

Strategies to Implement Safe and Ethical Adoption of Artificial Intelligence

Vini Ehsan, Kim A. Critchlow

College of Management and Human Potential, Walden University, Minneapolis, USA

Email: kim.critchlow@mail.waldenu.edu

How to cite this paper: Ehsan, V., & Critchlow, K. A. (2026). Strategies to Implement Safe and Ethical Adoption of Artificial Intelligence. *Open Journal of Business and Management*, 14, 3108-3157. <https://doi.org/10.4236/ojbm.2026.145153>

Received: July 17, 2026

Accepted: September 27, 2026

Published: September 30, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The absence of ethical guidelines and leadership accountability in AI deployment can result in reputational harm and regulatory noncompliance. Business leaders who struggle to implement AI technologies in ways that are both safe and ethically responsible risk professional credibility and financial instability. Grounded in responsible innovation theory, the purpose of this qualitative pragmatic inquiry research project was to identify and explore effective strategies business leaders use to develop and apply safe and ethical practices when adopting AI technologies for use. The participants were nine business leaders who implemented effective AI use strategies. Data were collected through semistructured interviews and secondary data from publicly available sources. Using thematic analysis, six themes were identified: 1) structured AI governance and policy development, 2) data governance, privacy, and confidentiality protection, 3) human oversight, AI constraint, and quality assurance, 4) phased pilots, controlled experimentation, and use case prioritization, 5) workforce training, communication, and change management, and 6) continuous monitoring, measurement, and adaptive improvement. A key recommendation is for business leaders to adopt AI with clear governance and ethical safeguards to prevent or minimize risks related to data exposure, biased outputs, employee misuse, poor decision quality, and loss of stakeholder confidence. The implications for positive social change may include AI practices that positively influence individuals, communities, organizations, institutions, and broader society because AI systems increasingly may shape decisions across many business sectors.

Keywords

Artificial Intelligence, Ethical AI Adoption, Responsible Innovation, AI Governance, Data Governance, Human Oversight, Workforce Training, Continuous Monitoring, Qualitative Pragmatic Inquiry

1. Introduction

Business leaders across industries are rapidly adopting artificial intelligence (AI) to drive innovation, optimize workflows, and gain competitive advantages. Despite these benefits, many business leaders struggle to implement AI technologies in ways that are both safe and ethically responsible. The lack of clear governance models, employee training protocols, and inclusive oversight mechanisms often leads to challenges such as algorithmic bias, diminished trust, and organizational resistance (Kumar et al., 2023). The absence of ethical guidelines and leadership accountability in AI deployment can result in reputational harm and regulatory noncompliance (Santos et al., 2022). As AI tools become more advanced and integrated into decision-making processes, developing strategic frameworks for responsible adoption is becoming essential to mitigate operational and societal risks.

2. Literature Review

The purpose of this research was to identify and explore the effective strategies business leaders use to promote the safe and ethical adoption of AI technologies. Stilgoe et al.'s (2013), (RI) theory served as the lens through which the phenomenon was examined. RI emphasizes anticipation, inclusion, reflexivity, and responsiveness as guiding principles for governing emerging technologies responsibly (Stilgoe et al., 2013). Applying RI to this project provided a framework to understand how business leaders operationalize ethical intent in strategies when adopting AI. Recent research highlights the importance of leadership in ensuring AI adoption aligns with ethical principles and sustainable outcomes (Uddin, 2023; Perera et al., 2024).

2.1. Theoretical Framework

RI theory is an established theoretical framework that emphasizes placing ethical reflection, societal engagement, and anticipatory governance at the center of the innovation processes. With RI, Stilgoe et al. (2013) underscored the fact that AI and other new technologies are organizational and social inventions, rather than neutral technologies, crafted by values, power, and decision-making. RI presumes technologies come with opportunities and risks and, hence, proactive strategies are adopted by organizations so that innovation benefits ethical accountability and societal good (Blok & Lemmens, 2015). Organizational leaders, according to this viewpoint, must adopt strategies ensuring AI integration does not lead to diluting trust, unfairness, or compliance but ensures benefits that are sustainable and inclusive.

Among the underlying principles of RI are its four key dimensions: anticipation, inclusion, reflexivity, and responsiveness (Stilgoe et al., 2013). Anticipation refers to the leader's responsibility of uncovering and managing potential risks before implementation. Inclusion highlights the importance of bringing diverse stakeholders on board in decision-making processes so that diverse opinions can

be considered. Reflexivity refers to the continuous critical self-examination of leaders on their own assumptions, values, and biases in AI uptake determination. Responsiveness highlights firms' capability of adapting plans according to new knowledge, ethical concerns, or social problems (Owen et al., 2013). These dimensions have direct implications for the real business issue of AI governance since business leaders will be compelled to entrench these principles in their operational and strategic decision-taking so as to avoid unintended harm such as algorithm bias, discrimination, and resistance among customers and workers (Uddin, 2023).

RI also identifies the role of outside forces, such as regulation, public trust, and the expectations of key parties, in establishing business strategy (Wirtz et al., 2023). Organizations adopting AI function under conditions of regulatory uncertainty, fast-paced technological progress, and rising societal scrutiny (Wirtz et al., 2023). Environmental pressures determine the shape of leaders' AI governance policy, investing in ethical oversight, and the application of monitoring processes. For example, the European Union's AI Act and the United States' National Institute of Standards and Technology (NIST, 2023) AI risk management framework have compelled business leaders to structure compliance systems and ethical review procedures beyond voluntary principles (Perera et al., 2024). Against this backdrop, RI provides a valuable framework for explaining the processes of leaders in meeting outside forces and, at the same time, managing innovation and compliance and social legitimacy.

Among the benefits of RI is that organizational strategies and ethical considerations get closely entwined. RI provides a systematic method of embedding accountability in innovation processes through institutional practices such as risk appraisal, dialogue workshops among stakeholders, and continually running monitoring (Von Schomberg, 2013). Findings from case studies indicate that organizations implementing RI-driven governance frameworks are more likely to demonstrate higher levels of public trust, employee acceptance, and sustained innovation performance (Batool et al., 2025).

Still, scholars also present criticisms and flaws of RI. Some argue the framework is abstract and difficult to operationalize in fast-paced business environments, in which time and resources for interacting with stakeholders or conducting ethical foresight analyses remain limited (Novitzky et al., 2020). Other scholars have noted that aligning research and innovation policy with societal values remains difficult in practice, particularly when organizations must translate broad responsible innovation principles into operational action (Novitzky et al., 2020). Regardless of these arguments, RI is an appropriate framework for this project due to the fact that it aligns ethical principles for organizational strategies and offers a lens through which to interpret leaders' capability of translating ethical intentions into operational practices.

The research was conceptualized to investigate the degree to which business leaders' AI adoption plans involve anticipation, inclusion, reflexivity, and responsiveness. Such a framework underlies the investigation of leaders' lived practices

and allows the examination and reflection of their practices at the level of deeper ethical innovation governance. It does this by taking the investigation both into what practices leaders use and into the degree to which those practices align with broader societal expectations and ethical ideals and thus allows it to form a strong conceptual basis from which to derive practical insights capable of resolving the applied business problem of AI being adopted safely and ethically in organizational contexts.

In current literature, RI has been gaining popularity in AI governance, particularly in business and organizational contexts. Uddin (2023), for example, demonstrated that ethical leadership based on principles of RI facilitates openness and trust in AI application, helping leaders align innovation and organizational values better. Similarities can be seen in Perera et al. (2024), who used notions of RI to interconnect AI governance and ESG objectives and demonstrated that firms embedding anticipation and reflexivity in decision-making practices can align better across both regulators' and stakeholders' expectations. These examples indicate how RI is not only theoretical but operational, challenging leaders to embed ethical foresight in day-to-day organizational plans.

Scholars have additionally used RI at cross-sector levels, again showing its applicability. Wirtz et al. (2023) studied the level at which services firms employ responsible AI through customer engagement and operational processes integrating RI principles, illustrating successful response to societal concerns improves long-run customer trust. Subsequently, Batool et al. (2025) undertook a systematic review and found RI serves multilevel AI governance's baseline, linking business practices at different levels of control and oversight, including broader regulative and societal levels. Together, these additions confirm the ongoing value of RI as an overarching framework of analysis and guidance for ethical tech adoptions across sectors and affirm its applicability here, considering my objective of investigating business strategies for the ethical and safe use of AI.

2.2. Business Problem Evidence

The increasing reliance on AI within organizations has produced a paradox for business leaders: while AI adoption promises innovation, operational efficiency, and competitive advantage, it simultaneously introduces new ethical, governance, and compliance risks. Government, industry, and academic literature confirm that the lack of clear strategies for the safe and ethical implementation of AI is a pressing, real-world concern. For instance, the U.S. Government Accountability Office (GAO, 2023a) has cautioned that without formal governance and oversight mechanisms, AI adoption could lead to algorithmic discrimination, cybersecurity vulnerabilities, and public distrust. Similarly, the National Institute of Standards and Technology (NIST, 2023) released the *AI Risk Management Framework* to guide responsible AI use, yet noted that few organizations have established comprehensive processes to operationalize ethical AI principles. These federal reports substantiate that ethical AI governance is not an abstract academic issue but a

national business priority requiring business leadership attention.

Industry data reinforced the conclusion that ethical AI governance is a national business priority. A 2024 Deloitte global survey on AI governance found that 56% of organizations reported concerns about unintended bias, privacy violations, and reputational risk from AI-driven decisions (Fancher et al., n.d.). Only 38% of respondents indicated they had formal ethical frameworks or dedicated governance teams in place. Similarly, a 2023 report by IBM revealed that while 85% of CEOs recognize AI ethics as a competitive advantage, fewer than 30% have translated those values into actionable leadership strategies (IBM, 2023). These findings align with recent peer-reviewed studies showing that organizations struggle to convert responsible innovation principles into structured policies and measurable outcomes (Uddin, 2023; Perera et al., 2024). Collectively, these sources affirm that ethical AI adoption is not yet standard practice despite its growing importance across business sectors.

The academic literature provided further evidence that this issue persists across industries and geographic contexts. Wirtz et al. (2023) emphasized that ethical AI governance is hindered by fragmented accountability structures and inconsistent oversight at the leadership level. Similarly, Batool et al. (2025) argued that while regulatory frameworks are evolving, business leaders often lack the internal capacity and technical literacy to implement those frameworks effectively. These challenges echo GAO's (2023a) findings that leadership guidance and accountability structures are vital to reducing the risks of misuse and inequitable outcomes. Thus, the intersection of technology, ethics, and leadership remains a central point of tension that demands applied inquiry.

Critically analyzing the existing research reveals a gap between theoretical frameworks and real-world application. The RI theory, introduced by Stilgoe et al. (2013), continues to shape scholarly discourse by emphasizing anticipation, inclusion, reflexivity, and responsiveness in technology governance. However, recent studies have pointed out that while RI provides strong conceptual guidance, few organizations have effectively operationalized it (Zhang et al., 2023). This misalignment highlights the practical business problem leaders recognize ethical AI as necessary but lack actionable strategies to integrate it into daily operations. The integration of RI theory within organizational policy, therefore, serves as a critical pathway for bridging ethical aspiration with managerial execution.

Government frameworks such as NIST's *AI Risk Management Framework* (2023) and the European Union's *AI Act* (2024) have accelerated global expectations for ethical compliance, signaling that businesses can no longer rely on voluntary self-regulation. Yet, as Fancher et al. (n.d.) and IBM (2023) demonstrated, implementation remains inconsistent, especially among small to mid-sized organizations lacking specialized AI oversight teams. These gaps validate the premise of this study: that business leaders require structured, ethical, and strategic approaches to AI adoption that balance innovation with moral responsibility.

While government agencies and international organizations are establishing reg-

ulatory standards to mitigate AI risks, current evidence indicates that implementation at the organizational level remains inconsistent and fragmented. *The Organisation for Economic Co-operation and Development* (OECD, n.d.) *AI Principles* emphasized transparency, human-centered values, and accountability as fundamental components of responsible AI, yet global compliance assessments show that only a minority of firms have instituted these principles in practice. Similarly, the European Union's *AI Act* (2024) classifies AI systems by risk level and requires organizations to ensure human oversight and fairness; however, a 2024 PricewaterhouseCoopers (PwC) report revealed that 63% of businesses across Europe and North America are still unprepared to meet these compliance obligations (PwC, 2024b). These findings collectively demonstrate that while policies and frameworks exist, business leaders frequently lack the strategies, resources, and technical literacy to operationalize ethical AI effectively (OECD, n.d.; PwC, 2024b).

Moreover, recent industry research underscores the economic and reputational costs of poor ethical governance in AI deployment. Poor ethical governance in AI deployment may increase reputational, compliance, and operational risks when organizations lack clear oversight, accountability, and risk management practices (GAO, 2023b; NIST, 2023). These outcomes reinforce the necessity for leadership-driven approaches to AI ethics. Perera et al. (2024) emphasized that when ethical leadership principles and responsible innovation frameworks are absent, AI-driven decision-making processes can exacerbate inequality and bias. Conversely, integrating ethical leadership practices enhances compliance, customer satisfaction, and organizational sustainability. Thus, the business problem extends beyond theoretical ethics; it has direct implications for competitiveness and organizational longevity in both private and public sectors.

Finally, critical evaluation of recent literature reveals that many organizations still rely on reactive rather than proactive approaches to AI ethics. Studies by Batoool et al. (2025) and Handiman et al. (2025a) found that businesses tend to address ethical concerns only after incidents occur, rather than embedding governance principles during the design and deployment phases. This reactive behavior aligns with findings from the U.S. Government Accountability Office (GAO, 2023b), which stressed that insufficient leadership accountability and lack of ethical foresight are key barriers to responsible AI governance. These combined insights reaffirm that the absence of structured leadership strategies remains a current and pressing business problem. Bridging this gap through targeted leadership frameworks grounded in responsible innovation (Stilgoe et al., 2013) is essential to ensuring that AI-driven transformation supports both organizational goals and public trust.

In conclusion, a convergence of evidence from government agencies, industry analyses, and peer-reviewed scholarship demonstrates that the responsible adoption of AI remains an unresolved and highly relevant business challenge. Despite the establishment of global frameworks such as NIST's *AI Risk Management Framework* (2023) and the European Union's *AI Act* (2024), organizations continue to

struggle with translating ethical principles into actionable strategies. Leaders face mounting expectations from regulators, consumers, and stakeholders to uphold fairness, transparency, and accountability in AI decision making, yet structural and knowledge-based gaps persist. This synthesis of evidence underscores an urgent research need to identify and document effective, repeatable strategies that business leaders can apply to govern AI responsibly. Filling this gap will not only advance the theoretical foundations of responsible innovation but also provide practical guidance for organizations seeking to balance technological advancement with ethical integrity, risk mitigation, and sustainable performance.

2.3. AI in Business Management

AI has transformed the landscape of modern business management by introducing unprecedented opportunities for operational efficiency, strategic decision-making, and data-driven innovation. Organizations across industries are increasingly relying on AI systems for predictive analytics, customer engagement, supply chain optimization, and risk management. The global AI market is projected to surpass \$1 trillion by 2030, underscoring the strategic importance of AI integration for competitive advantage (PwC, 2024a). As AI systems continue to evolve, leaders are redefining business models and decision-making frameworks to harness their potential. However, this transformation brings with it complex ethical, governance, and compliance challenges that require leaders to balance technological advancement with responsible management. The significance of this business topic lies in the duality of AI's role—it serves as both a driver of innovation and a source of ethical and operational risk (Uddin, 2023).

AI's integration into business operations has fundamentally changed how organizations approach performance and growth. According to the OECD (n.d.), 72% of large enterprises across OECD countries have adopted AI-enabled systems in at least one major function, such as finance, human resources, or logistics. While this adoption increases productivity, it also raises questions of accountability, bias, and fairness in algorithmic outcomes. AI's predictive and analytical capabilities depend on vast datasets that may perpetuate social inequities if not properly managed (Perera et al., 2024). Therefore, while AI enhances business decision making, it also challenges leaders to maintain ethical oversight and ensure alignment between technological decisions and organizational values.

The transformative impact of AI extends beyond automation and analytics; it is now redefining organizational structures and leadership roles. Modern enterprises are transitioning from traditional hierarchies to digitally empowered ecosystems, where AI-driven insights influence both strategic planning and employee decision making. Companies that fully integrate AI into core business functions report a 40% improvement in decision-making speed and a 25% reduction in operational inefficiencies. However, these advances necessitate new forms of leadership competence—leaders must possess not only technical literacy but also ethical discernment to interpret and act on AI outputs responsibly (Handiman et al.,

2025b). Without this balance, organizations risk delegating critical judgment to algorithms, leading to potential errors, reputational harm, or inequitable outcomes.

The rapid integration of AI has also redefined competitive advantage by emphasizing the role of data as a strategic asset. Businesses that successfully leverage AI often exhibit superior data governance and cross-functional collaboration capabilities, enabling them to transform raw data into actionable intelligence (Fancher et al., n.d.). Yet, this reliance on AI-driven insights introduces vulnerabilities related to data privacy, model transparency, and intellectual property. A growing body of literature warns that inadequate oversight in AI-enabled decision systems can lead to biased outcomes that undermine consumer trust and stakeholder confidence (GAO, 2023b). Consequently, business leaders are increasingly tasked with developing responsible innovation strategies that ensure AI is implemented in ways that enhance performance while safeguarding ethical integrity and public trust.

2.4. Ethical AI Adoption and Governance

The ethical adoption of AI is a central concern for modern business leaders. As AI systems become embedded in decision-making processes, their outputs influence hiring, lending, healthcare access, and even public safety. NIST (2023) emphasized that responsible AI use requires governance frameworks that ensure transparency, explainability, and accountability in AI systems. Without structured governance, AI can inadvertently reinforce discrimination, infringe on privacy, or generate decisions that lack human oversight (GAO, 2023b). To mitigate these risks, business leaders must implement governance mechanisms grounded in ethical theory and practical oversight.

RI theory, developed by Stilgoe et al. (2013), provides a useful conceptual framework for guiding ethical AI governance. RI emphasizes four dimensions—anticipation, inclusion, reflexivity, and responsiveness—that help leaders foresee potential ethical risks, engage diverse stakeholders, and adapt strategies in real time. Recent research shows that organizations embedding RI principles into AI governance improve both employee trust and regulatory compliance (Uddin, 2023; Wirtz et al., 2023). For example, companies adopting anticipatory governance models—where ethical impact assessments are performed during AI system design—demonstrate fewer incidents of bias and system failure (Batool et al., 2025). Thus, ethical governance is not only a moral imperative but also a strategic one, as it directly correlates with sustained business performance and stakeholder trust.

2.5. Organizational Challenges in Responsible AI Implementation

Despite regulatory progress and increased awareness, many organizations struggle to implement AI responsibly due to structural, cultural, and resource-based barriers. Fancher et al.'s (n.d.) *State of AI Governance Report* reveals that only 38% of organizations have established a formal AI ethics committee or governance body. Common barriers include limited technical expertise, fragmented accountability structures, and a lack of clear ethical standards across departments. Leaders

often face competing priorities, accelerating innovation while ensuring fairness and compliance, which creates tension between business imperatives and ethical obligations (GAO, 2023b; NIST, 2023).

Additionally, small and medium-sized enterprises (SMEs) encounter unique constraints in implementing ethical AI. Unlike large corporations, SMEs often lack access to specialized data governance resources or compliance officers (PwC, 2024a). As a result, many rely on third-party AI vendors whose models may not fully align with ethical or legal standards. These dependencies exacerbate risks related to data misuse and algorithmic opacity. To address these challenges, scholars suggest developing cross-functional AI ethics teams that include professionals from legal, technical, and human resources backgrounds to ensure holistic oversight (Perera et al., 2024). Still, implementation gaps persist, highlighting the need for leadership strategies that translate ethical frameworks into actionable policies.

Another significant challenge organizations face in implementing responsible AI is the lack of ethical literacy and workforce preparedness. Employees at all levels must understand not only how AI systems function but also the ethical implications of their use. However, research indicates that most organizations have not adequately invested in employee training related to AI ethics, bias detection, or responsible data handling (UNESCO, 2023). This lack of awareness often leads to unintentional misuse or overreliance on AI outputs without critical evaluation. According to the World Economic Forum (2024), 60% of companies identify “ethical skill gaps” among employees as a major barrier to effective AI governance. These gaps can result in misinterpretations of model predictions, inconsistent ethical decision-making, and difficulty in maintaining transparency. To overcome this, organizations must institutionalize continuous training programs and promote a culture of ethical awareness across all functions rather than confining AI ethics to technical teams.

Furthermore, accountability remains a persistent organizational challenge in responsible AI adoption. The diffusion of responsibility among multiple departments—data science, compliance, legal, and executive leadership—often blurs ownership of ethical decision-making (Wirtz et al., 2023). This fragmentation makes it difficult to assign liability when AI systems fail or produce harmful outcomes. Regulatory bodies such as the U.S. Government Accountability Office (GAO, 2023a) have called for clearer delineation of AI accountability roles, emphasizing the importance of oversight committees that integrate diverse expertise. Some organizations are addressing this by establishing “AI stewardship offices” or appointing Chief AI Ethics Officers responsible for ensuring ethical compliance throughout the AI lifecycle (OECD, n.d.). Nonetheless, creating a unified accountability framework remains complex, particularly in global enterprises operating under differing legal and cultural standards. Overcoming these obstacles requires leadership commitment to embedding ethical principles into the organization’s governance infrastructure, ensuring that every AI-related decision is grounded in responsibility and transparency.

2.6. Leadership Strategies for Ethical AI Integration

Ethical leadership plays a pivotal role in shaping responsible AI adoption. Leaders who model transparency, inclusiveness, and moral accountability influence organizational culture and set expectations for technology use. Uddin (2023) found that ethical leadership significantly correlates with employee trust in AI systems and willingness to engage in ethical innovation practices. Similarly, Handiman, Wijaya, and Setiawan (2025c) emphasize that leaders who integrate AI ethics into sustainability and corporate responsibility agendas strengthen long-term resilience and public legitimacy.

One effective leadership strategy involves fostering “ethical foresight,” which enables leaders to anticipate potential consequences of AI deployment before they occur (Zhang et al., 2023). This approach is aligned with the anticipatory dimension of RI theory. Leaders can operationalize this by instituting mandatory ethical risk assessments during AI project planning and engaging multidisciplinary review boards to evaluate system implications (Batool et al., 2025). Furthermore, leaders should cultivate continuous learning cultures where employees are trained to recognize and address ethical issues in data and algorithmic decision-making. By promoting reflexivity and inclusivity, leaders ensure that ethical AI adoption becomes embedded in the organization’s values and decision-making processes.

In addition to fostering ethical foresight, leaders must embrace adaptive and participatory leadership models that promote inclusivity and shared responsibility in AI governance. Adaptive leadership encourages flexibility and collaboration, enabling organizations to respond effectively to new ethical challenges as AI technologies evolve (Northouse, 2024). Research by Perera et al. (2024) indicates that when leaders engage employees, customers, and external stakeholders in discussions about AI ethics, organizations experience higher levels of transparency and innovation quality. Such engagement helps identify emerging concerns, such as algorithmic bias or privacy risks, before they escalate into reputational or compliance issues. Moreover, participatory governance practices—such as stakeholder workshops and open ethics forums—strengthen mutual accountability and align organizational objectives with societal expectations. These strategies help operationalize the inclusivity and responsiveness dimensions of responsible innovation, reinforcing ethical leadership as a continuous, dialogic process rather than a static policy.

Another essential leadership strategy is embedding measurable accountability structures into AI initiatives. Ethical leadership requires not only clear values but also tangible mechanisms for monitoring adherence to those values. Leaders can implement ethical performance indicators, such as fairness metrics, bias audits, and transparency benchmarks, to evaluate the integrity of AI systems across their lifecycle (NIST, 2023). Establishing these accountability systems signals to employees and external stakeholders that ethical standards are integral to business performance, not optional add-ons. According to Fancher et al. (n.d.), organizations that publicly report on their AI governance metrics achieve stronger stakeholder trust and regulatory readiness. Leaders who communicate openly about AI

successes and failures foster a culture of psychological safety and continuous improvement. This culture ensures that ethical AI integration is sustained not through compliance alone but through shared commitment to moral responsibility and organizational integrity.

2.7. Technological and Regulatory Alignment

Global regulatory developments are driving organizations to rethink their AI governance models. The European Union's *AI Act* (2024) and the U.S. *NIST AI Risk Management Framework* (2023) have introduced comprehensive guidelines to classify and mitigate AI-related risks. These frameworks require that AI systems be transparent, traceable, and human-centric. However, studies show that only a minority of firms are fully compliant with these emerging standards (OECD, n.d.; PwC, 2024b). The lag in compliance stems from the rapid pace of AI development, which often outstrips the ability of organizations to adapt governance systems in real time.

Regulatory alignment requires collaboration between industry, academia, and government. Organizations that proactively engage with regulatory bodies not only improve compliance but also shape evolving policies. Wirtz et al. (2023) note that companies participating in AI ethics consortia and industry partnerships experience fewer compliance failures and greater public trust. Moreover, alignment between organizational policy and external regulation fosters a shared understanding of what constitutes responsible AI. Such alignment ensures that businesses can innovate confidently within a framework that safeguards public interests and mitigates reputational risks.

Technological and regulatory alignment also requires organizations to integrate compliance mechanisms directly into AI system design, an approach often referred to as “compliance-by-design.” This proactive strategy embeds legal, ethical, and transparency requirements into the algorithmic lifecycle—from data collection to model deployment—rather than addressing them reactively (OECD, n.d.). Organizations adopting this approach are better positioned to meet international standards such as the ISO/IEC 42001:2023 for AI management systems, which establishes globally recognized benchmarks for risk assessment, documentation, and accountability (ISO, 2024). Moreover, regulatory alignment is not solely a technical process but a cultural one; it demands that leaders foster organizational awareness of regulatory trends, engage in cross-sector dialogue, and integrate continuous monitoring practices that ensure sustained compliance. As emerging technologies continue to outpace traditional policy cycles, organizations that align their technological innovation with evolving governance frameworks will maintain both ethical integrity and strategic resilience (UNESCO, 2023).

2.8. Cross-Industry Applications and Global Perspectives

The ethical adoption of AI is not confined to a single sector—it is a cross-industry challenge. In healthcare, AI systems improve diagnostic accuracy but raise ques-

tions about data consent and patient privacy. In finance, algorithmic credit scoring introduces efficiency but risks amplifying discrimination (GAO, 2023b). In manufacturing, AI-driven automation enhances productivity but disrupts workforce structures and creates ethical dilemmas around employment and re-skilling (OECD, n.d.). These examples illustrate that while AI applications differ across industries, the ethical challenges share a common foundation: accountability, transparency, and fairness.

Global perspectives further emphasize the importance of shared ethical standards. According to the United Nations Educational, Scientific and Cultural Organization (UNESCO, 2023), a unified global framework for AI ethics is essential to address transnational challenges such as data sovereignty and algorithmic bias. Countries that integrate RI principles into their national AI strategies—such as Canada, Singapore, and the Netherlands—demonstrate higher stakeholder confidence and innovation sustainability (UNESCO, 2023). This suggests that ethical AI adoption is not just a business necessity but a global imperative that transcends geographic and sectoral boundaries.

In the public sector, governments are increasingly leveraging AI for decision-making in areas such as social services, policing, and immigration management, yet these applications introduce unique ethical and governance risks. The U.S. Government Accountability Office (GAO, 2023a) warns that the absence of consistent ethical frameworks in public AI programs can lead to biased decision outcomes, decreased citizen trust, and violations of civil liberties. Similarly, the European Parliament and Council of the European Union's (2024) AI Act establishes a risk-based regulatory model to ensure transparency and human oversight in high-stakes applications. These developments highlight how ethical AI in government operations not only ensures compliance but also reinforces democratic accountability and public trust in institutions. Applying responsible innovation principles in these contexts encourages inclusivity and reflexivity—key elements for balancing technological efficiency with the protection of human rights.

In the education and research sectors, ethical AI adoption has become central to advancing equity and inclusion in learning environments. AI tools such as adaptive learning systems and predictive analytics can personalize instruction and improve student outcomes, but they also raise concerns about data privacy, algorithmic bias, and accessibility (UNESCO, 2023). Research by Zhang et al. (2023) suggests that when educational institutions integrate ethical foresight and stakeholder engagement into their AI governance frameworks, they can mitigate bias while promoting fairness and transparency in learning assessments. These findings reinforce that cross-industry adoption of AI ethics must include diverse perspectives—teachers, students, policymakers, and technologists—to ensure that innovation serves collective well-being rather than perpetuating systemic disparities.

Finally, multinational organizations face the challenge of harmonizing ethical AI practices across jurisdictions with differing legal and cultural expectations. The Organisation for Economic Co-operation and Development (OECD, n.d.) em-

phasizes that global companies must navigate a complex regulatory landscape that includes the European Union's AI Act, the United States' NIST AI Risk Management Framework, and China's generative AI regulations. To maintain consistency and trust, leading firms are adopting global AI ethics charters grounded in RI principles that align their internal governance with international norms (Batool et al., 2025). This convergence of global and organizational ethics underscores the importance of leadership in promoting shared accountability for AI technologies that transcend borders. In essence, responsible AI adoption represents not merely a compliance requirement but a universal commitment to transparency, justice, and sustainable innovation.

2.9. Business Implications of Ethical AI

The business implications of ethical AI adoption are profound. Organizations that implement responsible AI frameworks experience enhanced innovation performance, stronger stakeholder trust, and reduced regulatory risk. Ethical AI practices contribute to long-term value creation by aligning business goals with societal expectations (Perera et al., 2024). Conversely, unethical or poorly governed AI systems can lead to reputational, compliance, financial, and operational risks when organizations lack effective governance, accountability, and risk management practices (GAO, 2023a; NIST, 2023). For example, recent case studies in the technology sector revealed that companies facing ethical controversies related to AI transparency suffered significant declines in market valuation and employee morale.

Moreover, ethical AI directly impacts organizational resilience and competitiveness. Leaders who integrate ethics into AI governance not only prevent harm but also foster innovation through trust and inclusivity (Handiman et al., 2025c). Deloitte (2024) similarly emphasized that ethical maturity and business resilience are connected to responsible AI adoption because organizations that strengthen governance, oversight, and accountability are better positioned to manage AI-related risks while sustaining stakeholder confidence. This alignment between ethical responsibility and performance outcomes illustrates that responsible innovation is not a constraint but an enabler of sustainable growth.

Beyond compliance and reputation management, ethical AI adoption increasingly serves as a differentiator in competitive markets. Organizations that prioritize transparency and fairness in AI operations often experience higher customer loyalty and brand equity (Fancher et al., n.d.). Ethical AI signals to consumers that a company values accountability and social responsibility, which strengthens market positioning—particularly in industries such as finance, healthcare, and technology, where data privacy and trust are paramount (OECD, n.d.). Moreover, research suggests that companies implementing comprehensive AI ethics policies attract greater investor confidence, as responsible governance reduces the likelihood of litigation and long-term operational disruptions (PwC, 2024a). Thus, integrating ethical principles into AI design and decision-making processes not only

fulfills moral and regulatory obligations but also enhances financial sustainability and shareholder value.

Furthermore, ethical AI drives internal organizational transformation by fostering cultures of accountability and continuous learning. When employees perceive that their organization prioritizes fairness and transparency in its use of AI, they are more likely to engage in innovation and ethical risk-taking (Wirtz et al., 2023). This culture of trust improves cross-functional collaboration and supports the development of responsible data practices that mitigate bias and improve model accuracy. As Uddin (2023) emphasizes, ethical leadership in AI promotes a shared sense of purpose that aligns business performance with societal well-being. In this way, the long-term business implication of ethical AI extends beyond technological implementation—it redefines corporate success around principles of sustainability, equity, and public trust.

3. Methods

A qualitative research method with a pragmatic inquiry research design was used for the research. A qualitative approach was appropriate because it enabled an in-depth exploration of the lived experiences and perspectives of organizational leaders as they navigate the ethical adoption of AI. Qualitative methods allow researchers to capture complex, context-dependent insights (Creswell & Poth, 2018). A pragmatic inquiry research design complemented the qualitative method by focusing the research on real-world problem solving and the practical application of knowledge to address organizational challenges.

This method provided the flexibility needed to capture diverse viewpoints and uncover emerging themes that reflect how leaders interpret and respond to ethical challenges in AI implementation. It was therefore the most appropriate method for exploring strategies grounded in real-world organizational contexts.

The pragmatic inquiry design aligned with the focus on practical, real-world solutions. Pragmatism emphasizes finding what works in specific contexts and understanding how individuals translate beliefs into actions (Dewey, 1920; Morgan, 2014). Because my goal with this study was to identify strategies organizational leaders use in practice not merely theoretical constructs pragmatic inquiry provided the ideal structure. This design supports methodological flexibility and values multiple forms of evidence, allowing the researchers to incorporate interviews, observations, and document review to construct a comprehensive understanding of ethical AI adoption (see Patton, 2015). Pragmatic inquiry was therefore well aligned with the applied nature of the business problem.

Nine business leaders participated in the study. The selection criterion was that the participants implemented effective AI use strategies using observable evidence, such as documented governance outcomes or completed AI initiatives, to avoid selecting participants on the basis of the conclusion under investigation. The eligibility criteria were for participants with authority over AI-related initiatives and a minimum of two years of direct involvement in AI adoption, governance, policy de-

velopment, or AI-driven decision-making, and to have used effective strategies to develop and apply safe and ethical practices when adopting AI technologies for use. The primary data collection method consisted of semistructured interviews. Data collection followed a structured, step-by-step process, identified potential subject matter experts through LinkedIn, industry associations, leadership forums, and referrals from participants, and professional networks. Purposeful outreach was conducted using professional communication platforms and direct email invitations. Snowball sampling was used as a supplementary strategy, allowing participants to recommend other qualified leaders who met the eligibility criteria.

See **Table 1** below for a concise participant profile description covering industry, organization size, geographic setting, leadership role, and AI use context.

For the purposes of the table below and the article, the researchers used the following terms and definitions:

1) Enterprise

A large organization with formal governance, standardized business processes, multiple business units, and significant technology, operational, or administrative infrastructure.

2) Enterprise/Large

An organization operating at considerable scale, typically serving a large work force, customer base, or mission scope, and requiring structured governance, risk management, and cross-functional coordination.

3) Enterprise AI/Data

The organizational use of artificial intelligence and data assets within a governed environment that includes policies, security controls, data management practices, compliance requirements, and oversight processes to support business or mission objectives.

4) CPIC

Acronym for Capital Planning and Investment Control, a federal government governance framework used to evaluate, select, manage, and assess information technology investments. CPIC helps ensure that technology initiatives align with organizational strategy, deliver value, manage risk, and meet performance objectives.

Prospective participants were vetted to confirm that they met the eligibility criteria. After confirming eligibility and interest to participate in the research, the participants were required to sign an informed consent and then scheduled for the interview. An interview protocol was used to ensure consistency across interviews, reduce researcher bias, and support dependability in the data collection process. The interview was conducted remotely, via zoom and lasted anywhere between 45 minutes to one hour and one half hour. Secondary data from publicly available sources such as NIST AI Risk Management Framework and the Office of Management and Budget guidance on federal AI governance were also obtained and included to support triangulation of the data and to corroborate participant responses. A reflective journal was included throughout data collection and anal-

ysis. The reflective journal was used to document my assumptions, observations, evolving interpretations, and reflexive insights related to the research process. Reliability was ensured through the use of data saturation, member checking, and methodological triangulation of the data. Triangulation strengthened confirmability by ensuring that interpretations were grounded in evidence rather than researcher bias. Trustworthiness was strengthened and misinterpretations were reduced with the use of member checking that allowed participants to review the accuracy of interpretations and confirm that their perspectives have been captured accurately. Maintaining an audit trail including interview transcripts, reflective notes, coding decisions, and analytic memos allowed the research process to be transparent and repeatable. Dependability was further supported by achieving data saturation, which occurred when no new themes arose during interviews and the data corpus becomes sufficiently rich to support meaningful thematic analysis (Hennink et al., 2017). Data saturation was determined after the eighth interview; however, an additional interview was conducted to be sure there was no new data to emerge.

A cataloging and labeling system using Microsoft Excel was employed to organize interview transcripts, field notes, analytic memos, and secondary data sources. Each participant was assigned a unique pseudonym to protect confidentiality and ensure consistency across datasets. Excel was used to create structured data matrices that include a transcript excerpt column, initial code column, category column, theme column, and analytic notes column. A separate worksheet was used to function as a codebook, documenting code definitions, inclusion and exclusion rules, and example quotes for each code. Color coding was applied as the initial transcription method across worksheets to visually differentiate codes, categories, and emerging themes, supporting pattern recognition and systematic comparisons across participants. This structured cataloging approach supported traceability by allowing each theme to be traced back to coded excerpts and original transcript evidence.

Table 1. Participant profiles.

P1-P9	Industry	Org Size	Geographic setting (f-2-f, remote, domestic, int'l, or regional description)	Leadership Role	AI use context
1	Technology/ AI-enabled business environment	Not specified; organizational/ technical environment	Remote; domestic U.S.	Technical-side AI strategy and implementation leader	AI strategy, AI transformation roadmap, ethical risk assessment, data governance, AI capability development, bias detection, transparency, and business alignment
2	Technology/software/ AI-enabled systems	Small company	Remote; domestic U.S.	Director; leadership team contributor; technical lead/project leader	Internal AI use, intellectual property protection, client-facing AI-enabled products, LLM agents, collaborative autonomy, testing, safety, security, and reliability

Continued

3	Federal IT governance/CPIC/ portfolio management	Large federal/ public-sector environment	Remote; domestic U.S.	CPIC, IT governance, and IT portfolio management SME	AI governance, CPIC integration, portfolio reporting, AI community of practice, metrics, knowledge sharing, and governance board support
4	Federal government/ intellectual property/ public-sector technology	Large enterprise/ federal agency environment	Remote; domestic U.S.	CIO/executive decision maker for AI policy and deployment	AI policy, neural network classification, AI tools for patent and trademark operations, phased pilots, stakeholder feedback, and enterprise scaling
5	Enterprise/corporate technology governance	Not specified; enterprise environment with controlled user groups	Remote; domestic U.S.	Policy/governance leader involved in acceptable use and AI training	Approved AI tools, limited user community, prompt and document oversight, acceptable use policy, AI training, data labeling, and controlled expansion
6	Healthcare/health technology	Not specified; healthcare organization	Remote; domestic U.S.	Strategy and AI implementation participant working with development/technical teams	Healthcare AI, patient data protection, safe AI testing environments, human-in-the-loop review, AI constraints, QA/QC pipelines, dashboards, and model version monitoring
7	Federal/enterprise IT and AI governance	Large federal/ enterprise environment	Remote; domestic U.S.	Enterprise technology leader working with Chief AI Officer and AI governance structures	Microsoft Copilot, Office 365 AI features, AI council, use-case collection, data quality, authoritative data sources, zero-trust access, and enterprise scaling
8	Enterprise/corporate technology environment	Large enterprise	Remote; domestic U.S.	AI strategy and enterprise policy leader	Enterprise AI acceptable use policy, approved LLMs, tenant-contained AI tools, blocked unapproved AI domains, employee training, surveys, A/B testing, and customer satisfaction metrics
9	Research and development/ scientific enterprise environment	Large enterprise/ scientific or R&D user community	Remote; domestic U.S.	Enterprise AI and data governance implementation leader for R&D/business group	Translation of enterprise AI/data governance policies into R&D practices, AI tool risk assessment, procurement review, partnership review, governance approval, pulse surveys, annual ratings, sustainability guardrails, and right-sized model use

Note: Participant Profiles.

Braun and Clarke's (2006) six phase thematic analysis process was used to analyze the data. Codes were organized into preliminary categories using constant comparison across participants. Excel based matrices were used to compare codes across interviews and identify similarities and differences in participant responses.

Conceptual mind mapping was conducted through visual grouping of coded data to explore relationships among categories. Mind mapping was performed by grouping related categories visually within Excel using clustered sections and color coded matrices to explore relationships among codes and identify early theme structures. Mind mapping was performed by clustering related codes and categories visually using Excel grouped sections and color coded matrices to explore relationships and identify early theme structures. Candidate themes were developed by synthesizing related categories into broader patterns that represented shared strategies and experiences across participants. These candidate themes were reviewed iteratively against the coded data and the full dataset to ensure coherence, internal consistency, and clear distinction between themes. MAXQDA software was used to confirm and support the coding process, the identification of patterns, and relationships among themes.

Themes were refined and defined through ongoing analytic memo writing and comparison with current peer reviewed literature to ensure each theme was clearly articulated, grounded in participant data, and directly responsive to the research question.

The final report was produced by synthesizing the finalized themes into a clear and coherent narrative that directly answers the research question. The themes were related to the responsible innovation theoretical framework and current peer reviewed literature to strengthen scholarly alignment and ensure the findings are grounded in both participant experiences and established research. This was accomplished through the effort that focused on the key themes by explicitly mapping each theme to the responsible innovation conceptual framework dimensions of anticipation, inclusion, reflexivity, and through correlating each finalized theme with peer reviewed literature on ethical AI adoption and AI governance, including newly published studies that have emerged since the proposal stage, to strengthen currency and scholarly alignment. This process ensured themes remain grounded in participant data while also demonstrating consistency with the current evidence base and theoretical framework.

4. Results and Discussion

The purpose of this qualitative pragmatic inquiry research was to identify and explore effective strategies business leaders use to develop and apply safe and ethical practices when adopting AI technologies for use. Thematic analysis of data collected from nine participants revealed six major themes: 1) structured AI governance and policy development, 2) data governance, privacy, and confidentiality protection, 3) human oversight, AI constraint, and quality assurance, 4) phased pilots, controlled experimentation, and use case prioritization, 5) workforce training, communication, and change management, and 6) continuous monitoring, measurement, and adaptive improvement. The themes reflected the strategies participants used to govern AI adoption, protect sensitive data, maintain human accountability, test AI tools before scaling, prepare employees for responsible AI

use, and evaluate AI practices over time. The overarching research question guiding this study was: What effective strategies do business leaders use to develop and apply safe and ethical practices when adopting AI technologies for use?

Sources of data included semistructured interviews with nine participants and secondary contextual sources from publicly available documents, federal AI guidance, responsible AI frameworks, and industry publications related to ethical AI adoption. Each of the sources were reviewed to confirm or disconfirm the representation of the context of the statements made by the participants in the sources. Each participant held a leadership, technical, governance, or strategy role and had direct experience with AI adoption, AI governance, data governance, cybersecurity, technology implementation, or responsible technology oversight. The secondary sources provided additional context for interpreting participant responses and strengthened the credibility of the findings. Public documents, including the NIST AI Risk Management Framework and the Office of Management and Budget guidance on federal AI governance, similarly emphasized governance, data protection, human oversight, risk management, measurement, and continuous improvement as important mechanisms for responsible AI adoption.

Interview data were transcribed, reviewed for accuracy, and organized using participant pseudonyms to maintain confidentiality. Participants are referred to throughout the project as Participant 1 through Participant 9. MAXQDA was used to support manual coding, code organization, memo development, and thematic analysis of the interview transcripts. During the first phase of analysis, each transcript was reviewed multiple times to identify meaningful statements related to safe and ethical AI adoption. Initial codes were assigned to recurring concepts within participant responses. These codes reflected frequently mentioned concepts such as AI policy development, data protection, human oversight, controlled pilots, workforce training, stakeholder engagement, and effectiveness measurement.

Through an iterative process of reviewing and comparing coded data, initial codes were grouped into broader categories representing patterns across participant experiences. These categories were then refined into six major themes that addressed the research question. **Table 1** illustrates the relationship between representative codes and the themes that emerged during analysis. Following the coding process, thematic analysis was conducted using **Braun and Clarke's (2006)** six phase approach to qualitative data analysis. This process involved reviewing coded data segments, identifying patterns across participant responses, refining theme boundaries, and organizing findings into overarching themes that directly answered the research question.

The analysis revealed six major themes representing the strategies business leaders use to develop and apply safe and ethical practices when adopting AI technologies. The themes illustrate that responsible AI adoption requires a multifaceted organizational approach that integrates governance, data protection, human accountability, controlled implementation, workforce readiness, and continuous improvement. **Table 2** highlights the major themes, the number of participants who ref-

erenced each theme, and the number of coded references associated with each theme. Collectively, these findings indicate that effective AI adoption strategies are not limited to technical implementation. Instead, the findings show that business leaders must combine policy, process, people, technology, and measurement to ensure AI is adopted responsibly.

Table 2. Codes and corresponding themes identified during data analysis.

Theme	Representative codes
Structured AI governance and policy development	AI strategy components, enterprise AI framework, acceptable use policy, AI council and pilots, governance review, layered risk assessment, vendor risk assessment, policy translation
Data governance, privacy, and confidentiality protection	Data governance, intellectual property protection, tenant contained AI tools, data classification, data access control, confidential information protection, patient data protection, authoritative data ownership
Human oversight, AI constraint, and quality assurance	Human in the loop, AI constraint, ethical controls, validation process, subject matter expert review, QA and QC pipeline, model drift monitoring, version monitoring
Phased pilots, controlled experimentation, and use case prioritization	Phased approach, controlled rollout, safe testing environment, controlled AI exposure, use case prioritization, demand front loading, proof of concept testing
Workforce training, communication, and change management	AI user training, policy awareness, communication and training, stakeholder input, roadshows and workshops, user awareness gap, change management, AI democratization
Continuous monitoring, measurement, and adaptive improvement	Usage tracking, customer satisfaction metrics, A/B testing, pulse score surveys, policy revalidation, performance metrics, operational safeguards, sustainability guardrails

Note. Codes were derived from recurring concepts identified during thematic analysis of semi structured interview transcripts from nine participants. Codes were grouped into initial categories and later refined into themes using [Braun and Clarke's \(2006\)](#) thematic analysis process (see [Table 3](#)).

The findings align with the responsible innovation framework guiding this research project. [Stilgoe et al. \(2013\)](#) emphasized anticipation, inclusion, reflexivity, and responsiveness as core dimensions of responsible innovation. The participants' descriptions of AI governance, risk assessment, policy development, stakeholder engagement, human oversight, and continuous monitoring illustrate how business leaders operationalize these principles in practice. Anticipation appeared in participants' emphasis on risk assessment, phased pilots, data governance, and defined success criteria before broad AI deployment. Inclusion appeared in descriptions of AI councils, stakeholder engagement, user training, workshops, and cross-functional decision-making. Reflexivity appeared in participants' recognition that AI systems require human oversight, ethical review, and awareness of bias, model limitations, and misuse risk. Responsiveness appeared in continuous monitoring, policy revalidation, model performance review, user feedback, and adaptive improvement.

The results also align with established scholarship and public guidance on re-

sponsible AI adoption. NIST (2023) emphasized that trustworthy AI requires governance, risk mapping, measurement, and risk management across the AI lifecycle. The current findings confirm this guidance by showing that business leaders use governance structures, data protection controls, human review, controlled pilots, workforce preparation, and measurement systems to reduce AI risks. The findings extend existing knowledge by demonstrating how these principles are applied in real organizational contexts through AI councils, approved tool restrictions, tenant-controlled AI environments, subject matter expert validation, user training, pulse surveys, usage tracking, and sustainability considerations.

Taken together, these findings illustrate that safe and ethical AI adoption functions as both a governance process and an organizational change process. From a governance perspective, the themes demonstrate how organizations establish policies, data safeguards, approval processes, oversight structures, and risk controls. From an organizational change perspective, the findings show that successful AI adoption also depends on user awareness, workforce training, communication, culture, stakeholder engagement, and continuous feedback. This integrated perspective supports the value of applying responsible innovation to AI adoption because the framework explains how leaders anticipate risk, include stakeholders, reflect on ethical implications, and respond to changing technology conditions. For business practitioners, these findings suggest that responsible AI adoption requires more than purchasing or deploying AI tools. It requires an intentional system of governance, human accountability, organizational learning, and continuous improvement.

Table 3 presents the data saturation summary for this qualitative pragmatic inquiry and illustrates the progression of coding across the nine participant interviews. Data saturation was assessed by examining whether later interviews introduced new themes or meaningful insights beyond those already identified in earlier interviews. As shown in **Table 3**, the number of new representative codes decreased progressively across interviews, reflecting increasing convergence in participant responses. Later interviews reinforced the existing themes rather than producing new major themes, indicating that thematic consistency had emerged across the data set.

Table 4 illustrates that during the initial interviews, a higher number of new representative codes emerged, capturing a broad range of participant perspectives on AI governance, data protection, human oversight, controlled implementation, workforce readiness, and measurement. Eight new representative codes were identified in Interview 1, followed by five and four new representative codes in Interviews 2 and 3, respectively. As data collection and coding continued, the number of new representative codes declined, with three new representative codes identified in Interview 4, two in Interview 5, and one in Interview 6 (**Table 4**). This pattern indicates that the core concepts were identified early and that subsequent interviews increasingly reinforced existing findings rather than generating new major insights.

Table 3. Major themes identified from participant interviews.

Theme	Participants Referencing Theme	Number of References
Structured AI governance and policy development	9	41
Data governance, privacy, and confidentiality protection	7	17
Human oversight, AI constraint, and quality assurance	7	21
Phased pilots, controlled experimentation, and use case prioritization	6	11
Workforce training, communication, and change management	8	27
Continuous monitoring, measurement, and adaptive improvement	8	14

Note. Participants are referenced as Participant 1 through Participant 9. The number of references reflects the frequency with which participants discussed concepts associated with each theme during semi structured interviews, as reflected in the MAXQDA Code Matrix Browser.

Table 4. Data Saturation summary.

Interview number	New representative codes identified	Cumulative representative codes	Saturation status
Interview 1	8	8	No saturation
Interview 2	5	13	No saturation
Interview 3	4	17	No saturation
Interview 4	3	20	No saturation
Interview 5	2	22	Approaching saturation
Interview 6	1	23	Saturation emerging
Interview 7	0	23	Saturation achieved
Interview 8	0	23	No new major themes
Interview 9	0	23	No new major themes

Note. Data saturation was assessed by reviewing whether additional interviews produced new representative, final, codes, categories, or major themes. Interviews 7 through 9 reinforced existing themes and confirmed thematic consistency across participants.

By Interview 6, saturation was emerging because no new major themes were identified after that point. Interviews 7 through 9 confirmed thematic consistency, as participant responses aligned with the previously identified six themes. Alt-

though later interviews added depth, examples, and contextual variation, they did not produce additional major themes. For example, later participant responses strengthened the findings related to enterprise AI governance, approved AI tools, workforce communication, and sustainability considerations, but these insights fit within the existing themes rather than requiring the creation of new themes. This outcome is consistent with qualitative research standards indicating that saturation is reached when additional data collection produces redundancy rather than substantially new insights (Saldaña, 2021; Saunders et al., 2018).

In total, 23 representative codes were identified and consistently observed across the data set. The systematic decline in new representative codes and the stabilization of themes after Interview 6 provided evidence that the data set was sufficiently rich to address the research question. **Table 3** therefore provides a transparent audit trail of the saturation process and supports the trustworthiness of the findings by demonstrating that the six themes were grounded in repeated and consistent participant responses. Since the current project included 9 completed interviews, versus 10 which was the initial plan, the findings indicate that saturation was strongly supported within the existing data set. Data saturation was achieved by the seventh interview and confirmed by the eighth and ninth interviews, consequently, the tenth interview was not necessary and was not conducted.

4.1. Theme 1: Structured AI Governance and Policy Development

The first theme that emerged from the thematic analysis was structured AI governance and policy development. Analysis of the interview data revealed that participants consistently viewed formal governance as a foundational strategy for developing and applying safe and ethical practices when adopting AI technologies. All nine participants referenced some aspects of governance, policy, leadership oversight, risk review, or organizational control when describing responsible AI adoption. These strategies included establishing AI governance boards, creating acceptable use policies, aligning legal and technical requirements, reviewing use cases before implementation, assessing vendor risk, and ensuring executive sponsorship for AI-related decisions.

Participants described AI governance as an intentional organizational process through which leaders establish the rules, structures, and accountability mechanisms needed to guide responsible AI use. Rather than allowing AI adoption to occur informally or through individual employee experimentation, participants emphasized that leaders must define approved tools, acceptable uses, data handling expectations, monitoring requirements, and escalation procedures. Public AI governance guidance also supported this finding. The NIST AI Risk Management Framework identifies governance as a central function for managing AI risks across the AI lifecycle (NIST, 2023).

Participant 1 emphasized that AI governance begins with a strategic foundation. Participant 1 described effective AI adoption as requiring “the AI, AI, vision and transformation road map,” along with “ethical risk assessment, data govern-

ance, AI capability development and strategic alignment with the business objective.” This response illustrates how business leaders may approach AI governance as a structured strategy rather than as a single technology decision. Participant 1’s emphasis on strategy, ethics, risk, and business alignment reflects the responsible innovation dimension of anticipation because it demonstrates the importance of identifying risks and governance needs before broad AI deployment.

Participant 3 described the importance of governance in a complex organizational environment where different components may adopt AI at different levels of maturity. Participant 3 explained that leaders must consider “what type of governance are they building around it” and how AI is “being embedded into their governance processes.” This response demonstrates that structured AI governance helps organizations reduce fragmentation and ensure that AI adoption is coordinated across business units. Participant 3 also suggested that organizations could “build a community of practice” so representatives could share best practices and maintain communication. This emphasis on communication and shared governance aligns with the responsible innovation dimension of inclusion because governance requires collaboration among stakeholders rather than isolated decision-making.

Participant 4 described governance from the perspective of executive leadership and policy authority. Participant 4 stated, “In my role as the CIO at the Patent and Trademark Office, I was the decision maker and the recommender of policy and implementation or deployment of those AI policies and tools.” This response illustrates how governance connects strategic leadership, policy development, legal compliance, and operational implementation. Participant 4 also emphasized that governance must be grounded in stakeholder input, stating that “stakeholders were included in every phase of our deployments” and that the organization took their feedback. This finding confirms prior research indicating that AI adoption requires leadership accountability, formal oversight, and organizational controls to reduce ethical and operational risks (Mittelstadt et al., 2016; Wirtz et al., 2023).

Participant 7 further supported this theme by describing the formation of an AI council, the updating of organizational policies, and the use of pilots before enterprise deployment decisions. Participant 7 emphasized that AI governance required leadership review, use case prioritization, and evaluation before tools were expanded across the organization. This response demonstrates how AI councils and governance bodies can serve as mechanisms for reviewing AI opportunities, identifying risks, and determining whether AI tools should be scaled. Participant 7’s response aligns with the responsible innovation dimension of responsiveness because governance was described as a process that adapts as new tools, risks, and use cases emerge.

Participant 8 described AI governance through enterprise policy and acceptable use expectations. Participant 8 explained that enterprise AI policy should identify approved large language models, define approved methods of use, establish safeguards, and communicate monitoring expectations. Participant 8 also described

AI policy as a living framework because AI risks change rapidly. This response suggests that leaders should not treat AI governance as a one-time policy activity. Instead, governance must be reviewed and adjusted as AI technologies, employee behaviors, and organizational risks evolve.

Participant 9 extended this theme by describing how enterprise AI and data governance policies must be translated into practical procedures for specific business groups. Participant 9 explained that part of the governance role involved translating enterprise level policies into “meaningful” practices for the R&D group. Participant 9 also described a structured process in which new AI tools go through multiple layers of assessment, risk review, procurement review, partnership review, and governance approval before adoption. This response demonstrates how AI governance can be operationalized through layered assessments and decision gates that reduce vendor, procurement, technical, ethical, and organizational risk.

These findings align closely with the responsible innovation framework guiding this project. [Stilgoe et al. \(2013\)](#) described anticipation, inclusion, reflexivity, and responsiveness as key dimensions of responsible innovation. The structured AI governance theme reflects anticipation because participants emphasized identifying risks, policies, and controls before enterprise AI deployment. The theme reflects inclusion because participants discussed AI councils, stakeholder engagement, executive sponsors, legal and technical collaboration, and governance boards. The theme reflects responsiveness because participants emphasized updating policies, reviewing AI tools, and adapting governance practices as AI technologies and risks change.

The findings also align with current AI governance scholarship and public guidance. [NIST \(2023\)](#) emphasized that organizations should establish governance structures to manage AI risks and ensure trustworthy AI outcomes. The current findings confirm this guidance by showing that business leaders apply structured governance through policies, councils, acceptable use requirements, vendor reviews, use case prioritization, and executive oversight. The findings extend existing knowledge by illustrating how leaders translate high level AI governance principles into practical business processes that guide daily AI adoption decisions.

Collectively, these findings indicate that structured AI governance and policy development are essential to safe and ethical AI adoption. Participants emphasized that AI governance must be established before broad implementation and must continue as AI tools, organizational needs, and risks evolve. These findings suggest that business leaders seeking to adopt AI responsibly should create formal governance structures, define acceptable use policies, align AI practices with legal and technical requirements, review vendors and tools before use, and ensure that executive leaders remain accountable for AI adoption outcomes. This theme provides the foundation for the remaining findings because data protection, human oversight, controlled pilots, workforce training, and continuous monitoring all depend on a clear governance structure that defines how AI should be adopted and managed.

4.2. Theme 2: Data Governance, Privacy, and Confidentiality Protection

The second theme that emerged from the thematic analysis was data governance, privacy, and confidentiality protection. Analysis of the interview data revealed that participants consistently viewed data protection as a core requirement for safe and ethical AI adoption. Seven of the nine participants referenced data governance, privacy safeguards, confidential information protection, intellectual property protection, patient data protection, data quality, approved tool use, or access control when describing responsible AI strategies. Participants emphasized that organizations cannot adopt AI safely unless leaders first determine what data may be used, where data may be processed, who may access it, and which AI tools are approved for organizational use.

Participants described data governance as an essential safeguard because AI systems rely on organizational data to generate outputs, automate tasks, and support decision making. Several participants emphasized that data exposure, poor data quality, and unclear data ownership can create ethical, legal, operational, and reputational risks. Public guidance also supported this theme. The National Institute of Standards and Technology AI Risk Management Framework emphasizes that trustworthy AI depends on governance, mapping, measurement, and management practices that address risks to individuals, organizations, and society, including risks related to data quality, privacy, security, and reliability ([National Institute of Standards and Technology \(NIST\), 2023](#)). These findings also align with scholarship emphasizing that responsible AI requires privacy protection, accountability, transparency, and trustworthy data practices ([Floridi et al., 2018](#); [Jobin et al., 2019](#)).

Participant 1 emphasized the importance of determining what data should be exposed to AI systems and protecting sensitive information from inappropriate use. Participant 1 explained that leaders must decide “which data to expose” and emphasized that the organization should not expose “sensitive or personally identifiable information.” This response illustrates that data governance begins before AI tools are deployed. Leaders must define data boundaries, determine which information is appropriate for AI use, and prevent personally identifiable information from being entered into uncontrolled systems. Participant 1’s response aligns with the responsible innovation dimension of anticipation because it demonstrates proactive identification of data related risks before AI adoption is expanded.

Participant 2 emphasized intellectual property protection as a central data governance concern. Participant 2 explained that organizations must ensure they “do not leak IP” and should not place “project sensitive or IP sensitive information” into public large language models. Participant 2 further emphasized the importance of using company provided AI tools that include safeguards. This response demonstrates that ethical AI adoption requires leaders to protect proprietary business information as well as personal or regulated data. Participant 2’s emphasis on lim-

iting what information may be entered into AI systems confirms the need for acceptable use rules, approved tools, and employee awareness related to data protection.

Participant 4 also emphasized confidentiality protection when discussing AI adoption in a highly sensitive environment. Participant 4 explained that the organization had an obligation to keep patents and trademarks confidential and therefore did not allow examiners to use public chatbots. Participant 4 described the principle as “once in, never out,” meaning that information entered into the organization’s internal AI environment should not leave the controlled environment. This response illustrates how leaders can translate confidentiality requirements into operational safeguards. Rather than simply warning employees not to share confidential information, the organization developed internal AI capabilities and secure boundaries to prevent data leakage. This finding supports the professional practice implication that data governance requires both policy controls and technical controls.

Participant 6 emphasized data protection in the context of healthcare and sensitive patient information. Participant 6 described the importance of protecting veteran and patient data when AI tools are used in healthcare related settings. This response illustrates that data governance is especially important in contexts where AI systems may interact with health information, protected records, or vulnerable populations. Participant 6’s response also supports the finding that data protection is not only a compliance issue but also an ethical issue. Protecting patient and veteran data helps preserve trust, confidentiality, and responsible use of AI in settings where harm from data misuse could be significant.

Participant 7 described data governance through data quality, authoritative data sources, and zero trust access. Participant 7 emphasized that AI outputs are only as reliable as the data used to support them. Participant 7 discussed the importance of data cleansing, identifying authoritative sources, and applying zero trust access principles before using AI at scale. This response demonstrates that data governance involves more than protecting data from exposure. It also requires ensuring that data are accurate, reliable, properly governed, and accessible only to authorized users. Participant 7’s emphasis on data cleansing and zero trust access aligns with responsible AI guidance because poor data quality and weak access controls can lead to unreliable outputs and increased organizational risk.

Participant 8 described data governance through approved AI tools, enterprise tenant protection, and the blocking of unapproved AI domains. Participant 8 explained that organizations should select AI tools that keep organizational data inside the enterprise tenant and should restrict unapproved AI technologies that may expose business information. This response illustrates how leaders can implement data governance through technical safeguards and enterprise architecture decisions. Participant 8’s emphasis on tenant controlled tools demonstrates that safe AI adoption depends on selecting technologies that align with organizational privacy, security, and monitoring requirements.

Participant 9 extended this theme by describing the role of enterprise data governance and AI governance in research and development environments. Participant 9 explained that enterprise policies must be translated into practical guidance for specific business groups and that new AI tools may require risk assessment, procurement review, partnership review, and governance approval. This response demonstrates that data governance must be applied across the full AI adoption process, including internal tool selection, vendor review, external partnerships, and business group implementation. Participant 9's response also shows that data governance requires coordination between enterprise policy and local operational practice.

These findings align closely with the responsible innovation framework. The theme reflects anticipation because participants emphasized identifying data exposure risks, privacy risks, intellectual property risks, and patient data risks before broad AI implementation. The theme reflects reflexivity because participants recognized that AI adoption can create unintended consequences when data are inaccurate, sensitive, proprietary, or used outside approved boundaries. The theme also reflects responsiveness because participants described adjusting data controls, approved tool lists, and governance processes as AI technologies and business needs evolve.

The findings also align with current scholarship and public AI governance documents. [NIST \(2023\)](#) emphasized that AI risk management requires attention to data quality, privacy, security, transparency, and trustworthiness. [Floridi et al. \(2018\)](#) identified privacy, accountability, and beneficence as central concerns in ethical AI. [Jobin et al. \(2019\)](#) similarly found that privacy, transparency, accountability, and fairness appear consistently across global AI ethics guidelines. The current findings confirm this literature by showing that business leaders view data protection as a practical requirement for responsible AI adoption. The findings extend existing knowledge by illustrating how leaders operationalize data governance through approved tools, internal AI environments, zero trust access, data cleansing, tenant controlled systems, vendor review, and restrictions on sensitive data use.

Collectively, these findings indicate that data governance, privacy, and confidentiality protection are essential components of safe and ethical AI adoption. Participants emphasized that business leaders must protect sensitive data, define appropriate data use, prevent exposure of proprietary information, validate data quality, and ensure that AI tools operate within secure and approved environments. These findings suggest that organizations seeking to adopt AI responsibly should establish data classification rules, approved AI tool lists, access controls, vendor review procedures, data quality standards, and employee guidance on prohibited data use. This theme builds on the first theme because formal AI governance provides the structure needed to enforce data protection expectations across the organization.

4.3. Theme 3: Human Oversight, AI Constraint, and Quality Assurance

The third theme that emerged from the thematic analysis was human oversight, AI constraint, and quality assurance. Analysis of the interview data revealed that participants viewed human accountability as essential to safe and ethical AI adoption. Seven of the nine participants referenced human review, AI decision limits, subject matter expert validation, quality assurance, quality control, model monitoring, bias testing, or output validation when describing responsible AI strategies. Participants emphasized that AI should support human decision making but should not replace human judgment in sensitive, safety related, clinical, security, legal, or high impact business contexts.

Participants described human oversight as a safeguard that helps organizations prevent overreliance on AI outputs. Several participants emphasized that AI systems can produce inaccurate, biased, incomplete, or contextually inappropriate outputs if qualified humans do not review them. Public guidance and scholarly literature also supported this finding. The National Institute of Standards and Technology AI Risk Management Framework emphasizes that trustworthy AI requires risk management practices that address validity, reliability, safety, security, accountability, transparency, and human oversight ([National Institute of Standards and Technology \(NIST, 2023\)](#)). Similarly, [Mittelstadt et al. \(2016\)](#) emphasized that algorithmic systems create ethical risks when accountability is unclear or when organizations fail to maintain meaningful human review.

Participant 1 emphasized that AI systems should not operate without human involvement. Participant 1 stated that leaders must make sure “there is a human oversight involved and not to give the full control.” This response illustrates that human oversight is a core ethical safeguard in AI adoption. Participant 1 also discussed the need for subject matter experts to review and verify AI outputs to determine whether the AI results align with human level performance. This response demonstrates that oversight is not simply a symbolic requirement. Instead, human review must be built into the process to validate whether AI outputs are accurate, appropriate, and aligned with organizational expectations.

Participant 2 emphasized AI constraint as a strategy for reducing risk in safety and security related contexts. Participant 2 explained that the main strategy is to “give the LLM less power” and not depend on a large language model to determine what is safe or secure. Participant 2 further stated, “The problem was you let the LLM make a safety related decision. That should not have been in its power anyway.” This response illustrates how leaders can operationalize ethical AI adoption by limiting what AI systems are allowed to decide. Rather than relying on AI to make high risk decisions, leaders can design systems in which AI supports analysis while humans retain decision authority. This finding confirms the importance of accountability and human responsibility in AI-enabled decision-making.

Participant 4 described human oversight through stakeholder involvement and feedback during AI deployment. Participant 4 emphasized that safe AI adoption

required stakeholders throughout implementation and using their feedback to refine deployment decisions. This response demonstrates that human oversight can occur not only through technical review of AI outputs but also through organizational review of how AI affects users, business processes, and stakeholder expectations. Participant 4's response aligns with the responsible innovation dimension of inclusion because it shows that responsible AI adoption requires input from people affected by AI systems.

Participant 6 emphasized human oversight in healthcare settings, where AI output could influence diagnosis, treatment, or patient care. Participant 6 explained that AI should not be the final decision maker in healthcare matters because human professionals must remain accountable for decisions that affect patients. This response illustrates the importance of human judgment in high risk environments where AI errors could create serious harm. Participant 6 also described the use of quality assurance and quality control pipelines, dashboards, and performance monitoring to track accuracy and consistency across AI model versions. This response demonstrates that human oversight must be supported by structured quality processes that monitor whether AI tools continue to perform as intended over time.

Participant 7 also emphasized the importance of human in the loop validation. Participant 7 described reviewing AI outputs before relying on them for organizational decisions and emphasized that AI outputs should be validated by humans before broader use. This response illustrates how leaders can maintain human accountability while still benefiting from AI supported analysis. Participant 7's emphasis on validation also shows that AI adoption should include decision checkpoints where users confirm whether outputs are reliable, repeatable, and appropriate for the business context.

Participant 9 extended this theme by describing governance reviews and risk assessments before new AI tools are adopted. Participant 9 explained that AI tools in the organization may go through multiple layers of assessment, including enterprise review, procurement review, partnership review, and governance approval. Although this process supports broader governance, it also functions as quality assurance because new AI tools are evaluated before they are introduced into the operating environment. Participant 9's response demonstrates that oversight can occur before implementation as well as during use. Leaders can reduce AI risk by reviewing tools, vendors, ethical concerns, and operational fit before employees begin using the technology.

These findings align closely with the responsible innovation framework. The theme reflects reflexivity because participants critically examined the limitations of AI and recognized that AI outputs should not be treated as automatically reliable or ethically neutral. Participants acknowledged that AI systems can create risks if leaders fail to question their assumptions, outputs, data sources, or decision boundaries. The theme also reflects responsiveness because participants described quality assurance, quality control, model monitoring, and human review

as ways to adjust AI practices when accuracy, reliability, or performance concerns emerge.

The findings also align with current scholarship and public AI governance guidance. [Mittelstadt et al. \(2016\)](#) emphasized that algorithmic systems raise ethical concerns related to accountability, transparency, responsibility, and bias. [NIST \(2023\)](#) emphasized that trustworthy AI requires measurement and management of risks related to reliability, validity, security, and accountability. The current findings confirm this literature by showing that business leaders use human oversight and AI constraints to reduce the risk of unsafe or unethical AI decisions. The findings extend existing knowledge by illustrating how leaders operationalize oversight through human in the loop review, subject matter expert validation, decision limits, quality dashboards, model version monitoring, and preadoption governance review.

Collectively, these findings indicate that human oversight, AI constraint, and quality assurance are essential to safe and ethical AI adoption. Participants emphasized that AI should assist human decision making but should not replace human accountability in high risk contexts. These findings suggest that business leaders should define which AI use cases require human review, limit AI authority in sensitive decisions, establish subject matter expert validation processes, monitor model performance, and maintain quality assurance practices across the AI lifecycle. This theme builds on the previous themes because governance structures and data protection controls are strengthened when leaders also ensure that humans remain accountable for how AI tools are tested, interpreted, and used.

4.4. Theme 4: Phased Pilots, Controlled Experimentation, and Use Case Prioritization

The fourth theme that emerged from the thematic analysis was phased pilots, controlled experimentation, and use case prioritization. Analysis of the interview data revealed that participants viewed staged implementation as an important strategy for reducing risk when adopting AI technologies. Six of the nine participants referenced pilots, proof of concept testing, limited user groups, safe testing environments, controlled rollouts, use case review, or demand prioritization when describing effective strategies for safe and ethical AI adoption. Participants emphasized that organizations should not immediately deploy AI tools enterprise wide without first testing the tools in a controlled setting, evaluating risks, and determining whether the use case aligns with organizational goals.

Participants described phased implementation as a practical way to learn before scaling. Rather than treating AI adoption as a one-time deployment decision, participants emphasized that business leaders should test AI tools with smaller groups, gather user feedback, evaluate technical and ethical risks, and determine whether broader implementation is justified. Public guidance also supported this finding. The Office of Management and Budget emphasized that AI governance, innovation, and risk management should be connected activities, especially for AI uses

that may affect rights, safety, or public trust. Similarly, the National Institute of Standards and Technology AI Risk Management Framework emphasizes mapping, measuring, and managing AI risks before and during implementation (National Institute of Standards and Technology (NIST), 2023).

Participant 4 described phased implementation as a structured adoption process. Participant 4 explained that the organization used AI in “three ways, 3 phases,” beginning with proving the use case, then scaling the concept, and finally identifying an executive champion or funding owner for the service. This response illustrates how leaders can reduce adoption risk by first determining whether AI is useful and feasible before investing in broader deployment. Participant 4’s emphasis on proving the use case before scaling demonstrates that safe AI adoption requires disciplined decision making, not rapid implementation without evidence of value.

Participant 5 also supported this theme by emphasizing that AI tools should be introduced in a controlled manner rather than adopted casually across the organization. Participant 5 described the importance of ensuring that users understand which tools are approved and how AI may be used safely. This response suggests that controlled adoption is connected to governance and workforce readiness. Leaders must not only decide which AI tools to implement but also determine when and how employees should gain access to those tools. Controlled implementation can reduce misuse, support compliance, and allow leaders to identify training or policy gaps before wider rollout.

Participant 6 emphasized safe testing environments where employees could experiment with AI before the organization made larger investments. Participant 6 described the importance of allowing frontline workers, developers, and staff to test AI in a safe space before significant resources were committed. This response illustrates how controlled experimentation can support both innovation and risk reduction. Safe testing environments allow employees to explore AI capabilities while leaders evaluate whether the technology is accurate, secure, practical, and aligned with operational needs. Participant 6’s response also reflects the responsible innovation dimension of anticipation because testing occurs before full implementation.

Participant 7 described a pilot strategy involving limited access to Microsoft Copilot and Office 365 AI features. Participant 7 explained that a smaller group of users received access first so they could build familiarity before broader implementation. Participant 7 also emphasized that AI use cases were collected, reviewed, and prioritized with leadership, including the Chief AI Officer, before enterprise scaling. This response demonstrates how business leaders can use pilots and use case prioritization to prevent uncontrolled AI expansion. The pilot process provided a structured way to evaluate readiness, user experience, business value, and governance implications before determining whether AI tools should be scaled.

Participant 8 described controlled AI exposure and approved tool restrictions

as part of safe adoption. Participant 8 emphasized that organizations should provide employees with access to approved AI tools while restricting or blocking unapproved AI technologies. This approach allowed the organization to support productivity and innovation while maintaining control over data exposure, monitoring expectations, and acceptable use. Participant 8's response demonstrates that controlled experimentation is not only about limiting access; it is also about creating a safe pathway for employees to learn and use AI within approved boundaries.

Participant 9 extended this theme by describing the importance of front loading AI demand. Participant 9 explained that business groups should identify AI needs early, determine whether existing enterprise tools can meet those needs, and only pursue new AI technologies when necessary. Participant 9 also described governance review and risk assessment before new AI tools or vendors are introduced. This response shows that use case prioritization helps organizations avoid unnecessary or duplicative AI adoption. By evaluating demand early, leaders can determine whether the proposed AI use case is appropriate, whether existing tools are sufficient, and whether additional review is required before implementation.

These findings align closely with the responsible innovation framework. The theme reflects anticipation because participants emphasized testing AI tools, evaluating use cases, identifying risks, and limiting exposure before broad deployment. The theme also reflects responsiveness because pilot results, user feedback, and implementation outcomes can inform whether AI tools should be expanded, revised, delayed, or discontinued. In some cases, the theme also reflects inclusion because participants described involving users, frontline workers, developers, stakeholders, and executive sponsors in the testing and evaluation process.

The findings also align with business practice literature and public AI governance guidance. [Creswell and Creswell \(2023\)](#) emphasized that systematic inquiry involves collecting and analyzing evidence before drawing conclusions. Applied to organizational AI adoption, this principle supports the value of pilots, feedback loops, and structured evaluation before enterprise scaling. [NIST \(2023\)](#) emphasized the importance of measuring and managing AI risks throughout the AI lifecycle. The current findings confirm this guidance by showing that business leaders use pilots, safe test environments, limited access, use case review, and executive sponsorship to evaluate AI before broader adoption.

Collectively, these findings indicate that phased pilots, controlled experimentation, and use case prioritization are important strategies for safe and ethical AI adoption. Participants emphasized that organizations should test AI tools in controlled environments, prioritize use cases based on business need and risk, involve appropriate stakeholders, and evaluate pilot outcomes before expanding AI use. These findings suggest that business leaders should establish formal pilot processes, define success criteria, review proposed use cases, limit early access to approved users, and require governance review before enterprise deployment. This theme builds on the prior themes because AI governance, data protection, and human oversight are more effective when AI implementation occurs through con-

trolled and evidence based adoption rather than unrestricted organizational use.

4.5. Theme 5: Workforce Training, Communication, and Change Management

The fifth theme that emerged from the thematic analysis was workforce training, communication, and change management. Analysis of the interview data revealed that participants viewed employee awareness, user behavior, stakeholder engagement, and organizational culture as essential to safe and ethical AI adoption. Eight of the nine participants referenced training, communication, user readiness, policy awareness, stakeholder engagement, roadshows, workshops, change management, or organizational culture when describing responsible AI adoption strategies. Participants emphasized that AI implementation cannot be managed as a technology deployment alone. Instead, leaders must prepare employees to understand approved tools, prohibited uses, data risks, ethical expectations, and the organizational purpose behind AI governance controls.

Participants described workforce readiness as a practical safeguard against misuse, resistance, and uncontrolled adoption. Several participants emphasized that employees may use AI tools based on personal habits without understanding the risks of entering sensitive business information, intellectual property, client data, or regulated data into unapproved tools. Others emphasized that users may resist AI restrictions if leaders fail to explain why the controls exist. These findings align with responsible AI scholarship emphasizing that AI governance requires not only technical controls but also organizational learning, stakeholder engagement, and leadership communication (Uddin, 2023; Wirtz et al., 2023). Public AI guidance also supports this theme because responsible AI implementation requires organizations to define roles, communicate expectations, and prepare users to engage with AI systems responsibly (National Institute of Standards and Technology (NIST), 2023).

Participant 1 emphasized the importance of training and policy awareness in ethical AI adoption. Participant 1 noted that one challenge is ensuring employees complete required ethical training and understand organizational policies related to AI use. This response illustrates that policies alone are insufficient if employees do not understand them or do not know how to apply them in daily work. Participant 1's response suggests that training must be treated as part of the AI governance process because employees are the individuals who interact with AI tools, interpret AI outputs, and make decisions about what information to use.

Participant 3 described workforce readiness through knowledge sharing and community learning. Participant 3 explained that "the best avenue for getting people knowledgeable about AI and understanding it" is "knowledge share." Participant 3 also suggested that organizations could build a community of practice where representatives share lessons learned and best practices. This response demonstrates that communication can help reduce inconsistency across organizational units. Participant 3's emphasis on knowledge sharing aligns with the responsible

innovation dimension of inclusion because responsible AI adoption requires participation, communication, and shared learning across stakeholder groups rather than isolated technology decisions.

Participant 4 emphasized the importance of including stakeholders throughout AI deployment. Participant 4 explained that the organization safely implemented AI by ensuring that “stakeholders were included in every phase” and by taking their feedback. This response illustrates how change management can reduce adoption risk by involving the people affected by AI implementation. Rather than imposing AI tools on users, leaders can create opportunities for feedback, testing, and refinement. Participant 4’s response also suggests that stakeholder involvement can build trust and improve adoption because users are more likely to accept AI tools when they understand the purpose, risks, and expected benefits.

Participant 5 described user education as necessary because employees may bring personal AI habits into the workplace without understanding organizational data risks. Participant 5 emphasized that users need to understand what tools are approved, what information may be entered into AI systems, and what risks may arise from unapproved AI use. This response demonstrates that responsible AI adoption depends on employee literacy and behavioral change. Leaders must educate employees not only on how to use AI but also on when not to use it, what data not to share, and why organizational safeguards matter.

Participant 7 emphasized that AI should be implemented with employees and customers rather than to them. This response highlights the importance of participatory change management. Participant 7’s perspective suggests that AI adoption is more effective when leaders involve users early, communicate openly, and build trust through shared implementation. This approach aligns with the responsible innovation dimension of inclusion because employees and customers are treated as stakeholders in the AI adoption process rather than passive recipients of new technology.

Participant 8 described a key communication challenge: convincing users that AI restrictions were protective rather than barriers to productivity. Participant 8 explained that some users may view blocked tools, approved tool lists, and monitoring requirements as obstacles unless leaders clearly communicate that these controls protect the organization, employees, and customers. This response illustrates that change management must address perception as well as policy. If employees interpret AI governance as restrictive without understanding the protective purpose, they may resist controls or seek workarounds. Participant 8’s response therefore extends business practice knowledge by showing that ethical AI implementation requires leaders to explain the “why” behind governance.

Participant 9 further supported this theme by describing roadshows, all hands presentations, AI workshops, posters, internal champions, and stakeholder engagement as strategies to increase awareness and reduce siloed or noncompliant AI adoption. Participant 9 explained that lack of awareness was a challenge because some users attempted to adopt AI tools independently before understanding

enterprise standards. This response demonstrates that AI change management must be proactive, repeated, and visible across the organization. Participant 9's description of champions and outreach activities shows how leaders can create communication networks that help employees understand approved AI pathways and responsible use expectations.

These findings align closely with the responsible innovation framework. The theme reflects inclusion because participants emphasized engaging employees, customers, technical teams, business users, stakeholders, and internal champions in AI adoption. The theme also reflects responsiveness because training and communication must change as AI tools, policies, risks, and user behaviors evolve. Participants recognized that AI adoption occurs in a dynamic environment where employees need ongoing guidance rather than one time instruction.

The findings also align with current scholarship on responsible AI and organizational change. Wirtz et al. (2023) emphasized that responsible AI adoption requires trust, accountability, and organizational readiness. Uddin (2023) similarly highlighted the role of ethical leadership in shaping employee trust and responsible technology use. The current findings confirm this literature by showing that workforce training and communication are necessary for responsible AI adoption. The findings extend existing knowledge by demonstrating how leaders operationalize change management through communities of practice, required training, stakeholder feedback, roadshows, workshops, internal champions, approved tool communication, and user awareness campaigns.

Collectively, these findings indicate that workforce training, communication, and change management are essential components of safe and ethical AI adoption. Participants emphasized that employees must understand AI policies, data risks, approved tools, prohibited uses, and the ethical reasons behind organizational safeguards. These findings suggest that business leaders should implement recurring AI literacy training, role based policy education, stakeholder engagement sessions, communication campaigns, internal champions, and safe learning opportunities for employees. This theme builds on the earlier themes because governance, data protection, human oversight, and pilot testing are only effective when employees understand how to apply them in practice.

4.6. Theme 6: Continuous Monitoring, Measurement, and Adaptive Improvement

The sixth theme that emerged from the thematic analysis was continuous monitoring, measurement, and adaptive improvement. Analysis of the interview data revealed that participants viewed responsible AI adoption as an ongoing process rather than a one-time implementation decision. Eight of the nine participants referenced monitoring, measurement, policy review, usage tracking, surveys, A/B testing, customer satisfaction metrics, compliance reviews, quality dashboards, model version monitoring, or adaptive governance when describing safe and ethical AI adoption. Participants emphasized that AI systems, organizational risks, user be-

havior, and regulatory expectations change over time; therefore, leaders must continuously evaluate whether AI practices remain effective, ethical, secure, and aligned with business needs.

Participants described measurement as a practical mechanism for determining whether AI strategies are producing intended outcomes. Several participants emphasized that leaders should define success criteria, monitor AI use, evaluate user feedback, track model performance, and revise policies when risks or gaps emerge. Public guidance also supported this finding. The National Institute of Standards and Technology AI Risk Management Framework identifies governance, mapping, measurement, and management as core functions of AI risk management ([National Institute of Standards and Technology \(NIST\), 2023](#)).

Participant 1 emphasized monitoring and compliance review as part of responsible AI adoption. Participant 1 described the importance of reviewing whether AI implementation supports business objectives, protects sensitive information, and complies with organizational policies. Participant 1 also discussed comparing AI model outputs with subject matter expert evaluations to determine how closely AI results aligned with human review. This response illustrates that measurement must consider both operational value and ethical reliability. Participant 1's emphasis on compliance review and comparison with human performance aligns with the responsible innovation dimension of responsiveness because leaders must adjust practices when AI performance or policy compliance does not meet expectations.

Participant 2 described effectiveness measurement through testing, automation, and external feedback. Participant 2 explained that feedback from clients should come into the testing framework and that organizations may use automated testing indicators to determine whether systems are working as intended. This response demonstrates that AI effectiveness can be evaluated through both technical validation and user or client feedback. Participant 2's perspective suggests that continuous improvement requires leaders to connect AI monitoring to real world performance outcomes rather than relying only on internal assumptions about whether AI tools are successful.

Participant 3 emphasized the importance of establishing metrics early in the AI adoption process. Participant 3 stated that it was "the perfect time to come up with some really good metrics" that could be tracked over time to determine where progress was occurring and where course correction was needed. This response illustrates how measurement supports adaptive governance. By identifying metrics early, leaders can determine whether AI adoption is improving organizational performance, increasing risk, or requiring changes in training, policy, or implementation strategy. Participant 3's response aligns with the responsible innovation dimension of anticipation because measurement planning begins before AI adoption is fully mature.

Participant 4 described surveys, adoption metrics, and long term oversight as ways to measure AI effectiveness. Participant 4 explained that short term measures

included polling users and tracking how often AI products were being used, including whether employees used the tools daily and whether they provided positive or negative feedback. This response demonstrates that AI effectiveness is not measured only by technical performance. Leaders also need to understand whether users are adopting the tools, whether the tools are useful, and whether users perceive them as valuable. Participant 4's emphasis on usage and feedback suggests that AI measurement should include both operational and human factors.

Participant 6 emphasized monitoring model accuracy, consistency, and performance across versions. Participant 6 described the use of QA and QC pipelines, dashboards, and monitoring to ensure AI outputs did not drift away from human accuracy. This response illustrates that continuous monitoring is especially important when AI systems are updated, retrained, or modified over time. Participant 6's emphasis on version monitoring demonstrates that AI governance must extend beyond initial deployment. Leaders must continue to evaluate whether AI outputs remain reliable as models change. This finding is especially important in high risk environments where inaccurate or inconsistent outputs may affect decisions involving people, health, safety, or compliance.

Participant 7 described the importance of evaluating AI tools against defined success criteria before broader implementation. Participant 7 emphasized that leaders should establish success measures, review results, and determine whether an AI tool is ready to scale. This response demonstrates that measurement can inform go or no go decisions. Participant 7's perspective connects continuous monitoring to controlled experimentation because pilot results should guide whether AI is expanded, revised, or discontinued. This finding supports the broader conclusion that responsible AI adoption requires evidence based decision making rather than adoption based on novelty or pressure to keep pace with technology trends.

Participant 8 described AI policy as a living framework rather than a static document. Participant 8 explained that AI risks change quickly and that policies must be revisited as tools, behaviors, and threats evolve. This response illustrates adaptive improvement at the governance level. Continuous monitoring is not limited to technical systems; it also applies to policies, training, approved tool lists, and monitoring expectations. Participant 8's response aligns strongly with the responsible innovation dimension of responsiveness because it emphasizes that leaders must revise governance practices as new risks emerge.

Participant 9 extended this theme by describing pulse scores, surveys, annual user ratings, and sustainability considerations as part of AI governance and improvement. Participant 9 explained that the organization uses feedback mechanisms after awareness events and annual ratings to assess how enterprise AI and IT support the scientific user community. Participant 9 also discussed sustainability guardrails, including the need to evaluate vendors and encourage right sized model use so advanced models are not used unnecessarily for simple tasks. This response extends the theme by showing that continuous improvement may include not only performance and adoption metrics but also sustainability, resource

use, and responsible vendor evaluation.

These findings align closely with the responsible innovation framework. The theme reflects responsiveness because participants emphasized revising policies, adjusting controls, monitoring model performance, reviewing user feedback, and adapting AI practices as organizational needs and technology risks evolve. The theme also reflects anticipation because participants described defining metrics, success criteria, and evaluation processes before scaling AI tools. In addition, the theme reflects reflexivity because participants recognized that AI systems require ongoing review to identify limitations, drift, user misuse, or unintended consequences.

The findings also align with current public guidance and scholarship on AI governance. NIST (2023) emphasized that AI risk management includes measuring and managing risks throughout the AI lifecycle. The current findings confirm this guidance by showing that leaders use surveys, adoption metrics, usage tracking, compliance reviews, QA dashboards, model monitoring, policy revalidation, and feedback loops to evaluate AI practices. The findings extend existing knowledge by illustrating how business leaders operationalize continuous improvement through living policies, pilot success criteria, pulse surveys, customer satisfaction metrics, A/B testing, model version monitoring, and sustainability guardrails.

Collectively, these findings indicate that continuous monitoring, measurement, and adaptive improvement are essential to safe and ethical AI adoption. Participants emphasized that AI governance should not end when a tool is approved or deployed. Instead, business leaders should establish ongoing evaluation mechanisms to determine whether AI remains accurate, secure, useful, compliant, ethical, and aligned with organizational objectives. These findings suggest that organizations should define measurable success criteria, track usage and user feedback, monitor model performance, review policies regularly, evaluate vendor and sustainability risks, and revise governance practices as AI systems evolve. This theme completes the set of findings by showing that responsible AI adoption is a continuous cycle of learning, measuring, and improvement rather than a fixed implementation event.

5. Recommendations for Professional Practice

The findings of this qualitative pragmatic inquiry provide practical insights for business leaders seeking to develop and apply safe and ethical practices when adopting AI technologies. Participants emphasized that responsible AI adoption requires an integrated business approach that combines governance, data protection, human oversight, controlled implementation, workforce readiness, and continuous improvement. These findings are significant for professional practice because AI adoption is no longer only a technical issue. It directly affects risk management, regulatory readiness, customer trust, workforce behavior, operational performance, and organizational reputation. Business leaders who adopt AI without clear governance and ethical safeguards may increase risks related to data ex-

posure, biased outputs, employee misuse, poor decision quality, and loss of stakeholder confidence.

Structured AI governance and policy development emerged as a central contribution to professional practice. Participants emphasized that organizations need formal AI policies, acceptable use guidelines, AI councils, executive sponsorship, governance review, and cross-functional decision-making before AI tools are broadly adopted. This finding suggests that business leaders should not rely on informal experimentation or individual employee judgment to guide AI use. Instead, leaders should establish formal governance structures that define approved AI tools, acceptable uses, data restrictions, oversight responsibilities, escalation procedures, and accountability expectations. This finding aligns with the National Institute of Standards and Technology AI Risk Management Framework, which identifies governance as a core function for managing AI risks across the AI lifecycle ([National Institute of Standards and Technology \(NIST\), 2023](#)). For business practice, this means leaders should create governance bodies that include executive leadership, legal counsel, cybersecurity, data governance, technology teams, business unit leaders, and end user representation.

The findings also contribute to business practice by demonstrating that data governance, privacy, and confidentiality protection must be treated as central requirements for AI adoption. Participants identified data quality, intellectual property protection, tenant controlled AI tools, zero trust access, patient data protection, and prevention of data leakage as critical strategies. These findings suggest that business leaders should establish data governance protocols before expanding AI use across the organization. Leaders should define authoritative data sources, classify data by sensitivity, restrict confidential or regulated data from unapproved AI tools, and ensure vendors meet organizational data protection requirements. This recommendation is consistent with scholarship emphasizing that responsible AI requires privacy, accountability, transparency, and trustworthy data practices ([Floridi et al., 2018](#); [Jobin et al., 2019](#)). It is also consistent with AI management system principles, which emphasize policies, objectives, and processes for responsible AI development and use.

A further contribution to professional practice is the finding that human oversight, AI constraint, and quality assurance are essential for ethical AI use. Participants emphasized that AI should assist human decision-making but should not independently make final decisions in sensitive, safety-related, clinical, legal, security, or high-impact contexts. The findings suggest that business leaders should implement human-in-the-loop review, establish AI decision boundaries, require subject matter expert validation, and use quality assurance processes to monitor AI output accuracy. This recommendation confirms prior research indicating that algorithmic systems create ethical risks when accountability is unclear or when organizations fail to maintain meaningful human oversight ([Mittelstadt et al., 2016](#)). In practice, business leaders should classify AI use cases by risk level and determine which uses require human review, legal review, technical validation, execu-

tive approval, or continued monitoring before deployment.

The findings further contribute to practice by showing that phased pilots, controlled experimentation, and use case prioritization reduce implementation risk. Participants described controlled rollouts, safe testing environments, limited user groups, proof of concept activities, use case review, and predefined success criteria as effective strategies. These findings suggest that organizations should avoid broad enterprise deployment until AI tools have been tested in controlled settings. Leaders should use pilot programs to assess usability, data exposure, output quality, employee adoption, cost, operational value, and risk. Pilot results should guide whether AI tools are scaled, revised, delayed, or discontinued. This finding aligns with federal AI guidance emphasizing the connection between governance, innovation, and risk management, particularly for AI uses that may affect rights, safety, or public trust. For professional practice, the recommendation is that business leaders create a formal AI intake and pilot review process before approving enterprise wide adoption.

Another important contribution is the finding that workforce training, communication, and change management are required for responsible AI adoption. Participants described employees as both users and potential sources of risk. Several participants noted that employees may apply personal AI habits in the workplace without understanding organizational data restrictions, confidentiality risks, or approved tool requirements. Others emphasized that employees may resist AI restrictions unless leaders clearly explain that controls are designed to protect the organization, employees, customers, and stakeholders. Business leaders should therefore implement mandatory AI literacy training, acceptable use education, scenario based learning, and recurring communication about approved tools and prohibited uses. Leaders should also frame AI governance as an enablement strategy rather than a barrier to productivity. This finding extends professional practice by showing that ethical AI adoption is not only a governance or technology issue; it is also a workforce readiness and organizational change issue.

The final contribution is the finding that AI governance must be continuously monitored, measured, and improved. Participants described surveys, usage tracking, customer satisfaction measures, A/B testing, quality dashboards, compliance reviews, model performance tracking, policy revalidation, and sustainability considerations as mechanisms for measuring AI effectiveness. This finding suggests that organizations should treat AI governance as a living framework rather than a static policy. Leaders should define measurable success criteria before implementation and revise policies as AI tools, risks, regulations, and organizational needs evolve. This recommendation aligns with the responsible innovation framework, particularly responsiveness, because leaders must adjust practices when new risks, unintended consequences, or changing conditions appear (Stilgoe et al., 2013). It also aligns with responsible AI guidance that emphasizes ongoing governance, measurement, and risk management throughout the AI lifecycle.

For business and organizational leaders, these findings provide several action-

able recommendations. First, leaders should establish an enterprise AI governance structure before broad AI adoption. This structure should include executive sponsorship, legal representation, cybersecurity expertise, data governance leadership, technical expertise, business unit input, and user representation. Second, leaders should create an AI acceptable use policy that identifies approved tools, prohibited tools, data restrictions, employee responsibilities, monitoring expectations, and escalation procedures. Third, leaders should require data classification and data protection controls before AI tools are used with organizational information. Fourth, leaders should use phased pilots and safe testing environments to validate AI use cases before enterprise deployment. Fifth, leaders should require human oversight for high-risk or sensitive AI use cases. Sixth, leaders should establish ongoing measurement practices, including adoption metrics, output accuracy, user feedback, customer impact measures, business value indicators, compliance reviews, and model performance monitoring.

For the research scholar community, the findings provide an applied contribution by connecting the responsible innovation framework to real-world AI implementation practices. Responsible innovation includes the dimensions of anticipation, inclusion, reflexivity, and responsiveness (Stilgoe et al., 2013). The findings show how these dimensions appeared in business practice. Anticipation appeared through risk assessment, data governance, controlled pilots, and predefined success criteria. Inclusion appeared through AI councils, stakeholder engagement, user education, and cross-functional governance. Reflexivity appeared through human oversight, ethical review, quality assurance, and recognition of AI limitations. Responsiveness appeared through continuous monitoring, policy revalidation, model drift monitoring, user feedback, and adaptive improvement. These findings extend the research literature by showing how business leaders translate responsible innovation principles into practical actions, routines, and controls.

Overall, the findings strengthen professional practice by offering a practical model for safe and ethical AI adoption. Business leaders can use these findings to move from fragmented experimentation toward disciplined AI governance. Rather than adopting AI solely for efficiency or productivity, leaders should evaluate whether AI use is governed, secure, explainable, monitored, aligned with policy, and beneficial to stakeholders. Organizations that implement governance, data protection, human oversight, pilot testing, workforce training, and continuous improvement may reduce the risk of data exposure, improve accountability, increase employee confidence, strengthen compliance readiness, and protect organizational trust. As AI technologies continue to evolve, organizations that embed these strategies into their AI adoption processes will be better positioned to innovate responsibly, manage risk effectively, and sustain stakeholder confidence.

6. Implications for Social Change

The findings of this qualitative pragmatic inquiry may contribute to positive social change by encouraging business and organizational leaders to adopt AI in ways

that are more transparent, equitable, accountable, and protective of stakeholders. Participants emphasized that responsible AI adoption requires formal governance, data protection, human oversight, controlled implementation, workforce training, and continuous monitoring. These findings suggest that ethical AI adoption is not limited to internal business performance. Instead, responsible AI practices may influence individuals, communities, organizations, institutions, and broader society because AI systems increasingly shape decisions that affect employment, healthcare, finance, customer service, education, public services, and access to information.

At the individual level, the findings may support positive social change by helping reduce harm caused by biased, inaccurate, or poorly governed AI systems. AI-assisted decisions can affect people's opportunities, privacy, access to services, and trust in organizations. If business leaders implement the strategies identified in this study, individuals may experience fairer and more transparent treatment in AI-supported decisions. For example, formal governance, bias testing, human-in-the-loop review, and quality assurance practices can reduce the likelihood that AI systems will produce decisions based on incomplete, inaccurate, or biased data. Similarly, stronger data governance and privacy protections can help prevent unauthorized exposure of personal information, confidential records, and sensitive data. These actions support social change by protecting individual dignity, privacy, fairness, and informed participation in technology enabled environments.

At the community level, the findings may contribute to more responsible and inclusive use of AI systems. Communities are affected when organizations use AI to shape access to services, employment opportunities, financial products, healthcare support, or public-facing information. If leaders apply the findings of this study, communities may benefit from AI systems that are more carefully governed, monitored, and aligned with ethical expectations. Participants' emphasis on stakeholder engagement, workforce communication, controlled pilots, and continuous monitoring suggests that responsible AI adoption can help organizations identify unintended consequences before they create broader social harm. These practices may also increase community trust because organizations that communicate clearly, protect data, and maintain human accountability are better positioned to demonstrate that AI is being used responsibly.

At the organizational level, the findings may promote positive social change by encouraging leaders to move from informal AI experimentation to structured, principle-based governance. Organizations help shape social norms through the practices they adopt and reinforce. When organizations normalize AI governance, privacy protection, human oversight, training, and continuous improvement, they influence how employees and managers understand responsible innovation. Over time, these practices may create organizational cultures in which AI adoption is evaluated not only by speed, cost savings, or productivity but also by fairness, accountability, transparency, and stakeholder impact. This shift may help organizations become more trustworthy institutions and may reduce the risk of ethical failures that damage employees, customers, and the public.

The findings also have implications for institutions and professional practice. Business leaders, professional associations, policy influencers, and industry groups may use the findings to strengthen expectations for responsible AI adoption. High-level AI principles are common across public documents and ethical frameworks, but organizations often struggle to translate those principles into daily business practices. This study contributes to social change by identifying practical strategies leaders can use to operationalize ethical AI principles. These strategies include creating AI councils, developing acceptable use policies, protecting sensitive data, requiring human review for high-risk use cases, piloting tools before scaling, training employees, measuring AI performance, and revising policies as risks evolve. These practices may help bridge the gap between broad ethical principles and practical organizational action.

The findings may also inform policy conversations related to AI governance and accountability. Although this study focused on business leader strategies, the findings may be useful to policymakers and professional bodies seeking to understand how organizations can implement ethical AI in practice. Participants described several mechanisms that could inform broader policy development, including risk assessment, vendor review, approved AI tool lists, data classification, documentation, human oversight, employee training, and performance monitoring. When organizations demonstrate feasible approaches to responsible AI governance, those approaches may influence future standards, compliance expectations, and professional guidelines. In this way, the findings may support positive social change beyond the participating organizations by contributing to broader discussions about responsible technology governance.

For the research scholar community, this study may contribute to positive social change by advancing applied knowledge about how the responsible innovation framework can be used in organizational AI adoption. The findings show how anticipation, inclusion, reflexivity, and responsiveness appear in practice. Anticipation appeared through risk assessment, governance planning, data protection, and pilot testing. Inclusion appeared through stakeholder engagement, AI councils, workforce communication, and user training. Reflexivity appeared through human oversight, ethical review, and recognition of AI limitations. Responsiveness appeared through policy revalidation, monitoring, user feedback, and adaptive improvement. These findings may encourage future researchers to examine how responsible innovation principles can be translated into practical strategies across industries, organization sizes, and regulatory contexts.

Collectively, the implications for social change are tangible and multidimensional. The findings may help leaders adopt AI in ways that protect individuals, support communities, strengthen organizational accountability, and inform broader professional and policy practices. By promoting governance practices that reduce bias, improve transparency, protect data, require human oversight, and support continuous improvement, this study contributes to a more responsible model of technological change. For business leaders, ethical AI adoption may become a

pathway not only to improved organizational performance but also to greater public trust, fairness, and social benefit. For the research scholar community, the findings provide evidence that responsible innovation in AI is not only an abstract framework but also a practical foundation for positive social change in contemporary business environments.

7. Directions for Further Research

The findings of this study contribute to the understanding of strategies used by business leaders to develop and apply safe and ethical practices when adopting AI technologies. However, several opportunities exist for future research to expand upon these findings and further improve business practices. As AI continues to evolve rapidly across industries, additional research is necessary to deepen understanding of how organizations can effectively balance innovation, governance, and ethical responsibility (Stilgoe et al., 2013; Wirtz et al., 2023).

One recommendation for future research is to conduct studies across a broader range of industries and organizational sizes. This study focused on leaders with experience in AI adoption, many of whom operate within structured or regulated environments. Future research could examine how small and medium-sized businesses implement ethical AI practices, as these organizations often have fewer resources and less formal governance structures. Expanding the scope to include diverse organizational contexts would provide a more comprehensive understanding of how ethical AI strategies can be adapted across varying business environments (Creswell & Creswell, 2023).

Another area for future study is the longitudinal evaluation of AI governance strategies. This study provides a cross-sectional view of leadership practices at a specific point in time. However, AI technologies and associated risks continue to evolve. Longitudinal research could examine how organizations adapt their governance frameworks over time and how the long-term effectiveness of ethical AI practices impacts organizational performance, risk mitigation, and stakeholder trust. This aligns with the need for continuous monitoring and adaptation emphasized in AI governance frameworks (National Institute of Standards and Technology (NIST), 2023).

Future research should also explore the role of organizational culture in shaping ethical AI adoption. While this study identified leadership strategies and governance practices, further investigation is needed to understand how organizational values, employee behavior, and cultural norms influence the successful implementation of these strategies. Organizational culture has been identified as a critical factor in technology adoption and ethical decision-making, particularly in environments undergoing digital transformation (Schein, 2017).

Additionally, future studies could incorporate quantitative or mixed-methods approaches to complement qualitative findings. While this study provides rich, in-depth insights into leadership strategies, quantitative research could measure the effectiveness of specific AI governance practices, such as the impact of data governance policies, training programs, or human-in-the-loop controls on organizational

outcomes. Mixed-methods approaches are particularly valuable in providing both depth and generalizability in research findings (Creswell & Creswell, 2023).

Another important recommendation is to examine the role of emerging regulatory frameworks and policy developments in shaping organizational AI practices. As governments and regulatory bodies continue to develop guidelines and requirements for AI use, future research could investigate how organizations interpret, implement, and comply with these evolving standards. The increasing global focus on AI ethics highlights the need for alignment between policy and practice (Jobin et al., 2019).

The limitations identified in Section 1.4 of this study provide additional direction for future research. One limitation relates to the sample size and use of purposive sampling, which may limit the generalizability of the findings. Future studies could address this limitation by increasing sample size and incorporating a more diverse participant pool across industries, geographic regions, and organizational roles. This would enhance the transferability of findings and provide a broader perspective on AI adoption practices (Creswell & Poth, 2018).

Another limitation involves the reliance on self-reported data from participants, which may introduce bias or reflect subjective interpretations of organizational practices. Future research could address this limitation by incorporating additional data sources, such as organizational documents, performance metrics, or observational data, to triangulate findings and strengthen validity. Triangulation is a widely recognized strategy for improving credibility in qualitative research (Carter et al., 2014).

A further limitation relates to the rapid evolution of AI technologies, which may affect the long-term relevance of the findings. Future studies can address this by focusing on adaptable frameworks and continuously updating research to reflect emerging technologies and practices. Researchers may also explore specific AI domains, such as generative AI or autonomous systems, to provide more targeted insights into evolving challenges and strategies (Floridi et al., 2018).

8. Limitations

This research project had a few limitations. The number of business leaders included in the research project may have limited the generalizability of the findings outside the participant group. Whilst purposeful sampling is suitable for a qualitative study, the views depicted may not encompass all strategies adopted in different industries or locations. The research project may be limited by the research methodology and design used to conduct it.

9. Conclusion

The purpose of this qualitative pragmatic inquiry project was to identify and explore effective strategies business leaders use to develop and apply safe and ethical practices when adopting AI technologies for use. The overarching research question guiding this research project was: What effective strategies do business lead-

ers use to develop and apply safe and ethical practices when adopting AI technologies for use? Data collected from semistructured interviews with nine business and technology leaders revealed six major themes: structured AI governance and policy development; data governance, privacy, and confidentiality protection; human oversight, AI constraint, and quality assurance; phased pilots, controlled experimentation, and use case prioritization; workforce training, communication, and change management; and continuous monitoring, measurement, and adaptive improvement.

The findings of this research project suggest that organizational leaders seeking to adopt AI safely and ethically should implement integrated governance systems that combine policy, data protection, human accountability, controlled implementation, workforce readiness, and ongoing evaluation. Participants emphasized that responsible AI adoption should not occur through informal experimentation or isolated technical decisions. Instead, leaders should establish formal governance bodies, define acceptable use policies, protect sensitive data, limit AI authority in high-risk contexts, test AI tools before scaling, train employees, and monitor AI systems continuously. These strategies may help organizations reduce risks related to data exposure, biased outputs, unclear accountability, employee misuse, poor decision quality, and loss of stakeholder trust.

The findings of this research project align with the responsible innovation framework, which emphasizes anticipation, inclusion, reflexivity, and responsiveness. Anticipation was reflected in participants' emphasis on risk assessments, data governance, phased pilots, and predefined success criteria before broad AI deployment. Inclusion was reflected in participants' emphasis on AI councils, stakeholder engagement, cross-functional governance, user training, and communication with employees and customers. Reflexivity was reflected in participants' recognition that AI systems have limitations and require human oversight, ethical review, quality assurance, and validation. Responsiveness was reflected in participants' emphasis on policy revalidation, continuous monitoring, model performance tracking, user feedback, and adaptive improvement as AI tools and organizational risks evolve.

Overall, the findings of this research project contribute to the body of knowledge on responsible AI adoption by identifying practical strategies business leaders use to translate ethical AI principles into organizational practice. The findings indicate that safe and ethical AI adoption requires more than technical implementation or compliance language. Organizations that intentionally implement AI governance, data protection, human oversight, controlled pilots, workforce training, and continuous monitoring may be better positioned to innovate responsibly, manage risk effectively, improve employee confidence, strengthen compliance readiness, and sustain stakeholder trust. As AI technologies continue to influence business operations and decision-making, the strategies identified in this project may help business leaders adopt AI in ways that are not only efficient but also ethical, accountable, and socially responsible.

Author Contributions

Vini Ehsan was the primary researcher. Dr. Kim A. Critchlow was the mentor/reviewer/approver of the research being performed.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Batool, A., Zowghi, D., & Bano, M. (2025). AI Governance: A Systematic Literature Review. *AI and Ethics*, 5, 3265-3279. <https://doi.org/10.1007/s43681-024-00653-w>
- Blok, V., & Lemmens, P. (2015). The Emerging Concept of Responsible Innovation. Three Reasons Why It Is Questionable and Calls for a Radical Transformation of the Concept of Innovation. In B. J. Koops, I. Oosterlaken, H. Romijn, T. Swierstra, & J. van den Hoven (Eds.), *Responsible Innovation 2* (pp. 19-35). Springer International Publishing. https://doi.org/10.1007/978-3-319-17308-5_2
- Braun, V., & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology*, 3, 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- Carter, N., Bryant-Lukosius, D., DiCenso, A., Blythe, J., & Neville, A. J. (2014). The Use of Triangulation in Qualitative Research. *Oncology Nursing Forum*, 41, 545-547. <https://doi.org/10.1188/14.onf.545-547>
- Creswell, J. W., & Creswell, J. D. (2023). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (6th ed.). Sage.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative Inquiry and Research Design: Choosing among Five Approaches* (4th ed.). Sage.
- Deloitte (2024). *State of AI Governance Report 2024: Ethical Maturity and Business Resilience in AI Adoption*. Deloitte Insights. <https://www.deloitte.com/insights>
- Dewey, J. (1920). *Reconstruction in Philosophy*. Henry Holt & Company.
- European Parliament and Council of the European Union (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying down Harmonised Rules on Artificial Intelligence. *Official Journal of the European Union*, L2024, 1689. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>
- Fancher, D. Foster, J., & Hess, E. (n.d.). Perspective: What Are the Characteristics of an Effective Governance Program? Deloitte Insights. <https://www.deloitte.com/us/en/programs/chief-legal-officer/articles/generative-ai-governance-considerations-for-clos.html>
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V. et al. (2018). AI4People: An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, 28, 689-707. <https://doi.org/10.1007/s11023-018-9482-5>
- Handiman, U. T., Rahardja, U., Simarmata, J., Santoso, N. et al. (2025a). AI Ethics and Green Leadership: Towards a Responsible Tech-Driven Sustainability Framework. *International Journal of Economics and Management Systems*, 10, 184-194.
- Handiman, U. T., Wijaya, H., & Setiawan, R. (2025b). Sustainable Leadership and Responsible AI Integration in Organizations. *International Journal of Management Science and Information Technology*, 25, 14-29.
- Handiman, U. T., Wijaya, L., & Setiawan, D. (2025c). Integrating AI Ethics into Corporate Sustainability Leadership Models. *Technology in Society*, 75, Article ID: 102382.

- Hennink, M. M., Kaiser, B. N., & Marconi, V. C. (2017). Code Saturation versus Meaning Saturation: How Many Interviews Are Enough? *Qualitative Health Research*, 27, 591-608. <https://doi.org/10.1177/1049732316665344>
- IBM (2023). *CEO Guide to Ethical AI Adoption*. IBM Institute for Business Value. <https://www.ibm.com/thought-leadership/institute-business-value>
- International Organization for Standardization (ISO) (2024). *ISO/IEC 42001:2023-AI Management System-Requirements*. ISO.
- Jobin, A., Ienca, M., & Vayena, E. (2019). The Global Landscape of AI Ethics Guidelines. *Nature Machine Intelligence*, 1, 389-399. <https://doi.org/10.1038/s42256-019-0088-2>
- Kumar, R., Carabantes, J., & Binns, R. (2023). Responsible AI Governance: Bridging Technical and Organizational Perspectives. *Journal of Business Ethics and Technology*, 6, 102-119.
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The Ethics of Algorithms: Mapping the Debate. *Big Data & Society*, 3, 1-21. <https://doi.org/10.1177/2053951716679679>
- Morgan, D. L. (2014). Pragmatism as a Paradigm for Social Research. *Qualitative Inquiry*, 20, 1045-1053. <https://doi.org/10.1177/1077800413513733>
- National Institute of Standards and Technology (NIST) (2023). *AI Risk Management Framework (AI RMF 1.0)*. <https://www.nist.gov/itl/ai-risk-management-framework>
- Northouse, P. G. (2024). *Leadership: Theory and Practice* (10th ed.). SAGE Publications.
- Novitzky, P., Bernstein, M. J., Blok, V., Braun, R., Chan, T. T., Lamers, W. et al. (2020). Improve Alignment of Research Policy and Societal Values. *Science*, 369, 39-41. <https://doi.org/10.1126/science.abb3415>
- Organisation for Economic Co-Operation and Development (OECD) (n.d.). *OECD AI Policy Observatory*. OECD. <https://unaiwork.org/frameworks/oecd-ai-policy-observatory>
- Owen, R., Macnaghten, P., & Stilgoe, J. (2013). Responsible Research and Innovation: From Science in Society to Science for Society, with Society. *Science and Public Policy*, 39, 751-760. <https://doi.org/10.1093/scipol/scs093>
- Patton, M. Q. (2015). *Qualitative Research and Evaluation Methods* (4th ed.). SAGE.
- Perera, H., Lee, S. U., Liu, Y., Xia, B., Lu, Q., Zhu, L., Cairns, J., & Nottage, M. (2024). *Achieving Responsible AI through ESG: Insights and Recommendations from Industry Engagement*. <https://doi.org/10.1109/MS.2026.3689573>
- PricewaterhouseCoopers (PwC) (2024a). *AI Business Outlook 2024: Building Trust in the Age of Intelligent Enterprise*. PwC Research. <https://www.pwc.com>
- PricewaterhouseCoopers (PwC) (2024b). *Enterprise AI readiness assessment*. PwC Insights. <https://www.pwc.com/us/en/services/consulting/strategy/enterprise-functional-strategy/ai-readiness-assessment.html>
- Saldaña, J. (2021). *The Coding Manual for Qualitative Researchers* (4th ed.). SAGE Publications.
- Santos, M. J., Zhang, Y., & Aveling, E. (2022). Organizational Readiness for Ethical AI: A Qualitative Study of Leadership and Policy Gaps. *AI and Society*, 37, 589-605.
- Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B. et al. (2018). Saturation in Qualitative Research: Exploring Its Conceptualization and Operationalization. *Quality & Quantity*, 52, 1893-1907. <https://doi.org/10.1007/s11135-017-0574-8>
- Schein, E. H. (2017). *Organizational Culture and Leadership* (5th ed.). Wiley.

- Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a Framework for Responsible Innovation. *Research Policy*, 42, 1568-1580. <https://doi.org/10.1016/j.respol.2013.05.008>
- U.S. Government Accountability Office (GAO) (2023a). *AI: An Accountability Framework for Federal Agencies and Entities*. GAO-23-106294. <https://www.gao.gov>
- U.S. Government Accountability Office (GAO) (2023b). *AI: Key Practices for Responsible Adoption* (GAO-23-105451). <https://www.gao.gov/>
- Uddin, A. S. M. A. (2023). The Era of AI: Upholding Ethical Leadership. *Open Journal of Leadership*, 12, 400-417. <https://doi.org/10.4236/ojl.2023.124019>
- United Nations Educational, Scientific and Cultural Organization (UNESCO) (2023). *Recommendation on the Ethics of AI: Implementation Progress Report*. UNESCO. <https://www.unesco.org>
- Von Schomberg, R. (2013). A Vision of Responsible Research and Innovation. In R. Owen, M. Heintz, & J. Bessant (Eds.), *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society* (pp. 51-74). John Wiley & Sons.
- Wirtz, J., Kunz, W. H., Hartley, N., & Tarbit, J. (2023). Corporate Digital Responsibility in Service Firms and Their Ecosystems. *Journal of Service Research*, 26, 173-190. <https://doi.org/10.1177/10946705221130467>
- World Economic Forum (2024). *AI at Work: Reskilling and Ethical Literacy in the Age of Intelligent Automation*. World Economic Forum. <https://www.weforum.org>
- Zhang, S. X., Chen, J., He, L., & Choudhury, A. (2023). Responsible Innovation: The Development and Validation of a Scale. *Technovation*, 124, Article ID: 102754. <https://doi.org/10.1016/j.technovation.2023.102754>