



7 actions that enforce responsible AI practices

By Fanny Ip

As leaders deploy artificial intelligence to drive success across their industries, enforcing responsible AI practices is critical.

Responsible AI practices serve multiple goals:

- Positioning your organization as a leader in responsible AI usage
- Building trust inside and outside your organization
- Supporting positive impact on the individuals, communities, and environments you serve
- Priming your organization for compliance with evolving AI standards and regulations

Cultivating responsible practices requires more than written principles and verbal commitment — specific actions are needed to take responsible AI from concept to practice. The following recommendations and outlined action steps will help your organization enforce responsible and impactful AI practices as the technology evolves.

In Brief

- Enforcing responsible AI practices is essential for organizations to build trust, ensure positive impact, and comply with evolving regulations.
- Cultivating responsible AI practices must go beyond verbal commitment and written guidelines; specific action is necessary.
- The seven actions organizations can take are: promote safety and security, support validity and reliability, lead with explainability and transparency, establish accountability, build fair and unbiased systems, protect data and prioritize privacy, and design for human-centeredness.

7 action steps toward responsible AI

1. Promote safety and security

AI systems can cause unintended harm to individuals, communities, and the environment through malfunctions and misuse. By prioritizing safety and security, organizations can build trust, foster innovation, and maximize the benefits of AI.

Recommendation:

Design, develop, and deploy AI systems with robust safeguards to prevent harm, ensure security, and mitigate risks.

Actions to take:

- Conduct risk assessments at all AI lifecycle stages.
- Implement robust cybersecurity protocols.
- Perform rigorous testing and validation.
- Maintain human-in-the-loop (HITL) oversight of AI systems.
- Develop and test containment protocols.

2. Support validity and reliability

Inaccurate AI can lead to harmful outcomes and loss of trust. Ensuring that AI systems are accurate, reliable, and consistent in their performance produces trustworthy and valid outputs.

Recommendation:

Design, develop, and deploy AI systems that undergo rigorous testing, validation, and monitoring to ensure accuracy, reliability, and consistency throughout their lifecycle.

Actions to take:

- Use diverse, high-quality data.
- Apply robust validation techniques.
- Continuously monitor AI systems.
- Implement effective error detection.

3. Lead with explainability and transparency

AI systems that operate as “black boxes” can lead to mistrust, misunderstandings, and potentially harmful consequences. Providing clear explanations about decisions and processes to both technical and non-technical stakeholders and ensuring that AI systems are understandable and interpretable builds trust in AI, promotes fairness, and enables educated human oversight.

Recommendation:

Design, develop, and deploy AI systems that prioritize transparency by providing clear documentation, interpretable models, and accessible explanations for their outputs.

Actions to take:

- Use explainable AI (XAI) methods for clear explanations.
- Choose interpretable or easily explained AI models.
- Develop user-friendly interfaces and visualizations.
- Document AI design, data sources, and decisions transparently.

4. Establish accountability

AI systems have the potential to significantly impact individuals and society. Establishing clear lines of responsibility ensures those who create and deploy AI systems are accountable for their outcomes and impacts.

Recommendation:

Design and implement governance structures and processes that assign clear roles and responsibilities for AI decision-making, oversight, and redress of unintended consequences.

Actions to take:

- Define stakeholder roles and responsibilities.
- Monitor and audit AI systems for compliance.
- Implement accountability measures for non-compliance.
- Establish guidelines for legal compliance.

5. Build fair and unbiased systems

AI systems can perpetuate or amplify existing biases. Intentionally building AI systems that are designed and operate in a manner that is fair, unbiased, and non-discriminatory towards all individuals and groups is crucial for building trust in AI.

Recommendation:

Design, develop, and deploy AI systems that prioritize fairness by actively identifying and mitigating biases throughout the entire AI lifecycle.

Actions to take:

- Conduct thorough bias audits of data and algorithms.
- Implement bias mitigation techniques proactively.
- Evaluate AI system performance regularly across diverse groups.
- Engage with diverse stakeholders.

6. Protect data and prioritize privacy

Personal data misuse or breaches can erode trust. Safeguarding the privacy and confidentiality of individuals' data throughout the entire AI lifecycle is not only an ethical requirement but also a legal issue.

Recommendation:

Design, develop, and deploy AI systems that prioritize data privacy by implementing robust data protection measures and adhering to relevant regulations.

Actions to take:

- Implement robust data security measures.
- Minimize data collection to necessary amounts.
- Ensure informed consent and grant individuals control over their data.
- Adhere to data protection regulations.
- Establish data breach response protocols.

7. Design for human-centeredness

AI should serve humanity and contribute to a better future. Designing and deploying AI systems to augment human capabilities, empower individuals, and prioritize human well-being and agency ensures AI aligns with human values and dignity.

Recommendation:

Design, develop, and deploy AI systems that prioritize human values and needs throughout their lifecycle, ensuring that AI serves as a tool to empower humans, not replace or undermine them.

Actions to take:

- Engage end-users and affected communities.
- Design to augment human capabilities.
- Prioritize human oversight and control.
- Enhance user interaction and experience.
- Continuously evaluate impact on well-being.

Incorporating responsible AI practices into your organization's AI initiatives is not just a matter of compliance, but a commitment to ethical innovation. By taking the seven actions — promote safety and security, support validity and reliability, lead with explainability and transparency, establish accountability, build fair and unbiased systems, protect data and prioritize privacy, and design for human-centeredness — leaders can pave the way for trustworthy, impactful AI systems.



[huronconsultinggroup.com](https://www.huronconsultinggroup.com)

© 2024 Huron Consulting Group Inc. and affiliates. Huron is a global consultancy and not a CPA firm, and does not provide attest services, audits, or other engagements in accordance with standards established by the AICPA or auditing standards promulgated by the Public Company Accounting Oversight Board ("PCAOB"). Huron is not a law firm; it does not offer, and is not authorized to provide, legal advice or counseling in any jurisdiction. 24-7759