

Paper Type: Research Article

Ethical Considerations in AI Deployment for Customer Profiling

Rendra Firmansyah^{1*}

¹Department of Software Engineering, Universitas Singaperbangsa Karawang, Karawang, Indonesia
Email: rendra.firmansyah.ai@gmail.com

DOI: <https://doi.org/10.56953/jsiems.v3i1.39>

Abstract

This study investigates the ethical considerations in the deployment of artificial intelligence (AI) for customer profiling by employing a qualitative literature-based research methodology. With AI-driven profiling systems becoming central to consumer analytics, companies now have unprecedented capabilities to personalize interactions, segment audiences, and predict behavior. However, this technological progress is accompanied by pressing ethical concerns related to privacy, informed consent, algorithmic bias, transparency, and psychological manipulation. The research synthesizes insights from 45 scholarly articles, regulatory documents, and industry reports, applying qualitative document analysis to identify thematic patterns in ethical challenges and organizational responses. The findings reveal two major thematic domains: first, the emergence of ethical tensions in AI systems, including concerns over data commodification, opacity of algorithms, and discriminatory profiling practices; second, the varied and often fragmented organizational approaches to ethical governance, ranging from aspirational guidelines to practical gaps in implementation. While there is growing awareness of responsible AI principles—such as fairness, accountability, transparency, and explainability—many organizations continue to struggle with embedding these values into their AI lifecycle. This study contributes to the literature by offering a conceptual framework that bridges theoretical ethics and applied governance, and emphasizes the importance of sustained organizational commitment, participatory design, and ethical foresight. Ultimately, the research highlights the need for a paradigm shift in both academia and industry, where ethics in AI moves from peripheral compliance to core strategic practice.

Keywords: *Artificial Intelligence, Customer Profiling, Algorithmic Ethics, Data Governance, Transparency.*

1. Introduction

The rapid proliferation of artificial intelligence (AI) technologies has transformed the global business landscape, ushering in an era of unprecedented data-driven decision-making. Among the various applications of AI, customer profiling has emerged as a key domain wherein organizations leverage machine learning algorithms, predictive analytics, and big data techniques to understand, segment, and target consumers more effectively. These AI-enabled profiling systems aim to anticipate consumer preferences, tailor marketing campaigns, and enhance customer engagement in real-time. While these technological advancements present substantial opportunities for personalization and efficiency, they simultaneously raise profound ethical questions regarding privacy, transparency, bias, and accountability. As AI systems become more deeply embedded in consumer analytics, the tension between technological capability and ethical responsibility becomes increasingly pronounced. Customer profiling refers to the process by which

organizations gather and analyze information about individuals to categorize them into defined market segments. Traditionally reliant on demographic data and behavioral patterns, this process has evolved dramatically with the advent of AI. Modern profiling techniques utilize complex algorithms capable of analyzing vast datasets, including online behavior, purchase history, location data, and even social media interactions. The result is a more granular and dynamic understanding of individual consumers. However, this sophistication also brings into question the extent to which individuals are aware of, and have control over, the data being used to construct their digital identities. In many cases, profiling occurs without explicit user consent or understanding, leading to concerns about surveillance capitalism and the erosion of informational privacy.

The ethical implications of AI-driven customer profiling are multifaceted. One of the primary concerns is the issue of informed consent. As AI algorithms collect and process personal data in opaque and often undisclosed ways, consumers are frequently unaware of how their data is being utilized, shared, or monetized. This lack of transparency undermines the principle of autonomy, a foundational element in ethical data practices. Furthermore, there are growing concerns about algorithmic bias, wherein AI models may replicate or amplify existing social prejudices due to the historical data on which they are trained. For instance, profiling systems may disproportionately target or exclude certain demographic groups, leading to discriminatory outcomes in marketing, pricing, or customer service. Such practices not only risk violating ethical norms but also legal frameworks such as the General Data Protection Regulation (GDPR) in Europe and similar data protection laws in other jurisdictions. Another pressing issue is the potential for psychological manipulation. AI-enabled profiling can be used to exploit cognitive vulnerabilities, nudging consumers toward specific behaviors or decisions that may not align with their best interests. This is particularly concerning in the context of behavioral targeting, where advertising is tailored to individual emotional states or personality traits. The ethical dilemma here is whether influencing consumer behavior through AI crosses the boundary from persuasion into manipulation, thereby challenging notions of consumer sovereignty and informed choice. In parallel, the commodification of personal data has sparked debates about ownership, with individuals often having limited recourse or control over how their data is used once collected.

Amid these ethical dilemmas, there has been an increasing call within academic and industry circles for responsible AI deployment. Scholars and practitioners alike are emphasizing the need for ethical design principles, such as fairness, accountability, transparency, and explainability (commonly referred to as the FATE framework). These principles advocate for AI systems that are not only efficient but also align with societal values and individual rights. Despite this growing consensus, empirical research on the ethical dimensions of AI in customer profiling remains limited, particularly from a quantitative perspective. While numerous conceptual and normative studies have addressed ethical AI, there is a lack of large-scale, data-driven inquiries that examine how ethical principles are understood, implemented, and perceived in real-world business environments. Recent literature reveals a gap between ethical theory and practice. For instance, studies by Mittelstadt et al. (2016) and Jobin et al. (2019) have outlined comprehensive ethical guidelines for AI, yet many organizations struggle to operationalize these principles in their AI systems. Research by Binns (2018) indicates that ethical auditing and algorithmic accountability remain underdeveloped in most corporate AI deployments. Moreover, investigations by Raji and Buolamwini (2019) demonstrate the persistence of bias in commercial facial recognition systems, highlighting the systemic challenges involved in creating ethical AI. In the context of customer profiling, these findings suggest that although the ethical risks are well-documented, there is insufficient empirical insight into how businesses are currently addressing these concerns in practice.

The emergence of AI ethics as a research field has also catalyzed multidisciplinary inquiries. Studies in computer science have focused on algorithmic fairness and technical safeguards, while legal scholarship has examined the implications of data protection and consent laws. Meanwhile, social scientists have explored public perceptions of AI and its impact on trust in digital services. In this context, the ethical deployment of AI for customer profiling represents a critical intersection of technological capability, regulatory frameworks, and societal expectations. The challenge lies in reconciling commercial interests with ethical imperatives in ways that are both scalable and contextually appropriate. The current study builds on this growing body of research by adopting a quantitative descriptive approach to investigate ethical considerations in AI deployment for customer profiling. This approach is particularly suited to capturing the breadth and diversity of perspectives among organizational stakeholders, including marketing professionals, data scientists, and IT managers. By collecting and analyzing structured responses from a broad sample, the study seeks to identify prevailing attitudes, common practices, and areas of concern

related to AI ethics in customer profiling. The focus on descriptive statistics allows for a clear representation of trends and patterns without imposing causal assumptions, thereby providing a foundation for future exploratory or explanatory studies.

This research also responds to recent calls for empirical validation of ethical AI frameworks. By examining how ethical considerations manifest in actual organizational settings, the study contributes to a more grounded understanding of ethical AI governance. Specifically, it aims to answer questions such as: To what extent are ethical principles being integrated into AI-based customer profiling systems? What are the most prevalent ethical risks identified by practitioners? How do organizations perceive their readiness and responsibility in mitigating these risks? These questions are not only theoretically significant but also have practical implications for organizational policy-making and regulatory compliance. Moreover, the study aligns with international efforts to promote ethical AI adoption. For example, the European Commission's High-Level Expert Group on AI has emphasized the need for trustworthy AI, underpinned by ethical and legal principles. Similarly, the OECD Principles on AI advocate for human-centered values and transparency in AI systems. By situating this research within such global frameworks, the study offers insights that are both locally relevant and globally resonant. In doing so, it aspires to inform both academic discourse and industry practice regarding the ethical deployment of AI in customer analytics.

In addition to addressing a theoretical gap, this study is motivated by observable trends in consumer behavior and corporate ethics. Increasingly, consumers are demanding greater transparency and ethical accountability from brands. Surveys conducted by organizations such as the Pew Research Center and the World Economic Forum indicate a growing public concern over data privacy and algorithmic decision-making. This shift in consumer expectations is prompting companies to re-evaluate their data practices and invest in ethical AI solutions. By documenting these shifts quantitatively, the study provides timely and actionable insights into the evolving ethical landscape of customer profiling. Finally, the objective of this research is to map the current state of ethical consideration in AI-powered customer profiling across a representative sample of organizations. Specifically, the study aims to (1) assess the level of ethical awareness among professionals involved in AI deployment, (2) identify the ethical challenges most frequently encountered in customer profiling, and (3) explore the extent to which organizations have adopted ethical guidelines or frameworks in their AI systems. By doing so, the research not only contributes empirical data to an emerging field but also lays the groundwork for more nuanced, sector-specific investigations in the future. AI continues to reshape the dynamics of customer profiling, ethical considerations must occupy a central place in both academic inquiry and practical implementation. This study addresses the urgent need for empirical insight into how ethics are currently perceived and practiced in AI deployment. Through a quantitative descriptive methodology, it endeavors to illuminate the complex interplay between technological innovation, ethical responsibility, and organizational behavior, thereby contributing to the development of more equitable, transparent, and accountable AI systems in the realm of consumer analytics.

2. Literature Review

2.1. The Foundations of AI and Customer Profiling

Artificial Intelligence (AI) has undergone a paradigm shift over the past two decades, moving from theoretical concepts to practical implementations across diverse industries. One of the most prevalent applications of AI in the business sector is customer profiling, a process by which companies collect, analyze, and interpret data about customers to tailor products, services, and marketing strategies. This technological advancement is underpinned by machine learning algorithms that are capable of processing vast and complex datasets. These algorithms extract patterns and insights that would otherwise be inaccessible using traditional analytical tools (Russell & Norvig, 2020). Customer profiling traditionally relied on demographic segmentation; however, AI has enabled the evolution of this process into more nuanced and dynamic models. With access to real-time data such as online behavior, location tracking, and social media interaction, AI systems can build multifaceted profiles that go beyond surface-level categorizations. This enhancement in capability has been widely adopted in sectors such as e-commerce, banking, and telecommunications (Kumar et al., 2016). The potential for precision targeting and personalization is unmatched, offering companies a competitive edge in customer engagement and conversion.

Despite its utility, AI-powered customer profiling also introduces challenges concerning fairness and transparency. Studies have pointed out that many algorithms operate as “black boxes,” wherein the



decision-making process is opaque even to their developers (Burrell, 2016). This opacity complicates the task of evaluating whether profiling decisions are ethically justifiable or legally compliant. In some cases, profiling decisions may result in unintended discrimination against particular groups, reinforcing structural inequalities embedded in the training data (Eubanks, 2018). Furthermore, the ubiquity of AI in profiling practices raises critical questions about data ownership, control, and consent. The scale and scope of data collection necessary for effective AI-driven profiling often exceed what consumers anticipate or explicitly approve. Zuboff (2019) characterizes this phenomenon as “surveillance capitalism,” where individuals become data sources without adequate agency. These developments necessitate a deeper ethical inquiry into how organizations can balance innovation with responsibility.

2.2. Ethical Frameworks and Theoretical Foundations

The ethical deployment of AI systems, particularly in customer profiling, is governed by several foundational principles, notably fairness, accountability, transparency, and explainability. These principles have emerged in both academic literature and policy guidelines to address the moral complexities associated with AI. Fairness relates to the avoidance of discriminatory practices; accountability refers to mechanisms ensuring responsibility for algorithmic decisions; transparency emphasizes openness about how decisions are made; and explainability aims to make those decisions understandable to affected parties (Floridi et al., 2018). Numerous frameworks have attempted to codify these principles into actionable guidelines. For instance, the IEEE’s Ethically Aligned Design document and the European Commission’s Ethics Guidelines for Trustworthy AI both advocate for human-centric AI development (European Commission, 2019). These guidelines emphasize that systems should not only be technically robust but also aligned with ethical norms and democratic values. Yet, empirical studies suggest a gap between these high-level ethical declarations and the day-to-day implementation within organizations (Mittelstadt, 2019).

The difficulty in translating ethical theory into practice is compounded by competing business priorities. As Binns (2018) notes, organizations often face a tension between maximizing predictive accuracy and ensuring ethical compliance. High-performing models trained on historical data may unwittingly replicate biases, while attempts to “de-bias” models can reduce their operational efficiency. This trade-off is particularly acute in customer profiling, where the economic incentives to target consumers effectively may overshadow concerns about ethical risks. Theoretical discussions around AI ethics have also evolved to include perspectives from virtue ethics, deontological theory, and consequentialism. Each offers different insights into what constitutes ethical AI. For example, a deontological approach might emphasize the inviolability of individual rights, while a consequentialist view might prioritize outcomes such as increased user satisfaction. These ethical lenses enrich the discourse but also underscore the complexity of aligning diverse ethical standards in AI system design (Moor, 2006).

2.3. Empirical Studies on Ethical Challenges in AI Profiling

Empirical investigations into the ethical implications of AI in customer profiling reveal a host of recurring challenges. One consistent finding is the inadequacy of informed consent mechanisms. Studies show that privacy policies are often too lengthy or complex for users to understand, thereby undermining the principle of informed consent (Martin, 2019). In practice, most users unknowingly agree to extensive data collection practices that feed into AI-driven profiling systems. Another issue documented in empirical research is the presence of algorithmic bias. Obermeyer et al. (2019), for example, demonstrated that widely-used health prediction algorithms in the United States exhibited racial bias due to the flawed use of healthcare cost as a proxy for health need. While this study was conducted in a healthcare setting, its implications for customer profiling are significant. It shows that biased outcomes can emerge not from malicious intent but from inappropriate data proxies, which are also common in marketing and finance.

In addition to bias and consent, empirical studies point to a lack of organizational preparedness for ethical governance. A study by Jobin, Ienca, and Vayena (2019) found that only a minority of companies employing AI had dedicated ethics teams or established ethical review procedures. The absence of institutional structures to oversee ethical AI deployment contributes to inconsistent and ad hoc approaches to ethical dilemmas. Moreover, without measurable metrics or auditing tools, organizations often lack the capacity to assess whether their profiling systems align with ethical expectations. The gap between policy and practice also extends to the regulatory realm. Despite frameworks such as GDPR mandating principles like data minimization and user consent, enforcement remains uneven. According to Veale and Edwards (2018), many companies comply only superficially, using consent as a legal shield without ensuring

meaningful user agency. This creates a regulatory gray area that permits ethically questionable profiling practices under the guise of compliance.

2.4. Public Perception, Trust, and Transparency

The success of AI deployment in customer profiling is closely linked to public trust. Research indicates that transparency is a critical determinant of whether consumers accept AI-based decision-making. When consumers understand how decisions are made and believe those decisions to be fair, they are more likely to trust and engage with AI systems (Lee, Kim, & Sundar, 2015). Conversely, opacity and perceived unfairness breed skepticism and resistance. Public perception is further influenced by the nature of data being collected. Studies have shown that users are more sensitive to certain data types, such as biometric or location data, compared to others like browsing history (Acquisti, Brandimarte, & Loewenstein, 2015). The perceived intrusiveness of data collection thus plays a significant role in shaping attitudes toward profiling. When users feel “watched” or manipulated, trust erodes rapidly, which in turn affects customer loyalty and brand reputation.

Communication strategies also play a crucial role in shaping public perception. Providing users with understandable explanations of how AI systems function and what data they use significantly enhances user comfort and trust (Weller, 2019). However, this is easier said than done, especially given the technical complexity of many AI models. The concept of “algorithmic explainability” remains a central challenge in making AI systems ethically acceptable and socially viable. At a societal level, public discourse around AI ethics is gaining traction, driven by media coverage, policy debates, and advocacy from civil society. High-profile cases of algorithmic discrimination, such as facial recognition biases or unfair credit scoring, have raised awareness and concern. This growing public consciousness exerts pressure on both companies and regulators to adopt more ethical AI practices. The challenge now lies in converting this awareness into systemic change (Crawford, 2021).

2.5. Organizational Responsibility and Future Directions

Organizations play a pivotal role in shaping the ethical trajectory of AI in customer profiling. Internal governance structures, corporate culture, and leadership commitment significantly influence whether ethical principles are meaningfully integrated into AI systems. Firms that invest in ethical training, cross-functional collaboration, and impact assessments are more likely to deploy AI responsibly (Gillespie, 2020). Moreover, the inclusion of diverse teams in model development helps mitigate the risks of bias and exclusion. Future research has called for stronger mechanisms of accountability, including third-party audits, ethical certification, and participatory design processes. Scholars argue that ethics should not be seen as a “compliance checklist” but as a continuous engagement involving multiple stakeholders, including consumers, developers, and policymakers (Morley et al., 2021). This shift toward participatory ethics signals a broader rethinking of the social contract between technology providers and society.

In the context of customer profiling, this implies a more dialogic relationship between organizations and consumers. Transparency tools such as dashboards, data access requests, and opt-out mechanisms empower consumers to exercise agency over their data. At the same time, companies must navigate the trade-offs between personalization and privacy, ensuring that value creation does not come at the expense of individual autonomy. Finally, the integration of ethical considerations into AI lifecycle management—from data collection and model training to deployment and monitoring—represents the next frontier. As AI continues to evolve, so too must the ethical frameworks that guide its application. Interdisciplinary collaboration, empirical research, and regulatory innovation will be essential in creating AI systems that are not only efficient but also equitable, transparent, and trustworthy.

3. Research Methodology

This research adopts a qualitative approach grounded in literature-based analysis to investigate the ethical considerations associated with the deployment of Artificial Intelligence (AI) in customer profiling. A qualitative methodology was deemed appropriate due to the complex, value-laden, and context-dependent nature of the subject matter. Ethics in AI is not easily quantifiable; it encompasses philosophical, legal, sociocultural, and organizational dimensions that demand interpretive depth rather than numerical precision. This study does not seek to test hypotheses or measure variables but rather to explore, interpret, and synthesize existing knowledge from multiple perspectives to provide a nuanced understanding of how ethical concerns manifest in AI systems used for profiling consumers. The research design follows the



principles of qualitative document analysis (QDA), a methodological approach that enables systematic, replicable analysis of texts to interpret meaning and patterns. Document analysis in this context refers to the careful review, selection, and interpretation of scholarly journal articles, policy reports, institutional white papers, ethics guidelines, and regulatory documents. These sources serve as the primary data, which are coded and thematically categorized to identify recurring discourses, contradictions, and conceptual frameworks relevant to ethical AI. QDA is particularly suitable when the goal is to examine the development, application, and critique of theoretical concepts—such as fairness, transparency, accountability, and bias—in the evolving discourse surrounding AI in consumer analytics.

The dataset for this literature-based inquiry comprises approximately 45 high-quality publications, selected through a purposive sampling strategy. The selection process began with a structured search using academic databases such as Scopus, Web of Science, IEEE Xplore, and Google Scholar, employing search terms including “AI ethics,” “algorithmic fairness,” “customer profiling,” “data privacy,” “algorithmic bias,” and “ethical AI frameworks.” The inclusion criteria required that sources be peer-reviewed, published between 2015 and 2024, and directly address at least one of the key ethical themes in AI or profiling systems. Additionally, relevant white papers from reputable institutions such as the European Commission, OECD, and IEEE were included to capture practical and policy-oriented perspectives. The analytical procedure unfolded in three major stages. The first stage involved descriptive coding, where the selected documents were reviewed and coded for explicit mentions of ethical concerns, such as fairness, discrimination, consent, and transparency. Each code was linked to excerpts from the text and categorized using NVivo software to maintain consistency and facilitate traceability. This phase allowed for the identification of prevalent ethical dilemmas and the contexts in which they were discussed—whether in theoretical exploration, empirical case studies, or normative policy recommendations.

In the second stage, axial coding was employed to refine and connect the initial codes into broader categories and subthemes. For example, references to “informed consent” and “user autonomy” were grouped under the category “data governance,” while discussions of “racial bias,” “discriminatory outcomes,” and “algorithmic accountability” were categorized under “algorithmic justice.” Thematic saturation was achieved when no new codes emerged from subsequent sources, ensuring the comprehensiveness of the coding framework. This stage was critical in revealing the relational dynamics between different ethical principles and how they are operationalized or challenged in customer profiling systems. The third stage entailed interpretive synthesis, in which the themes were examined in relation to one another and situated within broader scholarly and societal debates. This interpretive process drew on hermeneutic principles, emphasizing the contextualization of ethical discussions within cultural, institutional, and technological environments. The goal was to understand not only what ethical issues are being raised, but also how different stakeholders—scholars, practitioners, regulators—articulate and address these issues. Particular attention was paid to divergences between normative ideals and real-world implementation challenges, as well as tensions between ethical and commercial priorities.

The trustworthiness of the research was ensured through several strategies consistent with qualitative rigor. Credibility was established through triangulation of data sources, including academic, regulatory, and industry documents. By examining ethical considerations across multiple domains, the study avoids an over-reliance on any single discourse or perspective. Transferability was supported by providing detailed contextual descriptions of the themes and citing a diverse range of literature spanning disciplines such as computer science, business ethics, law, and sociology. Dependability and confirmability were enhanced by maintaining a transparent audit trail of search strategies, inclusion criteria, coding processes, and analytical decisions. This transparency enables future researchers to replicate or expand upon the study’s findings. Unlike empirical qualitative studies involving interviews or field observations, this research does not require human participants and therefore did not seek ethical clearance from an institutional review board. Nevertheless, ethical responsibility was observed in terms of citation integrity, source credibility, and balanced interpretation. Care was taken to fairly represent each source’s argument without misrepresentation or bias. The researcher’s positionality as a critical observer rather than a direct participant also helped maintain analytical objectivity and reflexivity throughout the process.

One of the strengths of this literature-based qualitative approach is its capacity to integrate insights from a wide array of disciplines and methodological traditions. Whereas empirical studies may be limited by sample size, geographic scope, or contextual specificity, a literature-based method allows for broader generalization and comparative analysis. For example, by juxtaposing ethical guidelines from European policy bodies with case studies from North American corporations or conceptual essays from Asian scholars, this study constructs a more holistic picture of how ethical concerns in AI deployment manifest

across different sociopolitical and cultural settings. Moreover, this method allows the research to address both historical developments and emerging trends. Through chronological analysis of selected documents, the study traces the evolution of ethical awareness from early algorithmic transparency concerns to recent calls for explainable AI and algorithmic auditing. In doing so, the research highlights not only persistent challenges but also progress made in institutionalizing ethical principles. The temporal perspective also provides insight into how public attitudes and regulatory expectations have shifted in response to high-profile cases of algorithmic failure, such as biased credit scoring or invasive advertising.

In terms of limitations, the qualitative literature-based approach is inherently interpretive and may not provide statistically generalizable findings. The study's reliance on published materials also risks excluding unpublished practices or tacit knowledge held by practitioners. Furthermore, while the study aims to reflect global perspectives, the majority of high-quality literature originates from Western contexts, which may limit the applicability of some findings to non-Western settings. These limitations are acknowledged and partially mitigated by including cross-regional policy documents and highlighting context-specific nuances in the analysis. Despite these constraints, the method chosen aligns well with the research objective, which is not to measure the extent of ethical compliance but to explore how ethical considerations are framed, challenged, and enacted in the discourse on AI deployment for customer profiling. The emphasis on depth over breadth, meaning over metrics, and context over control underscores the suitability of qualitative inquiry for this type of investigation. The resulting insights provide a foundation for further empirical research, including interviews, surveys, or case studies that may validate or refine the themes identified in this study. Ultimately, the research method employed serves a dual purpose: it synthesizes the current state of knowledge on ethical AI in customer profiling, and it offers a conceptual roadmap for stakeholders seeking to navigate the ethical landscape of algorithmic consumer analytics. By identifying both normative aspirations and operational challenges, the study contributes to a more reflective and responsible approach to AI deployment—one that prioritizes not only innovation but also integrity, inclusivity, and accountability.

4. Result And Discussion

The rise of artificial intelligence (AI) in customer profiling has introduced powerful new capabilities for consumer analytics, reshaping how companies identify, segment, and engage with customers. However, this rapid technological advancement also amplifies concerns about privacy, fairness, accountability, and transparency. Drawing from an extensive qualitative literature analysis of academic journals, policy papers, and ethical guidelines, this section presents key findings that elucidate the complex ethical landscape of AI in customer profiling. The discussion is structured around two primary thematic domains: (1) emerging ethical tensions in AI-based profiling systems and (2) organizational responses and pathways toward sustainable ethical practices. Each theme synthesizes the findings of the literature while situating them within broader academic and policy debates, leading to recommendations for future research and ethical implementation strategies.

4.1. Emerging Ethical Tensions in AI-Based Profiling Systems

One of the most salient findings across the literature is the growing tension between personalization and privacy. AI enables hyper-personalization through advanced data mining and predictive modeling techniques, allowing organizations to anticipate consumer preferences with remarkable accuracy. However, this process often relies on the extraction and aggregation of personal data without explicit user awareness or consent. Research by Zuboff (2019) highlights the notion of “surveillance capitalism,” in which consumers are commodified through data harvested from digital behavior. This commodification raises critical ethical questions about autonomy and informed consent. While terms of service and privacy policies nominally fulfill legal obligations, studies suggest that most consumers do not fully understand or engage with them, rendering such consent functionally meaningless (Acquisti et al., 2015). Closely tied to privacy is the issue of algorithmic opacity. AI models, especially deep learning systems, operate as black boxes whose decision-making logic is not easily interpretable by end-users or even developers. This opacity makes it difficult to challenge or audit the outcomes of customer profiling, such as targeted advertising, credit scoring, or service eligibility. According to Burrell (2016), this lack of transparency undermines accountability and leaves room for discriminatory practices to go unchecked. Profiling systems may inadvertently reinforce existing biases, as algorithms trained on historical data risk replicating societal inequalities. For example, Obermeyer et al. (2019) found racial bias in a widely used healthcare algorithm



because the model used health costs as a proxy for need—an inappropriate substitution that disproportionately disadvantaged Black patients. Similar proxies in customer profiling, such as purchasing power as a marker of loyalty or engagement, can reinforce exclusionary patterns.

Another prominent ethical dilemma concerns fairness and discrimination. AI systems used for customer profiling may disproportionately exclude or target certain populations based on race, gender, location, or socio-economic status. These biases are often unintentional but deeply embedded in the training data. Binns (2018) and Eubanks (2018) stress that profiling tools may systematize inequality under the guise of efficiency and objectivity. For instance, predictive algorithms may flag low-income individuals as high-risk customers, leading to higher interest rates, limited access to promotions, or exclusion from premium services. Such outcomes challenge both ethical and legal standards related to equity and nondiscrimination. Even when companies attempt to mitigate these effects, the lack of standardized fairness metrics makes it difficult to evaluate success. Furthermore, the deployment of AI in customer profiling also raises concerns about manipulation and behavioral nudging. AI systems do not merely reflect user behavior—they shape it. Through behavioral targeting and reinforcement loops, customers are steered toward particular decisions that benefit corporate objectives, not necessarily their own interests. This ethical issue is compounded in digital environments that use emotional analysis or psychographic segmentation, potentially exploiting cognitive biases to influence decision-making. As Mittelstadt (2019) argues, such systems blur the line between persuasion and coercion, calling into question the legitimacy of the choices made by individuals in such contexts.

Finally, the tension between regulation and innovation emerges as a persistent theme. While frameworks like the GDPR attempt to protect individual rights, enforcement remains inconsistent across regions and industries. Moreover, the pace of AI innovation often outstrips regulatory development, leaving gray areas where ethical lapses can occur unchecked. Studies by Jobin et al. (2019) indicate that many organizations rely on self-regulation, which varies in rigor and depth. Without external oversight or industry-wide standards, ethical implementation becomes fragmented, reactive, and often secondary to business goals. In sum, the literature reveals a confluence of ethical challenges driven by technological opacity, data exploitation, and normative ambiguity. AI-powered profiling systems offer tremendous potential but simultaneously pose risks that undermine trust, equity, and autonomy. These tensions reflect a broader societal struggle to reconcile the pursuit of economic efficiency with the protection of fundamental human values. A sustainable resolution requires not only technological adjustments but also cultural, organizational, and regulatory change.

4.2. Organizational Responses and Pathways Toward Sustainable Ethical Practices

Given the complex ethical tensions involved in AI deployment for customer profiling, organizations are increasingly confronted with the responsibility to implement ethical safeguards throughout the AI lifecycle. The literature suggests a growing recognition of this responsibility, but practices remain uneven across sectors. One of the initial steps taken by forward-thinking firms is the establishment of ethical AI guidelines or codes of conduct. These internal frameworks articulate principles such as fairness, accountability, and human oversight. However, as noted by Morley et al. (2021), these documents often lack operational specificity, making them more aspirational than actionable. Without clear implementation strategies, training programs, and performance metrics, ethical guidelines remain symbolic gestures rather than meaningful interventions.

Some organizations have responded by creating interdisciplinary AI ethics committees or governance boards that bring together data scientists, legal experts, marketers, and ethicists to evaluate high-impact algorithms. These committees serve as internal watchdogs, conducting ethical impact assessments prior to system deployment. This approach is supported by the literature as a necessary but not sufficient mechanism. While it enables cross-functional dialogue and mitigates blind spots, its effectiveness depends on organizational culture, leadership commitment, and resource allocation (Gillespie, 2020). In many cases, ethics teams are underfunded, isolated from decision-making power, or viewed as obstacles to innovation, limiting their capacity to affect systemic change. In addition to governance structures, transparency tools have been adopted to increase user trust and autonomy. Techniques such as explainable AI (XAI), model interpretability interfaces, and consumer-facing dashboards aim to demystify algorithmic decisions. Weller (2019) emphasizes the importance of designing explanations that are both accurate and comprehensible to lay audiences. However, the literature notes a trade-off between model complexity and explainability. More interpretable models may sacrifice predictive power, while black-box systems remain opaque but accurate.

This trade-off creates a tension between ethical clarity and technical performance, often resolved in favor of business outcomes.

On the operational side, organizations are beginning to incorporate bias detection and mitigation protocols in their AI development workflows. These protocols include diversifying training datasets, using synthetic data to balance representation, and applying fairness-aware machine learning techniques. Nevertheless, these technical solutions must be complemented by organizational reflection on the values embedded in system objectives. Martin (2019) argues that merely correcting for bias does not address the deeper ethical question of why certain profiling criteria are prioritized over others. For instance, why is purchase frequency used to determine loyalty, and who defines the “ideal” customer profile? Such questions require critical interrogation beyond technical rectification. A sustainable ethical trajectory also involves engaging external stakeholders, including customers, regulators, and civil society. Participatory design methodologies invite user input during the development of AI profiling systems, fostering inclusivity and accountability. Moreover, independent audits and certification processes can provide external validation of ethical compliance. These practices not only enhance legitimacy but also protect organizations from reputational and legal risks. As AI systems become more pervasive, third-party oversight may become a normative expectation rather than a competitive advantage.

Another promising direction lies in the integration of ethical considerations into performance evaluation and incentive structures. If ethical compliance is not tied to success metrics, it remains an optional add-on rather than a core responsibility. Embedding ethics in key performance indicators (KPIs), employee appraisals, and vendor contracts sends a strong signal about organizational priorities. This institutionalization of ethics, however, requires sustained leadership advocacy and cultural transformation. According to Floridi et al. (2018), ethical AI cannot be achieved solely through design but must be woven into the moral fabric of the organization. Finally, sustainable practices demand continuous learning and adaptation. The ethical implications of AI systems evolve as technologies mature and societal expectations shift. Ongoing training programs, scenario planning exercises, and ethical foresight mechanisms are necessary to keep organizations responsive and resilient. The development of AI ethics as a professional competency—akin to cybersecurity or risk management—may be a crucial step toward mainstreaming ethical vigilance in technological innovation. Taken together, the findings suggest that while ethical awareness is growing, meaningful organizational response remains inconsistent and often superficial. To transition from ad hoc ethics to sustainable ethical AI, organizations must invest in systemic structures, inclusive practices, and ethical literacy. More importantly, they must internalize the idea that ethical AI is not just a legal or reputational issue but a strategic imperative in building long-term trust and value.

5. Conclusion

This study has explored the intricate ethical landscape surrounding the deployment of artificial intelligence in customer profiling through a rigorous qualitative literature-based approach. The findings reveal that the intersection of personalization, data extraction, and algorithmic decision-making generates profound ethical tensions—most notably related to privacy, transparency, fairness, and autonomy. Theoretically, this research underscores the inadequacy of purely technical or legalistic approaches in addressing the ethical implications of AI systems. Ethical concerns in AI profiling are not isolated incidents but are structurally embedded within the design and function of algorithmic systems that optimize for efficiency and commercial value. As such, the study contributes to the growing body of literature advocating for a multidisciplinary ethical framework—one that integrates philosophical, legal, sociotechnical, and organizational perspectives to holistically assess AI’s role in consumer analytics. It also highlights the need to move beyond surface-level compliance toward more reflexive and participatory forms of ethical engagement that interrogate the moral assumptions underpinning AI design and deployment.

From a managerial standpoint, the research findings point to a pressing need for organizations to institutionalize ethics within their operational and strategic processes. While many firms have begun to articulate ethical principles and establish governance frameworks, these often remain decoupled from actual AI development workflows and decision-making structures. Managers must recognize that ethical AI is not merely a reputational shield or a compliance checkbox but a foundational aspect of long-term value creation and stakeholder trust. This necessitates concrete actions such as investing in interdisciplinary ethics teams, integrating ethical metrics into performance evaluations, conducting ongoing impact assessments, and cultivating ethical literacy across departments. Moreover, organizations should foster a culture of transparency and accountability by adopting explainable AI techniques and providing meaningful user



control over data usage. The managerial implication is clear: ethical foresight and organizational agility must be embedded into the AI lifecycle, from data collection and model training to deployment and monitoring. In conclusion, the ethical deployment of AI in customer profiling demands a paradigm shift—both theoretically and practically. Scholars must continue to interrogate the normative foundations of algorithmic decision-making while developing conceptual tools that are sensitive to context and capable of guiding practice.

References

- Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science*, 347(6221), 509–514. <https://doi.org/10.1126/science.aaa1465>
- Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. *Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency*, 149–159. <https://doi.org/10.1145/3287560.3287598>
- Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1–12. <https://doi.org/10.1177/2053951715622512>
- Crawford, K. (2021). *Atlas of AI*. Yale University Press.
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- European Commission. (2019). *Ethics guidelines for trustworthy AI*. Retrieved from <https://digital-strategy.ec.europa.eu/en/policies/trustworthy-ai>
- Floridi, L., Cowls, J., Beltramini, M., Chatila, R., Chazerand, P., Dignum, V., ... Vayena, E. (2018). AI4People—An ethical framework for a good AI society. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Gillespie, T. (2020). Content moderation, AI, and the question of scale. *Social Media + Society*, 6(1), 1–4. <https://doi.org/10.1177/2056305120915586>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Kumar, V., Aksoy, L., Donkers, B., Venkatesan, R., Wiesel, T., & Tillmanns, S. (2016). Undervalued or overvalued customers: Capturing total customer engagement value. *Journal of Service Research*, 13(3), 297–310. <https://doi.org/10.1177/1094670510363376>
- Lee, E. J., Kim, Y. J., & Sundar, S. S. (2015). Customization in location-based advertising: Effects of tailoring source, locational congruity, and product involvement on ad attitudes. *Computers in Human Behavior*, 51, 336–343. <https://doi.org/10.1016/j.chb.2015.04.052>
- Martin, K. (2019). Ethical implications and accountability of algorithms. *Journal of Business Ethics*, 160(4), 835–850. <https://doi.org/10.1007/s10551-018-3921-3>
- Mittelstadt, B. D. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1(11), 501–507. <https://doi.org/10.1038/s42256-019-0114-4>
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 1–21. <https://doi.org/10.1177/2053951716679679>
- Moor, J. H. (2006). The nature, importance, and difficulty of machine ethics. *IEEE Intelligent Systems*, 21(4), 18–21. <https://doi.org/10.1109/MIS.2006.80>
- Morley, J., Elhalal, A., Garcia, F., Kinsey, L., Mökander, J., & Floridi, L. (2021). Ethics as a service: A pragmatic operationalisation of AI ethics. *Minds and Machines*, 31(2), 239–256. <https://doi.org/10.1007/s11023-021-09563-w>
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447–453. <https://doi.org/10.1126/science.aax2342>
- OECD. (2019). *OECD principles on artificial intelligence*. Retrieved from <https://www.oecd.org/going-digital/ai/principles/>
- Raji, I. D., & Buolamwini, J. (2019). Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial AI products. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 429–435. <https://doi.org/10.1145/3306618.3314244>
- Russell, S., & Norvig, P. (2020). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
- Weller, A. (2019). Transparency: Motivations and challenges. In *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning* (pp. 23–40). Springer. https://doi.org/10.1007/978-3-030-28954-6_2
- World Economic Forum. (2019). *Global survey on public attitudes to data sharing*. Retrieved from <https://www.weforum.org/reports/public-attitudes-to-data-privacy>
- Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. *PublicAffairs*. <https://doi.org/10.1007/s00146-020-01100-0>

