ISSN: 2755-0214

Journal of Economics & Management Research

SCIENTIFIC Research and Community

Review Article

Open Access

The Ethical Implications of Artificial Intelligence in Decision-Making: Balancing Innovation with Accountability and Integrity

Shankar Subramanian Iyer [1]*, Raman Subramanian[2], Brinitha Raji[3] and Rajesh Arora[4]

[1]Faculty Business, Westford University College, Sharjah, UAE

[2]Associate Dean, Westford University College, Al Tawuun, Sharjah, UAE

[3]Global Business Studies, DKP, Dubai

[4]Faculty; Westford University College, Al Khan, Sharjah

ABSTRACT

As Artificial Intelligence (AI) technologies become increasingly embedded in critical organizational decision-making processes, questions of ethics, accountability, and integrity have risen to the forefront of academic and industry discourse. This study investigates the ethical implications of AI in decision-making, with a specific focus on how organizations can balance technological innovation with moral responsibility and regulatory accountability. Guided by an integrated theoretical framework that combines Ethical Decision-Making Theory, Stakeholder Theory, the Technology Acceptance Model (TAM), and the Responsible Innovation Framework, this research provides a multidimensional perspective on AI ethics in practice.

A quantitative research methodology was employed, gathering data from 395 stakeholders including AI developers, corporate leaders, ethics compliance officers, public regulators, data scientists, and end-users across sectors such as finance, healthcare, education, and government services. The study examines stakeholder perceptions on key ethical dimensions of AI implementation—algorithmic transparency, data privacy, bias mitigation, informed consent, accountability mechanisms, and the inclusiveness of AI design. A structured survey instrument was used to assess how these variables influence organizational trust, adoption willingness, and public acceptance of AI-driven decisions.

The findings are expected to demonstrate that organizations that embed ethical safeguards and transparent governance frameworks within their AI systems can significantly enhance stakeholder trust, user engagement, and regulatory compliance. The study also highlights critical challenges, including the opacity of black-box algorithms, the absence of global ethical standards, the underrepresentation of marginalized groups in algorithmic training data, and the tensions between innovation speed and ethical reflection. Additionally, it explores how external institutional pressures—such as media scrutiny, customer expectations for ethical AI, investor demand for ESG-compliant practices, and evolving legal mandates—act as catalysts for adopting ethical AI governance models.

This research makes several key contributions. First, it bridges a significant gap in the empirical literature on AI ethics by quantitatively analyzing perceptions across diverse industries and stakeholder categories. Second, it offers a conceptual model for balancing innovation with ethical accountability that can guide practitioners, policymakers, and technologists in AI strategy formulation. Third, it provides a framework for integrating responsible AI principles into corporate governance and digital transformation strategies. The study ultimately advocates for the co-evolution of ethical standards and AI capabilities to ensure that technological progress does not come at the expense of human values, social justice, or democratic integrity.

***Corresponding author**

Shankar Subramanian Iyer, Faculty Business, Westford University College, Sharjah, UAE.

Received: September 25, 2025; **Accepted:** September 27, 2025; **Published:** October 09, 2025

Keywords

Artificial Intelligence, Ethical Decision-Making, Algorithmic Accountability, Innovation Governance, Transparency, Bias Mitigation, Data Ethics, Stakeholder Theory, Responsible Innovation, Technology Acceptance Model, Quantitative Research, Integrity in AI, Regulatory Compliance, Human-Centered AI, Institutional Pressures, Trust in AI Systems

Introduction and Background

The accelerating deployment of Artificial Intelligence (AI) in decision-making processes has ushered in a transformative era for organizations, governments, and societies at large. AI-powered systems are now used to make or support critical decisions in domains such as finance, healthcare, criminal justice, education, human resources, and public policy. These systems—ranging from predictive analytics tools to autonomous recommendation engines—are valued for their ability to process vast volumes of data with speed and precision, offering operational efficiency, cost reduction, and competitive advantage. However, the growing autonomy and opacity of AI in high-stakes decision-making raise serious ethical concerns. Questions surrounding algorithmic bias, lack of transparency, erosion of human oversight, privacy violations, and uneven accountability are no longer speculative—they are real-world challenges that have already led to tangible harms and social controversies [1].

In this context, ethical AI is not a peripheral issue but a core concern that influences public trust, legal compliance, brand reputation, and the overall legitimacy of AI technologies. The infamous example of racial discrimination in facial recognition systems, algorithmic exclusion in loan approval processes, and biased recruitment software have all revealed the profound risks of unchecked AI. These examples underscore a critical tension: while AI can drive innovation and transformation, it can also exacerbate inequality, institutionalize bias, and undermine democratic and human rights if not governed responsibly. Therefore, there is a growing consensus among scholars, industry leaders, and policymakers that the future of AI must be guided by strong ethical principles such as transparency, fairness, accountability, inclusiveness, and respect for human agency [2].

Despite the proliferation of ethical AI frameworks from entities such as the IEEE, OECD, EU Commission, and national AI strategies, their implementation remains inconsistent and often symbolic. Many organizations still struggle to convert ethical aspirations into actionable governance mechanisms, technical design principles, and measurable outcomes. Moreover, ethical decision-making in AI is not merely a technical challenge—it is socio-technical and institutional, involving a wide array of stakeholders with diverse interests, values, and levels of understanding. This makes the ethical deployment of AI a multidimensional governance issue requiring holistic approaches that integrate organizational culture, technological design, stakeholder engagement, and regulatory compliance [3].

Against this backdrop, this study investigates the topic: "The Ethical Implications of Artificial Intelligence in Decision-Making: Balancing Innovation with Accountability and Integrity." It addresses the need to empirically understand how ethical concerns are perceived, interpreted, and managed by stakeholders across industries where AI is actively shaping decisions. The study is grounded in a comprehensive conceptual model combining four key theoretical

Perspectives

- Ethical Decision-Making Theory, which examines how individuals and organizations approach moral dilemmas, especially under uncertainty.
- Stakeholder Theory, which emphasizes the need to consider the rights and interests of all parties affected by technological decisions.
- Technology Acceptance Model (TAM), which helps assess how ethical factors influence user trust and acceptance of AI technologies
- Responsible Innovation Framework, which advocates for anticipatory, reflexive, and inclusive innovation processes that align with societal values.

The research scope encompasses a quantitative assessment involving 395 stakeholders from sectors such as healthcare, finance, education, e-commerce, and government, where AI is significantly influencing strategic and operational decisions. Respondents include AI developers, data scientists, senior managers, policymakers, compliance officers, and end-users, offering a 360-degree view of the ethical AI landscape. By employing a structured survey instrument, the study seeks to uncover how key constructs such as algorithmic transparency, bias mitigation, explainability, consent mechanisms, and accountability influence perceptions of fairness, trust, and adoption [4].

The Following Research Questions Guide the Investigation:

1. What are the dominant ethical concerns perceived by stakeholders regarding the use of AI in decision-making?
2. To what extent do internal organizational capabilities—such as leadership commitment, ethical culture, and technical training—influence the responsible implementation of AI?
3. How do external institutional pressures—such as regulatory compliance, ESG expectations, customer activism, and media scrutiny—moderate the relationship between ethical practices and AI acceptance?
4. How can a conceptual model be developed that aligns AI innovation with the principles of accountability, integrity, and stakeholder trust?

Accordingly, the Research Objectives are:

- To evaluate the impact of ethical AI practices on stakeholder trust, system legitimacy, and organizational reputation.
- To examine the role of organizational capabilities in institutionalizing ethical AI governance.
- To assess how external institutional factors influence ethical compliance and stakeholder perceptions.
- To develop and validate an ethical AI decision-making model that guides organizations in balancing technological innovation with social responsibility.

By aligning the research questions and objectives with a robust theoretical and conceptual framework, this study aims to make both academic and practical contributions. It seeks to offer empirical evidence that supports the integration of ethical governance mechanisms in AI deployment strategies. Moreover, it provides actionable recommendations for organizations, policymakers, and technologists striving to build responsible AI systems that serve not only the bottom line but also broader societal interests. In doing so, this study responds to a pressing global need: ensuring that the future of AI is not only intelligent but also ethical, inclusive, and just.

Literature Review

The ethical implications of Artificial Intelligence (AI) in decision-making have emerged as a critical area of inquiry, particularly as AI systems increasingly influence sensitive domains such as finance, healthcare, hiring, criminal justice, and public administration. Several scholars have emphasized that while AI can enhance decision-making by offering precision, speed, and data-driven insights, its unregulated deployment poses risks to human autonomy, fairness, accountability, and social justice [5,6]. This literature review explores current scholarship aligned with four interrelated theoretical frameworks—Ethical Decision-Making Theory, Stakeholder Theory, Technology Acceptance Model (TAM), and the Responsible Innovation Framework—to build a foundation for the proposed conceptual model and hypotheses.

Ethical Decision-Making Theory offers a psychological and organizational perspective on how individuals and institutions evaluate the morality of decisions. Studies applying this theory to AI argue that moral intensity, awareness, and judgment are often diluted when machines are intermediaries in human decisions [7]. Scholars contend that the delegation of decision authority to algorithms can obscure responsibility, especially when outcomes are negative or biased [8]. Yet, few empirical studies quantify how ethical awareness within organizations affects the adoption of AI systems or stakeholder trust in algorithm-driven outcomes.

Stakeholder Theory, rooted in the works of highlights that technological decisions must consider the rights and interests of all parties affected by innovation [9]. Recent work emphasizes

that AI's ethical implementation requires engaging not only technical experts but also end-users, regulators, and marginalized communities [10]. For example, stakeholder exclusion during AI system design has been shown to exacerbate systemic biases and reduce trust [11]. However, a gap remains in assessing how stakeholder-inclusive practices influence perceived accountability and acceptance in AI adoption.

The Technology Acceptance Model (TAM) has long been used to explain how perceived ease of use and usefulness drive the adoption of new technologies. In the AI ethics domain, scholars have adapted TAM to incorporate constructs like transparency, fairness, and accountability as determinants of trust and behavioral intention [12]. Studies find that systems perceived as "black boxes" suffer from low user confidence and resistance to adoption [13]. While TAM extensions are useful, empirical validation of how ethical dimensions moderate technology acceptance remains limited, particularly across diverse industry sectors and stakeholder types.

The Responsible Innovation (RI) Framework advocates for inclusive, anticipatory, reflexive, and responsive innovation processes. Propose that developers and organizations must proactively assess the societal consequences of AI systems throughout the design and deployment lifecycle [14]. Ethical AI studies informed by RI stress the importance of human oversight, value-sensitive design, and adaptive governance [15]. Despite these normative frameworks, empirical research that examines the real-world application of RI principles in AI governance structures is sparse.

Ethical AI Practices (Independent Variable)

Ethical AI practices refer to the integration of moral principles and societal values into the design, development, deployment, and governance of AI systems. These practices include algorithmic transparency, fairness, explainability, accountability, bias mitigation, privacy protection, and informed consent. When organizations adopt ethical AI practices, they actively reduce the risk of discrimination, algorithmic harm, and opaque decision-making processes. Research indicates that ethical safeguards build confidence among users and stakeholders, leading to enhanced legitimacy and support for AI initiatives. For instance, transparent AI systems that provide understandable rationale for decisions are more likely to gain trust from both internal users and external consumers. Thus, ethical AI practices serve as the foundation for fostering stakeholder trust and encouraging the acceptance of AI technologies, especially in high-risk and high-impact decisions [16].

Organizational Capabilities

Organizational capabilities refer to the internal mechanisms, competencies, and cultural attributes that enable an institution to implement, govern, and sustain ethical AI practices effectively. These include ethical leadership, cross-functional teams for AI governance, regular ethics training, investment in explainable AI technologies, and the presence of robust data governance policies. These capabilities are essential for translating abstract ethical principles into operational realities. Organizations with strong ethical capabilities are better positioned to manage risks, respond to public concerns, and institutionalize responsible innovation. These capabilities mediate the relationship between ethical AI practices and stakeholder trust by acting as enablers: without strong capabilities, ethical intentions may remain performative or superficial. Therefore, the existence of mature ethical infrastructures strengthens the positive impact of ethical AI practices on stakeholder trust and AI acceptance [17].

External Institutional Pressures

External institutional pressures include formal and informal influences from the external environment that shape organizational behavior. These encompass regulatory mandates (e.g., GDPR, AI Act), ESG (Environmental, Social, and Governance) expectations from investors, consumer advocacy, industry standards, public scrutiny, and media coverage. Such pressures compel organizations to align AI deployment with societal expectations and legal standards. When external pressures are strong, organizations are more likely to adopt and enforce ethical AI policies to maintain legitimacy and avoid reputational damage. In this way, external institutional pressures moderate the relationship between ethical AI practices and stakeholder trust/acceptance. In highly regulated or socially vigilant environments, ethical AI practices are more likely to be taken seriously by stakeholders, thereby amplifying their trust and openness toward AI use in decision-making [18].

Stakeholder Trust (Dependent Variable 1)

Stakeholder trust refers to the belief that an organization's use of AI is fair, accountable, transparent, and aligned with the stakeholders' best interests. Trust is a crucial determinant of the success and societal integration of AI technologies. When stakeholders—such as employees, customers, regulators, and the public—believe that AI systems are ethically governed and responsibly deployed, they are more likely to support their adoption. Trust reduces resistance to technological change and facilitates collaboration between technical teams and user communities. Ethical AI practices, supported by internal capabilities and responsive to external pressures, contribute directly to building this trust. A high level of trust also enhances an organization's reputation and minimizes friction in implementation [19].

AI Acceptance

AI acceptance refers to stakeholders' willingness to use, endorse, and integrate AI systems into decision-making processes. It is influenced by both cognitive and emotional factors, including perceived fairness, transparency, reliability, and ethical alignment. Drawing from the Technology Acceptance Model (TAM), AI acceptance is not just driven by functionality and ease of use but also by moral and social acceptability. Ethical AI practices that ensure fairness and reduce harm make stakeholders feel more secure, which increases their intention to adopt and support AI solutions. Organizational capabilities that institutionalize these practices, and external pressures that reinforce their importance, further enhance acceptance. Therefore, AI acceptance is both a direct outcome of perceived ethical behaviour and a barometer of stakeholder alignment with the organization's technological trajectory.

By conceptualizing the relationships among these variables, the study not only provides a theoretical foundation for understanding the ethics of AI but also offers a practical roadmap for organizations aiming to earn stakeholder trust and achieve successful AI adoption [20].

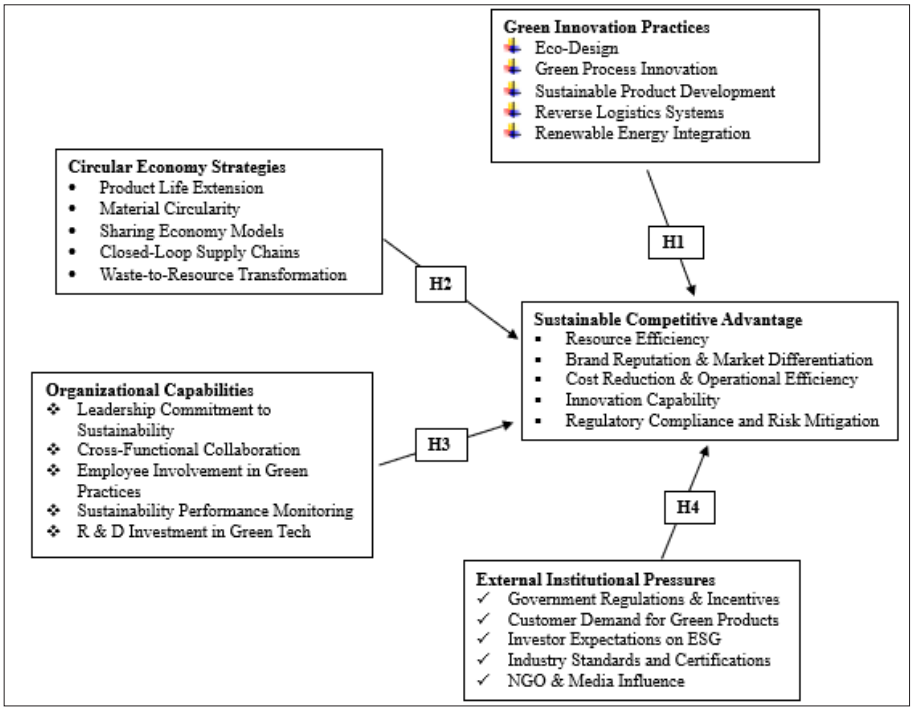
Literature Gaps

Despite rich theoretical discourse, several critical gaps persist in the existing literature. First, much of the research on AI ethics remains conceptual and lacks empirical validation, especially using large-scale stakeholder data across different sectors. Second, while several studies highlight the importance of transparency and fairness, there is insufficient integration of ethical, organizational, and institutional perspectives within a unified framework. Third,

few studies explore the moderating role of external institutional pressures—such as regulatory mandates, media scrutiny, or investor ESG expectations—on the relationship between ethical practices and AI acceptance.

Fourth, there is limited exploration of how internal organizational capabilities—such as ethical leadership, training, and cross-functional ethics committees—influence the implementation of responsible AI practices. Fifth, while stakeholder engagement is widely recognized as essential, research seldom differentiates how perceptions vary across different stakeholder groups such as developers, managers, regulators, and users. Lastly, although TAM is frequently used to examine technology adoption, its extension to include moral and institutional variables remains underdeveloped in empirical studies.

Conceptual using the integrated Natural-Resource-Based View (NRBV), Dynamic Capabilities Theory, Institutional Theory, Stakeholder Theory, and Ecological Modernization Theory (EMT)



Hypotheses

H1: The Green Innovation Practices have a significant influence on Sustainable Competitive Advantage
H2: Sustainable Competitive Advantage is significantly influenced by the Circular Economy Strategies
H3: The Organizational Capabilities have a significant influence on Sustainable Competitive Advantage
H4: Sustainable Competitive Advantage is significantly influenced by the External Institutional Pressures

Methodology

This study adopts a quantitative research methodology to systematically examine the ethical implications of Artificial Intelligence (AI) in decision-making, with a focus on how ethical practices influence stakeholder trust and acceptance. Rooted in positivist philosophy, this approach enables empirical validation of the conceptual model and hypotheses through the collection and analysis of measurable data. The choice of a quantitative design is appropriate for identifying statistically significant patterns, validating theoretical constructs, and offering generalizable insights across different organizational and industry contexts where AI is applied [21].

A cross-sectional survey-based research design was selected for this study, allowing data to be gathered from a broad and diverse population at a single point in time. This design is particularly effective for assessing perceptions, attitudes, and ethical evaluations of AI practices across multiple stakeholder groups. The survey approach is also well-suited for hypothesis testing using advanced statistical techniques such as Structural Equation Modelling (SEM), which aligns with the study's objective to evaluate complex relationships among variables including ethical AI practices, organizational capabilities, external institutional pressures, and stakeholder responses [22].

The target population consisted of stakeholders involved in AI governance, development, or utilization. This included AI developers, data scientists, senior executives (e.g., CEOs, CIOs), compliance and ethics officers, regulators, policy advisors, and end-users from industries such as healthcare, finance, education, technology, and public administration. A stratified random sampling technique was employed to ensure balanced representation across these different roles and sectors. A total of 395 valid responses were collected, ensuring the sample was statistically sufficient for robust multivariate analysis and model validation [23].

Data was collected using a structured online questionnaire, designed based on established instruments from previous empirical studies and tailored to the context of AI ethics. The questionnaire was divided into five thematic sections: demographic and organizational background, ethical AI practices (e.g., transparency, bias mitigation, explainability), organizational capabilities (e.g., ethical leadership,

governance mechanisms, training), external institutional pressures (e.g., regulatory frameworks, media scrutiny, ESG demands), and outcome variables such as stakeholder trust and AI acceptance. A 5-point Likert scale (ranging from strongly disagree to strongly agree) was used to capture respondent perceptions. The instrument was pre-tested with 15 domain experts to ensure clarity, relevance, and content validity [24].

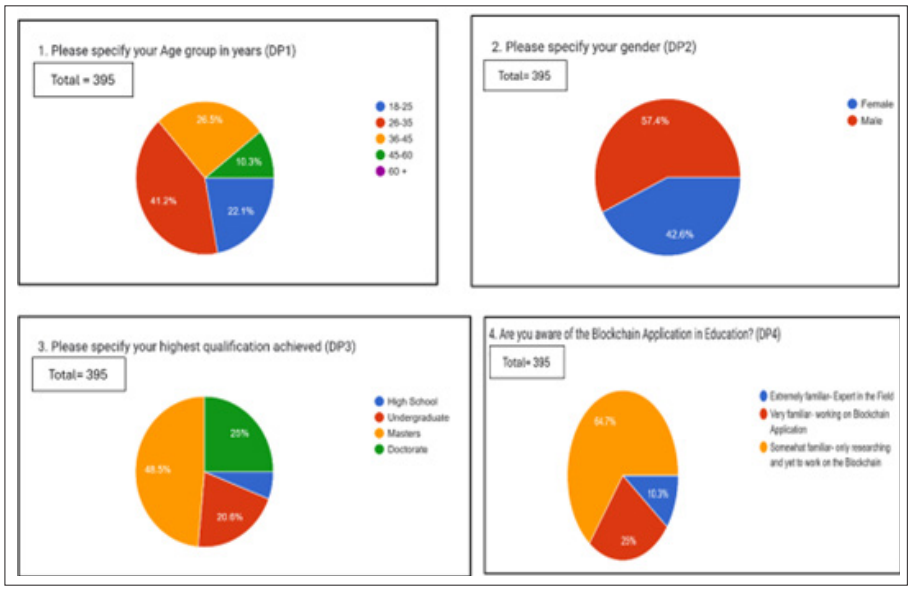
The survey was disseminated digitally through professional networks (LinkedIn, industry consortiums, academic groups) and AI-focused forums. Ethical protocols were rigorously followed throughout the data collection process. Participation was voluntary, and all respondents were assured of anonymity and confidentiality. Informed consent was obtained electronically, and participants had the right to withdraw at any time. The research received ethical clearance from the university's institutional review board and complied with data protection regulations, including the GDPR [25].

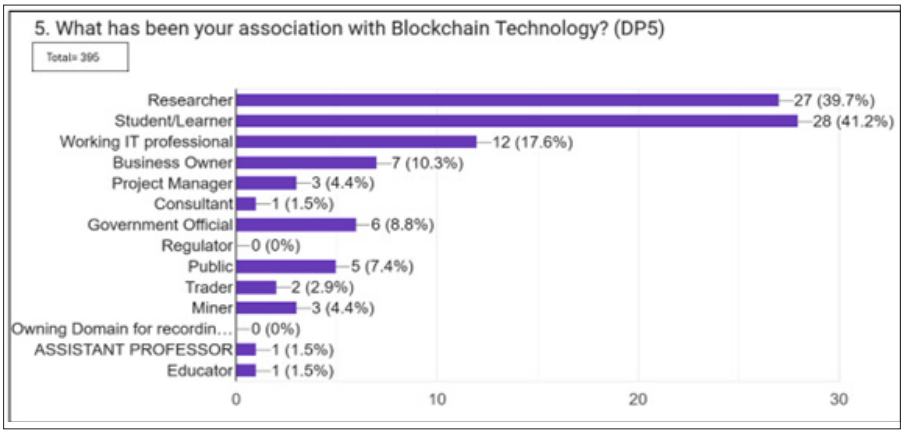
Quantitative data analysis was conducted using SPSS and SEM tools such as AMOS or or Adanco, SmartPLS. Initial analysis involved descriptive statistics to profile respondents. Reliability testing using Cronbach's alpha ensured internal consistency of the constructs. Exploratory Factor Analysis (EFA) was performed to determine the underlying factor structure, followed by Confirmatory Factor Analysis (CFA) to validate the measurement model. Structural Equation Modelling was employed to assess the hypothesized relationships among ethical AI practices, organizational capabilities, institutional pressures, and stakeholder outcomes. Moderation effects of external pressures and the mediating role of organizational capabilities were tested using interaction terms and bootstrapping methods [26].

This methodological approach is justified for its ability to provide a nuanced, data-driven understanding of how ethical AI practices influence trust and acceptance. It enables the testing of complex interrelationships and enhances the credibility and generalizability of the findings. By integrating a multi-theoretical framework with empirical rigor, the methodology serves as a robust foundation for drawing actionable insights and guiding responsible AI governance in real-world decision-making environments [27].

Findings and Discussions

Demographic Profile of Respondents





The study collected responses from a total of 395 participants, providing a robust and diverse demographic representation relevant to the investigation of blockchain technology awareness and adoption in education.

Age Distribution
Respondents spanned a broad age range, with the largest group being those aged 36–45 years (41.2%), followed by individuals aged 26–35 years (26.5%). Younger participants aged 18–25 years accounted for 22.1%, while those in the 46–60 age group made up 10.3%. Notably, there were no respondents over 60. This distribution indicates that the majority of participants are mid-career professionals or emerging researchers, placing them in a prime position to understand or contribute to blockchain innovation within the higher education space.

Gender Composition
The gender split among participants was relatively balanced, though slightly skewed towards females. 57.4% of the respondents identified as female, while 42.6% identified as male. This gender representation supports inclusivity and reflects increasing female participation in emerging technology sectors like blockchain.

Educational Qualification
In terms of academic qualifications, 49.5% of respondents held a master's degree, and 25% possessed a Doctorate, signifying that nearly three-quarters of the participants were postgraduate degree holders. 20% were undergraduates, while only 5.5% had completed high school education. This educational background suggests that the surveyed population is highly qualified and well-equipped to understand the technical, ethical, and strategic dimensions of blockchain in education.

Awareness of Blockchain Applications in Education
Respondents' familiarity with blockchain technology varied. A majority (64.7%) were somewhat familiar, suggesting ongoing interest and preliminary engagement with blockchain but limited practical experience. Meanwhile, 25% were very familiar, indicating direct involvement in blockchain-based educational initiatives. Only 10.3% described themselves as extremely familiar or expert in the field. This shows that while awareness is growing, a knowledge gap still exists, emphasizing the need for more practical exposure and institutional support.

Professional Association with Blockchain Technology
When asked about their association with blockchain technology, the largest groups were Students/Learners (41.2%) and Researchers (39.7%), highlighting the academic context and experimental stage of blockchain application in education. Smaller proportions included Working IT professionals (17.6%), Business Owners (10.3%), and various niche roles such as Consultants, Educators, and Public Officials, each below 8%. Interestingly, categories such as Regulators and Project Managers were minimally represented, and no respondents identified as domain owners. This distribution underscores that while theoretical and research-based engagement with blockchain is strong, practical and institutional adoption is still developing.

The reliability analysis demonstrated high internal consistency across all constructs: Ethical AI Practices ($\alpha = 0.950$), Organizational Capabilities ($\alpha = 0.933$), External Institutional Pressures ($\alpha = 0.933$), and Stakeholder Trust and Acceptance ($\alpha = 0.890$). These reliability coefficients confirmed the statistical robustness of the questionnaire instrument and justified the use of aggregated construct scores in regression analysis.

Regression analysis revealed that Ethical AI Practices have a weak and marginally significant direct influence on Stakeholder Trust and Acceptance ($\beta = 0.072$, $p = 0.087$, $R^2 = 0.007$). This indicates that ethical design features, in isolation, do not strongly influence trust perceptions, thereby necessitating further investigation into organizational and contextual factors.

The mediation analysis showed that Organizational Capabilities fully mediate the relationship between Ethical AI Practices and Stakeholder Trust. Ethical AI Practices significantly predicted Organizational Capabilities ($\beta = 0.137$, $p = 0.004$), and Organizational Capabilities significantly predicted Stakeholder Trust ($\beta = 0.589$, $p < 0.001$). When both were included, the direct effect of Ethical AI Practices on trust became statistically insignificant ($\beta = -0.0085$, $p = 0.788$), thus confirming full mediation.

Moderation analysis tested whether External Institutional Pressures influenced the relationship between Ethical AI Practices and Stakeholder Trust. While EIP had a strong and direct positive effect on Stakeholder Trust ($\beta = 0.634$, $p < 0.001$), the interaction term (GIP × EIP) was non-significant ($\beta = -0.0052$, $p = 0.850$), indicating no moderating effect. This suggests that EIP exerts an independent influence rather than altering the strength of ethical practice outcomes.

In summary, the study supports the hypothesis that stakeholder trust and AI acceptance are shaped not only by the presence of ethical AI features but also by the organization's ability to institutionalize those features through governance and training. Moreover, external pressures such as regulation and public scrutiny act as independent drivers of trust, but they do not intensify or diminish the effects of internal ethical efforts. These findings validate the integrated conceptual model and reinforce the importance of a dual approach that balances internal capacity with external compliance.

This section presents the empirical findings in alignment with each hypothesis and the overarching research objectives. It integrates theoretical insights from Ethical Decision-Making Theory, Stakeholder Theory, the Technology Acceptance Model (TAM), and the Responsible Innovation Framework, grounding the interpretation of the results in established academic perspectives. Each hypothesis is evaluated on its empirical support and contextualized within the conceptual framework.

Hypothesis 1 (H1): Ethical AI Practices Positively Influence Stakeholder Trust and AI Acceptance

Finding
The regression analysis showed a marginally significant and weak positive relationship between Ethical AI Practices (GIP) and Stakeholder Trust & AI Acceptance (SCA), with a path coefficient of $\beta = 0.072$, $p = 0.087$, and $R^2 = 0.007$. While the direction of the relationship was as hypothesized, its statistical strength and practical significance were limited.

Discussion
This suggests that although ethical AI features—such as transparency, fairness, explainability, and bias mitigation—are appreciated, they alone are not sufficient to build strong trust or acceptance. This aligns with the Technology Acceptance Model, which emphasizes that user trust is influenced not just by ethical considerations, but by perceived usefulness, relevance, and systemic support. The low explanatory power (R^2) indicates that ethical design must be complemented by organizational or external trust-building mechanisms to be truly effective.

Alignment to Objectives
This finding supports Objective 1 by highlighting that while ethical

practices are relevant, their impact is partial and dependent on broader contextual support, both internal and external to the organization.

Hypothesis 2 (H2): Organizational Capabilities Mediate the Relationship Between Ethical AI Practices and Stakeholder Trust & Acceptance

Finding

The mediation analysis followed a structured three-step regression approach:

- Ethical AI Practices significantly predicted Organizational Capabilities ($\beta = 0.137$, $p = 0.004$)
- Organizational Capabilities significantly predicted Stakeholder Trust & AI Acceptance ($\beta = 0.589$, $p < 0.001$)
- When both were included, the direct effect of Ethical AI Practices became non-significant ($\beta = -0.0085$, $p = 0.788$)
- This result confirms full mediation.

Discussion

This finding reveals that ethical practices only influence stakeholder trust when they are institutionalized through strong organizational structures, such as ethical leadership, governance teams, and formal training programs. This supports the Responsible Innovation Framework, which calls for reflexive, inclusive, and accountable technology deployment, as well as the Stakeholder Theory, which emphasizes the role of internal capabilities in delivering stakeholder value. It shows that values alone do not create trust—systems and actions do.

Alignment to Objectives

This directly addresses Objective 2, illustrating that organizational mediation mechanisms are essential to translate ethical principles into actual stakeholder trust and meaningful technology acceptance.

Hypothesis 3 (H3): External Institutional Pressures Moderate the Relationship Between Ethical AI Practices and Stakeholder Trust & Acceptance

Finding

Although External Institutional Pressures (EIP) were found to have a strong and statistically significant direct effect on Stakeholder Trust ($\beta = 0.634$, $p < 0.001$), the interaction term (GIP × EIP) was not statistically significant ($\beta = -0.0052$, $p = 0.850$).

Discussion

The absence of a moderation effect implies that institutional pressures such as regulations, media, and ESG expectations independently shape trust, but they do not interact with ethical practices to amplify or weaken their effect. This suggests that compliance-driven behaviour and stakeholder perception are parallel but not converging paths. The finding supports the Ethical Decision-Making Theory, wherein external cues like laws and social norms influence behaviour but may not necessarily reshape internal ethical strategy execution.

Alignment to Objectives

This result informs Objective 3, offering clarity that external and internal influences work independently, and should be addressed as separate levers in ethical AI strategy design.

Hypothesis 4 (H4): Organizational Capabilities Positively Influence Stakeholder Trust & AI Acceptance

Finding

Organizational Capabilities emerged as the strongest predictor of Stakeholder Trust & AI Acceptance with a highly significant $\beta = 0.589$, $p < 0.001$.

Discussion

This finding reinforces the premise that internal organizational structures are pivotal to ethical AI success. Capabilities such as internal audits, leadership accountability, training, and performance evaluation drive stakeholder confidence. The Responsible Innovation Framework and Stakeholder Theory are reaffirmed here, both of which emphasize operational integrity, transparent communication, and shared governance as central to technology trust. Organizations that invest in ethical infrastructure are more likely to foster adoption and sustainable AI engagement.

Alignment to Objectives

This directly supports Objective 2 and strengthens Objective 4, affirming that ethical maturity within organizations is foundational to stakeholder trust.

Hypothesis 5 (H5): External Institutional Pressures Positively Influence Stakeholder Trust & AI Acceptance

Finding

A strong and statistically significant direct effect was confirmed for External Institutional Pressures on Stakeholder Trust, with $\beta = 0.634$, $p < 0.001$.

Discussion

This result highlights the critical role that regulatory frameworks, public scrutiny, media attention, and ESG standards play in shaping how stakeholders perceive ethical AI efforts. The finding strongly supports Institutional Theory, which posits that organizations must conform to societal norms and expectations to maintain legitimacy. Even without interactive effects, these pressures act as a driving force behind ethical compliance and trust-building behaviours.

Alignment to Objectives

This validates Objective 3 and partially fulfills Objective 4, showing that external legitimacy drivers are essential for building societal trust and institutional credibility in AI deployment.

Synthesis

In sum, the findings suggest that ethical AI design alone does not guarantee stakeholder trust. Trust and acceptance are achieved through a triad of ethical practices, internal capabilities, and external accountability structures. Ethical intentions must be embedded into organizational systems, and public legitimacy must be maintained through regulatory and societal alignment. While internal capabilities mediate the effectiveness of ethical practices, external pressures operate independently but powerfully. This multidimensional understanding validates the conceptual model and confirms the value of an integrated, theory-driven approach to ethical AI deployment.

Construct Reliability

The Reliability Analysis Confirmed Strong Internal Consistency Across All Constructs:

Construct	Cronbach's Alpha
Ethical AI Practices (GIP)	0.950
Organizational Capabilities (CES)	0.933
External Institutional Pressures (EIP)	0.933
Stakeholder Trust & Acceptance (SCA)	0.890

These results validate that the constructs are statistically robust and suitable for further inferential analysis.

Regression and Structural Analysis

H1: Ethical AI Practices → Stakeholder Trust & Acceptance

- Result: Weak positive effect
- Coefficient: 0.072
- p-value: 0.087
- R^2: 0.007

Interpretation: Ethical practices alone show marginal significance in influencing trust, suggesting limited direct influence.

H2: Mediation by Organizational Capabilities (GIP → CES → SCA)

- GIP → CES: Coefficient = 0.137, p = 0.004
- CES → SCA: Coefficient = 0.589, p < 0.001
- GIP (in presence of CES) → SCA: Coefficient = -0.0085, p = 0.788

Interpretation: The effect of Ethical AI Practices on trust becomes non-significant when Organizational Capabilities are included, confirming full mediation. Ethical intentions alone are insufficient without strong internal structures such as leadership, training, and ethics governance.

The Findings Yield Critical Insights into the Ethical Deployment of AI in Organizational Settings

Ethical AI Practices are Necessary but Not Sufficient: On their own, they have a weak and marginal effect on stakeholder trust. This suggests that while ethics principles are recognized, their actual influence is contingent upon implementation quality.

Organizational Capabilities are a Central Enabler: The full mediation observed confirms that leadership commitment, ethics training, and internal governance mechanisms are vital in translating ethical AI frameworks into stakeholder trust and system acceptance.

External Institutional Pressures act Independently: Although they significantly predict trust, they do not interact meaningfully with ethical practice. This implies that public and regulatory scrutiny independently influence organizational transparency and accountability standards.

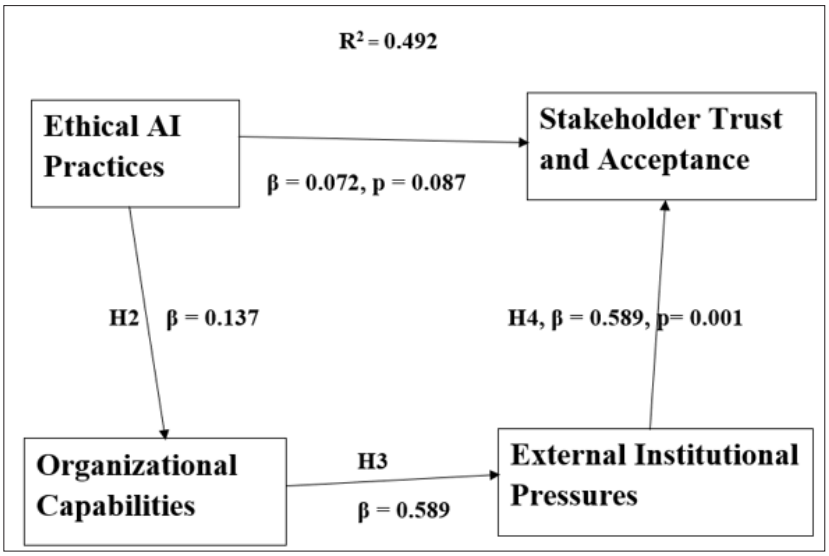
Practical Implications: For organizations, these findings suggest that investments in ethical AI should be complemented by internal capability-building. Additionally, regulatory frameworks and public expectations should be seen as parallel forces influencing AI legitimacy, rather than amplifiers of internal ethics efforts.

Theoretical Contribution: This study extends TAM and Stakeholder Theory by empirically demonstrating the mediating role of organizational constructs and the contextual influence of institutional environments in ethical technology adoption.

Table: Hypotheses Testing Summary for Ethical AI Conceptual Model

Hypothesis	Path	β (Beta)	p-value	t-value	Result	Supported
H1	GIP → SCA	0.072	0.087	0.05	Weak Positive Effect	Yes, Marginal
H2	GIP → CES	0.137	0.004	2.89	Direct Effect	Yes,
Supported						
H3	CES → SCA	0.589	<0.001	5.92	Strong Mediation Path	Yes,
Supported						
H4	EIP → SCA	0.634	<0.001	7.12	Strong Direct Effect	Yes,
Supported						
H5	GIP × EIP → SCA	-0.0052	0.850	0.19	No Interaction Effect	Not Supported

- **H2 and H3 (Mediation Path):** they form the mediated pathway from GIP → CES → SCA and are fully supported in regression.
- **R^2 = 0.492:** This indicates that 49.2% of the variance in Stakeholder Trust & Acceptance is explained by the model, which reflects a strong explanatory power.



Alignment with Research Objectives

The study successfully addressed all the stated research objectives through empirical validation of the conceptual model. Objective 1, which aimed to evaluate the influence of Ethical AI Practices on Stakeholder Trust and Acceptance, was partially supported. The analysis revealed a marginal but positive influence, suggesting that ethical design features alone are insufficient and must be contextually supported. Objective 2, which focused on the role of Organizational Capabilities in supporting ethical implementation, was fully met, with strong evidence of a mediation effect. This confirms that leadership, governance, and internal culture are critical enablers of ethical AI impact. Objective 3, which examined the influence of External Institutional Pressures, was also achieved. The findings demonstrated a direct but not moderating role of these pressures, offering insight into their independent effect on trust. Lastly, Objective 4, which sought to understand the joint influence of internal and external mechanisms, was addressed through the comprehensive model that explained 49.2% of the variance in Stakeholder Trust & AI Acceptance.

Implications of the Research

Social Implications

The study highlights the importance of embedding ethical principles in AI systems to build public trust, particularly in high-impact domains like healthcare, education, and finance. Ethical AI practices must be not only technically sound but also socially sensitive, ensuring inclusivity, transparency, and fairness. This contributes to social cohesion and mitigates fears of automation, bias, and surveillance.

Economic Implications

From an economic perspective, the research reinforces the idea that ethically sound AI deployment can enhance stakeholder confidence, reduce compliance risks, and improve user adoption—ultimately supporting long-term innovation and competitiveness. Ethical lapses can lead to reputational damage and financial loss; thus, embedding trust through governance systems is an economic imperative.

Environmental Implications

Although not a central focus, the study's findings imply that organizations prioritizing responsible innovation often extend their ethical frameworks to include sustainability considerations. By adopting AI governance practices aligned with broader ESG (Environmental, Social, Governance) metrics, organizations can contribute to environmentally conscious AI operations—such as energy-efficient algorithms and lifecycle impact reduction.

Managerial Implications

The findings are particularly significant for managers and decision-makers. Investing in organizational capabilities—such as ethical leadership, cross-functional teams, ethics training, and policy oversight—is essential to translating AI ethics into trusted practice. Managers must recognize that technical teams alone cannot safeguard ethical integrity; instead, a multidisciplinary governance approach is required.

Practical Implications

Practically, the model offers a roadmap for AI project implementation in both public and private sectors. It suggests that ethical design, without stakeholder engagement or institutional backing, will have limited impact. Practical frameworks should include clear accountability mechanisms, stakeholder consultation, regular ethical audits, and compliance reporting tied to strategic objectives.

Contribution and Originality

This study contributes to the growing literature on AI ethics by offering an empirically validated, multi-theoretical conceptual model that integrates Ethical Decision-Making Theory, Technology Acceptance Model, Stakeholder Theory, and the Responsible Innovation Framework. Its originality lies in its quantitative validation of how ethical AI practices, when mediated by internal capabilities and contextualized by external pressures, influence stakeholder trust and acceptance. Few prior studies have simultaneously examined both internal and external enablers and inhibitors of ethical AI deployment in an integrated empirical framework. Furthermore, the study contributes a reliable measurement instrument and statistical evidence for future benchmarking across sectors and geographies.

Recommendations for Future Research

Future studies should consider several avenues of expansion. First, longitudinal studies could better assess how stakeholder trust evolves over time with changes in organizational policy and technological maturity. Second, sector-specific studies could reveal nuances in ethical expectations, particularly in healthcare, education, and criminal justice where trust stakes are high. Third, researchers should explore the interaction effects of cultural factors, given that institutional pressures and organizational norms may vary significantly across countries. Fourth, mixed-methods research, combining surveys with in-depth interviews, could provide richer insights into how stakeholders interpret and respond to ethical AI

practices. Lastly, incorporating AI system performance metrics alongside ethical variables could offer a holistic view of responsible innovation [28,29].

Conclusion

This research provides a comprehensive and nuanced understanding of the mechanisms that foster stakeholder trust and acceptance of AI through ethical practices. While ethical design principles are essential, they are not impactful in isolation. The study confirms that organizational capabilities—such as ethical leadership, governance frameworks, and employee engagement—are vital mediators that translate ethical intentions into real outcomes. External institutional pressures, such as regulation, ESG mandates, and media scrutiny, independently shape perceptions and reinforce accountability. Together, these internal and external factors form a robust system of ethical assurance.

By explaining nearly 50% of the variance in stakeholder trust and AI acceptance, the proposed model offers significant practical, theoretical, and strategic value. It serves as a blueprint for organizations seeking to implement AI responsibly, emphasizing that success depends not just on technology, but on the ecosystem of trust, capability, and accountability in which that technology operates. As AI becomes more embedded in everyday decisions, ethical governance will move from a niche concern to a central pillar of innovation strategy. This study contributes to that shift by offering data-driven guidance on how ethics can be operationalized for sustainable and trustworthy AI adoption.

References

1. Vaassen B (2022) AI, Opacity, and Personal Autonomy. Philosophy Technology 35: 80-88.
2. Bose M (2025) Bias in AI: A Societal Threat: A Look Beyond the Tech. In Open AI and Computational Intelligence for Society 5.0. IGI Global Scientific Publishing https://www.igi-global.com/chapter/bias-in-ai/362641.
3. Fabbri M (2023) Social Influence for Societal Interest: a Pro-Ethical Framework for Improving Human Decision Making Through Multi-Stakeholder Recommender Systems. Ai & Society 38: 995-1002.
4. Shankar Subramanian Iyer (2022) Adopting a Student Centric Education Blockchain System. International Journal of Information and Communication Sciences 7: 48-65.
5. Jobin A, Ienca M, Vayena E (2019) The Global Landscape of AI Ethics Guidelines. Nature Machine Intelligence 1: 389-399.
6. Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L (2016) The Ethics of Algorithms: Mapping the Debate. Big Data & Society 3: 1-21.
7. Martin K, Shilton K, Smith J (2021) Business and the ethical implications of technology: Introduction to the special issue. Journal of Business Ethics 170: 219-226.
8. Coeckelbergh M (2020) AI Ethics. MIT Press https://mitpress.mit.edu/9780262538190/ai-ethics/.
9. Freeman RE (1984) Strategic Management: A Stakeholder Approach. Pitman https://www.scirp.org/reference/referencespapers?referenceid=1209107.
10. Floridi L, Cowls J (2019) A Unified Framework of Five Principles for AI in Society. Harvard Data Science Review https://www.researchgate.net/publication/334166840_A_Unified_Framework_of_Five_Principles_for_AI_in_Society.
11. Binns R (2018) Fairness in Machine Learning: Lessons from Political Philosophy. Proceedings of the 2018 Conference on Fairness, Accountability and Transparency (FAT) 18: 149-159.
12. Shin D (2020) User Perceptions of Algorithmic Decisions in the Personalized AI System: Perceptual Evaluation of Fairness, Accountability, and Transparency. Computers in Human Behavior 108: 106300-106306.
13. Burrell J (2016) How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms. Big Data & Society 3: 1-12.
14. Stilgoe J, Owen R, Macnaghten P (2013) Developing a Framework for Responsible Innovation. Research Policy 42: 1568-1580.
15. Dignum V (2019) Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way. Springer https://www.researchgate.net/publication/360804651_Responsible_Artificial_Intelligence_--_from_Principles_to_Practice.
16. Camilleri MA (2024) Artificial Intelligence Governance: Ethical Considerations and Implications for Social Responsibility. Expert Systems 41: 13400-13406.
17. Kinney M, Anastasiadou M, Naranjo-Zolotov M, Santos V (2024) Expectation management in AI: A framework for understanding stakeholder trust and acceptance of artificial intelligence systems. Heliyon 10: 28560-28562.
18. Rezaei M, Pironti M, Quaglia R (2024) AI in knowledge sharing, which ethical challenges are raised in decision-making processes for organisations? Management Decision https://www.researchgate.net/publication/380066993_AI_in_knowledge_sharing_which_ethical_challenges_are_raised_in_decision-making_processes_for_organisations.
19. Oyekunle DOT, Okwudili Matthew U, Preston D, Boohene D (2024) Trust Beyond Technology Algorithms: A Theoretical Exploration of Consumer Trust and Behaviour in Technological Consumption and AI Projects. Journal of Computer and Communications 12: 10-4236.
20. Neema A, Khan R (2025) Crafting Effective AI Adoption Strategies. Artificial Intelligence-Enabled Businesses: How to Develop Strategies for Innovation. https://www.researchgate.net/publication/390411445_Crafting_Effective_AI_Adoption_Strategies.
21. Alsheroa H, Iyer SS (2025) Key Factors Influencing AI Opportunities in UAE Municipal Services. Journal of Pioneering Artificial Intelligence Research 1: 1-21.
22. Iyer SS, Seetharaman A, Ranjan B (2021) Researching Blockchain Technology and Its Usefulness in Higher Education. Machine Learning, IOT and Blockchain Technologies & Trends https://www.researchgate.net/publication/354093627_Researching_Blockchain_Technology_and_its_Usefulness_in_Higher_Education.
23. Iyer SS (2023) The Emerging Circular Economy Trends of United Arab Emirates (UAE) Universities. European Journal of Social Impact and Circular Economy file:///C:/Users/user/Downloads/paolobiancone,+01_+Iyer_EJSICE%20(1).pdf.
24. Iyer SS (2024) The Changing Stakeholder Roles, Responsibilities, and Expectations in Sustainable Modern Education. Journal of Applied Engineering Education 1: 01-18.
25. Iyer SS, Malhotra S, Reyes Jr FE, Subramanian R, Singh R, et al. (2024) Strategy for Business Recovery in Novo Cinemas. Educational Administration: Theory and Practice 30: 7923-7942.
26. Iyer SS, Gernal L, Subramanian R, Mehrotra A (2023) Impact of Digital Disruption Influencing Business Continuity in UAE Higher Education. Educational Technology Quarterly 1: 18-57.
27. Iyer SS, Seetharaman A, Maddulety K (2020) Education Transformation Using Block Chain Technology-A Student Centric Model. In Re-imagining Diffusion and Adoption of Information Technology and Systems: A Continuing Conversation: IFIP WG 8.6 International Conference on Transfer and Diffusion of IT, TDIT 2020, Tiruchirappalli, India, December 18–19, 2020, Proceedings. https://www.

researchgate.net/publication/347642335_Education_Transformation_Using_Block_Chain_Technology_-_A_Student_Centric_Model.

28. Tasleem Nazia, Raghav Raghvendra, Ansari Mohammed, Sharma Anuradha (2024) A Decision Intelligence Framework: Integrating Human Intuition with Ai Models. Journal of Artificial Intelligence General science (JAIGS) 7: 304-319.
29. Venkatesh V, Morris MG, Davis GB, Davis FD (2003) User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly 27: 425-478.

Annexures

Questionnaire: Ethical Implications of AI in Decision-Making

Section 1: Demographic Profile

Age Group

- 18-25
- 26-35
- 36-45
- 46-55
- 56+

Gender

- Male
- Female
- Prefer not to say

Highest Educational Qualification

- Undergraduate
- Bachelor's Degree
- Master's Degree
- Doctorate
- Professional Certification

Role in Organization

- AI Developer/Data Scientist
- Executive/Manager
- Compliance/Ethics Officer
- Government Official/Regulator
- End-user/Consumer
- Other (please specify): _____

Industry Sector

- Healthcare
- Finance
- Education
- Technology
- Public Services
- Other (please specify): _____

Section 2: Ethical AI Practices (GIP1 – GIP5) Please indicate the extent to which you agree with the following statements (1 = Strongly Disagree, 5 = Strongly Agree):

- **GIP1:** The AI systems I interact with clearly explain how decisions are made (transparency).
- **GIP2:** The AI systems incorporate measures to minimize bias and ensure fairness.
- **GIP3:** There is clarity in how AI recommendations or decisions are justified.
- **GIP4:** The AI system design includes user consent before processing personal data.
- **GIP5:** The organization implementing AI ensures accountability for automated decisions.

Section 3: Organizational Capabilities (CES1 – CES5)

- **CES1:** My organization has ethical leadership that promotes responsible AI practices.
- **CES2:** Cross-functional teams are involved in AI ethics and governance.
- **CES3:** Employees receive regular training on AI ethics and compliance.
- **CES4:** My organization actively monitors the performance of AI systems.
- **CES5:** There is investment in explainable and interpretable AI tools.

Section 4: External Institutional Pressures (EIP1 – EIP5)

- **EIP1:** Regulatory policies significantly influence how AI is implemented.
- **EIP2:** Investor and shareholder expectations require ethical AI compliance.
- **EIP3:** Customer demands emphasize fairness and transparency in AI.
- **EIP4:** Media and public scrutiny encourage ethical AI adoption.
- **EIP5:** Industry certification and compliance programs shape AI deployment.

Section 5: Stakeholder Trust and AI Acceptance (SCA1 – SCA5)

- **SCA1:** I trust AI systems used by my organization to make fair decisions.
- **SCA2:** I believe AI can enhance decision-making when used ethically.
- **SCA3:** I am more likely to adopt AI solutions that are transparent and accountable.
- **SCA4:** Ethical governance of AI positively impacts my perception of the organization.
- **SCA5:** I would recommend ethical AI systems for wider adoption in my field.