

DOI: https://doi.org/10.48009/2_iis_2026_137

Governing ethical AI in organizations: Stakeholder perspectives on barriers and enablers

Chandni Bansal, *O.P. Jindal Global University*, cbansal@jgu.edu.in

Krishan Kumar Pandey, *O.P. Jindal Global University*, kkpandey@jgu.edu.in

Rajni Goel, *Howard University*, rgoel@howard.edu

Anuj Sharma, *O.P. Jindal Global University*, anuj.sharma@jgu.edu.in

Abstract

In today's world, the need for ethical AI is at an all-time high due to the fact that AI systems are becoming more autonomous and less transparent, as well as the number of people impacted by their decision-making processes growing exponentially with those that are the most vulnerable. This qualitative study examined the views of stakeholders regarding the barriers and facilitators to implementing ethical artificial intelligence in technology organizations. Using thematic analysis to analyze 39 semi-structured interviews conducted with AI developers, organizational leaders, ethicists, and community members impacted by AI across 12 organizations, five main themes were identified: (1) The Principle-Practice Gap, where although organizations articulate their ethical principles, they are unable to implement these due to resource and operational demands; (2) Structural and Cultural Barriers to Ethical Implementation of AI, including the absence of governing bodies, the absence of organizational leadership support, and the presence of conflicting organizational priorities; (3) Algorithmic bias and fairness as persistent implementation challenges; (4) Privacy, transparency, and governance complexities; and (5) Pathways to ethical implementation, which included the utilization of governance frameworks, the diversity of the team, and the involvement of stakeholders to increase the likelihood of ethical implementation of AI. Findings suggest that the successful implementation of ethical AI will depend upon the organization-wide transformation of culture and values beyond just implementing technical solutions for ethical AI, which further will be dependent upon the authentic commitment of leadership and the inclusion of ethics throughout every aspect of the development life cycle. These findings contribute significantly to the body of knowledge in AI governance, organizational best practices, and policy development for promoting responsible AI.

Keywords: Ethical AI, AI governance, stakeholder perspectives, responsible innovation, thematic analysis

Introduction

Despite the many advances in the development and application of Artificial Intelligence (AI), there are many challenges facing the development of AI in ways that respect the rights and dignity of humans (Bansal et al., 2024). The development and application of AI are creating new and unprecedented challenges to the ethical development and deployment of AI technologies. In response to these challenges, many well-established and widely recognized ethical frameworks and guidelines have been developed to guide the development of responsible AI. However, despite the widespread recognition of the need for ethical AI, there remains a significant disconnect between the stated ethical principles of organizations and the actual

implementation practices of organizations (Ali et al., 2023). Despite the awareness of the importance of ethical AI, organizations are struggling to convert general ethical principles into specific actions, and organizations are under pressure to meet other organizational goals and objectives, are constrained by limited resources, and are uncertain about the feasibility of implementing ethically-based practices (Bahangulu & Owusu-Berko, 2025). Understanding how stakeholders in organizations view the challenges of developing and implementing ethically-based AI, the barriers that they encounter, and the ways in which they may facilitate the development of ethically-based AI remain underdeveloped.

Research studies have been conducted to identify theoretical frameworks and policy-related recommendations to guide the development of ethically-based AI (Floridi & Cows, 2022). However, the research base identifying how organizations are implementing ethically-based AI, the barriers encountered by organizations when attempting to implement ethically-based AI, and the organizational characteristics that either support or hinder the ethically-based implementation of AI are limited (Ali et al., 2023). The gap between theory-based research on ethical AI and practice-based research on implementation challenges represents a key blind-spot in the research base. This study fills the gap by using a qualitative approach to examine stakeholder views of ethical AI implementation, to identify the barriers to ethically-based AI implementation, to identify the organizational characteristics that support ethically-based AI implementation, and to understand strategies organisations use to handle innovation-ethical challenges related to AI implementation. For the same, this study investigates the following research questions:

RQ1. What barriers impede ethical AI implementation in an organisation?

RQ2. What organizational factors support the implementation of ethical AI?

RQ3. How do organisations navigate innovation-ethical challenges related to AI implementation?

Literature Review

Ethical AI Principles and Frameworks

There exists consensus around the key ethical principles that should guide AI development. An examination of the key ethical initiatives related to AI development reveals commonalities among fairness, transparency, accountability, privacy, and human-centeredness (Doniyor et al., 2025). The principles of beneficence and non-maleficence for creating systems that benefit society without causing harm, forms the foundation of all frameworks (Floridi et al., 2018). Additionally, the principles of autonomy and justice address issues of equity and respect for human agency (Laitinen & Sahlgren, 2021). Furthermore, explicability, encompassing both the ability to understand technical processes and the ability to hold decision-makers accountable has emerged as a distinct principle in AI-related contexts (Beaudouin et al., 2020). However, translating these principles into actionable advice that practitioners can follow has proven difficult (Akinrinola et al., 2024). General principles provide little help to practitioners when they are making concrete decisions related to selecting algorithms, prioritizing fairness metrics, or determining whether to include particular features, creating what scholars refer to as the "principle-practice gap" (Correa et al., 2025).

Organizational Challenges to Ethical AI Adoption

The effective implementation of ethical AI involves more than just developing new technologies. Organizations also need to address the wide range of structural, cultural, and governance-related challenges which currently prevent them from effectively adopting ethical AI. The overwhelming majority of researchers have found that organizational impediments to the effective adoption of ethical AI are generally much larger than any technical limitations (Bansal, 2025). An area of particular concern related to the adoption of ethical AI is the persistent disconnection between what organizations state about their ethical AI principles and their day-to-day operational practices (Ali et al., 2023; Jain & Menon, 2023), known as

the principle-practice gap. Although many organizations have made public commitments to ethical AI principles, in many cases these commitments are purely rhetorical and have not been formally incorporated into the organization's AI development and deployment processes.

Organizational factors include several reasons why ethics professionals find it difficult to institutionalize ethical AI. For example, organizations tend to favor rapid product development over slow, thorough development; they often evaluate the success of products based on short-term productivity metrics that do not allow for evaluation of long-term impacts; and, they frequently reorganize their teams which disrupts knowledge sharing and collaborative decision making among team members (Setiawati et al., 2025; Palumbo et al., 2024; Birhane et al., 2022; Dai & Wang, 2021). In addition, individuals who express concerns regarding the ethical impact of an organization's actions (especially those from marginalized communities) may be subject to additional personal and professional risks when they attempt to speak out against the organization's decision-making process (Murikah et al., 2024; McCradden et al., 2020). These findings demonstrate that for organizations to implement ethical AI successfully does not simply depend upon improving the technology used for AI development, but rather upon creating positive changes at all levels within an organization by shifting their leadership priorities, governance models, and overall organizational culture (Eyinade et al., 2025; Bansal et al., 2023).

Additionally, if organizations fail to incorporate systematic fairness considerations into their internal processes and decision-making procedures, then their AI systems will likely perpetuate any societal biases which exist today. Therefore, organizations must develop governance mechanisms that extend well beyond simple ethical principles to clearly define responsibility within the organization, standardize review processes, and establish clear accountability structures (Muceku & Muceku, 2025). Unfortunately, however, many organizations still do not have formal mechanisms in place for reporting on AI-assisted decisions in a way that is transparent or provides meaningful oversight of AI-based decisions by stakeholders (Musrifah & Hasanah, 2025). As reliance on AI continues to grow exponentially in terms of influencing decision-making at individual and social levels, explainability and transparency are becoming increasingly critical to building and maintaining public trust in AI applications and protecting individuals' rights regarding their personal data (Shrestha & Joshi, 2025). Nonetheless, despite growing recognition of importance of transparency and explainability in AI systems, developing explainable and transparent AI remains a complex problem due to both the technical difficulty of contemporary AI systems and organizational reluctance to alter entrenched development practices (Chaudhary, 2024).

Collectively, the body of literature suggests that adoption of ethical AI is primarily an organizational challenge rather than a technical one. Therefore, addressing broader issues such as governance, accountability, fairness, transparency and organizational culture requires organizations to transition from theoretical/declarative commitments to institutionalized governance structures, leadership support for responsible AI development, cross-functional collaboration and tangible accountability structures that can translate abstract ethical goals into actionable responsible AI practices.

Research Gap

Although prior research has examined various aspects of governance for ethical AI through conceptual frameworks, technical methods for mitigating bias in AI systems, regulatory developments and responsible AI principles (Muceku & Muceku, 2025; Chaudhary, 2024), very little research has examined how collective experiences among diverse stakeholders including developers of AI systems, organizational leaders, ethics professionals, policymakers and community representatives shape organizational realities of ethical AI adoption. Consequently, relatively little empirical evidence exists regarding how organizational structure and governance models affect how organizations implement ethical AI principles in their daily activities. To fill this gap in the current research base regarding the collective organizational experiences among diverse stakeholders related to ethical AI implementation practices, the current study employed a multi-stakeholder qualitative methodology to explore barriers and enablers to the adoption of ethical AI

practices in organizations. Additionally, the study examines organizational dynamics as it relates to transforming ethical AI principles into action-oriented responsible AI practices. Overall, by examining collective experiences among multiple stakeholder groups across diverse organizations, the study provides a comprehensive understanding of how organizations operationalize ethical AI principles in practice. In addition, the study identifies organizational pathways that enable organizations to successfully adopt responsible AI practices.

Methodology

Research Design

This qualitative study utilized thematic analysis to identify themes that reflected stakeholders' perceptions of ethical AI implementation. Thematic analysis, a widely applied and flexible qualitative method, allowed for the systematic identification and analysis of patterns in the narratives of the various stakeholders (Christou, 2022; Braun & Clarke, 2021), and thus developed themes that represented important concepts and experiences relative to the challenges and facilitators of implementing ethical AI. The method was suitable for this research question as it allowed for a very deep examination of stakeholder's perceptions of the ethical challenges and implementation factors related to AI in their organizational contexts without imposing a priori theoretical framework, allowing the themes to develop inductively from the data. Semi-structured interviews also provided rich narrative data to enable the researcher to examine the experiences, reasoning, and perspectives of the participants in detail relative to the ethical challenges of AI within their organizational contexts.

Participant Recruitment and Sampling

A purposeful sampling strategy was employed to recruit 39 participants from 12 organizations (8 technology companies, 2 government agencies, 2 research institutions) that varied in size, AI maturity, and ethical implementation sophistication. The tenure of the participants in their current role ranged from 1-10 years. Participants were recruited using purposive sampling to ensure representation from key stakeholder groups involved in ethical AI development and governance. Potential participants were identified through professional networks, organizational contacts, and publicly available information regarding AI initiatives. Individuals were eligible if they had direct experience in AI development, deployment, governance, ethics, or policy. Recruitment continued until thematic saturation was achieved. Each participant gave informed consent and was aware that participation was voluntary.

The organizational characteristics are presented in Table 1. The sample included a mix of different types of organizations. The two largest technology companies (n=2; 9 participants) represented organizations with the most developed AI capabilities and existing governance framework along with formalized ethics departments. The two medium-sized technology companies (n=2; 7 participants) represented organizations with formalized yet evolving approaches to ethical AI. The four start-up/scale-up companies (n=4; 10 participants) represented organizations that were at a nascent stage of development that were beginning to establish ethical practices and governance frameworks. The two government agencies (n=2; 7 participants) represented organizations from the public sector that addressed regulatory requirements and policy implications of AI. Lastly, the two research institutions (n=2; 6 participants) provided research-focused and academic perspectives. This variety of organizational types allowed for exploration of how differing organizational context and resources impact an organization's ability to develop its capacity for ethical AI implementation practices. As summarized in Table 1, the sample included participants with roles like AI practitioners, data scientists, machine learning engineers, organizational leaders, ethics professionals, policy advisors, and community representatives. This diversity enabled the study to capture multiple stakeholder perspectives across the AI ecosystem.

Table 1. Organizational Characteristics and Participant Distribution

Org ID	Organization	Type	AI Maturity	n	Participant Roles
01	Tech Corp A	Large Tech Company	Mature	5	AI practitioners; data scientists; organizational leaders
02	Tech Corp B	Large Tech Company	Mature	4	AI practitioners; ethics professionals
03	Tech Corp C	Medium Tech Company	Established	4	ML engineers; data scientists; organizational leaders
04	Tech Corp D	Medium Tech Company	Established	3	Data scientists; ethics professionals
05	Tech Corp E	Startup / Scale-up	Developing	3	AI practitioners; data scientists
06	Tech Corp F	Startup / Scale-up	Developing	3	ML engineers; data scientists
07	Tech Corp G	Startup / Scale-up	Emerging	2	Organizational leaders
08	Tech Corp H	Startup / Scale-up	Developing	2	AI practitioners; data scientists
09	Government Agency 1	Government	Developing	4	Leaders; ethics professionals; policy advisors
10	Government Agency 2	Government	Developing	3	Leaders; policy advisors
11	Research Institute 1	Research Institution	Research-focused	3	Ethics researchers; community representatives
12	Research Institute 2	Research Institution	Research-focused	3	Ethics researchers; community representatives

Data Collection

A semi-structured interview methodology was used to collect data. Semi-structured interviews allow for consistency in the way the interviews are conducted, but also allows for the ability to ask follow-up questions. Interview guides were created based upon the research objectives and prior knowledge gathered from the literature on ethical AI, AI governance, responsible innovation, and organizational adoption of ethical AI. Open-ended questions were asked during the interviews to encourage participants to discuss their perceived barriers and enablers to implementing ethical AI at their organization. Additionally, the question asked about stakeholders and how they interact with one another as it relates to the implementation of ethical AI. The broad areas of inquiry and examples of relevant interview questions are listed in Table 2.

Table 2. Illustrative Semi-Structured Interview Guide

Research Objective	Representative Interview Question
Understanding ethical AI adoption	How does your organization define and approach ethical AI adoption?
Identifying organizational barriers	What are the major challenges your organization faces in implementing ethical AI principles?
Exploring governance mechanisms	What governance structures or policies guide ethical AI development in your organization?
Understanding organizational culture	How does leadership influence ethical AI decision-making and implementation?

Research Objective	Representative Interview Question
Examining bias and fairness	What practices are used to identify and mitigate algorithmic bias and ensure fairness?
Stakeholder engagement	How are employees, customers, regulators, or affected communities involved in ethical AI decisions?
Future improvements	What changes would most effectively strengthen ethical AI adoption within organizations?

While the interview guide provided some standardization across the interviews, there were times when the researcher needed to ask additional probes to gain further insight into the participants' experiences. While this method of collecting data provided for rich contextual understanding, it ensured that all of the interviews would address the research objectives. The duration of each interview averaged 30-35 minutes. All of the participants granted permission for their interviews to be audio recorded, which facilitated transcription of each interview into written form for qualitative thematic analysis.

In addition to the semi-structured interviews, secondary data was also gathered through publicly available organizational documents (i.e. ethics guidelines, AI policy statements, governance structures) and access to previously published peer-reviewed articles containing case studies. Primary and secondary data from both sets of data sources were used together to enhance the reliability of the results by using the method of triangulation.

Data Analysis Procedure

The data was analysed using Braun & Clarke's six stages of thematic analysis (Byrne, 2022; Braun & Clarke, 2019). This approach was selected because the study sought to identify and interpret patterns of meaning across diverse stakeholder experiences rather than develop a formal explanatory theory. Consequently, the study does not adopt a Grounded Theory methodology but instead employs reflexive thematic analysis to generate rich, contextually grounded themes related to ethical AI adoption. In stage one, the data was analysed through a process called familiarity, wherein the researcher read and re-read the data transcripts and organizational documents, documenting first impressions regarding the data. Stage two of the analysis involved coding the data, identifying conceptual units that represented barriers to implementation, enablers of implementation, and stakeholder views. Stage three of the analysis involved aggregating the codes into themes that captured commonality in the data. Codes were systematically aggregated into themes using thematic maps to illustrate the hierarchical nature of the themes. Stage four of the analysis involved revising the themes through a detailed review of each theme to determine consistency within the theme, and to distinguish between themes. Stage five of the analysis involved naming and defining the themes by clearly articulating the range and content of each theme. Stage six of the analysis involved synthesizing the themes and the supporting data extracts, and integrating the results of this research with existing literature.

Interview transcripts were imported into the qualitative analysis workflow and analysed iteratively following Braun and Clarke's six-phase reflexive thematic analysis. The researchers first familiarized themselves with the data through repeated reading of interview transcripts and accompanying field notes to develop an in-depth understanding of participants' experiences. During the second phase, relevant sections of text were coded using an initial descriptive code that represented recurring ideas associated with implementing ethical AI. The majority of the code generation was inductive, thus permitting the researchers to build upon the data collected from the field, but still directed by the study's objective. After identifying the first round of codes, those codes were then evaluated, compared and categorized into higher-level

conceptual categories that represented broad similarities within those codes. The researcher engaged in an iterative cycle of evaluating and revising categories into potential themes, or areas of commonality among the various stakeholder groups. Potential themes were continuously reviewed against all data collected during the interviews to assess internal consistency, external distinctiveness and alignment to the research questions. Those themes were revised continually until there was a consistent thematic framework that accurately captured participants' perspectives regarding the implementation of ethical AI.

To enhance credibility of the results, the researcher cross-referenced emerging themes to both organizational documents and published case studies. No secondary data source had an independent analysis to generate additional themes; instead, the secondary data were analysed to confirm, contextually and add meaning to the interpretations drawn from interview data. As such, triangulation of evidence across multiple types of evidence validated consistency across multiple data sets and increased reliability and validity of the findings. Specifically, Table 3 summarizes how each type of data was used during the analytical process.

Table 3. Integration of Data Sources in the Thematic Analysis Process

Data Source	Purpose	Role in Analysis	Contribution to Findings
Semi-structured interviews	Primary empirical evidence	Initial coding and theme generation	Identification of barriers, enablers, and organizational dynamics
Organizational documents	Contextual validation	Verification of governance practices, policies, and organizational processes	Enhanced contextual understanding and credibility of themes
Published case studies	External triangulation	Comparison of emerging themes with documented organizational experiences	Strengthened interpretation and transferability of findings

The use of multiple data sources enabled triangulation across various forms of evidence and lessened reliance on a single form of evidence. The presence of consistencies between participant narratives from interviews, organizational documentation, and published case studies added credence to identified themes and significantly increased trustworthiness and robustness of findings.

Analysis occurred throughout the analytical process in an iterative manner to facilitate continued development of codes and themes. With each interview analyzed, new codes were evaluated against existing coding frameworks to determine if additional theoretical insights emerged. Approximately after conducting 35th interview, no significantly new codes or themes emerged from the data. Thus, thematic saturation had essentially been reached. Four additional interviews were conducted to validate the stability, consistency, and completion of the thematic framework across various stakeholder groups. In accordance with recommendations for reflexive thematic analysis, sample adequacy was thus determined by sufficiency and richness of data to answer research questions rather than pre-determined sample sizes (Braun & Clarke, 2021; Hennink et al., 2017).

NVivo 14 software supported systematic coding and identification of themes in completing the analysis. The analytical process followed a systematic sequence as depicted in Figure 1. This data flow diagram demonstrates how research questions guided data collection from diverse stakeholder groups, which was then systematically coded and organized into progressively refined themes, interpreting and integrating the results, to establishing the ultimate conclusion concerning the necessary organizational changes for implementing ethical AI.

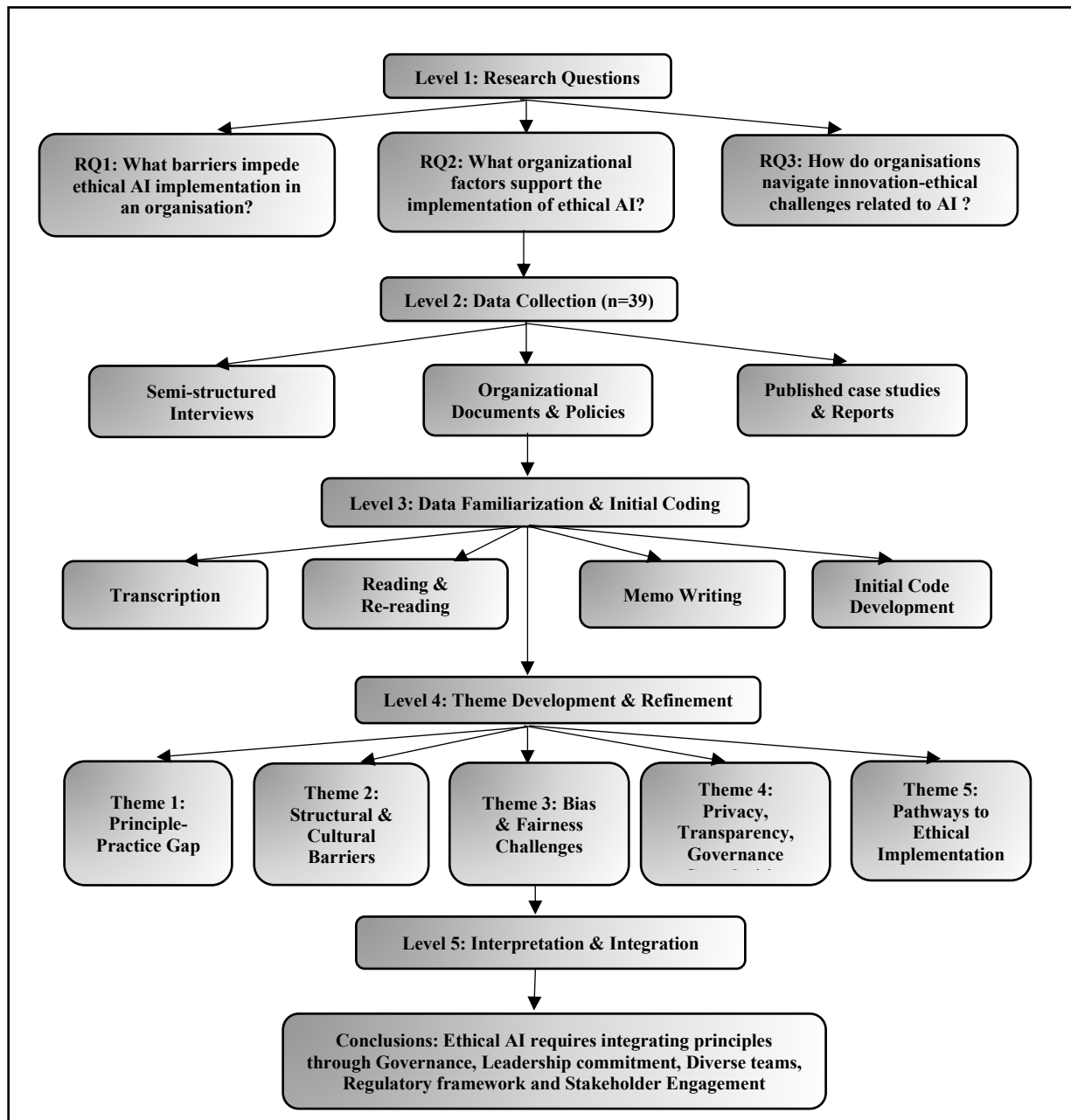


Figure 1. Data Flow Diagram

Trustworthiness and Rigor

The researcher employed a number of strategies to enhance the trustworthiness of the research process and to increase its methodological rigor. For example, the prolonged engagement provided the opportunity for extended data collection over eight months, which allowed the researcher to build a greater level of understanding and to develop trust with the participants (Dado et al., 2023). The researcher employed researcher triangulation in order to analyze the data collected from the interviews, where multiple researchers analysed the data collected during each interview (Jonsen & Jehn, 2009; Farmer et al., 2006).

Where there were any disagreements among the researchers, the researchers discussed the issue until they reached a consensus, which reduced the amount of individual bias present in the data. Along with this, organizational documents and published materials were also searched, which provided the researcher with multiple perspectives on similar phenomena. Additionally, the researcher employed member checking, which was used to validate the researcher's interpretation of the participant's responses (Thomas, 2017; Braun & Clarke, 2023). Twelve participants (approximately 27% of the sample) reviewed the summaries of the preliminary findings and provided the researcher with feedback and validated the researcher's interpretations. The researcher conducted negative case analysis to locate data that contradicted the emergent themes (Finlay, 2021). The researcher actively sought out contradictory data to ensure that an accurate and nuanced understanding of the phenomenon being studied. The researcher applied thick description to enable the reader to assess the applicability of the research findings to their own setting (Younas et al., 2023). The researcher maintained an audit trail to track the decision-making process of the researcher throughout the research process, the changes to the coding scheme over time, and the development of the themes. An audit trail provided transparency to the researcher's analytic process and allowed other researchers to assess the analytic rigor of the researcher independently.

Ethical Considerations

Prior to beginning the research process, the researcher secured ethical approval from the institutional review board. All participants signed an informed consent form that clearly stated the purpose of the research study, the ways in which the researcher intended to use the data collected, and the measures taken to protect the confidentiality of the data. Each participant was informed that he/she had the right to withdraw from the research study at any time without penalty or reprisal. The researcher took steps to protect the confidentiality of the data collected from the participants by deleting all identifying information from the data, and by using pseudonyms to refer to all participants in all reports of the research. The researcher also encrypted the audio recordings of the interviews and stored the recordings on a server that was password-protected and only accessible to the research team. The researcher deleted all identifying information from the transcripts and notes related to each participant.

Results

The thematic analysis generated five interrelated themes that collectively address the study's three research questions. Themes 1-4 primarily capture the organizational and technical barriers to ethical AI adoption (RQ1), while Themes 4 and 5 further explore the organizational mechanisms and pathways that facilitate responsible AI implementation (RQ2 and RQ3). Each theme is supported by representative quotations drawn from multiple stakeholder groups to illustrate the diversity of perspectives across the dataset. Table 4 provides a summary of the identified themes, the corresponding research question(s), the stakeholder groups from which the evidence emerged, and the nature of the supporting empirical evidence.

Table 4. Summary of Themes, Research Question Alignment, and Supporting Evidence

Theme	Research Question Addressed	Primary Stakeholder Groups	Illustrative Evidence
Theme 1: Principle-Practice Gap	RQ1	AI Practitioners, Leaders	Multiple quotations illustrating disconnect between ethical principles and implementation

Theme	Research Question Addressed	Primary Stakeholder Groups	Illustrative Evidence
Theme 2: Structural & Cultural Barriers	RQ1	Leaders, Ethics Professionals	Organizational constraints, resource limitations
Theme 3: Bias & Fairness Challenges	RQ1	AI Practitioners, Researchers	Bias mitigation challenges, fairness concerns
Theme 4: Privacy, Transparency & Governance	RQ1 & RQ2	Policy Advisors, Government Agencies	Governance mechanisms, accountability, explainability
Theme 5: Pathways to Ethical AI Adoption	RQ2 & RQ3	All stakeholder groups	Governance, leadership, diversity, stakeholder engagement

The themes presented below were developed primarily from the interview data and subsequently examined against organizational documents and published case studies. This triangulation process enabled the researchers to verify the consistency of the findings, enrich the organizational context, and identify areas where documentary evidence supported or expanded participants' perspectives. Across all five themes, the secondary sources largely corroborated the interview findings, thereby strengthening the credibility of the thematic analysis.

Theme One: The Principle-Practice Gap

A disconnect exists among all stakeholders regarding the stated ethical AI principles and how these are implemented in daily AI development and deployment. All stakeholders agree that ethical AI principles (fairness, transparency, accountability and privacy) are important. However, few of these are routinely practiced within organizations. An example of this is provided by an organizational leader: "Everyone agrees that ethical AI is very important; however, it is much easier to write about it than to take the steps necessary to convert those principles into daily decision-making processes on individual projects." Stakeholders cited several organizational barriers to implementing ethical AI, including lack of adequate training, lack of sufficient practical guidance, lack of organizational support, and competing business objectives. Also, many stakeholders stated that while organizations' policies include well-documented ethical AI guidelines; they lack the formal governance structures necessary to enforce the consistent application of these guidelines throughout the entire life cycle of an AI system. An example of this is provided by an AI developer who commented: "We have excellent ethical AI guidelines established within our organization. However, once the commercial pressure mounts and the urgency to release the product increases, the ethical concerns are put off until a later version. The reality is that rarely do we go back to correct or improve the prior version." Similar comments were made by numerous stakeholders indicating that many organizations' performance metrics focus primarily on rapid innovation, competitive advantage in the marketplace, and generating revenue rather than incorporating responsible AI development. One example of this is provided by an ethics professional who commented: "Generally speaking, ethics is seen as something that can be revisited at a later time. When commercial pressures increase, ethical concerns are generally delayed rather than resolved."

Theme Two: Structural and Cultural Barriers to Implementation

One of the executives interviewed about the disconnect between stated ethical standards and actual practices stated that he believes that one of the reasons that the disconnect occurs is that ethics is difficult to measure. He stated, "We can measure lots of things, such as accuracy, speed of deployment, and cost. However, we cannot easily measure whether an AI system is sufficiently ethical." As a result of being unable to measure ethics, the executives have difficulty implementing ethics through traditional business processes. As a result, although executives state that they are committed to ethics, they do not prioritize ethics when making decisions. As such, the rhetoric of ethics remained prominent in organizational discourse, while ethics itself

remained a low priority in the actual operations of the organizations. Ethics professionals working in organizations reported that they were marginalized in their organizations, and that their recommendations regarding ethics were typically treated as suggestions, rather than as mandatory requirements that would take precedence over business imperatives. An ethics professional said: "There is no clear ownership of ethical AI. Everyone agrees it's important, but nobody is ultimately responsible for ensuring it happens." Therefore, when ethics professionals identified an ethics-related issue in an organization, their recommendation to address that issue was frequently ignored or dismissed in favour of pursuing business objectives.

Secondary barriers were team restructuring and organizational volatility. Practitioners said that because of the constant turnover of employees due to frequent team reorganization, employees lost their relationship-building time and the ability to continue the work they were doing, especially in the areas of developing and promoting ethical practices. Ethics workers that were able to build trust and establish a network of colleagues that supported them in advocating for ethics were split up by reorganization and lost all of their institutional knowledge and momentum toward promoting ethical practices. The reoccurring nature of this process created an environment that prevented the organization from establishing a consistent set of ethical practices. Just as employees were starting to become proficient and gain some experience in ethics, reorganization would take place and move those same employees to new teams, and they would have to start over again.

The amount of commitment to ethics provided by senior leadership either contributed to or hindered ethical implementation. In organizations that have senior leadership that has expressed and demonstrated a strong commitment to ethics (i.e., having senior leadership that will provide leadership support for the development of ethics), significant progress toward establishing an ethics program has been achieved. Conversely, organizations that have a lack of commitment to ethics or do not treat ethics as a priority may be unable to commit sufficient resources to develop an ethics program, or they may have difficulty making trade-offs that involve ethical issues. This issue is supported by an AI Practitioner, who said: "Developers often work in isolation. Ethical issues are discussed much later, when changing the system becomes expensive and time-consuming."

Theme Three: Algorithmic Bias and Fairness as Persistent Implementation Challenges

Bias and fairness in machine learning algorithms were cited by every participant as the greatest implementation challenges. Many organizations have experienced high-profile instances of bias in machine learning algorithms, where biased outcomes were produced by the algorithm. Practitioners recognized that bias could stem from many factors (e.g., training data, algorithm design, lack of diversity among developers) and therefore require a multi-faceted approach to mitigate bias. However, many practitioners were disappointed to find out that implementing effective methods to mitigate bias was far more complicated than anticipated. A data scientist told me: "We know about fairness metrics, we know how to measure bias, but once we attempt to implement fairness constraints on our models, we see that the performance of our models degrade, and we are pressured to relax our fairness constraints... it becomes this negotiation about how much discrimination we are willing to accept to maintain our performance."

Practitioners reported that a practical barrier to fairness was the mathematical incompatibility of different fairness metrics. Practitioners responsible for fairness implementations described the confusion about which fairness metric to maximize, knowing that each metric (statistical parity, equal opportunity) involves trade-offs. As such, absent organizational guidance, individual practitioners made these important decisions informally. In the words of an AI Practitioner: "There is rarely a perfect definition of fairness. Improving fairness for one group can sometimes create unintended consequences for another." Additionally, participants reported that although technical solutions to fairness are needed, they are insufficient. An ethics professional stated: "We can develop more fair algorithms, but if the underlying data reflects historical

discrimination, or if the problem was poorly formulated, technical fairness patches will not correct the basic injustice." This realization—that fairness requires not only algorithmic interventions, but also a review of problem formulation and data origin—was rarely developed in most organizations.

Theme Four: Privacy, Transparency, and Governance Complexities

Privacy, transparency and data governance were identified as the largest implementation challenges. Practitioners consistently reported that organizations collect massive amounts of data to use for AI development purposes; however, most organizations have little to no mechanisms in place to ensure that individuals have meaningful control over how their data is being utilized. A privacy professional stated: "Our consent processes are technologically compliant, but practically meaningless. People are clicking through, without realizing what they are agreeing to...we are technically complying with regulations, but we are violating the spirit of privacy protection." Participants across organizations reported that privacy considerations are mandatory, but receive significantly less attention than bias mitigation, and frequently come into conflict with business objectives that demand full access to data. Transparency and Explainability have continued to present challenges to organizations (Karim et al., 2023; Chen et al., 2021). Organizations have been inconsistent in making trade-off decisions regarding explainability, and have often defaulted to prioritizing accuracy over explainability. Furthermore, participants have noted that providing technical explanations does not necessarily equate to real accountability. An ethics worker stated: "Even when we make our algorithms more transparent, if we don't have a system in place for people to complain and hold executives accountable for the decisions being made, transparency becomes performative." The idea that transparency in the absence of mechanisms for real accountability provide little to no ethical value was reflective of a larger understanding that successful integration of technical solutions and organizational practices that facilitate real accountability is necessary to create a functioning governance structure.

Two participants from Organizations 09 and 10 stated that ethical AI challenges are most prevalent within public-sector decision-making environments. Unlike private sector organizations, government agencies stated that algorithmic decisions impact citizen rights, public services and administrative decisions; thus, require substantially greater amounts of transparency and accountability. One participant stated, "In government every automated recommendation has to be defensible due to direct effects on citizens. Transparency is therefore required rather than optional in order to maintain public trust." Another participant stated that bias in public sector AI systems can negatively affect vulnerable populations; therefore, there is an ongoing need for continuous auditing and explanations mechanisms.

While participants identified several persistent barriers to ethical AI adoption, they also described organizational practices that helped overcome these challenges.

Theme Five: Pathways to Ethical Implementation

Unlike many of the previous themes related to barriers to the adoption of ethical AI, participants also reported several organizational mechanisms that facilitate successful implementation of ethical AI. Governance structures, leadership commitment, multidisciplinary collaboration and stakeholder engagement were identified as the primary enabling factors for implementing ethical AI across stakeholder groups. Many stakeholders believed that organizations with formalized governance mechanisms (including ethics committees, accountability frameworks and AI policies) can institutionalize ethical consideration into regular AI development. For example, an organizational leader indicated: "With a formal AI governance committee, ethical issues are addressed throughout a project rather than being dealt with in isolation at the end." Additionally, an ethics professional stated: "Policies are only as good as the supporting governance structures that hold teams accountable for making ethically informed decisions." Leadership commitment was viewed similarly as foundational to developing an organization culture that views responsible AI as important. Consistently, participants identified that when senior managers actively support ethical AI initiatives, project teams are more likely to invest sufficient time and resources to ensure

fairness, transparency and accountability in their work. An AI developer stated: “When leaders view ethics as a strategic initiative versus a compliance exercise, teams naturally integrate ethics into their daily work.” Participants emphasized the importance of multidisciplinary collaboration and identified that involving developers, lawyers, business professionals and ethicists allows organizations to recognize ethical risk from diverse viewpoints. One participant commented: “Different disciplines challenge one another’s assumptions which leads to more equitable decisions about AI.”

Stakeholder engagement was seen as critical to increasing transparency and building trust. Participants felt that including users, regulators and affected parties throughout the AI life cycle allow organizations to identify ethical concerns before they become major problems. A community representative indicated: “People who will be impacted by AI systems frequently identify issues that developers never consider. Therefore, it is necessary to involve them early in the process.” Overall, these findings demonstrate that effective implementation of ethical AI occurs through a combination of appropriate governance mechanisms, leadership commitment to ethics, collaborative and inclusive approaches to decision-making, and ongoing engagement with relevant stakeholders – thereby allowing organizations to translate ethics into practice in their use of AI.

Overall, the five themes together illustrate that the adoption of ethical AI goes far beyond technical implementation. Rather, it is shaped by a range of organizational structures (governance), leadership commitment and relationships with stakeholders. In addition, these findings are discussed in terms of prior literature.

Discussion and Implications

The findings demonstrate that although there is widespread agreement on ethical AI principles, the translation of these principles into practice is a significant barrier for many organizations. The gap between the “ideal” and the “real” is caused by organizational structural (i.e., lack of incentives for ethics), procedural (i.e., lack of mechanisms to measure/monitor/enforce ethics), and behavioral (i.e., lack of leadership commitment) constraints. These findings are consistent with prior research indicating that organizational constraints are the major impediment to ethical AI implementation, and not technological constraints.

Both structural and cultural barriers (e.g., poor governance, unstable teams, and ambivalence toward leadership) demonstrate that developing ethical AI will require more than just technology-based solutions. Instead, developing ethical AI will require broader, systemic changes to an organization’s culture, decision-making processes, and governance mechanisms. Such changes will only occur if organizational leaders commit organizational resources to ethics and are willing to make trade-offs between ethics and short-term optimization of performance.

Given the fact that despite decades of research and advancements in available tools, many organizations continue to experience algorithmic bias and fairness issues indicates that technical solutions alone will not solve the problem of bias. The reason for this is that the cause of bias is due to: (1) biased data, (2) the way in which problems are formulated, and (3) the lack of diversity in developer teams. Organizations wishing to avoid bias must utilize a variety of methods including improving data quality, reformulating the problem in different ways, utilizing diverse development teams, and establishing governance mechanisms to monitor and enforce fairer and less biased systems (Garcia et al., 2024). Additionally, fairness-based technical solutions must also be accompanied by accountability mechanisms and stakeholder engagement to ensure that solutions focus on metrics that reflect human impacts (Huriye, 2023).

Findings associated with privacy, transparency, and governance have shown that all three of these areas are currently poorly developed in most organizations (Denton et al., 2021; Zhang, 2025). Many organizations have come to realize the importance of privacy and transparency, but most of them prioritize functional performance or business goals over these two critical areas. Transparency without accountability (i.e., transparency as to how algorithms operate without mechanisms allowing people to contest unfair decisions made by the use of those algorithms), produces an illusion of accountability and provides no real protections to individuals (Moats & Trifuljesko, 2025). Real accountability requires that governance mechanisms combine both technical transparency (i.e., explain how algorithms operate) and organizational accountability (i.e., allow individuals to understand the basis for decisions, allow individuals to request explanations, and allow individuals to pursue remedies when they are harmed).

The findings from Organizations 09 and 10 illustrate that Ethical AI governance has more relevance in the public sector than in private sectors due to the fact that algorithmic decision-making impacts citizen's rights, access to public services and regulatory actions. Additionally, these observations correlate with current international efforts to create transparency, accountability, explainability, and human oversight for risky AI applications. The concerns raised by our participants mirror the OECD AI Principles (OECD, 2024); UNESCO's Recommendation on the Ethics of Artificial Intelligence (UNESCO, 2021); the NIST AI Risk Management Framework (NIST, 2023); and the risk-based governance mechanism utilized by the EU AI Act (European Union, 2024) that promote strong governance mechanisms for high impact AI systems.

As such, the triangulation of participant responses with organizational documentation and published case studies strengthened the validity of the results through a demonstration of strong agreement among multiple sources of evidence. Although participant responses provided a detailed description of individual experiences; the organizational documentation supported the understanding of the organization's practices and reinforced the fact that many of the identified issues, including governance gaps, lack of transparency and lack of organizational leadership commitment to ethical AI use, have similarly been documented in organizational policies and published examples of organizational implementations.

Research Contributions

This research makes four unique contributions to the literature on the adoption of ethical AI. First, this study empirically examines organizational barriers/enabling factors to adopt ethically developed AI from multiple stakeholder groups. Second, this study offers theoretical support for expanding governance models for ethical AI. Ethical AI cannot be adopted solely through technical measures. Instead, its success relies heavily on organizational structures and cultures that promote and protect ethical consideration during AI development. In doing so this study adds a socio-technical approach to prior discussions of responsible AI implementation by emphasizing the organizational aspects. The findings presented here offer action-based guidance for organizations interested in implementing ethical AI. The barriers/enabling mechanisms identified in this research provide organizations with tangible options for enhancing their AI governance frameworks, promoting a corporate culture that emphasizes ethics in AI development, increasing transparency and encouraging stakeholder participation throughout the full lifecycle of an AI system. Fourth, this study contributes to policy debates by illustrating the need for institutional support, regulatory clarity, standardized governance mechanisms for responsible AI deployments. Insights gained from this study could help policymakers and regulatory agencies design contextualized governance frameworks that promote trustworthy and accountable AI adoptions.

Limitations

This research has several important limitations. First, while the sample was diverse in terms of the roles of the stakeholders and types of organizations involved, the sample was limited to organizations based in developed countries, which limits the potential for generalizing the findings to organizations in the Global South, where institutions may vary significantly, and regulatory environments may differ substantially. Second, the interviews were conducted in such a way that the respondents' perceptions of the organization's behavior were based on their recollections of the organization's actions and decisions, which may be subject to social desirability bias (i.e., respondents may report their organizations behaving in ways that are more positive than actually occurred). Third, the study examined the perspectives of stakeholders at a single point in time, and as such does not permit an examination of how the implementation of ethical AI evolves over time. Fourth, while thematic analysis provides a rich and detailed understanding of the perspectives of stakeholders, the approach lacks statistical generalizability to larger populations. Fifth, the focus of this research was on organizations in technology sectors, and therefore, there may be limitations to the extent to which the findings can be generalized to other sectors and domains with differing regulatory requirements and structures of stakeholders.

Conclusion

Overall, this qualitative study examining the perspectives of stakeholders on the implementation of ethical AI found that although there is now consensus on the ethical principles to guide the development of AI, there is still an incredibly significant gap between these principles and the practice of implementing them in organizations. Five broad themes emerged from the data: the gap between principles and practice, which reflects systemic organizational characteristics, structural and cultural barriers that prevent implementation, algorithmic bias and fairness as persistent implementation challenges, privacy, transparency, and governance complexities, and pathways to ethical implementation, in which some organizations can move forward in implementing ethical AI through the establishment of governance mechanisms, diverse teams, and commitment by leaders to the implementation of ethical practices. In summary, the results show that the development of ethical AI is dependent on the implementation of systemic organizational changes that incorporate ethics into each phase of the lifecycle of AI development. For organizations to create effectively designed ethically developed artificial intelligence, organizations must establish formal governance mechanisms to ensure decision-making processes are structured to address ethical considerations. They must allocate sufficient resources to support the development of ethically developed artificial intelligence. Leaders must be held accountable for ensuring ethically positive outcomes when making decisions using artificially intelligent technology. Stakeholders who will be impacted by artificially intelligent technologies must be engaged during all stages of development of artificially intelligent technologies. Regulatory entities must establish regulatory frameworks that reward or incentivize organizations to develop and deploy ethically developed artificial intelligence. Future research must explore what organizational characteristics contribute to sustained commitment to ethically developed artificial intelligence and identify the factors that contribute to success in creating ethically developed artificial intelligence. Developing ethically developed artificial intelligence depends on recognizing that ethics form a base on which innovation worthy of public trust is built and serves humanity at every level of society and culture.

References

- Akinrinola, O., Okoye, C. C., Ofodile, O. C., & Ugochukwu, C. E. (2024). Navigating and reviewing ethical dilemmas in AI development: Strategies for transparency, fairness, and accountability. *GSC Advanced Research and Reviews*, 18(3), 050-058. <https://doi.org/https://doi.org/10.30574/gscarr.2024.18.3.0088>
- Ali, S. J., Christin, A., Smart, A., & Katila, R. (2023). Walking the walk of AI ethics: Organizational challenges and the individualization of risk among ethics entrepreneurs. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (pp. 217-226). <https://doi.org/10.1145/3593013.3593990>
- Bahangulu, J. K., & Owusu-Berko, L. (2025). Algorithmic bias, data ethics, and governance: Ensuring fairness, transparency and compliance in AI-powered business analytics applications. *World Journal of Advanced Research and Reviews*, 25(2), 1746-1763. <https://doi.org/10.30574/wjarr.2025.25.2.0571>
- Bansal, C. (2025). AI ethics and sustainability: Accelerating paradigm shifts toward sustainable development. *Journal of Strategy & Innovation*, 36(1), 200537. <https://doi.org/10.1016/j.jsinno.2025.200537>
- Bansal, C., Pandey, K. K., Goel, R., Sharma, A., & Jangirala, S. (2023). Artificial intelligence (AI) bias impacts: Classification framework for effective mitigation. *Issues in Information Systems*, 24(4), 367-389. https://doi.org/10.48009/4_iis_2023_127
- Bansal, C., Pandey, K. K., Goel, R., & Sharma, A. (2024). Analysis of barriers to AI banking chatbot adoption in India: An ISM and MICMAC approach. *Issues in Information Systems*, 25(4), 417-441. https://doi.org/10.48009/4_iis_2024_133
- Beaudouin, V., Bloch, I., Bounie, D., Cl  men  on, S., d'Alch  -Buc, F., Eagan, J., ... & Parekh, J. (2020). Flexible and context-specific AI explainability: a multidisciplinary approach. *arXiv preprint arXiv:2003.07703*. <https://doi.org/10.2139/ssrn.3559477>
- Birhane, A., Ruane, E., Laurent, T., S. Brown, M., Flowers, J., Ventresque, A., & L. Dancy, C. (2022). The forgotten margins of AI ethics. In *Proceedings of the 2022 ACM conference on fairness, accountability, and transparency* (pp. 948-958). <https://doi.org/10.1145/3531146.3533157>
- Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health*, 11(4), 589-597. <https://doi.org/10.1080/2159676x.2019.1628806>
- Braun, V., & Clarke, V. (2021). Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and psychotherapy research*, 21(1), 37-47. <https://doi.org/10.1002/capr.12360>
- Braun, V., & Clarke, V. (2023). Toward good practice in thematic analysis: Avoiding common problems and be(com)ing a knowing researcher. *International journal of transgender health*, 24(1), 1-6. <https://doi.org/10.1080/26895269.2022.2129597>
- Byrne, D. (2022). A worked example of Braun and Clarke's approach to reflexive thematic analysis. *Quality & quantity*, 56(3), 1391-1412. <https://doi.org/10.1007/s11135-021-01182-y>

- Chaudhary, G. (2024). Unveiling the black box: Bringing algorithmic transparency to AI. *Masaryk University Journal of Law and Technology*, 18(1), 93-122. <https://doi.org/https://doi.org/10.5817/mujlt2024-1-4>
- Chen, J., Storch, V., & Kurshan, E. (2021). Beyond fairness metrics: Roadblocks and challenges for ethical ai in practice. *arXiv preprint arXiv:2108.06217*. <https://doi.org/10.2139/ssrn.4941367>
- Correa, N. K., Santos, J. W., Galvão, C., Pasetti, M., Schiavon, D., Naqvi, F., ... & Oliveira, N. D. (2025). Crossing the principle–practice gap in AI ethics with ethical problem-solving. *AI and Ethics*, 5(2), 1271-1288. <https://doi.org/10.1007/s43681-024-00469-8>
- Christou, P. A. (2022). How to use thematic analysis in qualitative research. *Journal of Qualitative Research in Tourism*, 3(2), 79-95. <https://doi.org/10.4337/jqrt.2023.0006>
- Dado, M., Spence, J. R., & Elliot, J. (2023). The case of contradictions: How prolonged engagement, reflexive journaling, and observations can contradict qualitative methods. *International Journal of Qualitative Methods*, 22. <https://doi.org/10.1177/16094069231189372>
- Dai, E., & Wang, S. (2021). Say no to the discrimination: Learning fair graph neural networks with limited sensitive attribute information. In *Proceedings of the 14th ACM international conference on web search and data mining* (pp. 680-688). <https://doi.org/10.1145/3437963.3441752>
- Denton, E., Hanna, A., Amironesei, R., Smart, A., & Nicole, H. (2021). On the genealogy of machine learning datasets: A critical history of ImageNet. *Big Data & Society*, 8(2). <https://doi.org/10.1177/20539517211035955>
- Doniyor, U., Hong, S., & Kim, J. (2025). Research Ethics of Artificial Intelligence in the Digital Transformation Era: Fairness, Accountability, and Human-Centred Design. *International Journal of Contents*, 21(4), 10-19. <https://doi.org/10.5392/IJoC.2025.21.4.010>
- European Union. (2024). *Official Journal of the European Union*, L, 2024/1689. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- Eyinate Wasiu, Ezeilo Onyinye Jacqueline, & Ogundeji Ibidapo Abiodun. (2025). Navigating the complexities of ethical AI and Algorithmic accountability in modern technological practices. *Computer Science & IT Research Journal*, 6(6), 371-381. <https://doi.org/10.51594/csitrj.v6i6.1962>
- Farmer, T., Robinson, K., Elliott, S. J., & Eyles, J. (2006). Developing and implementing a triangulation protocol for qualitative health research. *Qualitative health research*, 16(3), 377-394. <https://doi.org/10.1177/1049732305285708>
- Finlay, L. (2021). Thematic analysis: the ‘good,’ the ‘bad’ and the ‘ugly.’ *European journal for qualitative research in psychotherapy*, 11, 103-116. <https://doi.org/10.24377/ejqrp.article3062>
- Floridi, L., & Cowls, J. (2022). A unified framework of five principles for AI in society. *Machine learning and the city: Applications in architecture and urban design*, 535-545. <https://doi.org/10.1002/9781119815075.ch45>
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Vayena, E. (2018).

- AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and machines*, 28(4), 689-707. <https://doi.org/10.1007/s11023-018-9482-5>
- Garcia, A. C. B., Garcia, M. G. P., & Rigobon, R. (2024). Algorithmic discrimination in the credit domain: what do we know about it? *AI & society*, 39(4), 2059-2098. <https://doi.org/10.1007/s00146-023-01676-3>
- Hennink, M. M., Kaiser, B. N., & Marconi, V. C. (2017). Code saturation versus meaning saturation: how many interviews are enough?. *Qualitative health research*, 27(4), 591-608. <https://doi.org/10.1177/1049732316665344>
- Huriye, A. Z. (2023). The ethics of artificial intelligence: examining the ethical considerations surrounding the development and use of AI. *American Journal of Technology*, 2(1), 37-44. <https://doi.org/10.58425/ajt.v2i1.142>
- Jain, L. R., & Menon, V. (2023). AI algorithmic bias: Understanding its causes, ethical and social implications. In *2023 IEEE 35th international conference on tools with artificial intelligence (ICTAI)* (pp. 460-467). IEEE. <https://doi.org/10.1109/ICTAI59109.2023.00073>
- Jonsen, K., & Jehn, K. A. (2009). Using triangulation to validate themes in qualitative studies. *Qualitative research in organizations and management: an international journal*, 4(2), 123-150. <https://doi.org/10.1108/17465640910978391>
- Karim, M. R., Islam, T., Shajalal, M., Beyan, O., Lange, C., Cochez, M., ... & Decker, S. (2023). Explainable AI for bioinformatics: methods, tools and applications. *Briefings in bioinformatics*, 24(5), bbad236. <https://doi.org/10.1093/bib/bbad236>
- Laitinen, A., & Sahlgren, O. (2021). AI systems and respect for human autonomy. *Frontiers in artificial intelligence*, 4, 705164. <https://doi.org/10.3389/frai.2021.705164>
- McCradden, M. D., Joshi, S., Mazwi, M., & Anderson, J. A. (2020). Ethical limitations of algorithmic fairness solutions in health care machine learning. *The Lancet Digital Health*, 2(5), e221-e223. [https://doi.org/10.1016/s2589-7500\(20\)30065-0](https://doi.org/10.1016/s2589-7500(20)30065-0)
- Moats, D., & Trifuljesko, S. (2025). Valuable actions and actionable values: Tinkering with principles and practices in AI ethics. *The Sociological Review*, 73(6), 1368-1387. <https://doi.org/10.1177/00380261251316661>
- Muceku, G., & Muceku, H. (2025). Ethics of using data in automated decision-making (Comparative perspectives on transparency, fairness, and institutional accountability in the age of AI). Article 10.58944/vcev7439. <https://doi.org/10.58944/vcev7439>
- Murikah, W., Nthenge, J. K., & Musyoka, F. M. (2024). Bias and ethics of AI systems applied in auditing-A systematic review. *Scientific African*, 25, e02281. <https://doi.org/10.1016/j.sciaf.2024.e02281>
- Musrifah, F., & Hasanah, I. A. (2025). Ethical implications of AI-driven recruitment: a multi-perspective study on bias and transparency in digital hiring platforms. *Journal of Management and Informatics*, 4(1), 599-616. <https://doi.org/10.51903/jmi.v4i1.140>
- National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management

Framework (AI RMF 1.0) (NIST AI 100-1). *U.S. Department of Commerce*.
<https://doi.org/10.6028/NIST.AI.100-1>

Organisation for Economic Co-operation and Development. (2024). *OECD AI principles*.
<https://oecd.ai/en/ai-principles>

Palumbo, G., Carneiro, D., & Alves, V. (2024). Objective metrics for ethical AI: a systematic literature review. *International Journal of Data Science and Analytics*, 20(2), 247-267.
<https://doi.org/10.1007/s41060-024-00541-w>

Setiawati, R., Kusmara, A., & Kuusk, T. (2025). Managing transparency fairness accountability in AI for sustainable human resource management. In *2025 4th International Conference on Creative Communication and Innovative Technology (ICCIT)* (pp. 1-7). IEEE.
<https://doi.org/10.1109/ICCIT65724.2025.11166968>

Shrestha, A. K., & Joshi, S. (2025, January). Toward ethical AI: A qualitative analysis of stakeholder perspectives. In *2025 IEEE 15th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 00022-00029). IEEE. <https://doi.org/10.1109/CCWC62904.2025.10903879>

Thomas, D. R. (2017). Feedback from research participants: are member checks useful in qualitative research? *Qualitative research in psychology*, 14(1), 23-41.
<https://doi.org/10.1080/14780887.2016.1219435>

United Nations Educational, Scientific and Cultural Organization. (2021). *Recommendation on the ethics of artificial intelligence*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>

Younas, A., Fàbregues, S., Durante, A., Escalante, E. L., Inayat, S., & Ali, P. (2023). Proposing the “MIRACLE” narrative framework for providing thick description in qualitative research. *International Journal of Qualitative Methods*, 22.
<https://doi.org/10.1177/16094069221147162>

Zhang, J. (2025). AI ethics and regulations: Ensuring trustworthy AI. *International Journal of Artificial Intelligence for Science (IJAI4S)*, 1(2). <https://doi.org/10.63619/ijai4s.v1i2.004>