



Responsible AI guidelines for decision makers

Best practices for intelligent apps

Contents

Introduction

Addressing the need for responsible AI 3

Chapter 1

Barriers to using AI responsibly 4

Chapter 2

How to use the eight guidelines 6

Chapter 3

Eight guidelines for responsible AI in intelligent apps 7

Guidelines 1–3: Assess and prepare 8

Guidelines 4–6: Design, build, and document 12

Guidelines 7–8: Validate and support 15

Conclusion 18

Introduction

Addressing the need for responsible AI

For many organizations, AI presents an opportunity to innovate or modernize intelligent apps and unlock unprecedented efficiency, personalization, and insight. Using AI, machine learning, and data analytics, intelligent apps are poised to transform the way businesses create and deliver value to their customers.

However, organizations need to ensure that their eagerness to embrace AI doesn't overshadow their responsibility to use it ethically. As your organization begins its intelligent app strategy, you must be vigilant about your innovations' societal, environmental, and reputational impacts. Responsible use of AI requires a careful balance.

This balance is crucial in an era where data-driven intelligence can redefine industries and user experiences. You must be able to use AI capabilities while avoiding actions that could cause harm. For many organizations, there are several barriers to striking that balance.

An IDC survey found that, in the next 12 months

“**//** roughly a third of respondents believe that organizations will prefer to buy AI software from a vendor or use in-house support alongside vendor-supplied AI software for specific use cases or application areas.”¹

This preference indicates a growing demand for AI as a strategy for success.

Barriers to adopting AI for intelligent app innovation

Barrier	Description	Mitigations
Complexity	Creating intelligent apps involves navigating intricate technologies such as machine learning and generative AI. Concerns arise about the scarcity of skilled talent, the challenges in integrating these technologies with existing systems, the need to secure relevant and high-quality data, and the best way to understand AI's extensive capabilities and limitations.	Use integrated, fully managed tools and services that simplify complex AI and data analytics processes.
Fear of the unknown	Some organizations face issues around potential customer backlash and a negative perception of AI and labor issues. Additionally, there is a worry that substantial investments in AI might not guarantee user acceptance, especially as some users who are slower to embrace new technology may doubt the effectiveness and precision of AI capabilities.	Partner with a trusted tech vendor that takes actionable steps to research and understand AI applications and their potential impacts on human populations.
Compliance and security risks	There are concerns regarding data security, privacy, and safeguarding, especially in regulated sectors like healthcare and financial services. Furthermore, the need for explainable AI, adherence to responsible AI practices, and the compliance issues linked to the difficulty in tracing AI decision-making processes are full of risks.	Use cloud-native tools and services with built-in security and compliance to ensure data, app, and AI systems are protected at every level of production and deployment.

For many organizations, meeting these challenges will be a new experience. Despite the potential risks associated with integrating AI, most executives feel the benefits are worth the leap. The key to overcoming these perceived risks is implementing guidelines that empower you to use AI responsibly with an awareness of the potential impacts of your intelligent apps.

Benefits of innovating with AI

- ✓ **Go to market quickly with differentiated intelligent apps and services**
Build and deploy solutions and digital experiences that drive long-term business results and establish your organization as an industry leader.
- ✓ **Increase efficiencies while reducing tech debt**
Find ways to optimize resources and eliminate unnecessary costs.
- ✓ **Attract and empower developer talent**
Give your teams the tools to build quickly, regardless of skill level.

Responsible AI principles

These are the six principles that Microsoft has identified as being critical for AI development and use.

- 1 Fairness**
AI systems should treat all people fairly.
- 2 Reliability and safety**
AI systems should perform reliably and safely.
- 3 Privacy and security**
AI systems should be secure and respect privacy.
- 4 Inclusiveness**
AI systems should empower everyone and engage people.
- 5 Transparency**
AI systems should be understandable.
- 6 Accountability**
People should be accountable for AI systems.



68%

“
68% of executives believe that the benefits of generative AI outweigh the risks, compared with just 5% that feel the risks outweigh the benefits.”²

—Gartner

How to use the eight guidelines

Ultimately, responsible AI is about driving a cultural shift within an organization. Truly operationalizing responsible AI practices requires broad leadership, governance, processes, and talent changes.

The eight guidelines outlined in this e-book focus on intelligent app development and delivery processes to provide a formalized way to anticipate and mitigate the risks of AI systems. Since every use case brings a unique context and set of challenges, the guidelines aren't intended to be used as a checklist or to prescribe specific design choices.

Furthermore, your organization must view intelligent app development through the lens of a rapidly changing regulatory environment. As such, these guidelines shouldn't be viewed as a tool for achieving regulatory or legal compliance. Teams should always work with relevant internal departments to ensure that the AI apps they create adhere to all applicable laws and regulations within the jurisdictions where the app will be developed, used, or marketed.

These guidelines will help you create a deeper synergy between performance, organizational goals, and values. Throughout the app development lifecycle, leaders should remember your organization's values and AI ethical principles to the extent your organization has formalized them. Inevitably, issues will arise that require intelligent app teams to make judgment calls. In these situations, you should rely on values and principles to guide your thinking. You should also seek diverse perspectives from within and outside your organization to provide you with input from all relevant stakeholders.

“

As the costs of risks associated with AI rise, the ability to assess those risks and to engage workers at all levels in defining and implementing controls will become a new source of competitive advantage.”³

— McKinsey

These guidelines have been organized according to the key phases of the app development lifecycle while recognizing that intelligent app development often cycles through these phases iteratively:

- 1 **Assess and prepare**
- 2 **Design, build, and document**
- 3 **Validate and support**

Eight guidelines for responsible AI in intelligent apps

Responsible AI ensures that apps are designed, developed, and deployed in a way that respects human values, rights, and dignity. Responsible AI also helps to prevent or mitigate the potential harms of AI, such as bias, discrimination, privacy violations, or malicious manipulation.

Building and modernizing apps with responsible AI guidelines

These eight guidelines can help app development teams infuse the principles of responsible AI into their processes. Review the table below to see how the guidelines align with the responsible AI principles above.

Phase	Guideline	Action
Assess and prepare	1 Merit assessment	Assess the merit of developing the app considering organizational values and business objectives.
	2 Team creation	Assemble a team that reflects diverse perspectives with clearly defined roles and responsibilities.
Design, build, and document	3 Impact analysis	Assess the potential app impact by including input from domain experts and potentially impacted groups.
	4 Risk discovery	Evaluate the data and system outcomes to ensure inclusivity and discover potential risk.
	5 Risk mitigation	Design AI apps to mitigate potentially negative impacts on society and the environment.
	6 Human control	Incorporate features to enable human control and ensure accountability.
Validate and support	7 Validation	Validate intelligent app performance for reliability and safety, testing for unplanned failures as well as foreseeable misuse unique to AI apps.
	8 Communication	Communicate design choices, performance, limitations, and safety risks to end users.

Applying the guidelines

Each guideline comes with a series of questions to help you engage your teams with responsible AI. The questions inspire critical thinking and help teams proactively surface risks. Certain questions will apply to some apps more than others, but they're designed to be broadly applicable. The questions are not all-inclusive but reflect best practices drawn from Microsoft and Boston Consulting Group's experience developing and delivering responsible AI apps. Several illustrative use cases demonstrate how intelligent app development can use these questions to deliver responsible solutions.

Across all phases of app development (e.g., assess and prepare), you should regularly convene your app development team, discuss relevant questions, and review prior answers. If the team cannot answer one of the questions sufficiently, the team should invest time in developing a specific action plan. App development teams should view these best practices as a starting point for discussion and planning.

Guidelines 1–3

Assess and prepare

Guideline #1: Merit assessment

Action: Assess the merit of developing the app considering organizational values and business objectives.

Questions to ask:

- › What are the primary use cases and benefits of the proposed intelligent app? Which uses are explicitly out of scope?
- › What is the desired business outcome for this intelligent app? How will the business impact be measured?
- › How might the operation of the intelligent app and the use of its outputs for business decisions encroach on core organizational values?

//

Designing and testing human interaction with AI systems as early as possible in the development process, even before teams invest in engineering, can help avoid costly failures and redesign."⁴

— Microsoft Research

Guideline #2: Team creation

Action: Assemble a team that reflects diverse perspectives with clearly defined roles and responsibilities.

Questions to ask:

- › Do you have a diverse (e.g., gender, age, ethnicity), multidisciplinary team with a range of functional expertise?
- › What perspectives or expertise are missing and how can you introduce them, including sources outside the team or organization?
- › Is the team structured so that domain experts can impact relevant design choices?

Guideline #3: Impact analysis

Action: Assess the potential app impact by including input from domain experts and potentially impacted groups.

Questions to ask:

- › What are the foreseeable modes of failure for this app? What edge scenarios could lead to failure and harm?
- › What are the societal and environmental implications of foreseeable app failure, misuse, or malicious attack?
- › What are the app's potential unplanned uses?
- › What external SMEs or groups could provide input informing design choices that would reduce the risk of negative societal impact and harm to individuals directly or indirectly affected by the intelligent app?

Assess and prepare scenario: Automating loan approval decisions

Intelligent app use case

A financial institution wants to automate its loan application approval process to increase efficiency and reach new customer bases. The intelligent app leader immediately recognizes that they must proceed with care because of the intelligent app's potential to impact individuals' economic well-being and quality of life.

Applying guideline #1: Assess the merit of developing the intelligent app considering organizational values and business objectives

Before assembling the team, the intelligent app leader discusses potential approaches with a data scientist and a user researcher. After an hour of whiteboarding, they conclude with the following assessment:

- › The use case is designed to predict an applicant's earning potential over the life of the loan under review, and thus the likelihood of repayment.
- › The business impact should be measured with three KPIs: default rate against a historical baseline, increase in overall loan volume, and speed of approval/rejection decision.
- › There may be race, gender, or other biases reflected in the historical data. Exploratory data analysis will need to include steps to bias assess the dataset composition so that appropriate mitigations can be implemented.
- › There is a significant risk that proxy variables for race, gender, and other protected categories could reinforce biases and impact outcomes, violating company values around fairness and the law. The intelligent app team will have to engage on these topics proactively.

Applying guideline #2: Assemble a team reflecting diverse perspectives and with clearly defined roles and responsibilities

The intelligent app leader builds the team based on the initial assessment. The team should include:

- › Data scientists with experience applying fairness tools to machine learning models.
- › A lawyer with expertise in the Equal Credit Opportunity Act to make sure the team fully understands the regulatory environment.
- › Two loan officers with extensive experience helping various customers navigate the application process. This will enable the team to design the intelligent app to augment the decision-making capabilities of loan officers and integrate their feedback in real time.
- › An intelligent app leader who knows that system security and privacy are critical since the system will deal with sensitive Personally Identifiable Information (PII). The leader must gain approval to bring on an external consultant with cybersecurity and AI expertise to guide optimal system architecture, data storage, and differential privacy across the build.
- › A user researcher to ensure customer needs are at the forefront, from idea inception to final intelligent app, so that customer understanding and satisfaction can be reconciled with business goals.
- › A designer to improve intelligent app usability and accessibility across a diverse user base.
- › A diverse team reflecting a variety of backgrounds and lived experiences.

Applying guideline #3: Assess potential intelligent app impact by including input from domain experts and potentially impacted groups

During a preliminary discussion, the loan officers share their experiences engaging with the financial institution's current customer base. As they detail interactions with various customers over the years, the intelligent app leader realizes that, hopefully, the AI will engage new customer demographics with needs and expectations that differ from those captured in historical data.

- › The team engages an economist with expertise in banking relationships within the demographic groups to whom the bank might expand its activities.
- › The team works with a vendor to deploy a survey for potential borrowers in new customer communities to understand how greater access to credit might impact earning potential—in ways both consistent and inconsistent with communities overrepresented in the historical data—and thus the likelihood of repayment.

The loan officers note that even with a new customer base, one thing is likely to remain constant—having a loan application declined is an unpleasant and potentially painful experience. The denial can be delivered more respectfully by a skilled and experienced professional. Based on this insight, the intelligent app team decides that all rejection decisions will be communicated to applicants by a loan officer.

Spotlight on intelligent app use cases: Financial services and healthcare

Intelligent apps have a wide range of use cases in industries like financial services and healthcare, where personal and sensitive information require highly ethical treatment, privacy standards, and regulatory compliance.

Financial institutions: AI algorithms analyze vast amounts of transaction data in real time, accurately detecting anomalies and fraudulent activities to help identify suspicious transactions and patterns indicative of fraud.

Insurance companies: By analyzing historical claims data, customer profiles, and external data sources, AI algorithms can flag suspicious claims for investigation, helping insurers mitigate losses from insurance fraud.

Healthcare fraud detection: AI models analyze medical records, billing data, and claims history to identify irregularities and behaviors like billing fraud, prescription fraud, and healthcare identity theft.

Guidelines 4–6

Design, build, and document

Guideline #4: Risk discovery

Action: Evaluate the data and system outcomes to ensure inclusivity and minimize the risk of fairness harms.

Questions to ask:

- › What fairness metrics (e.g., statistical parity, equalized odds), tests, and shipping criteria will you use? How will the intelligent app team validate that the training data, including data collected via APIs, captures the different groups and types of people likely to be impacted by the system's output?
- › How will the intelligent app team measure whether the AI app's outcomes are consistent with the chosen objective (i.e., avoid target leakage), fairness metrics, tests, and shipping criteria across a wide variety of potentially impacted groups or intersections of groups?
- › How will the intelligent app team ensure continued adherence to fairness metrics, tests, and criteria post-deployment?

Guideline #5: Risk mitigation

Action: Design AI apps to mitigate potentially negative impacts on society and the environment.

Questions to ask:

- › Suppose negative impacts (e.g., from system failure, unplanned use, abuse, attack, or simply side effect of normal use) are possible. What design processes (e.g., human-centered design) and choices can reduce, mitigate, or control them?
- › What design choices will help minimize the adverse environmental impact of the intelligent app outputs and related decisions? What design choices are critical to ensuring legitimate and transparent data collection and respect for user privacy?

Guideline #6: Human control

Action: Incorporate features to enable human control and ensure accountability.

Questions to ask:

- › How is the team designing the intelligent app to empower humans by augmenting their decision making, streamlining tasks, or otherwise making them more effective? Which decisions or functions require human oversight as a critical component of the AI app?
- › What mechanisms (e.g., interpretability) will support end-user comprehension of the system to enable continuous audit, monitoring, and human intervention?
- › What intelligent app features allow users to customize AI performance?
- › What channels will the intelligent app utilize to collect live feedback?
- › What intelligent app features will ensure inclusive experiences for people with disabilities?

Intelligent app use cases in retail, e-commerce, travel, and hospitality can have broad applications requiring inclusivity and transparency.

Smart retail

Connected intelligent apps are used in retail environments to enhance the shopping experience by tracking inventory levels, providing location-based offers to shoppers, and offering self-checkout systems that streamline the payment process.

E-commerce platforms

Intelligent apps enable scalable transaction processing and inventory management to ensure smooth and efficient online shopping experiences.

Travel and hospitality

Airlines, hotels, and booking platforms can use intelligent apps to handle transaction processing at scale for reservations, ticketing, and payment processing.

Design, build, and document scenario: Retail demand forecasting

Intelligent app use case

A fashion retail chain hopes to transform its in-store inventory management with AI. Using historical sales data, the company wants to optimize the amount of inventory held in stores to maximize sales per square foot. The intelligent app leader has clarified the business objective and assembled a team, and they are now shifting focus to the design, build, and document stage of work.

Applying guideline #4: Evaluate data and system outcomes to minimize the risk of fairness harms

The intelligent app leader brings the team together for a discussion around fairness. A poorly designed intelligent app could lead to service discrepancies across different demographics within the customer base, violating the company's values.

In brainstorming potential challenges, a data scientist on the team notes that customer feedback for the retail chain varies dramatically across neighborhoods in large metropolitan areas. It appears that stockouts are more common in some locations than others, resulting in certain products being unavailable when customers are ready to buy. Pivoting from that comment, the team aligns on the following approach:

- › The team decides to check for parity of service level based on stockouts reported in its historical data. Cross-referencing the results with census data will allow the team to assess whether stockouts reported are correlated with certain neighborhoods and demographic groups.
- › To prevent unacceptably low service levels in specific locations, the intelligent app team establishes a minimum inventory level for each SKU at each location to ensure minimum service levels that will avoid the possibility that stockouts would impact certain demographic groups more than others.

Applying guideline #5: Design AI apps to mitigate the potential negative impact on society and the environment

The team's sustainability expert asks the group if system usage might have some second-order environmental effects. Optimizing inventory levels could maximize profit at the store level, but smaller and more frequent inventory replenishment would rely on greater air and ground cargo traffic. At the country level, the environmental consequences of that cargo traffic could be significant. Furthermore, since significant amounts of excess inventory would need to be returned by stores, optimized inventory levels could create additional cargo traffic.

- › Using internal logistics data, the intelligent app team builds a feature that highlights tradeoffs between inventory levels and transportation emissions at the store and region level.

Applying guideline #6: Incorporate features to enable human control

A specialist in human-centered design observes that the retailer often learns about new trends from its frontline workers, particularly store managers. The historical data has limited predictive power for trend spotting. By enabling human control over the inventory system and augmenting store managers' decision-making abilities, stores could make on-the-fly adjustments to match changing consumer preferences. Based on further exploration, the team decides to:

- › Design a feedback mechanism through which store managers can indicate emergent trends and newly popular products at a specific location, thus allowing for the pooling of insights across the country to help spot trends and adjust inventory levels accordingly.

Guidelines 7–8:

Validate and support

Guideline #7: Validate

Action: Validate intelligent app performance for reliability and safety, testing for unplanned failures as well as foreseeable misuse unique to intelligent apps.

Questions to ask:

- › How will the team validate the AI app's performance against agreed-upon business KPIs, metrics, tests, and criteria?
- › How will the team validate the AI app's performance against technical standards and benchmarks?
- › What are the target environment and conditions under which this intelligent app can be expected to function properly and safely?
- › How will the system be tested and evaluated for safe and effective operations (e.g., graceful failure) in both business-as-usual and edge-case scenarios?
- › What are the mechanisms for continuously monitoring business, technical, and fairness performance?
- › What actions are needed to detect and prevent model drift to ensure machine learning models don't degrade over time?

Guideline #8: Communication

Action: Communicate design choices, performance, limitations, and safety risks to end users.

Questions to ask:

- › What information and instructions should the intelligent app team provide to the end user(s) to enable safe and reliable use?
- › How will the team ensure end users understand the intelligent app's primary use case, underlying assumptions, and limitations?
- › How will the system outputs be communicated in a way that helps end users understand how the system works?

Validate and support scenario: Predictive lead times in manufacturing

Intelligent app use case

An industrial goods manufacturer is experiencing repeated delays in the delivery of parts purchased from suppliers, which has caused disruptions to the manufacturing schedule and late deliveries to customers.

The manufacturer wants to develop an intelligent app to estimate lead times for supplier-procured components based on historical data to protect key relationships. Providing purchasing managers with an early warning of potential delays will let them proactively engage with suppliers and adjust the manufacturing schedule to avoid missed delivery deadlines.

Applying guideline #7: Validation

Although the intelligent app is expected to positively impact operations and customer relations, erroneous lead time estimates could also create more work for purchasing managers, further damaging supplier relationships and negatively impacting the bottom line.

Furthermore, because assembly is a labor-intensive process requiring specialized skills and safety certifications at different stages, ad hoc manufacturing schedule adjustments could put factory employees at risk.

The team leader gathered the app development team to pose a series of questions to reach alignment on the approach to validating app performance and robustness to unplanned failures. The meeting concluded with an agreement that included:

- › Testing the system's outputs against historical supplier-promise dates to determine the model's ability to flag potentially delayed shipments before it's too late.
- › Inputting a wide range of operational scenarios, including varied suppliers and component types, as well as edge scenarios (e.g., parts and suppliers not found in the historical data).
- › For each scenario tested that requires adjustments to the manufacturing schedule, developing associated work schedules to be validated by factory floor leadership for feasibility and safety.

During the discussion of operational scenarios, a team data scientist noted that factors related to the ongoing COVID-19 pandemic—such as limited trucking capacity and economic shutdowns in certain states—would not be captured in historical data but could impact lead times. To capture these insights, the intelligent app team integrated COVID caseloads in geographies proximate to suppliers to capture the pandemic's potential impact on the manufacturer.

In conversations with factory floor leadership, the team learned how scheduling had recently changed to minimize the risks of COVID-19 exposure. As a result, they concluded that any changes to the manufacturing schedule would have to be consistent with the new scheduling policies.

Applying guideline #8: Communication

The team's UX lead pushes the intelligent app team to consider how best to augment the purchasing managers' current decision-making process.

The team aligns on the following steps:

- › Calculate confidence intervals alongside estimated lead times to enable purchasing managers to leverage system outputs responsibly.
- › Design the dashboard to sort components for a certain product by predicted delay, focusing the end users' attention on prioritizing key products, and engaging problem suppliers early.
- › Make sure the system never automatically updates customers on delivery dates, a task that program managers themselves should continue to perform.
- › Build an additional supplier-level, as opposed to component-level, dashboard to equip the company for strategic engagement with suppliers that consistently struggle to deliver parts on time.
- › Output a draft work schedule based on proposed changes that will support the purchasing manager's decision making regarding adjustments to the manufacturing schedule. The factory floor leadership must sign off on the revised work schedule to ensure adherence to worker safety standards.

Having identified and engaged with the intelligent app's end users earlier, the team conducts several meetings and trainings to ensure that these features will be used effectively. It also designs modular training on tools to be integrated into future onboarding programs for purchasing managers.

Intelligent apps can be applied to improve efficiency in manufacturing, agriculture, and supply chain logistics—but in industries that involve machinery and transportation, intelligent apps must also include strict safety and accountability measures to protect workers' well-being.

Industrial IoT

Connected intelligent apps are used for remote monitoring and control of machinery and equipment, improving efficiency, predictive maintenance, and asset tracking.

Supply chain and logistics

Transaction processing at scale enables efficient processing of purchase orders, inventory management, shipping transactions, and tracking of goods.

Agriculture and farming

Intelligent apps like soil moisture sensors, weather stations, and automated irrigation systems help optimize crop intelligent production and reduce resource waste in agriculture.

Conclusion

Business leaders can be change agents in their organizations' cultural transformations. They can also play an important role in catalyzing and facilitating the complex and sometimes complicated conversations necessary to develop intelligent apps responsibly. Under such circumstances, team members should be made to feel comfortable raising questions around sensitive issues and identifying gaps in expertise and experience. Furthermore, they must be comfortable and confident with the tools and systems they use to innovate new apps or modernize their existing ones.

AI-powered apps are poised to change the way businesses serve customers, discover and process information, and personalize experiences. As you launch your intelligent app strategy, always keep responsible AI front and center so you don't risk overshadowing your innovative ideas with unintended consequences.

Intelligent app scenarios for transforming how you innovate, discover, and serve

- › Provide assistance on any channel with chatbots and virtual assistants
- › Find information in real time with an intelligent search engine
- › Detect suspicious behavior in real time to reduce fraud
- › Personalize interactions with AI-powered recommendation engines
- › Stay ahead of the curve with trend analysis and market research
- › Provide more options for communication with voice and speech recognition
- › Anticipate consumer demands with competitive intelligence and market monitoring
- › Let customers help themselves using AI-powered self-service options
- › Enhance decision making with automated report generation

Infuse your apps with AI using trusted cloud technology and advanced analytics.

Explore Azure Innovate >

Contact Sales >

Learn more about the Microsoft approach to responsible AI >

¹ IDC, [IDC Forecasts Revenue for Artificial Intelligence Software Will Reach \\$307 Billion Worldwide in 2027, Doc #US51345023, October 2023.](#)

² Gartner Press Release, [Gartner Poll Finds 45% of Executives Say ChatGPT Has Prompted an Increase in AI Investment, May 3, 2023.](#)

<https://www.gartner.com/en/newsroom/press-releases/2023-05-03-gartner-poll-finds-45-percent-of-executives-say-chatgpt-has-prompted-an-increase-in-ai-investment>

GARTNER is a registered trademark and service mark of Gartner, Inc. and/or its affiliates in the U.S. and internationally and is used herein with permission. All rights reserved.

³ [Confronting AI risks, McKinsey](#)

⁴ [Advancing human-centered AI: Updates on responsible AI research—Microsoft Research](#)