Microsoft

Grow Your Business with AI You Can Trust

Considerations for business leaders to plan their AI transformation with confidence

Table of contents

Introduction

AI solutions present a profound opportunity for organizations of every size and in every industry to grow revenue, reduce costs, enhance employee wellbeing, and operate more efficiently. It's no surprise, then, that business leaders are under pressure to adopt AI solutions as quickly as possible and avoid falling behind.

For all the hype around AI, there's as much concern about negative side effects of the technology. In the following sections, we outline several considerations business leaders can plan for to help unlock the promise of this new technology and avoid unintended consequences.

Establishing responsible and secure AI practices for your company helps you safely implement AI tools. And as global AI regulation increases, investing in responsible AI now better equips you to meet new regulatory requirements as they come. Taking a thoughtful approach to the implementation of responsible and secure AI in your business can help leaders embrace AI and innovation.

of business leaders said they were "**equally concerned and excited**" about generative AI.[1]

We're committed to Trustworthy AI and building industry-leading supporting technology. With capabilities that improve security, safety, and privacy, we continue to enable customers to use and build trustworthy AI solutions.

Learn more

Practice responsible AI

Consider responsible AI principles

Public policy and industry best practices are still catching up to the latest developments in AI technology, which leaves business leaders searching for reliable guidance on how to implement AI systems that have a positive impact on businesses, individuals, and society.

Taking the time to think through a responsible AI approach for your organization can help you and your teams move forward with confidence and protect against unintended risks. We've outlined six responsible AI principles for you to consider as you plan and build your own approach. These principles include:

- **Privacy and security**
- **Accountability**
- **Transparency**
- **Reliability and safety**
- **Inclusiveness**
- **Fairness**

Privacy and security

AI systems should adhere to the same privacy and security standards that businesses apply to their most sensitive data.

Prioritize the security of your infrastructure and the privacy of your data.

Know where data is located, how it's used, and confirm that it's secure at rest and in transit. Check that AI tools adhere to your company's values for privacy and security.

When you implement stringent data permissions and assign user permissions based on roles and group memberships, only authorized individuals will have access to sensitive information—reducing the risk of internal breaches.

Additionally, you can maintain security and meet compliance requirements with a governance solution, which includes retaining and logging interactions with AI apps, helping detect any regulatory or organizational policy violations when using those apps, and investigating incidents once they arise.

AI implementation must also comply with all local laws and regulations about data use and privacy. When it comes to data security, it's best to be overcautious and to work with security tool providers with a track record of reliability.

An example of privacy and security in practice:

Use of an AI tool to analyze customer communication to resolve a support ticket could involve access to sensitive or identifiable customer data. Understanding local laws and regulations, adhering to your organization's own high standards for security, and enforcing adequate controls can help keep that sensitive data private.

Reliability and safety

Reliability and safety mean AI systems perform as expected, without errors or interruptions. AI tool developers are responsible for making sure their product provides accurate outputs through testing and documentation, but a system of oversight helps verify if the tool is delivering on its intended use. Systems should also undergo regular monitoring, maintenance, feedback, and evaluation processes to identify new uses, troubleshoot and resolve issues quickly, and improve the AI system over time.

Regularly conduct stress testing.

Stress testing prepares an AI system to handle the types of uses and volume of use it's intended to handle without producing errors or becoming vulnerable to risks.

Red teaming is a type of stress testing that involves simulating real-world attacks and using the techniques hackers commonly use to gain access to secure systems. In 2018, Microsoft established our dedicated AI Red Team, and we've expanded the team's mission to map risks outside of traditional security risks, including risks from non-adversarial users compromising responsible AI standards. For example, red teaming a generative AI tool may involve testing whether a user can generate content that stereotypes a marginalized group using the tool. An AI model can also be red teamed to identify potential misuses, scope its capabilities, and understand its limitations. The insights can then be applied to future versions of the model to ensure it will operate reliably and safely.[2]

Conduct due diligence on reliability and safety measures of an AI system at purchase and conduct regular stress testing to identify risks afterward.

Careful review of documentation helps organizations understand what steps the AI system provider has taken to facilitate the reliable and safe use of their system and helps organizations comply with all requirements to operate the system safely.

An example of reliability and safety in practice:

An AI tool is used to model financial outcomes and report on performance. Testing is conducted regularly to ensure the AI tool reliably produces accurate results, avoiding adverse impact on the organization's financial health.

Accountability

In many instances, responsible AI is human-centered. Establishing a clear system of oversight helps your people control the AI tools you implement and stay accountable for the outcomes those tools produce.

Establish a system of oversight that clearly defines roles and responsibilities at every stage of the AI journey.

Implementing a system of oversight that conducts impact testing and responds to impact results, keeps people at the center. This helps protect against potential adverse impacts, and helps ensure adequate controls are implemented at every stage.

Ensure AI tools are fit for purpose.

Regularly assess that AI tools provide the right solutions for the problems they were intended to solve—and determine how your organization will respond if a tool fails to serve its intended purpose.

An example of accountability in practice:

Use of an AI tool to review legal contracts involves oversight by an individual with adequate context and expertise to verify compliance with applicable laws and regulations, and to sign off on the final outcome of the AI-supported review.

Inclusiveness

At its core, inclusivity requires that AI tools be accessible to people of all abilities. That means the tools that business leaders create or procure should follow accessible design principles and comply with the European accessibility standard, EN 301 549; Section 508 of the U.S. Rehabilitation Act; and the Web Content Accessibility Guidelines (WCAG).

Identify opportunities to build inclusivity in your organization with AI.

For example, recipients of Microsoft grants are creating a hiring platform for neurodiverse applicants, building better and more affordable braille displays for students with visual impairment, and creating a web app to help individuals with speech disabilities communicate more effectively.

Transparency

Transparency is a building block of trust. To achieve and maintain transparency, always be clear about how and when AI is being used, as well as its capabilities and limitations.

Be open about how AI is being implemented and used across your organization.

Stakeholders and employees may feel more confident using AI tools when they understand how the tool arrives at its conclusion, but also are clear regarding its limitations. This transparency helps build their capability for using AI-supported tools and knowing when to supplement its outputs with information outside the tool's scope.

Customers want to know when they're interacting with an AI tool, when AI is being used in decision-making, or when an audio or visual asset has been generated or manipulated with AI. Business leaders should consider how that information will be disclosed to customers.

An example of transparency in practice:

Generative AI is used to create content for a marketing campaign, and the organization identifies which elements are created by AI. A specialist reviews the AI-created content to verify its accuracy so it doesn't mislead customers about the features or capabilities of the advertised product.

Fairness

Fair AI implementation allocates opportunities, resources, and information equitably among the people who use and are impacted by AI.

Ensure that your AI systems provide a similar quality of service and delivery of resources and opportunities to all who use it or are affected by it, across demographic groups.

When using AI tools that describe, depict, or otherwise represent people, minimize the potential for stereotyping or demeaning people—especially those of marginalized groups—to promote fairness.

Include members of different backgrounds, experiences, education levels, and perspectives on the team that manages AI implementation, and identify statistical bias in datasets to help drive fairness in an AI system. Human review by subject matter experts in decisions that use AI can also help protect against biased outcomes.

An example of fairness in practice:

An AI tool is used to review applications and identify priority candidates in the hiring process with oversight from a human resources representative. This person confirms that the tool assesses information accurately, without statistical bias, and has the final review in decision-making.

Make informed decisions about AI and safety

Implementing AI responsibly minimizes risks while allowing your business to benefit from the potential of its various uses. Use the following questions as conversation starters with your team as you begin to think through your AI implementation.

Privacy and security

- □ Is the data accessed by AI systems secured according to your organization's policies for handling sensitive data?
- □ Are you in control of your data, including where it's stored and how it's used?
- □ Is your data secured at all times, including when it is in transit from one system to another?
- □ Do you have quality security tools in place to defend against third-party access or cyberattacks?
- □ Do you have threat identification and response tools in place in case of cyberattacks?

Inclusiveness

- □ Have you confirmed that the tools you intend to use meet accessible design principles?
- □ Do the tools comply with the European accessibility standard, EN 301 549?
- □ Do the tools comply with Section 508 of the U.S. Rehabilitation Act?
- □ Do the tools comply with the Web Content Accessibility Guidelines (WCAG)?

Reliability and safety

- □ Have the tools you intend to use been adequately tested to minimize errors?
- □ Do you have a plan in place to remediate any failures that occur?
- □ Will the tools be regularly monitored for reliability issues?
- □ Are you prepared to comply with all requirements to operate the tools safely?

Transparency

- □ Have you educated stakeholders about how this implementation will work, including its capabilities and limitations, or do you have a plan to do so before they start to use AI tools?
- □ Do you have a plan to communicate with employees about how your organization is going to be using AI and how its outputs should be interpreted?
- □ Have you determined how and when you'll notify customers that they are interacting with AI or viewing AI-generated content?

Accountability

- □ Have you assessed the impact this implementation would have on your employees, organization, and customers?
- □ Have you established a system of oversight and response in case of potential negative impacts?
- □ Have you implemented data governance and management best practices?
- □ Have you determined who will have oversight of AI tools and ensured they have adequate training and control?
- □ Have you made sure this solution is fit for its intended purpose?

Fairness

- □ Have you ensured this implementation will provide the same quality of service to all affected by it?
- □ Have you tested the system's outputs to make sure it will fairly allocate resources and opportunities across demographic groups?
- □ Are outputs of the system free of stereotypes and negative portrayals of marginalized groups?

Protect your data and AI systems

Secure your AI tools now and in the future

Implemented incorrectly, any technology that has access to sensitive data can present a security risk for businesses. Because AI systems require a large amount of proprietary data, it's critical to prioritize security from the start when considering AI implementation or procuring AI solutions.

At Microsoft, we launched our Secure Future Initiative, bringing together our learnings to address and prepare for the increasing scale and heightening risks of cyberattacks in the age of AI. The Secure Future Initiative identifies three principles that Microsoft upholds to help secure the entire digital ecosystem:

1. **Secure by design**
   Security comes first when designing any product or service.
2. **Secure by default**
   Security protections are enabled and enforced by default, require no extra effort, and aren't optional.
3. **Secure in operations**
   Security controls and monitoring will be continuously improved to meet current and future threats.

Security is the essential underpinning of any AI implementation. When you ensure basic security hygiene practices, you protect your data, your people, and your devices from more than 98% of cyberattacks.[3]

Effective practices for security hygiene include:

- Enabling multifactor authentication to protect against compromised user passwords and help provide extra resilience for identities.
- Applying Zero Trust principles, which involves explicit verification, use of least privileged access, and assumption of breach, to limit the impact of an attack.
- Using extended detection and response and antimalware to automatically block attacks and gain insight into the security operations software for faster response.
- Ensuring systems are up to date with the latest versions of firmware, operating systems, and applications.
- Implementing the right protections for critical data, which requires knowledge of which data is most important and where it's located.

Manage your AI systems with governance

A strong governance model helps build the solid foundation for responsible AI implementation. It's the role of governments and regulatory bodies to maintain baseline requirements to minimize adverse effects of AI use on society. Commercial organizations also have an ethical responsibility to create a governance structure to manage their own development or use of AI systems according to their organizational values, local laws and regulations, and the greater good.

Establishing your own governance

When creating your organization's system of governance, remember that the purpose of governance is to adhere AI solutions to company policy and responsible AI principles through a series of policies and procedures. This includes applying policies for assessing and implementing third-party AI solutions, coordinating stakeholder involvement and education, and producing documentation to inform employees, customers, and other users about AI tools.

Risks

Map

Mapping risks is the first stage of governing AI and should inform decisions about a tool's safety, reliability, and fitness for purpose. Mapping risks involves running AI impact assessments and privacy and security reviews—including red teaming and stress testing.

Measure

Measuring risks involves developing metrics by which to assess any identified risks and testing planned mitigation methods to determine how effective they will be.

Manage

Managing risks requires organizations to consistently monitor performance. At this stage, you should identify opportunities for user agency and educate stakeholders about responsible use. Human review and oversight should be included in the management process, as well as best practices for transparency according to the responsible AI principles.

Realize the potential of AI

How companies have transformed with safety in mind

Used responsibly and securely, AI can improve business operations, help your organization address your most pressing challenges, and establish trust and confidence with your customers. Learn from real-word examples and impact to understand how.

Establish a responsible AI council

Wipro, with 240,000 employees operating in 65 countries, is uniquely placed to evaluate the potential of AI tools to transform how people work. Wipro employees have been trained on the principles of GenAI and the organization remains committed to using the technology to help deliver value to customers faster and improve outcomes across the business.

To bring the power of AI to the workforce, Wipro leaders first established an AI council to establish the best approach to launching an AI program internally. In bi-weekly sessions, Wipro leaders drafted a responsible, persona-driven model for how to implement Microsoft 365 Copilot across their global business. Establishing distinct personas, including sales employees, developers, and those in the CTO organization, Wipro discovered that various Microsoft 365 Copilot modules complemented distinct roles.

Learn more

Protect vital research and sensitive data

Oregon State University (OSU), an R1 research-focused university, places a high priority on safeguarding its research to maintain its esteemed reputation. It must foster an open environment conducive to collaboration with other institutions and researchers while simultaneously ensuring that all their research remains secure and compliant with relevant standards.

In response to an incident, OSU created their Security Operations Center. OSU integrated Microsoft 365 A5 licensing and committed to a Zero Trust approach to cybersecurity by widely deploying Microsoft Sentinel and Microsoft Defender. OSU estimates that they achieved five years of maturity in roughly two years due to the rollout of the security capabilities technology and Microsoft support and consulting that helped them maximize the use of that tooling.

Learn more

IFAD
Investing in rural people

Extracting the right data at the right time

With a global mission to eradicate poverty and hunger in some of the world's most fragile rural communities in developing countries, IFAD (International Fund for Agricultural Development) works hand in hand with governments to fund projects and innovative practices that enable small-scale producers to develop sustainable agricultural practices, transform food systems, and build resilience in the face of some of the biggest global challenges. IFAD needed a better way for its global teams and partners to access and share critical information from and to the remote locations they work in.

Using Microsoft Azure OpenAI Service, Azure AI Search, and Power BI, the organization built Omnidata, a centralized analytics platform that connects data, dashboards, visualizations, and analytics powered by machine learning and AI. The solution gives all personnel fast, direct access to critical data in every region IFAD serves, plus training in analytics and machine learning so they can actively create new tools to address specific day-to-day challenges.

Learn more

avanade

Reduce complexity and increase security

Avanade, a joint venture between Accenture and Microsoft, is a consultancy firm that helps drive digital transformation at organizations across the globe. Avanade maintains a large and complex data estate. The company uses data and analytics for a variety of activities across the organization, from powering decision-making to ensuring the best outcomes for their clients.

Previously, it used multiple resources including Microsoft Azure Synapse Analytics, Azure SQL Database, and Power BI to manage data and analytics. This required the IT team to duplicate data every time a user needed to pull information from one tool into another. Now, Avanade is moving to Microsoft Fabric, which incorporates all its preferred tools. This eliminates the need to duplicate data, reducing complexity, saving time, and providing a better employee experience.

Learn more

Advance the sustainability of AI

Just as security is an essential foundation of responsible AI, sustainability is crucial to the conscientious use of it. A powerful tool in understanding and reducing environmental impact, AI can help businesses advance their sustainability goals and help private businesses and public institutions better understand and pursue environmental conservation, resource management, and climate change mitigation.

With AI data management tools and reporting, organizations can bring visibility into their sustainability activities so they can record, report, and reduce their environmental impact. AI can provide the insight necessary to make more informed decisions and stay on track toward your sustainability goals.

At the same time, we recognize the resource intensity of these applications and the need to address environmental impact from every angle.

Business leaders can make their own datacenters more sustainable or work with providers that are already taking these actions to reduce the environmental impact of the datacenters that fuel their AI solutions.

At Microsoft, we're deeply invested and are increasing our focus in three main areas in line with our sustainability commitments: optimizing datacenter energy and water efficiency, advancing low-carbon materials, and improving the energy efficiency of AI and cloud services—all with the goal of empowering our customers and partners with tools for collective progress.

Begin your AI transformation

By carefully considering responsible AI practices and prioritizing security throughout your organization, you can explore the potential of AI to help grow your business.

Our commitments and capabilities make it possible for you to accelerate AI transformation with confidence, and you can trust Microsoft to put your AI security, privacy, and safety first.

Learn more about Microsoft AI to begin your journey.

Discover how Microsoft is committed to responsibly developing and advancing AI.

Explore how Microsoft helps you safeguard AI with comprehensive security and governance solutions.

Sources

[1] "What Business Leaders Really Think About Generative AI," INSEAD, April 11, 2024, https://knowledge.insead.edu/leadership-organisations/what-business-leaders-really-think-about-generative-ai.

[2] "2024 Microsoft Responsible AI Transparency Report," Microsoft, accessed July 10, 2024, https://www.microsoft.com/corporate-responsibility/responsible-ai-transparency-report.

[3] Quy Nguyen, "Basic cyber hygiene prevents 98% of attacks," Microsoft Tech Community, September 18, 2023, https://techcommunity.microsoft.com/t5/security-compliance-and-identity/basic-cyber-hygiene-prevents-98-of-attacks/ba-p/3926856.