accountability, and privacy and builds these into AI systems and platforms.
- **Workforce:** It empowers leadership to elevate RAI as a critical business imperative and provides all employees with training to give them a clear understanding of responsible AI principles and how to translate these into actions. Training the broader workforce is paramount for ensuring RAI adoption.

The purpose of this paper is to provide information and evidence that a responsible AI approach fosters innovation by aligning AI deployment with organizational standards and societal expectations, resulting in sustainable value for organizations and their customers.

Introduction

According to IDC's February 2024 *Worldwide Semiannual Artificial Intelligence Systems Spending Guide, Version 1,* which tracks AI software, hardware, and services across industries and use cases, enterprises worldwide are expected to invest $232 billion on AI solutions in 2024.

AI solutions are transforming a diverse range of industries, from finance and manufacturing to agriculture and healthcare, by enhancing operations and reshaping the nature of work. Enterprises' application of generative AI (GenAI), which is rapidly unfolding, can revolutionize customer experiences, boost employee productivity, enhance creativity and content creation, and accelerate process optimization.

However, AI also creates real risks and unintended consequences. AI systems can inadvertently perpetuate or amplify societal biases due to biased training data or algorithmic design. AI systems are often trained on large amounts of data collected from various sources. AI program outputs may run into copyright infringement concerns. AI hallucinations are incorrect or misleading results that AI models generate. These errors can be caused by a variety of factors, including insufficient training data, incorrect assumptions made by the model, lack of context, or biases in the data used to train the model. So lack of grounding can cause the model to generate outputs that, while seemingly plausible, are factually incorrect, irrelevant, or nonsensical and further deplete trust.

As AI technologies become increasingly sophisticated, the security risks associated with their use and the potential for misuse also increase. For example, hackers/bad actors can control GenAI foundation model output by poisoning the grounding data. Or they could use prompt injection attacks that disguise malicious instructions as user inputs, tricking the large language model (LLM) into overriding developer instructions with the goal of manipulating the model to produce a desired response. Jailbreaking, a technique that attempts to bypass or subvert the safety filters and restrictions built into LLMs, is also popular with the bad actors.

According to IDC's March 2024 *Microsoft — Responsible AI Survey* (n = 2,309) (sponsored by Microsoft), which gathered insights on organizational attitudes and the state of responsible AI, 91% are currently using AI technology at their organization and expect more than 24% improvement in customer experience, business resilience, sustainability, and operational efficiency because of AI in 2024. Respondents who use responsible AI solutions say that it has helped with data privacy, customer experience, confident business decisions, brand reputation, and trust.

91%

are currently **using AI technology at their organization** and expect more than 24% improvement in customer experience, business resilience, sustainability, and operational efficiency because of AI in 2024.

AI brings not only unprecedented opportunities to businesses but also an incredible responsibility. To ensure trust and fairness with their customers and stakeholders, as well as adhere to emerging governmental regulations (e.g., the EU AI Act), organizations need to be focused on responsible AI.

The EU AI Act, which aims to govern the way companies develop, use, and apply AI, was approved in May 2024 and went into effect in August 2024. The legislation applies a risk-based approach to regulating AI, which means that different applications of the technology are regulated differently depending on the level of risk they pose to society.

For AI applications deemed to be "high risk," for example, strict obligations have been introduced. Such obligations include adequate risk assessment and mitigation systems, high-quality training data sets to minimize the risk of bias, routine logging of activity, and mandatory sharing of detailed documentation on models with authorities to assess compliance.

The EU AI Act has implications that go far beyond the EU. It applies to any organization with any operation or impact in the EU, which means the AI Act will likely apply to you no matter where you're located. Oversight of all AI models that fall under the scope of the Act — including general-purpose AI systems — will fall under the European AI Office, a regulatory body established by the Commission in February 2024.

Essentially, organizations need to be responsible at the core and proactively operationalize AI governance across the project life cycle, support collaborative risk management, and adhere to evolving AI regulations and their policies and values.

As consumers become more aware of AI's impact, they demand greater transparency and responsible use of AI. Many organizations are integrating responsible AI into their CSR strategies, recognizing that responsible AI practices can enhance their reputation and contribute to societal well-being. Businesses are adopting responsible AI to mitigate risks associated with AI, such as biases, security vulnerabilities, and unintended consequences. This proactive approach helps in safeguarding their operations and reputation. Companies that prioritize responsible AI are often seen as leaders in innovation. By addressing social and moral concerns, they can differentiate themselves in the market and attract more customers and partners.

There is a growing trend of collaboration between technologists, legal experts, and other stakeholders to develop comprehensive responsible AI frameworks. This interdisciplinary approach ensures that diverse perspectives are considered in AI development. With increasing regulations like the EU's AI Act and the U.S. AI Bill of Rights, companies are prioritizing responsible AI practices to ensure compliance and avoid legal repercussions. These trends highlight the industry's recognition of the critical role responsible AI plays in ensuring sustainable technological advancement.

IDC defines RAI as the practice of designing, developing, and deploying AI in a way that ensures fairness, reliability and safety, privacy and security, inclusiveness, transparency, and accountability. To create trust in AI, organizations must move beyond defining RAI principles and put those principles into practice. AI governance is essentially the set of processes, policies, and tools that bring together diverse stakeholders across data science, engineering, IT, compliance, legal, and business teams to ensure that AI systems are built, deployed, used, and managed to maximize benefits and prevent harm. AI governance allows organizations to align their AI systems with business and legal requirements throughout every stage of the machine learning (ML)/generative AI life cycle.

IDC defines RAI as the practice of designing, developing, and deploying AI in a way that ensures fairness, reliability and safety, privacy and security, inclusiveness, transparency, and accountability.

Key Findings from the Survey

According to IDC's *Microsoft — Responsible AI Survey,* over 30% of the respondents note that **lack of governance and risk management solutions is the top barrier to adopting and scaling AI** (see **Figure 1,** next page).

Equally important to note is that more than 75% of the respondents who use responsible AI solutions say that it has **helped with data privacy, customer experience, confident business decisions, brand reputation, and trust** (see **Figure 2,** page 9). Basically, by being proactive and using RAI tools and technologies to identify, mitigate, and monitor risks throughout the AI life cycle, they can mitigate unintended negative consequences.

As organizations are buying, developing, and deploying AI in a wide variety of solutions, they are also grappling with the need to develop responsible AI policies, procedures, and practices. According to the survey, **organizations are still in the early days of developing and following a comprehensive responsible AI practice on a worldwide level.** While AI is not new and organizations have been using AI-powered solutions for a while, **only the more AI-mature organizations have been proactive about embracing it responsibly.** GenAI has been a catalyst to broader AI adoption but has also brought a lot more issues around data security, IP leakage, hallucinations, copyright infringement, and threats from bad actors. On a regional basis, EMEA, Latin America, and Asia/Pacific lag behind North America in terms of governance structures and technology used to enforce governance. **Lack of human capital, data availability, funding, trust concerns, and regulations have been the key inhibitors to AI maturity** (see **Figure 3,** page 10).

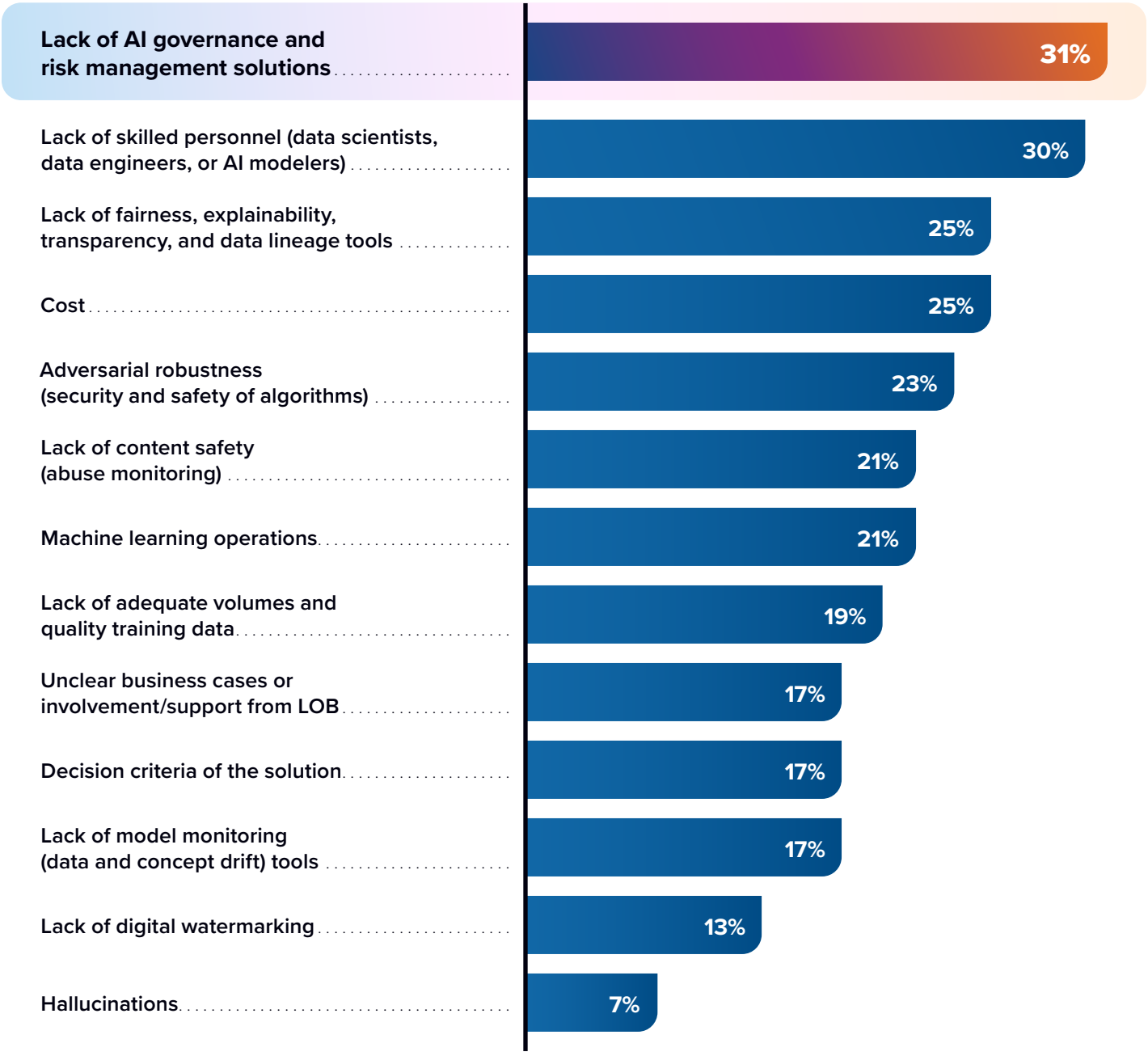
A systematic approach requires proven tools, frameworks, and methodologies, enabling organizations to move from principles to practice with confidence. Establishing a responsible AI approach that is robust, fair, and maintained on an ongoing basis can also enable organizations to communicate and collaborate with confidence. North America has been at the forefront with early adopters and more AI-mature organizations.

FIGURE 1

Top Barriers to AI Adoption

What have been your top barriers to adopting AI?

(Percentage of respondents)

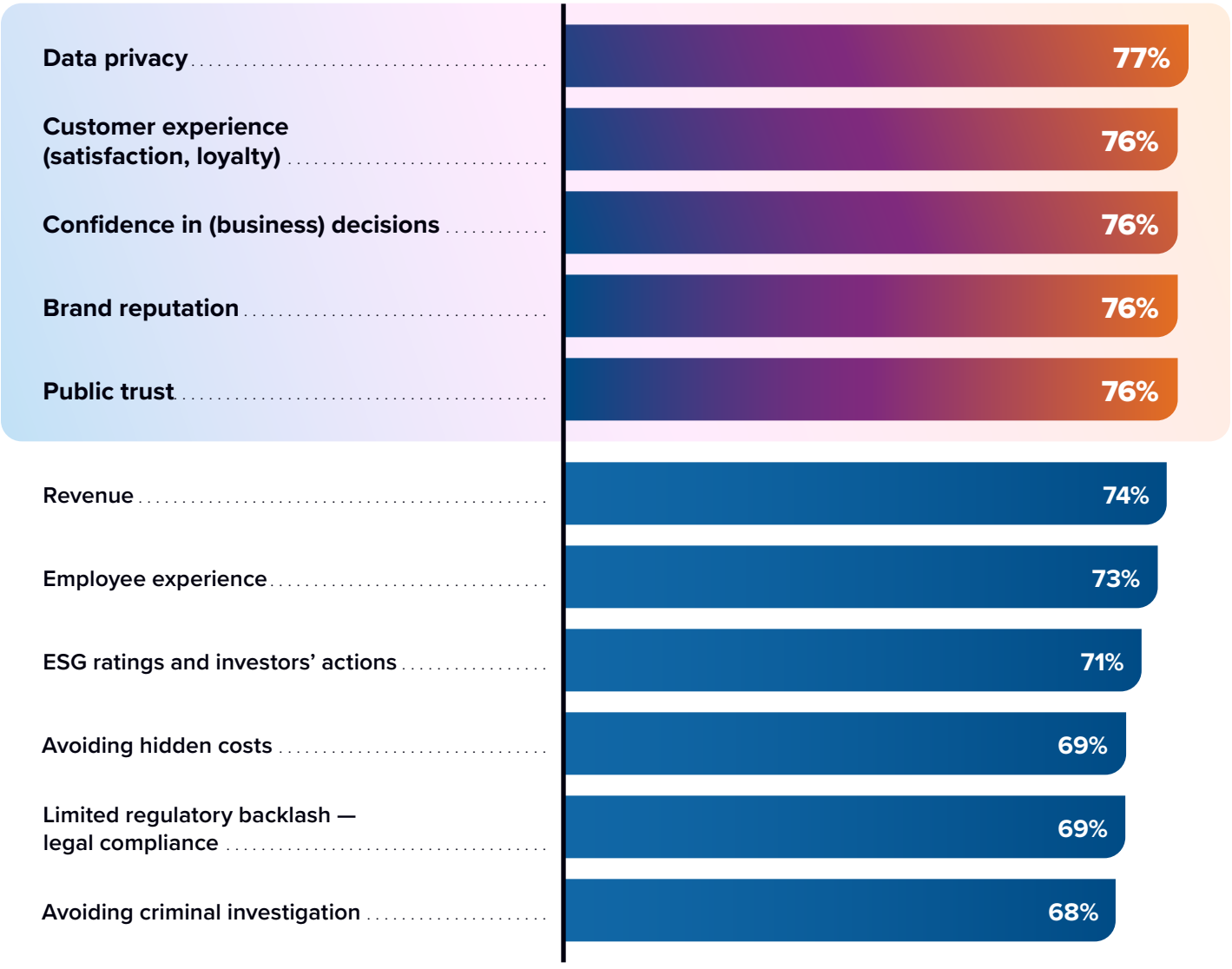


Notes: Data is managed by IDC's Global Primary Research Group. Data is weighted by IT spending by country.
Multiple responses were allowed. Use caution when interpreting small sample sizes.
n = 2,562; Source: IDC's *Microsoft — Responsible AI Survey*, March 2024

FIGURE 2

Level of Impact of Organization's Responsible Use of AI Solutions

How impactful do you consider your organization's responsible use of AI solutions in preserving each of the following?
(Percentage of respondents)



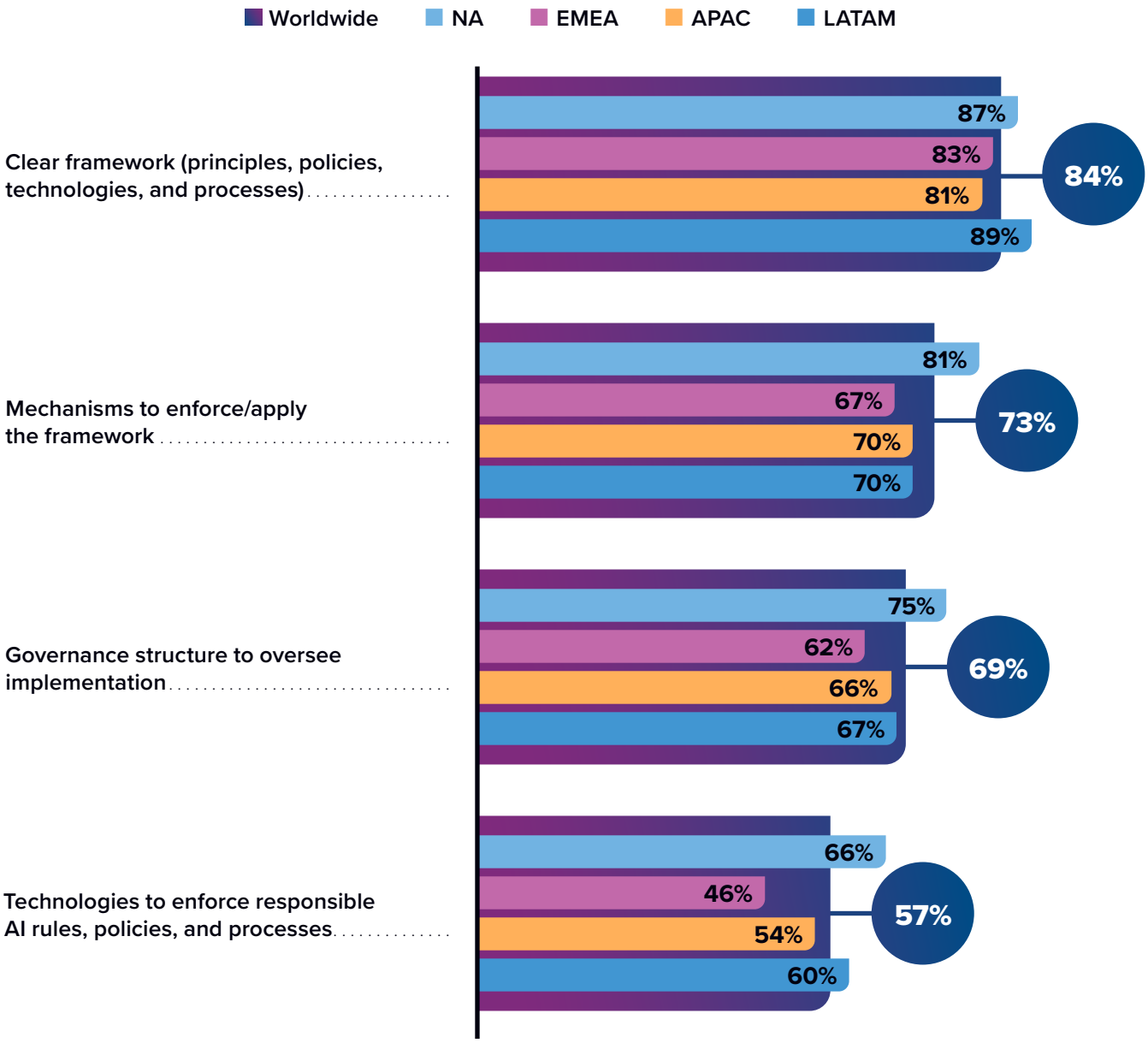
Notes: Data is managed by IDC's Global Primary Research Group. Data is weighted by IT spending by country.
Use caution when interpreting small sample sizes. Scores are based on a scale of 1–5 (1 = not impactful, 5 = very impactful).
n = 2,562; Source: IDC's *Microsoft — Responsible AI Survey*, March 2024

FIGURE 3

Governance Frameworks in Place: Worldwide and Regional Split

Which of the following are currently in place at your organization?

(Percentage of respondents)



Notes: Data is managed by IDC's Global Primary Research Group. Data is weighted by IT spending by country.
Multiple responses were allowed. Use caution when interpreting small sample sizes.
n = 2,562; Source: IDC's *Microsoft — Responsible AI Survey*, March 2024

For an accessible version of the data in this figure, see Figure 3 Supplemental Data in the Appendix.

Responsible AI Tooling

IDC is seeing that organizations are using a variety of tools to ensure responsible AI, ranging from software-based monitoring tools to including human oversight (also known as human in the loop).

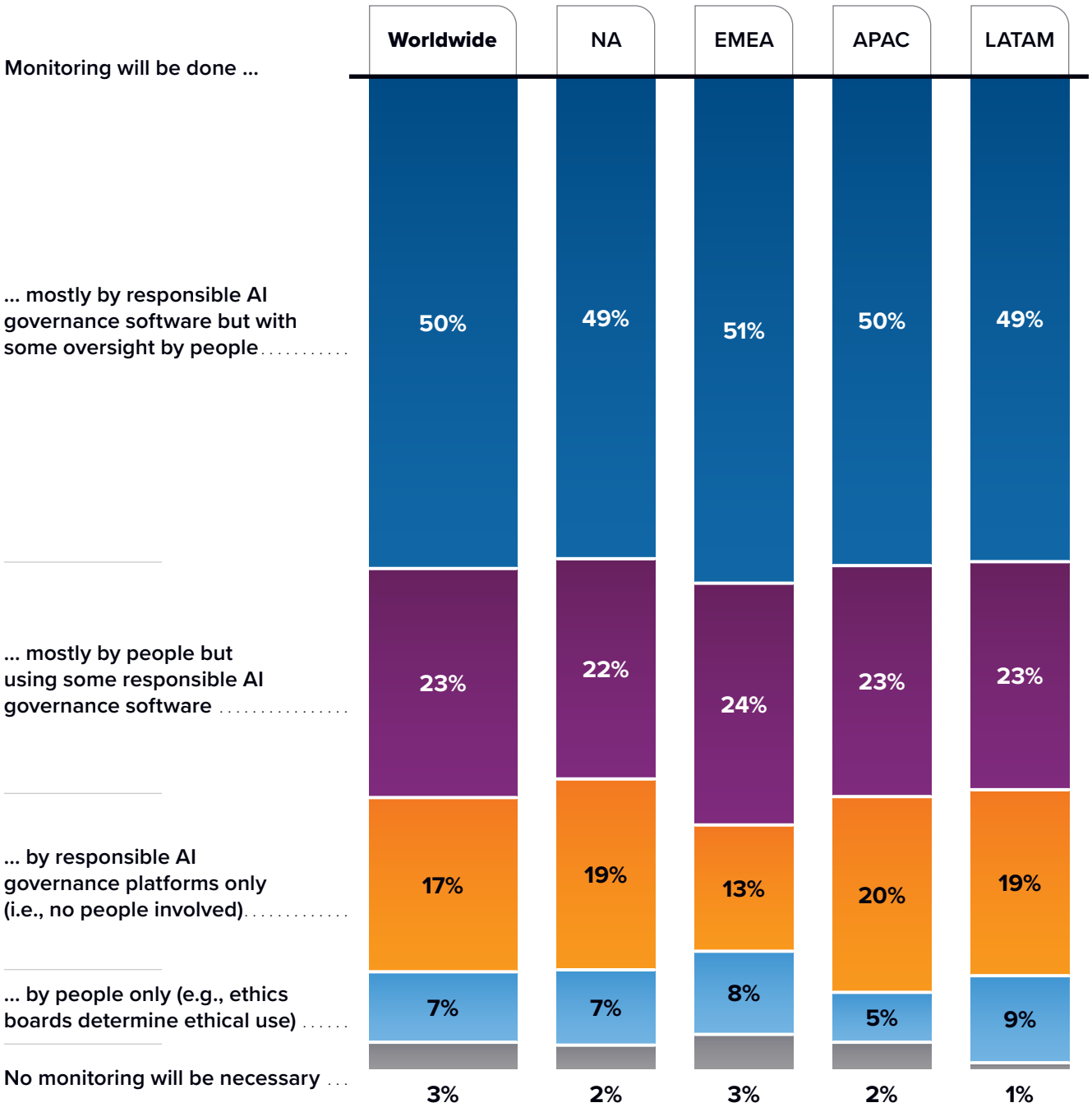
The tools for monitoring and checking output from AI range from content filtering and abuse monitoring to bias checking and from visual explainability to groundedness detection. This area of software is rapidly evolving, and IDC expects to see a larger set of vendors offering solutions in this area over the next 12–18 months. **Figure 4** (next page) shows how organizations are thinking about the use of technology combined with human oversight as the RAI tools and technologies are rapidly evolving. **Figure 5** (page 13) shows how organizations will be allocating their budget to include responsible AI software. Considering their lack of both AI skills and tools to support their RAI requirements, about one-third of the respondents plan to leverage professional services support along with RAI software.

FIGURE 4

Asset Mix for Monitoring After an AI System Has Gone Live: Worldwide and Regional Split

To ensure responsible use of AI by your organizations over the next 12–18 months, please indicate the most likely mix of assets to be used for monitoring after an AI system has gone live.

(Percentage of respondents)



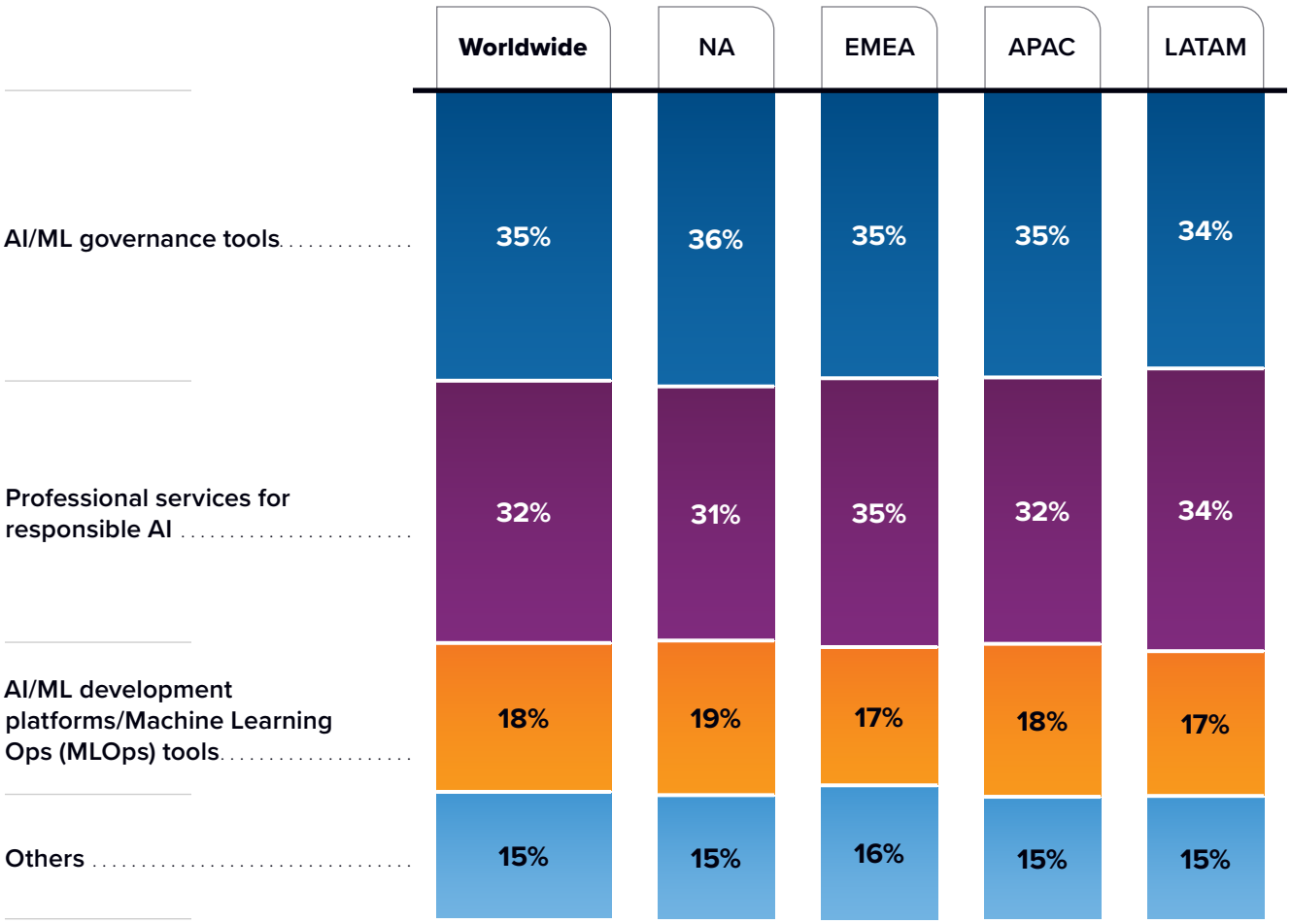
Notes: Totals may not sum up to 100% due to rounding. Data is managed by IDC's Global Primary Research Group. Data is weighted by IT spending by country. Use caution when interpreting small sample sizes.n = 2,562 (worldwide), n = 611 (NA), n = 819 (EMEA), n = 832 (APAC), n = 300 (LATAM); Source: IDC's *Microsoft — Responsible AI Survey*, March 2024

For an accessible version of the data in this figure, see Figure 4 Supplemental Data in the Appendix.

FIGURE 5

AI Organization's Budget Allocation, 2024

What percentage of your AI organization's spend in 2024 will be for each of the following?
(Percentage of respondents)



Base = respondents that indicated organization's plan to spend more than $1 on their AI projects in 2024.
Notes: Totals may not sum up to 100% due to rounding. Data is managed by IDC's Global Primary Research Group. Data is weighted by IT spending by country. Use caution when interpreting small sample sizes. n = 2,555 (worldwide): n = 611 (NA), n = 819 (EMEA), n = 830 (APAC), n = 300 (LATAM); Source: IDC's *Microsoft — Responsible AI Survey*, March 2024

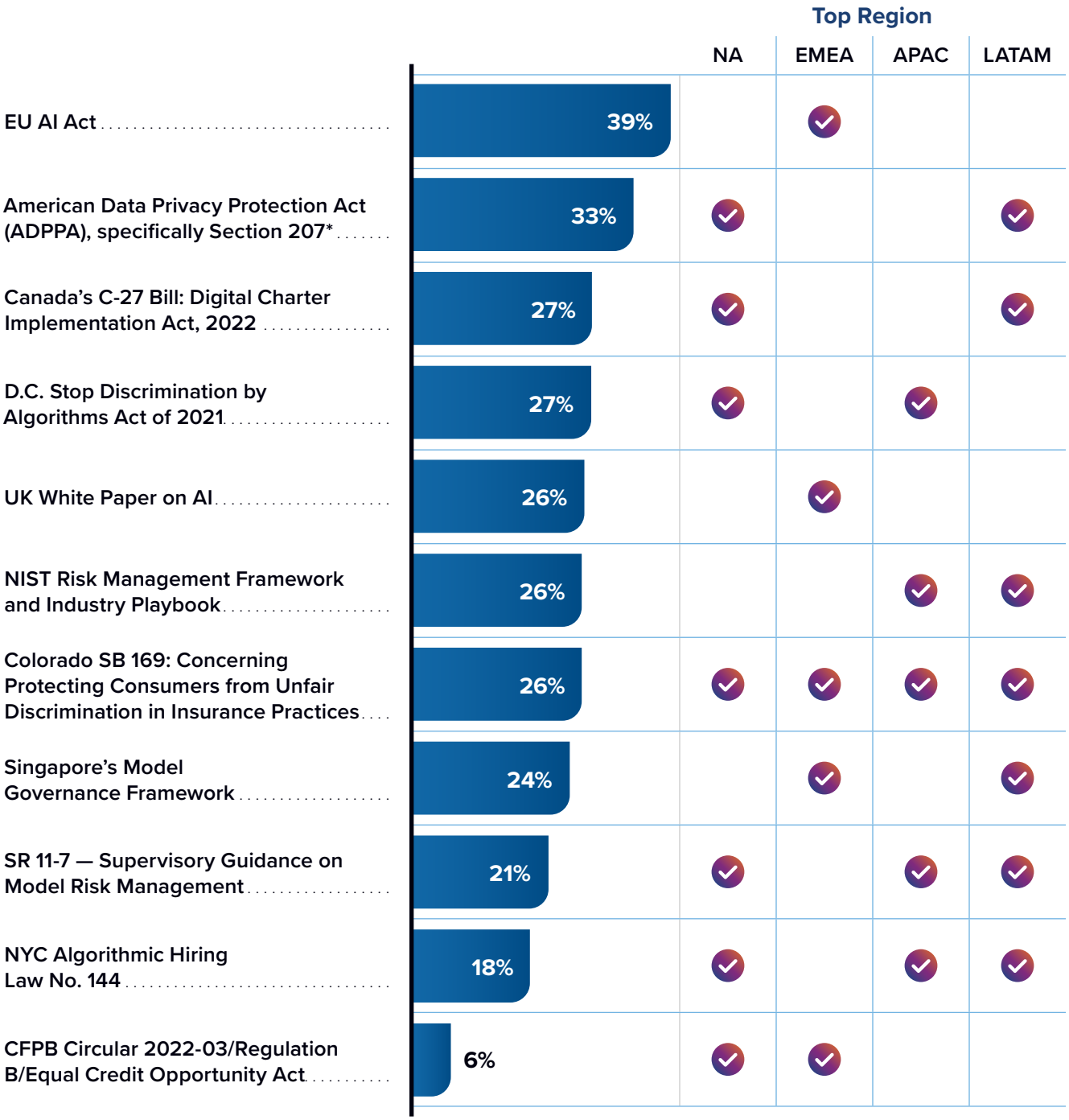
For an accessible version of the data in this figure, see Figure 5 Supplemental Data in the Appendix.

The AI regulatory landscape is dynamic, and currently the EU AI Act and American Data Privacy and Protection Act are critical regulations for organizations to adhere to (see **Figure 6**, next page). It is important to note that while the regulations will increase, organizations will continue to spend on AI solutions but do it responsibly using professional services and governance tools and technologies (see **Figure 7,** page 15).

FIGURE 6

AI Regulations Critical for Organizations' AI Implementations

Which of the following emerging AI regulations are critical for your organization's AI implementations?
(Percentage of respondents)



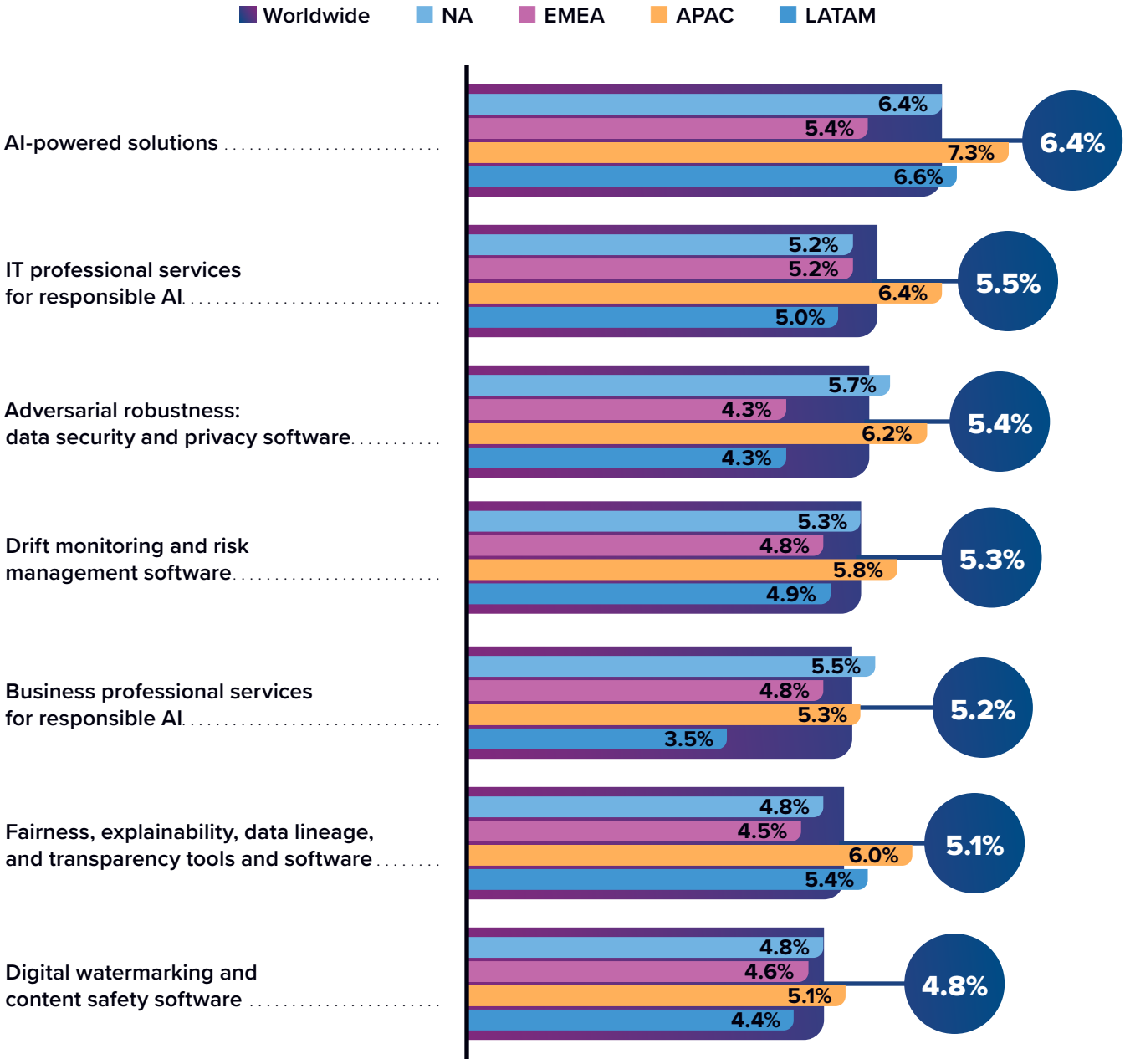
* Which requires an algorithm design evaluation and algorithmic impact assessment. Notes: Data is managed by IDC's Global Primary Research Group. Data is weighted by IT spending by country. Multiple responses were allowed. Use caution when interpreting small sample sizes.
n = 2,562; Source: IDC's *Microsoft — Responsible AI Survey*, March 2024

FIGURE 7

Influence of Worldwide Increase in AI Regulations on an Organization's Responsible AI Spend Plans in the Next Two Years: Worldwide and Regional Split

For each of the following areas, how would a worldwide increase in AI regulations influence your organization's responsible AI spend plans in the next two years?

(Mean — percentage of increase)



n = 2,562; Source: IDC's *Microsoft — Responsible AI Survey*, March 2024

Notes: Data is managed by IDC's Global Primary Research Group. Data is weighted by IT spending by country. Use caution when interpreting small sample sizes.

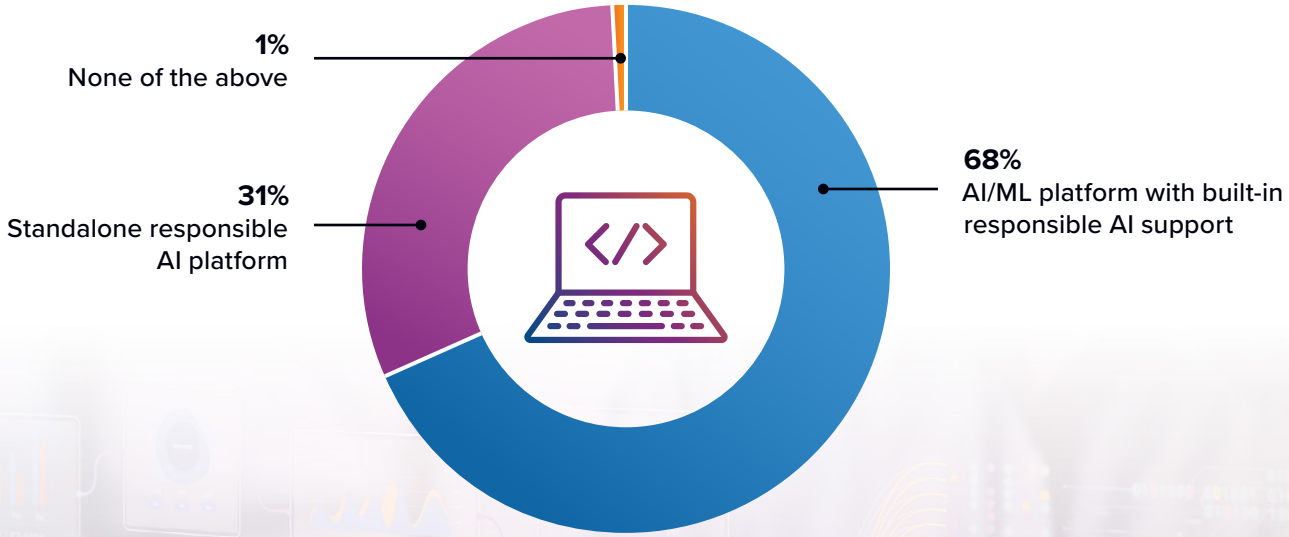
For an accessible version of the data in this figure, see Figure 7 Supplemental Data in the Appendix.

Over two-thirds of the respondents are planning to use AI/ML platforms with built-in RAI support (see **Figure 8**), and 39% report that the platform should provide dashboards to assess, monitor, and drive timely actions and multi-persona collaboration (see **Figure 9,** next page).

FIGURE 8

Type of Responsible AI Software Used/Planning to Be Used

What type of responsible AI software is your organization using/planning to use?
(Percentage of respondents)



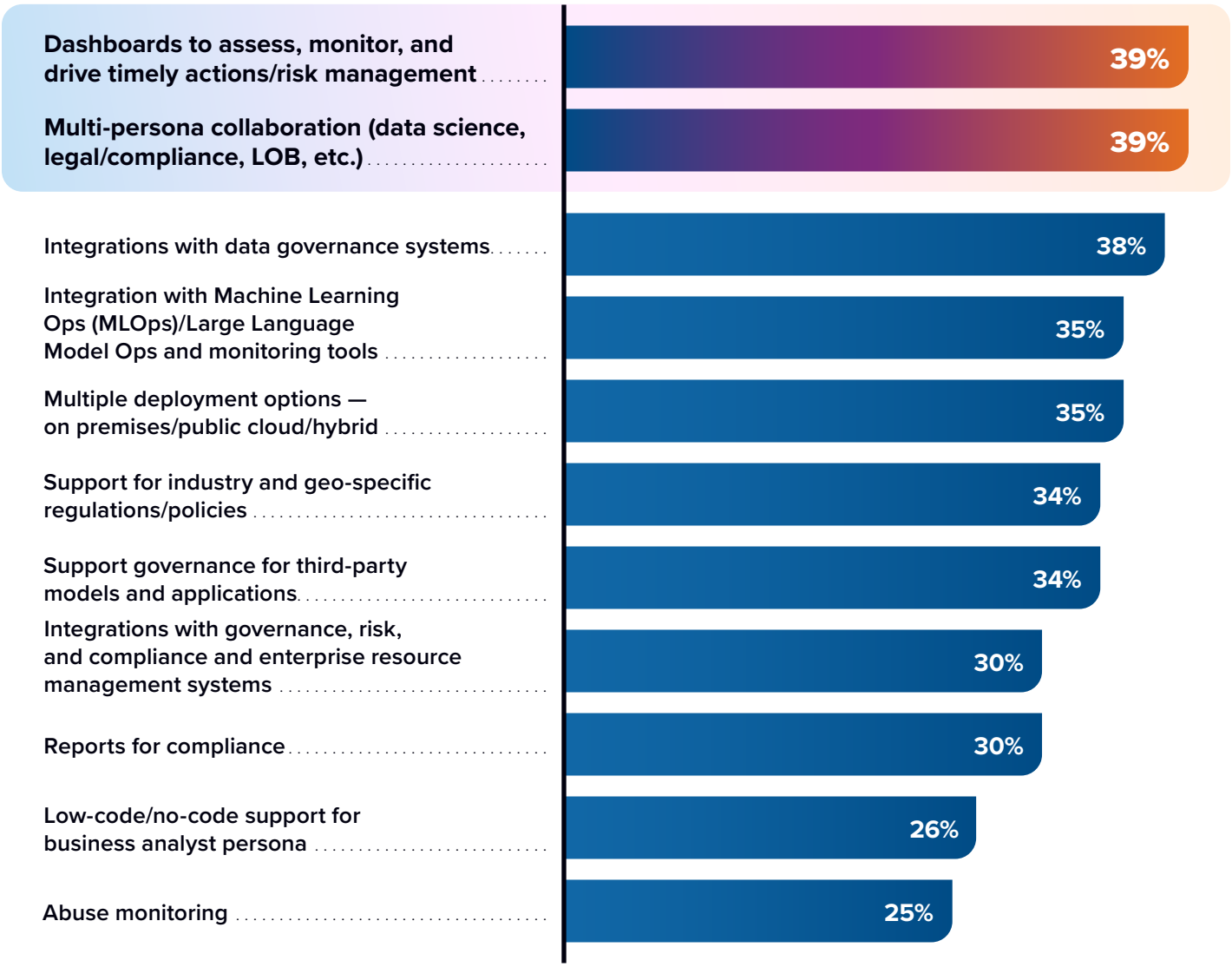
n = 2,562; Source: IDC's *Microsoft — Responsible AI Survey*, March 2024
Notes: Data is managed by IDC's Global Primary Research Group. Data is weighted by IT spending by country. Use caution when interpreting small sample sizes.

FIGURE 9

Critical Capabilities of a Responsible AI platform

What do you think are the critical capabilities of a responsible AI platform?

(Percentage of respondents)



n = 2,562; Source: IDC's *Microsoft — Responsible AI Survey*, March 2024
Notes: Data is managed by IDC's Global Primary Research Group. Data is weighted by IT spending by country.
Multiple responses were allowed. Use caution when interpreting small sample sizes.

Additional Insights from the Study

AI Adoption

IDC estimates that the use of AI is growing rapidly in excess of 40% and is projected to maintain its remarkable momentum, driven by the increasing adoption of AI across various industries (see *Worldwide Artificial Intelligence Platforms Software Forecast, 2024–2028: AI Integration Accelerates, Fueling Technological Breakthroughs and Business Transformations,* IDC #US52386424, July 2024) IDC research estimates the worldwide economic impact of generative AI by the end of 2033 to be close to $10 trillion.

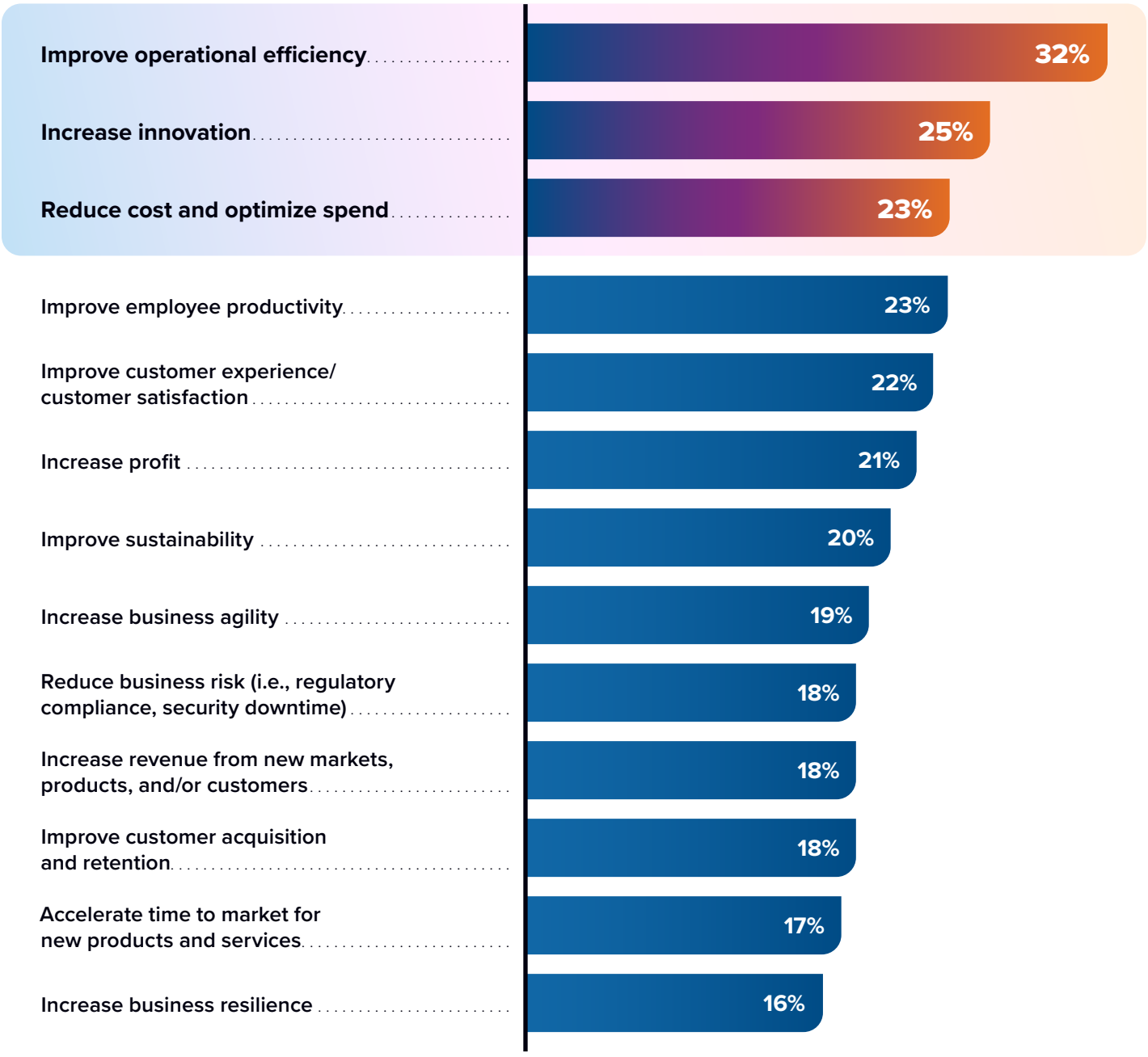
Some key facts to note from IDC's March 2024 *Microsoft — Responsible AI Survey* are:

- Over 77% of organizations across the world are either **exploring potential use cases or investing significantly in generative AI technologies.**
- 91% are currently **using AI technology.**
- 63% of organizations **have an AI strategy tied to their business objectives,** which includes a measurement strategy to evaluate success.
- **Improving operational efficiency, increasing innovation, and reducing cost** are the top business objectives for AI initiatives (see **Figure 10**, next page).

FIGURE 10

Top 3 Business Objectives for Investing in AI

Please rank your organization's top 3 business objectives for investing in AI.
(Percentage of respondents)



n = 2,562; Source: IDC's *Microsoft — Responsible AI Survey*, March 2024
Notes: Data is managed by IDC's Global Primary Research Group. Data is weighted by IT spending by country.
Multiple responses were allowed. Use caution when interpreting small sample sizes.

Important Use Cases

Organizations are using AI for a wide range of use cases, including:

- Software development
- Automating IT tasks
- Fraud detection and cybersecurity
- Product and service innovation
- Automating business processes
- Call center conversation summarization and categorization
- Conversational analysis and intelligence on call center transcripts

IDC expects rapid expansion of AI use cases to help businesses innovate and stay competitive and relevant. It is interesting to note that organizations are prioritizing AI investments in IT operations, IT service management, and machine learning operations (see **Figure 11,** next page). This is aligned with the need to drive foundational efficiencies so that they can scale the AI adoption for line-of-business use cases that transform customer and employee experiences.

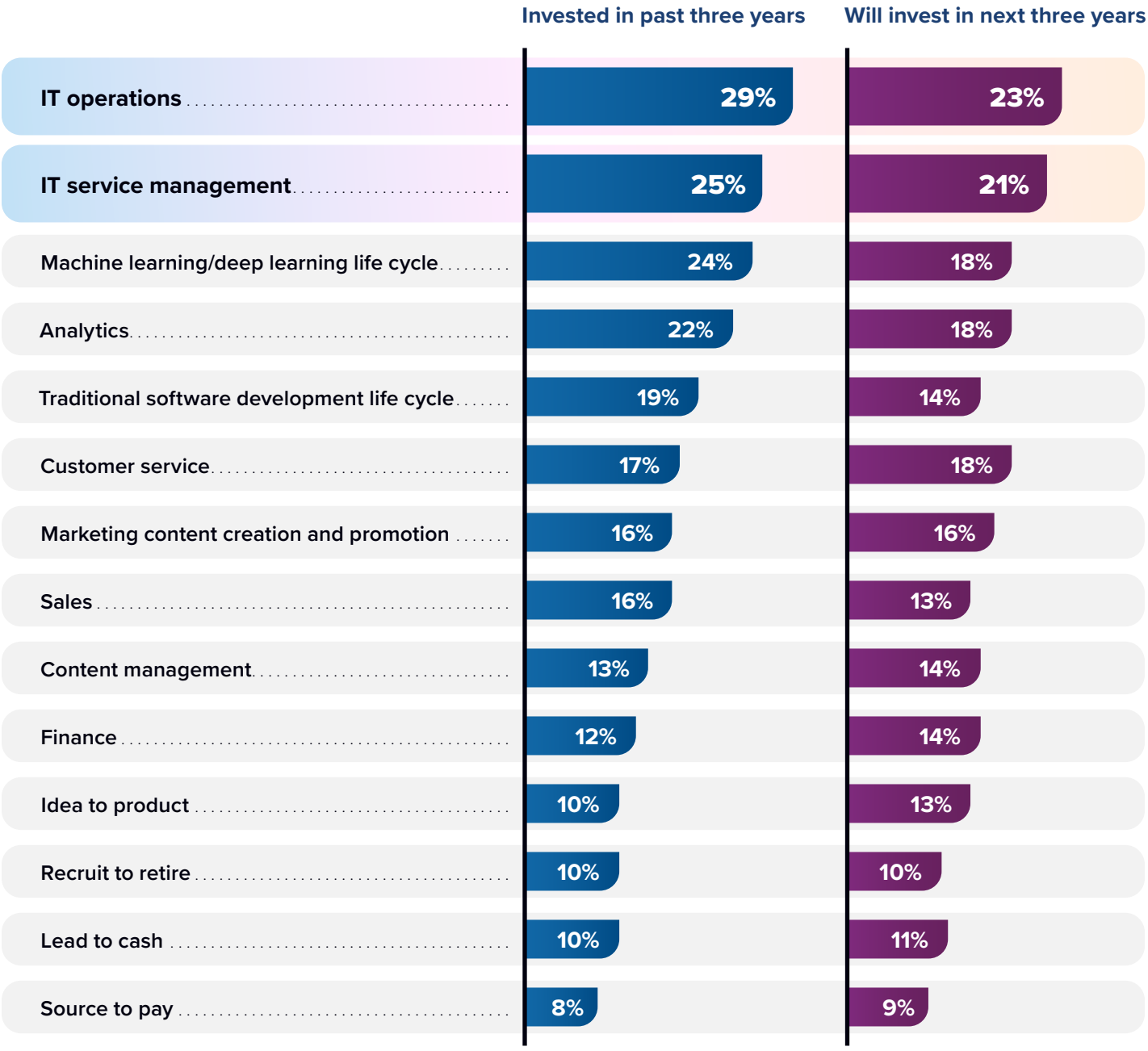
Over the next three years, IT operations and IT service management will be the areas in which organizations invest in AI the greatest (see next page).

FIGURE 11

Business/IT Processes for Which an Organization Will Be Investing in AI

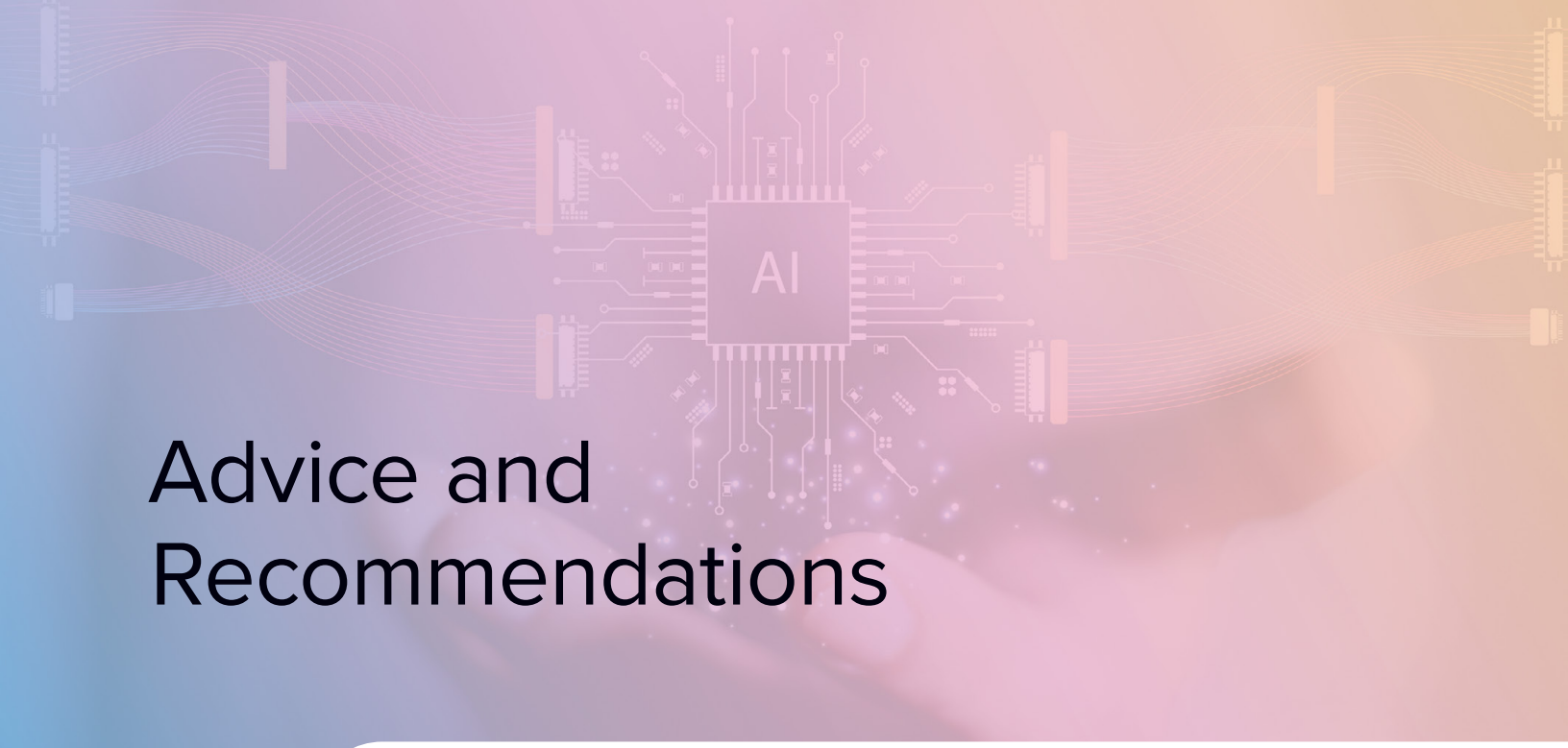
For which of these business/IT processes will your organization be investing in AI?

(Percentage of respondents)



n = 2,309 (respondents currently using AI technology); Source: IDC's *Microsoft — Responsible AI Survey*, March 2024
Notes: Data is managed by IDC's Global Primary Research Group. Data is weighted by IT spending by country.
Multiple responses were allowed. Use caution when interpreting small sample sizes.

For an accessible version of the data in this figure, see Figure 11 Supplemental Data in the Appendix.

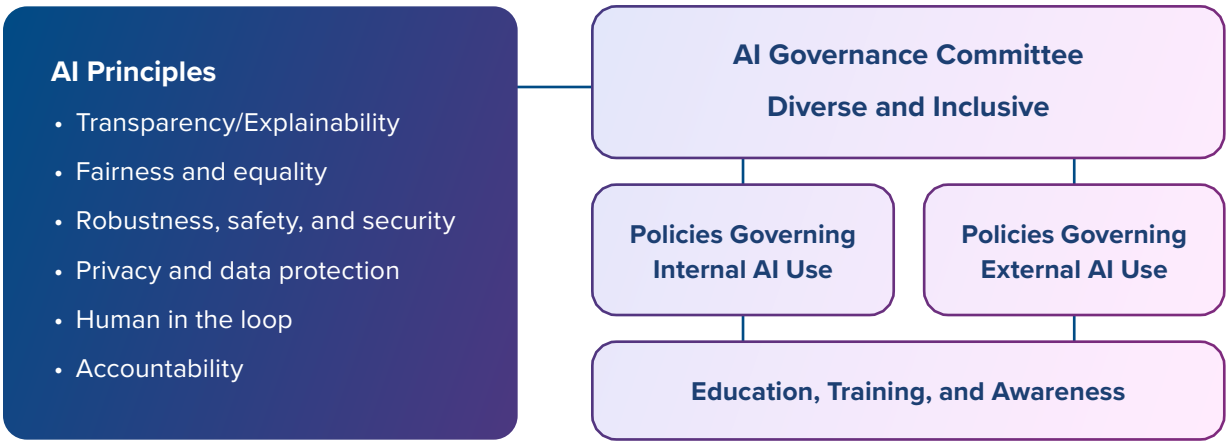


Advice and Recommendations

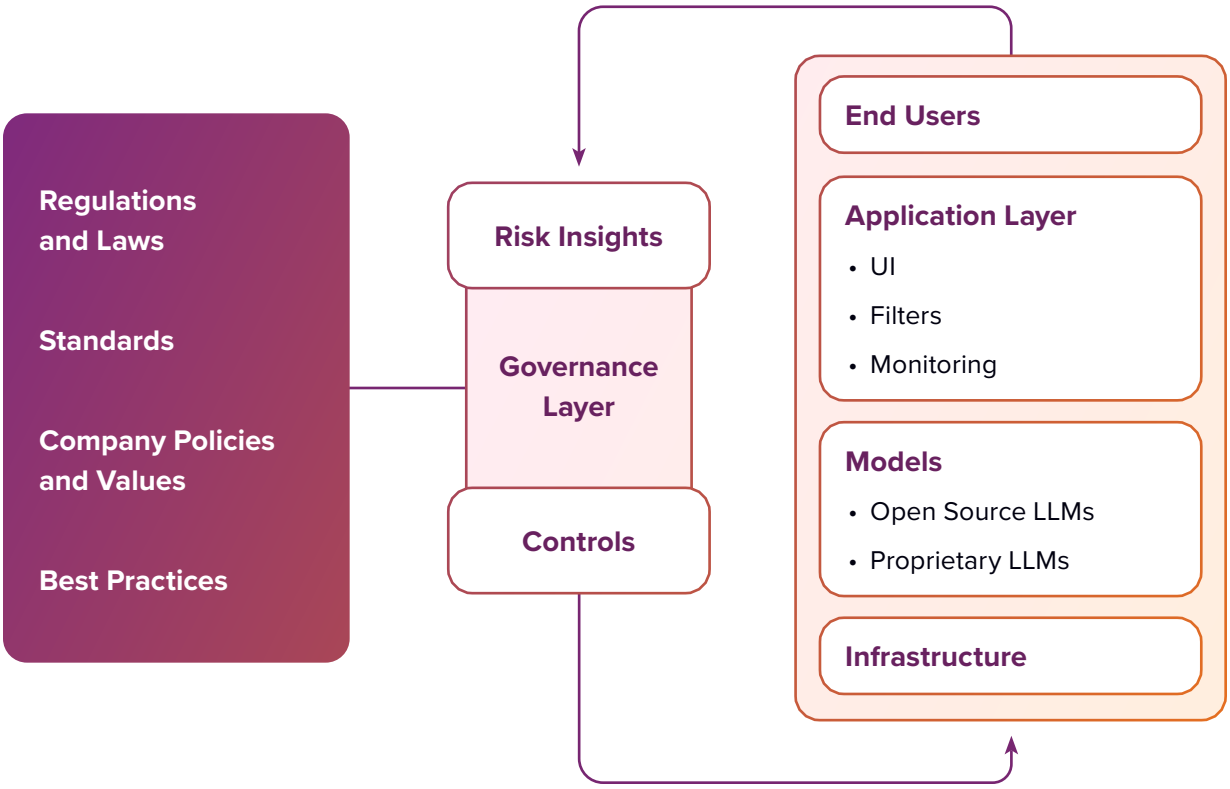
Organizations do business with organizations that they can trust. There is an incredible urgent need for organizations to operationalize AI governance across the project life cycle, support collaborative risk management, and adhere to regulations and their policies and values. Organizations need to be responsible at the core, leveraging the framework in **Figure 12** (next page).

FIGURE 12

Framework for Organizations to Be Responsible at the Core



Operationalize AI Governance



Source: IDC, 2024

As such, every organization should do the following:

- **Establish its AI principles:** This entails commitment to developing technology responsibly and work to establish specific application areas the organization will not pursue. For example, many prohibit the use of facial recognition technology for building AI solutions.

- **Avoid creating or reinforcing unfair bias:** AI algorithms and data sets can reflect, reinforce, or reduce unfair biases. Although not simple, and considering they differ across cultures and societies, every organization should seek to avoid unjust impacts on people, particularly those related to sensitive characteristics such as race, ethnicity, gender, nationality, income, sexual orientation, ability, and political or religious belief.

- **Build and test for safety:** Every organization should develop and apply strong safety and security practices to avoid unintended results that create risks of harm. It should test AI technologies in constrained environments and monitor their operation after deployment.

  The organization should design or adopt AI systems that provide appropriate opportunities for feedback, relevant explanations, and appeal.

  Every organization should incorporate privacy design principles.

- **Establish an AI Governance Committee:** Establish an AI Governance Committee that can help reduce the abuse and misuse of artificial intelligence. For the organization to adhere to its AI principles, it is critical that it has diverse (across different functions from legal and compliance to security to data team and from HR to marketing and finance) and inclusive (different genders, cultures, abilites, and racial backgrounds) representation in the AI Governance Committee:

  - **Define organization's policies for governing internal and external AI use:** These policies are crafted to align with legal requirements and organizational values, ensuring that AI technologies are used responsibly.
  - **Promote transparency and explainability:** Encourage the development of AI systems that are transparent about their decision-making processes and can be easily explained to nontechnical stakeholders.
  - **Implement diverse testing criteria:** Ensure AI models are tested against diverse data sets to minimize bias and verify their reliability across various scenarios and populations.
  - **Conduct regular AI audits:** Schedule periodic audits of AI systems to assess compliance with internal policies and external regulations, iterating on the systems as necessary to address discovered issues.

- **Prioritize privacy and data protection:** Reinforce privacy and data protection measures in AI operations to safeguard against unauthorized data access and ensure user trust.
- **Invest in AI training:** Allocate resources for regular training and workshops on responsible AI practices and concepts along with use that aligns with corporate policies for the entire workforce, including the executive leadership approach.

As we all know, it is not enough to just define principles and policies, but it is critical to leverage an iterative process to operationalize AI governance:

- Organizations need to keep abreast of global AI regulations. The EU AI Act, a landmark rule that aims to govern the way companies develop, use, and apply AI, was approved in May 2024 and went into effect in August 2024. The legislation applies a risk-based approach to regulating AI, which means that different applications of the technology are regulated differently depending on the level of risk they pose to society.
- For AI applications deemed to be "high risk," for example, strict obligations have been introduced. Such obligations include adequate risk assessment and mitigation systems, high-quality training data sets to minimize the risk of bias, routine logging of activity, and mandatory sharing of detailed documentation on models with authorities to assess compliance.
- The EU AI Act has implications that go far beyond the EU. It applies to any organization with any operation or impact in the EU, which means the AI Act will likely apply to you no matter where you're located. This will bring much more scrutiny to tech giants when it comes to their operations in the EU market and their use of EU citizen data. Companies that breach the EU AI Act could be fined from 35 million euros ($41 million) or 7% of their global annual revenue — whichever amount is higher — to 7.5 million euros or 1.5% of global annual revenue. The size of the penalties will depend on the infringement and size of the company fined. That's higher than the fines possible under the GDPR, Europe's strict digital privacy law. Companies face fines of up to 20 million euros or 4% of annual global turnover for GDPR breaches.
- Oversight of all AI models that fall under the scope of the Act — including general-purpose AI systems — will fall under the European AI Office, a regulatory body established by the Commission in February 2024.

Establish an end-to-end governance layer framework that includes the following:

- **Infrastructure governance:** Running AI systems on infrastructure that has appropriate security and privacy controls built into it is the only surefire way to mitigate one of the most critical risks of generative AI systems to organizations: leakage of sensitive data or IP.

- **Model governance:** Policies and processes that control the design, development, and deployment of AI models is nothing new. Many organizations have been doing some form of model risk management for years. In the era of generative AI, however, enterprise model governance looks very different because most enterprises aren't building their own foundation models. Instead, they are relying on third-party foundation model providers — for example, OpenAI and Anthropic. These third-party providers are increasingly investing in tools and processes to manage and mitigate privacy, safety, and security risks at the model level — these investments include foundation model evaluations to better quantify model behavior and "alignment" approaches like reinforcement learning through human feedback and constitutional AI, which reduce the likelihood of common failure modes and improve model steerability.

  These safeguards, however, are not tailored to any particular use of foundation models, nor are they grounded in a specific industry or organization's risk tolerance and compliance needs. Enterprises that have a low risk tolerance or specific concerns related to a particular application of foundation models are finding that the model governance of third-party providers is not sufficient for their needs. In these scenarios, you could explore the use of open source models to enhance your control or implement stronger layers of governance on top of and underneath the model in the other layers of the GenAI stack.

- **Application layer governance:** The application layer provides the user interface for generative AI APIs, and so there is a tremendous opportunity to insert governance controls into this layer to prevent a foundation model from being used in dangerous or noncompliant ways.

  By their nature, GenAI systems are more flexible and difficult to predict than traditional software engineering, which presents new challenges for application builders. For example, GenAI applications are vulnerable to prompt injections and misuse by malicious users. It is also easy for GenAI applications to return outputs that are harmful or in violation of governance policies and requirements. These issues can generally be dealt with by input/output governance, where safeguards (e.g., automatic content moderation) are added around foundation model API calls to reduce risks. Adding these kinds of governance controls to the application layer of the generative AI stack

is a very effective way to reduce the risk of these systems; however, if an organization isn't building its own generative AI applications, it still doesn't have control over this layer.

- **End-user governance:** Without direct control over models or the application layer, what capacity do you have to govern these systems and mitigate their most egregious risks? For most enterprises, the first line of defense against generative AI risk is end-user governance — governing the ways that end users are allowed to interact with generative AI systems.

  Many enterprises responded to the generative AI revolution by implementing the bluntest instrument when it comes to end-user governance: turning off end-user access. Of course, turning off access to GenAI chatbots is an effective way to make sure that your employees aren't exposing your organization to risk via usage; however, it also blocks your organization from realizing the many benefits and obtaining value from these tools.

  Examples of end-user governance that allows for safe and responsible exploration of generative AI include:

  - Adopting a code of conduct that defines how users are and are not allowed to interact with generative AI tools
  - Logging end-user interactions and monitoring for risky or edge case inputs and outputs
  - Implementing human-in-the-loop reviews that prevent generative AI outputs from being used without human feedback or input and enabling users to share effective prompts with one another so they can become better at successfully using generative AI tools

Conclusion

AI is being regarded as a critical enabler of businesses' strategic priorities. Scaling AI can deliver high performance for customers, shareholders, and employees, but organizations must overcome common hurdles to apply AI responsibly and sustainably. AI adoption can bring with it new, dynamic, organizational, and social issues. Failure to manage these issues can have a significant impact at a human and societal level, leaving organizations exposed to financial, legal, and reputational repercussions. Basically, embracing AI responsibly is a must and not an option.

While many organizations have taken the first step and defined AI principles, translating these into practice is far from easy, especially with few standards or regulations to guide them. Successful organizations understand the importance of taking a systematic approach from the start, addressing these challenges in parallel, while others underestimate the scale and complexity of change required. A systematic approach requires proven tools, frameworks, and methodologies, enabling organizations to move from principles to practice with confidence and supporting the professionalization of AI. Establishing an RAI approach that is robust, fair, and maintained on an ongoing basis can also enable organizations to communicate and collaborate with confidence.

Being responsible can become more beneficial, especially as governments, regulatory bodies, and international standard-setting bodies consider new rules of the road and standards for the development and deployment of AI.

The biggest barrier lies in the complexity of scaling AI responsibly — an undertaking that involves multiple stakeholders and cuts across the entire enterprise and ecosystem. IDC's *Microsoft — Responsible AI Survey* revealed that over 50% of respondents do not have a fully operationalized and integrated RAI governance structure and tools and technologies to enforce responsible AI adoption. As new requirements emerge, they must be baked into product development processes and connected to other regulatory areas, such as privacy, data security, and content.

By shifting from a reactive AI compliance strategy to the proactive development of mature responsible AI capabilities, organizations will have the foundations in place to adapt as new regulations and guidance emerge. This way, businesses can focus more on performance and competitive advantage and deliver business value with social and moral responsibility.

Being responsible can become more beneficial, especially as governments, regulatory bodies, and international standard-setting bodies consider new rules of the road and standards for the development and deployment of AI.

Definitions

Generative AI

Generative AI is a branch of computer science that involves unsupervised and semi-supervised algorithms that enable computers to create new content using previously created text, audio, video, images, and code in response to short prompts. Generative AI powers foundational models, which are a class of machine learning models that are trained on diverse data and can be adapted or fine-tuned for a wide range of downstream tasks. The era of the large-scale model was sparked by the emergence of transformer model architecture in 2017, namely the large language model. Generative AI requires significant amounts of data to build and operate models, and it requires access to significant data technologies to build or train models.

While GenAI technologies are relatively new, predictive and prescriptive AI based on various types of machine learning has been providing solutions to problems for over a decade. The combination of predictive, prescriptive, and generative AI is promising unprecedented productivity improvements and business transformation opportunities for organizations across the world.

Responsible AI

As noted previously, responsible AI is the practice of designing, developing, and deploying AI in a way that prioritizes fairness, reliability and safety, privacy and security, inclusiveness, transparency, and accountability. Responsible AI focuses on developing and using AI solutions in a manner consistent with societal laws, government regulations, organizational values, and user expectations. Planning, oversight, and governance are key aspects of responsible AI. Responsible AI aims to ensure that AI use in the organization is human centered, trustworthy, fair, explainable, privacy preserving, secure, documented, and governed.

Responsible AI Attributes

The key attributes and pillars of a responsible AI policy framework are explained in the sections that follow.

Accountability

Can the AI system and the people who designed and implemented the system be held accountable for the decisions made?

With more power comes more responsibility. As AI capabilities are being leveraged for making critical decisions such as medical treatments, it is important that we include humans in the loop around the AI system to ensure the best results. The chief data officer and chief trust officer (or equivalent roles) must collaborate to assess their business-specific regulations charter, review the problem at hand, and define the solution on a case-by-case basis subject to the business risk and potential business impact. These are the roles in the organization that have the authority and responsibility to be accountable for ensuring responsible AI use and operation.

Explainability and Transparency

Is the AI system transparent, and can the output of the AI system be explained?

AI systems need to be transparent — they should be able to safely report key attributes of the AI models, including the data and algorithms used to train the model, bias mitigations performed, model, and its assets. Explainability refers to the ability to understand how the decisions, conclusions, or outputs from the AI system are made. Key personas involved with transparency and explainability include data scientists, auditors, and decision-makers. Arriving at meaningful explanations of the AI models reduces uncertainty and helps quantify their accuracy. It is important to establish the right balance between explainability and improved trust in AI models.

Fairness

Is the AI system fair?

AI systems should be fair and unbiased to avoid any unintentional unfair treatment of certain groups. AI systems should use training data and models that are free of bias. Apart from unwanted bias during training from training data, bias can also creep in because of incorrect model build, selection, or deployment. The AI system needs to have correct checks and balances to ensure that the system doesn't discriminate based on gender, race, color, orientation, faith, or anything else. Again, this is part of the chief trust officer's responsibilities within the organization, and this person is charged with making sure that any AI output is fair and unbiased.

Inclusiveness

Is the AI system inclusive of all genders, races, appearances, languages, abilities, and experiences?

AI systems should be developed using inclusive and accessible practices to be inclusive of all human beings without excluding any groups of people intentionally or unintentionally.

Privacy and Security

Can the AI system protect the privacy and security of the data/users?

AI systems should follow established security and privacy practices to protect AI models from adversarial attacks, secure user data, ensure user privacy, and mitigate risks. This is part of the chief security officer's job and, in many ways, is the same as what the security organization is or should be doing for the rest of the company. In this particular case, the same principles, rules, guidelines, and approaches can be applied to AI systems in the same manner as any other applications.

Robustness and Security

Is the AI system robust and safe?

AI systems should be safe and secure, not vulnerable to tampering or compromising the data they are trained on. They also need to be robust without any performance degradation over time. These systems also need to have appropriate monitoring and human-in-the-loop processes to ensure operational safety. Again, this is part of the chief security officer's job and, in many ways, is the same as what the security organization is or should be doing for the rest of the company. In this particular case, the same principles, rules, guidelines, and approaches can be applied to AI systems in the same manner as any other applications.

Appendix 1: Supplemental Data

This appendix provides an accessible version of the data for the complex figures in this document. Click "Return to original figure" below each table to get back to the original data figure.

FIGURE 3 SUPPLEMENTAL DATA

Governance Frameworks in Place: Worldwide and Regional Split

	Worldwide	NA	EMEA	APAC	LATAM
Clear framework (principles, policies, technologies, and processes)	84%	88%	83%	81%	89%
Mechanisms to enforce/apply the framework	73%	81%	67%	70%	70%
Governance structure to oversee implementation	69%	75%	62%	66%	67%
Technologies to enforce responsible AI rules, policies, and processes	57%	66%	46%	54%	60%

n = 2,562; Source: IDC's *Microsoft — Responsible AI Survey*, March 2024
Notes: Data is managed by IDC's Global Primary Research Group. Data is weighted by IT spending by country.
Use caution when interpreting small sample sizes.

Return to original figure

Appendix 1: Supplemental Data (continued)

FIGURE 4 SUPPLEMENTAL DATA

Asset Mix for Monitoring After an AI System Has Gone Live: Worldwide and Regional Split

	Worldwide	NA	EMEA	APAC	LATAM
Monitoring will be done mostly by responsible AI governance software but with some oversight by people	50%	49%	51%	50%	49%
Monitoring will be done mostly by people but using some responsible AI governance software	23%	22%	24%	23%	23%
Monitoring will be done by responsible AI governance platforms only (i.e., no people involved)	17%	19%	13%	20%	19%
Monitoring will be done by people only (e.g., ethics boards determine ethical use)	7%	7%	8%	5%	9%
No monitoring will be necessary	3%	2%	3%	2%	1%

Notes: Totals may not sum up to 100% due to rounding. Data is managed by IDC's Global Primary Research Group. Data is weighted by IT spending by country. Use caution when interpreting small sample sizes. n = 2,562 (worldwide), n = 611 (NA), n = 819 (EMEA), n = 832 (APAC), n = 300 (LATAM); Source: IDC's *Microsoft — Responsible AI Survey*, March 2024

Return to original figure

FIGURE 5 SUPPLEMENTAL DATA

AI Organization's Budget Allocation, 2024

	Worldwide	NA	EMEA	APAC	LATAM
AI/ML governance tools	35%	36%	35%	35%	34%
Professional services for responsible AI	32%	31%	33%	32%	34%
AI/ML development platforms/ Machine Learning Ops (MLOps) tools	18%	19%	17%	18%	17%
Others	15%	15%	16%	15%	15%

Base = respondents that indicated organization's plan to spend more than $1 on their AI projects in 2024.
Notes: Totals may not sum up to 100% due to rounding. Data is managed by IDC's Global Primary Research Group. Data is weighted by IT spending by country. Use caution when interpreting small sample sizes. n = 2,555 (worldwide): n = 611 (NA), n = 819 (EMEA), n = 830 (APAC), n = 300 (LATAM); Source: IDC's *Microsoft — Responsible AI Survey*, March 2024

Return to original figure

Appendix 1: Supplemental Data (continued)

FIGURE 7 SUPPLEMENTAL DATA

Influence of Worldwide Increase in AI Regulations on an Organization's Responsible AI Spend Plans in the Next Two Years: Worldwide and Regional Split

	Worldwide	NA	EMEA	APAC	LATAM
AI-powered solutions	6.4%	6.4%	5.4%	7.3%	6.6%
IT professional services for responsible AI	5.5%	5.2%	5.2%	6.4%	5.0%
Adversarial robustness: data security and privacy software	5.4%	5.7%	4.3%	6.2%	4.3%
Drift monitoring and risk management software	5.3%	5.3%	4.8%	5.8%	4.9%
Business professional services for responsible AI	5.2%	5.5%	4.8%	5.3%	3.5%
Fairness, explainability, data lineage, and transparency tools and software	5.1%	4.8%	4.5%	6.0%	5.4%
Digital watermarking and content safety software	4.8%	4.8%	4.6%	5.1%	4.4%

n = 2,562; Source: IDC's *Microsoft — Responsible AI Survey*, March 2024
Notes: Data is managed by IDC's Global Primary Research Group. Data is weighted by IT spending by country. Use caution when interpreting small sample sizes.

Return to original figure

Appendix 1: Supplemental Data (continued)

FIGURE 11 SUPPLEMENTAL DATA

Business/IT Processes for Which an Organization Will Be Investing in AI

	Invested in past three years	Will invest in next three years
IT operations	29%	23%
IT service management	25%	21%
Machine learning/deep learning life cycle	24%	18%
Analytics	22%	18%
Traditional software development life cycle	19%	14%
Customer service	17%	18%
Marketing content creation and promotion	16%	16%
Sales	16%	13%
Content management	13%	14%
Finance	12%	14%
Idea to product	10%	13%
Recruit to retire	10%	10%
Lead to cash	10%	11%
Source to pay	8%	9%

n = 2,309 (respondents currently using AI technology); Source: IDC's *Microsoft — Responsible AI Survey*, March 2024
Notes: Data is managed by IDC's Global Primary Research Group. Data is weighted by IT spending by country.
Multiple responses were allowed. Use caution when interpreting small sample sizes..

Return to original figure

About the IDC Analysts

Ritu Jyoti
Group Vice President/General Manager,
Worldwide Artificial Intelligence, Automation,
Data and Analytics Research Practice, IDC

Ritu Jyoti is group vice president/general manager of the Worldwide Artificial Intelligence, Automation, Data and Analytics Research Practice with IDC's Software Market Research and Advisory Practice. Ms. Jyoti is responsible for leading the development of IDC's thought leadership for AI research and management of the worldwide AI, automation, data and analytics software research team. Her research focuses on the state of enterprise AI efforts and global market trends for the rapidly evolving AI and ML including GenAI innovations and ecosystems. Ms. Jyoti also leads insightful research that addresses the needs of AI technology vendors and provides actionable guidance to them on how to crisply articulate their value proposition, differentiate, and thrive in the AI era.

More about Ritu Jyoti

Dave Schubmehl
Research Vice President,
Conversational Artificial Intelligence and
Intelligent Knowledge Discovery, IDC

Dave Schubmehl is research vice president for IDC's Conversational Artificial Intelligence and Intelligent Knowledge Discovery research. His research covers information access and artificial intelligence (AI) technologies around conversational AI technologies including speech AI and text AI, machine translation, embedded knowledge graph creation, intelligent knowledge discovery, information retrieval, unstructured information representation, knowledge representation, deep learning, machine learning, unified access to structured and unstructured information, chatbots and digital assistants, and rich media search in SaaS, cloud, and installed software environments. This research analyzes the trends and dynamics of the Text and Audio AI software markets and the costs, benefits, and workflow impact of solutions that use these technologies.

More about Dave Schubmehl

Message from the Sponsor

Microsoft Azure

Microsoft is dedicated to enabling every person and organization to use and build AI that is Trustworthy, which means AI that is private, safe, and secure.

We use our own best practices from decades of research and learnings from building AI products at scale to provide industry-leading commitments and capabilities. Trustworthy AI is only possible when you combine our policy commitments with our product capabilities so you can achieve your AI transformation with confidence. Trust Microsoft for commitments and capabilities that put your AI privacy, safety and security first.

The study was commissioned and sponsored by Microsoft. This document is provided solely for information and should not be construed as legal advice.

Learn more at www.azure.com

IDC Custom Solutions

IDC Custom Solutions produced this publication. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis that IDC independently conducted and published, unless specific vendor sponsorship is noted. IDC Custom Solutions makes IDC content available in a wide range of formats for distribution by various companies. This IDC material is licensed for external use and in no way does the use or publication of IDC research indicate IDC's endorsement of the sponsor's or licensee's products or strategies.

IDC Research, Inc.
140 Kendrick Street, Building B, Needham, MA 02494, USA
T +1 508 872 8200

idc.com @idc @idc

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications, and consumer technology markets. With more than 1,300 analysts worldwide, IDC offers global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries. IDC's analysis and insight helps IT professionals, business executives, and the investment community to make fact-based technology decisions and to achieve their key business objectives.