Ethical Considerations of AI-Driven Decision-Making: Addressing Bias and Accountability in Business Practices

Antonieta Lima

ABSTRACT

The swift rate and depth of AI assimilation into the fabric of business operations have a considerable impact on the process of decision-making among managers. However, this also leads to the surfacing of considerable ethical dilemmas, especially those pertaining to formulative biases that are unidentified by the algorithms, as well as the loss of corporate accountability. This being an exhaustive analysis of AI ethics in the world of business, this paper seeks to highlight the impediments brought about by the assimilation of data bias into AI systems, especially along the avenues of corporate accountability. Moreover, it also seeks to provide an exhaustive understanding of the mitigators that currently exist, along with the EU AI Act, which entered into effect in 2024.

Keywords: Artificial intelligence ethics, Algorithmic bias, Corporate accountability, EU AI act, Machine learning governance, Responsible innovation.

ISVOUGA, Santa Maria da Feira, Portugal

Correspondence to: Antonieta Lima, lima.antonietamaria@gmail.com

Cite this as: Lima A. Ethical considerations of AI-driven decision-making: addressing bias and accountability in business practices. Premier Journal of Business and Management 2026;7:100014

DOI: https://doi.org/10.70389/PJBM.100014

Peer Review:

Received: 29 January 2026

Last revised: 05 May 2026

Accepted: 05 May 2026

Version accepted: 4

Published: 13 May 2026

Ethical approval: N/a

Consent: N/a

Funding: N/a

Conflicts of interest: N/a

Author contribution: Antonieta Lima–Writing-original draft

Guarantor: Antonieta Lima

Provenance and peer-review: Unsolicited and externally peer-reviewed

Data availability statement: N/a

Introduction

The emergence of artificial intelligence is greatly revolutionizing the way humans relate to technologies and the operations carried out, especially because of its crucial role in leading to the emergence of the fourth industrial revolution.[1,2] However, because of the extensive emergence of machine learning, neural networks, and generative technologies, artificial intelligence has shifted from being merely hypothetical with exploratory uses to becoming an essential tool in conducting global commerce. However, the use of artificial intelligence is still a challenge, especially because of the high failure rates that are claimed to be because of, as opposed to because of, a lack of technological advancements, because of its complex ethical undertones.[3,4]

The current shift in paradigm represents a transition from "assisted AI" to "autonomous AI," where the systems possess autonomous agency in order to be able to make decisions with minimal human intervention. This has, therefore, led to critical questions being raised with regard to individual and societal wellbeing.[5,6] Such autonomy in action can be seen in automated high-frequency trading, dynamic pricing algorithms, and automated human resource management systems in business operations. However, the bigger the systems are, the bigger the scale of their moral implications that are attached to their outputs.

Among the most prominent issues is the fact that the transparency and accountability of the decision-making processes at play are still not very impressive.[7] The truth is that, as deep learning models become more complex, it is still not very clear, to a certain extent, why a certain model arrived at a certain decision. The more mysterious the algorithms become, the more difficult it is for the stakeholders, including employees and end-users, as well as those who are responsible for monitoring the use of artificial intelligence, such as government officials, to grasp the underlying nature of decisions such as being "credit-worthy" or being in a job/being "sick."[8,9] There is also the problem of velocity of harm, thanks to the speed of artificial intelligence.

This paper aims to look at the fundamental principles of ethics, such as autonomy, beneficence, non-maleficence, and justice, as well as transparency and accountability, as they relate to business practice in today's business world.[10,11] The burning question that this article aims to answer is how corporations can seize the competitive advantage that AI has to offer without transgressing the ethical codes that form the foundation of trust within society. This study will .also present a comprehensive overview of how bias within algorithms can be identified and cured, as well as how accountability needs to be reordered when the actor is no longer a human.

While previous studies have attempted to categorize [Topic], there remains a lack of conceptual clarity. This review seeks to fill that gap. Specifically, it departs from a standard systematic reporting format to focus on narrative synthesis and theoretical integration, allowing for a broader interpretation of diverse methodological approaches in the literature.

The economic Framework: Externalities and the Social Cost of Carbon

The branch of AI ethics relates to human behavior in AI designing and usage, and AI autonomous entities.[10,12] In addressing these issues, business organizations are increasingly adopting multi-dimensional ethical frameworks in which classical ethical theories, such as deontology and utilitarianism, are used and adapted to new speeds of digital logic.[13–15] The fundamental goal of these frameworks is to create "guardrails of trust" for AI to perform in a manner that is consistent with human integrity and values.[16,17]

The Principle of Autonomy

In the form of AI, autonomy is defined as "the right of self-determination of the individual through well-informed decisions, and the duty of AI to support and not undermine self-determination of humans."[8] In business, this concept of autonomy is challenged in terms of the "nudge" effect and judgment removal.

Firstly, recommendation systems, like those used in marketing or workforce management, inherently have the ability to affect the autonomy of the user. For

instance, "in the context of a project management tool that utilizes AI, the user is given a 'forced choice' that reduces their autonomy, as the recommendation influences the likelihood that users will choose one option over another, altering the autonomy of users within the system—the manager may choose to rely on the tool's recommendations due to personal liability, thereby passing their autonomy over to the tool itself, effectively passing their authority over to the tool."[13] Secondly, with regard to critical applications like healthcare or legal services, the autonomy of the user is being compromised to the point that decisions are being made to the extent that human professionals are merely the "rubber stamp" of the decisions made by the machine.[18] The autonomy of the user can be preserved by keeping the "human-in-the-loop" or "human-on-the-loop" configurations in order to ensure that the employee remains the final authority with regard to personal and professional decisions through informed consent.[19,20]

Transparency and the "Black Box" Problem

One of the most significant obstacles in building ethical AI is the idea of a "black box," in which the reasoning process of a deep learning algorithm is impossible for the human user to understand.[12,21] A lack of transparency is not only difficult to achieve from a technological standpoint, but it is also a moral failing, and making decisions without the ability to explain an action to an unhappy party is impossible.[11,22]

Transparency in business AI systems is typically categorized into three types:

1. Traceability: is the capacity to trace the data sources, training data, and/or the system deployment history of the AI system.[23]
2. Explainability (XAI): is the technical capacity to express the rationale for achieving a particular outcome.[24]
3. Communication: is the organizational responsibility to communicate to the user that he or she is being controlled by the AI system, and also the availability of avenues of recourse to the user.[25]

These levels of abstraction constitute the foundation on which trust cannot be attained. The feeling of explainability must be created among the stakeholders so that the professionals can query and modify the results produced by the AI system, while keeping the AI system in the proper perspective as a tool of development of humankind and not vice versa.[8,26]

Accountability and the Responsibility Gap

Responsibility gaps are those situations wherein the machine is capable of creating a problem, but no individual, and no legal person or persons, can be held accountable for it.[13,27] The traditional corporate management model is heavily based on responsibility chains, but this is altered with AI, which can lead to the phenomenon of "distributed agency," wherein the final cause of the outcome is the intricate relationship between data scientists, data vendors, and the autonomous element itself.[15,28]

Nevertheless, it cannot be solelyheld responsible for ethical behavior, since it cannot meet the moral prerequisites for personhood, which entail having empathy, consciousness, and free agency.[21] It is not an area where there is delegation of accountability. Instead, there is a shared accountability between man and organization through modified processes.[29,30] This entails the incorporation of "ethics committees" and "chief AI officers" who will be responsible for its performance over its lifecycle.[31,32]

Justice, Fairness, and Non-Maleficence

The principles of justice and non-maleficence state that AI systems should do no harm and that the benefits and burdens of these systems should be equitably shared.[8,33] In a business context, this means that there should be a strict prevention of discrimination. The concept of justice in AI can often be divided into:

- Procedural justice: the treatment of the algorithms is carried out in a fair and consistent manner for all.[34]
- Distributive justice: the observation of whether the outcomes of AI systems are putting some demographic groups at a disadvantage.[35]

Non-maleficence also includes technical robustness and safety; that is, an ethical AI system must be immune to adversarial attacks and "data poisoning," which could cause it to behave in a harmful and erratic way in a business setting.[36,37] A brief comparison of ethical principles in human vs. AI decision-making is shown in Table 1.

The Digital Frontier: CBDCs and the Future of Money

This is because an AI system has the ability to learn on the basis of historical data, and it is expected that there will be underlying prejudices toward different groups of people. It is not only the case that AI systems are reflective of existing social barriers in society; it is also the case that AI systems are likely to proliferate, exacerbate, and codify prejudices, biases, etc. Thus, the

Table 1 | Comparison of ethical principles in human versus AI decision-making

Principle	Category	Business Application & Focus
Justice	Procedural	Ensuring algorithms follow fair and consistent rules for every user.
	Distributive	Monitoring outcomes to ensure no demographic group is disproportionately disadvantaged.
Non-maleficence	Technical robustness	Building systems resilient to adversarial attacks and "data poisoning."
	Safety	Preventing harmful or erratic behavior to ensure the system does no harm.

issue associated with AI systems, as characterized by "bias in, bias out," is likely to emerge as a major threat in the case of AI-based business processes.[26,38,39]

Mechanisms of Bias: Data Poisoning and Proxy Variables

Bias enters the AI lifecycle through several channels. Existing bias is based on discriminations that were created in the data that was collected and fed to the AI model (existing hiring patterns). Technical bias, on the other hand, is based on constraints that are associated with the programs and machines, while emergent bias is based on bias that emerges from user experience and possible inherent user bias.[26,40]

Another harmful problem relates to a phenomenon called "proxy variables." Sometimes, information like race, gender, or religion may be filtered out of a data pool. Nevertheless, AI systems will still be able to deal with correlated information like zip codes, shopping habits, or educational history, which will still be used as a proxy for the original information.[29,41] Because of this data, it will be impossible to identify the bias in AI systems' decisions.[42,43]

Recruitment and Human Resources: The Digital Gatekeeper

Hiring software using AI relies on NLP and machine learning to filter candidates through resume screening, affective states in video interviews, and psychometric testing, among other channels.[11,44] Biases in the software have been quick to penalize certain sections of society. For example, an AI trained on the resumes of "successful" employees over the past 10 years would likely favor traits that are common in the representative majority, such as playing certain sports or attending certain universities, while penalizing "non-traditional" career paths.[17]

The biases in R&S result in the creation of a digital gatekeeper that is capable of systematically rejecting qualified talent even before any human recruiter gets a glimpse of the candidate profile. It was not only justice that was at stake, but corporate performance is also being negatively impacted by the lack of diversity, and innovation is being stifled.[24,31] Active mitigation strategies involve a cross-functional approach where AI developers work with HR professionals to ensure that the training data is "de-biased" and monitoring of selection ratios is done.[45,46] This is shown in Figure 1.

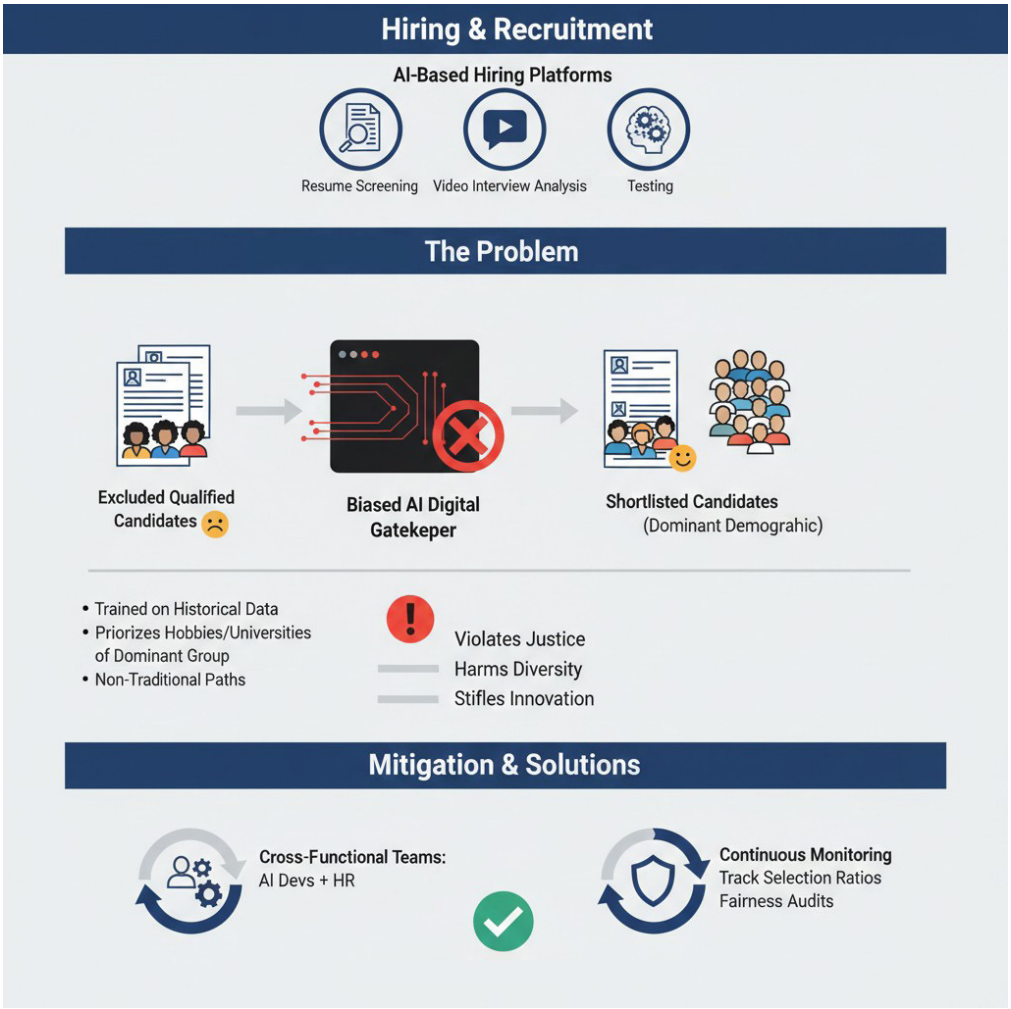

Fig 1 | AI bias in hiring and recruitment

To move beyond abstract principles, organizations should implement a RACI (Responsible, Accountable, Consulted, Informed) matrix across the AI lifecycle:

- Data Preparation: Data Engineers (R), Chief Data Officer (A). *Artifact:* Data Sheets for Datasets.
- Model Development: Data Scientists (R), Lead AI Architect (A). *Artifact:* Model Cards.
- Deployment/Monitoring: Ops Team (R), Business Unit Head (A). *Artifact:* Live Audit Logs.

Governance effectiveness is measured via Fairness Gap Thresholds (e.g., ensuring disparate impact ratio stays within 0.8–1.2) and Drift Detection Lag (time elapsed between statistical drift and automated alert).

Financial Services and Lending: Socioeconomic Impact

In finance, through credit decisions facilitated by an algorithm, there is significant scope for socioeconomic groups to be incorporated based on non-traditional data points.[13,18] In other words, if an algorithm is employed that is based on historical repayment trends from a period of systemic redlining, there is significant scope for individuals from socioeconomically marginalized groups to be deemed financially less creditworthy than they actually are.

There is still significant scope for socioeconomic groups to become excluded from the market. If an algorithm is based on data from previous repayment periods, there is significant scope for individuals from socioeconomically marginalized groups to be deemed financially less creditworthy than they actually are. Moreover, the results from the insurance and online business practices, in terms of the use of the aforementioned types of dynamic pricing strategies, may lead to what could be defined as "price discrimination."[43,47] In "price discrimination," the less fortunate members of our society can pay inflated prices or insurance rates due to the data that the algorithm assigns to them in relation to their likelihood of mishap or trouble, which creates a vicious cycle wherein the least fortunate members of our society can further be taken advantage of through the very means and processes that hope to increase efficiency in our business and economic practices, which can be seen in Figure 2.[48,49]

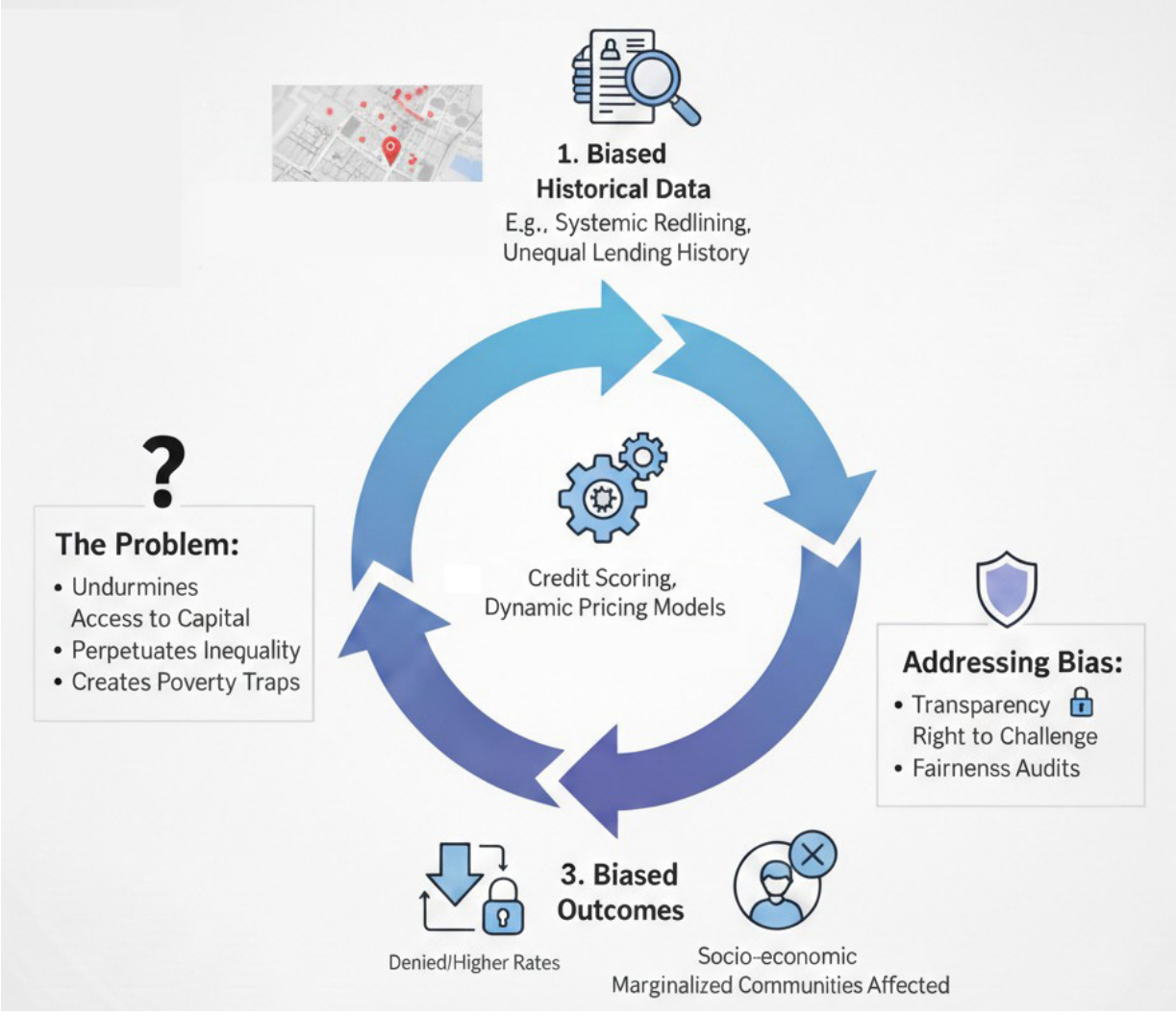

Fig 2 | The feedback loop of algorithmic bias

DOI: https://doi.org/10.70389/PJBM.100014 | Premier Journal of Business and Management 2026;7:100014

Barriers to Action: The Political Economy of Climate Change

Instead of being a one-trick solution that uses technology, it requires an ongoing commitment on the part of organizations to address AI ethics risks throughout their entire lifecycle, from start to end-of-life reuse or recycling.[11,15,19]

Technical Mitigation

The HR and technological solutions to the detection and neutralization of bias within the algorithmic process are categorized at three stages:

1. Pre-processing: The "vector space correction" or "weighting" method is applied to handle the unbalanced data set before entering the model.[19,29]
2. In-processing: At this stage, mathematical constraints are incorporated within the learning process. The aim is to incorporate "fairness constraints" within the learning process itself.[19]
3. Post-processing: Modifying the end results to ensure equal results for all groups.[34,50,51]

In addition to these, "explainable AI" must be incorporated. "Explainable AI" will allow us to explain to stakeholders "local explanations" (why a particular individual was denied) and "global explanations" (how it works in general), which will allow humans to monitor an algorithm if it is using flawed and discriminatory reasoning in its decisions.[24,25]

Selecting a fairness metric is a governance decision, not just a technical one.

- Demographic Parity: Ensures equal outcomes across groups, though it may sacrifice predictive accuracy.
- Equalized Odds: Requires equal true positive and false positive rates; essential for high-stakes hiring or lending.
- Predictive Parity: Focuses on the precision of the model. The choice between these often results in "Fairness Trade-offs," where optimizing for one metric mathematically precludes another. Organizations must document the rationale for their chosen metric based on sectoral legal constraints (GDPR Article 22).

Governance and Auditing

Besides code-level factors, there is also a need to create a cultural paradigm shift to "ethics-by-design." The latter is based on the notion that one can think about ethics as a set of fundamental requirements, such as security and performance metrics.[22–24] Some of the organizational factors include:

- Cross-functional ethics committees: data scientists leading reviews along with legal, HR, and outside ethicists on high-stakes AI projects.[15,41]
- Impact assessments: conducting binding "algorithmic impact assessments" before deployment to consider potential risks to basic rights.
- Human-in-the-loop: establishing clear processes to ensure that human experts review high-impact decisions to ensure that the algorithm is not considered a "sovereign" decision tool, but rather a tool to assist.[8,24]

Auditing and Continuous Monitoring

However, static ethics guidance will not suffice, and "continuous 'algorithmic auditing'" has to be done. Moreover, "internal auditing is needed "to flag 'concept drifts,' that is, when the model's behavior or bias changes over time."[15] "A gold standard" of corporate auditing has come to be recognized as "an independent third-party audit… to guarantee that the system under consideration not only complies with internal ethics rules but also 'outside' the current 'law'."[21,38] It is also necessary that "outcome fairness" and "procedural fairness" are ensured in the auditing efforts.[35,47]

The Fiscal-Monetary Nexus: A New Era of Coordination

The environment within which artificial intelligence is regulated is ever-evolving from being non-regulated, with ethics surrounding the practice, to being fully regulated with comprehensive AI regulations. This is what the "guardrails of trust" aim to do by fully regulating the practice within international businesses and governance.

The EU AI Act (2024)

The European Union AI Act, which was officially enacted in mid-2024, is considered to be the world's first significant horizontal rule on AI.[23] The AI Act has been considered to be a regime that introduces a tiered and risk-based classification system to determine the nature of compliance required for a particular AI system:

- Unacceptable risk: The legislation clearly prohibits the use of AI applications that are considered a risk in terms of safety, livelihood, and basic human rights. This could include the social scorekeeping of the government, subliminal manipulation techniques, and some types of predictive policing techniques.
- High risk: Infrastructures, educational settings, law enforcement agencies, or employment settings that require resume analysis require the highest level of specification in terms of requirements. When a business uses this system, the assessment of the impact in terms of basic human rights is already being addressed, along with the recording of the technical documentation that requires a large level of human intervention.[21,40]
- Limited and minimal risk: Chatbots/AI content generation, which is considered a lower-tier system or deepfakes, requires primary-level transparency, which is simply informing the user that they are interacting with an AI, but minimal risk applications, such as spam filters, are left unchecked.[40]

The classification of risks is depicted in Figure 3 below.

Regulatory Challenges

Moreover, it should be noted that the implementation of the AI Act would not be an isolated process; hence, it would not be implemented in isolation. Rather, it would be implemented in a parallel and concurrent manner with the GDPR, and in such a case, a complex situation termed "regulatory double bind" for businesses would arise.[34] For example, in accordance with the AI Act, it is recommended to train high-risk AI systems with "high-quality data" in order to prevent

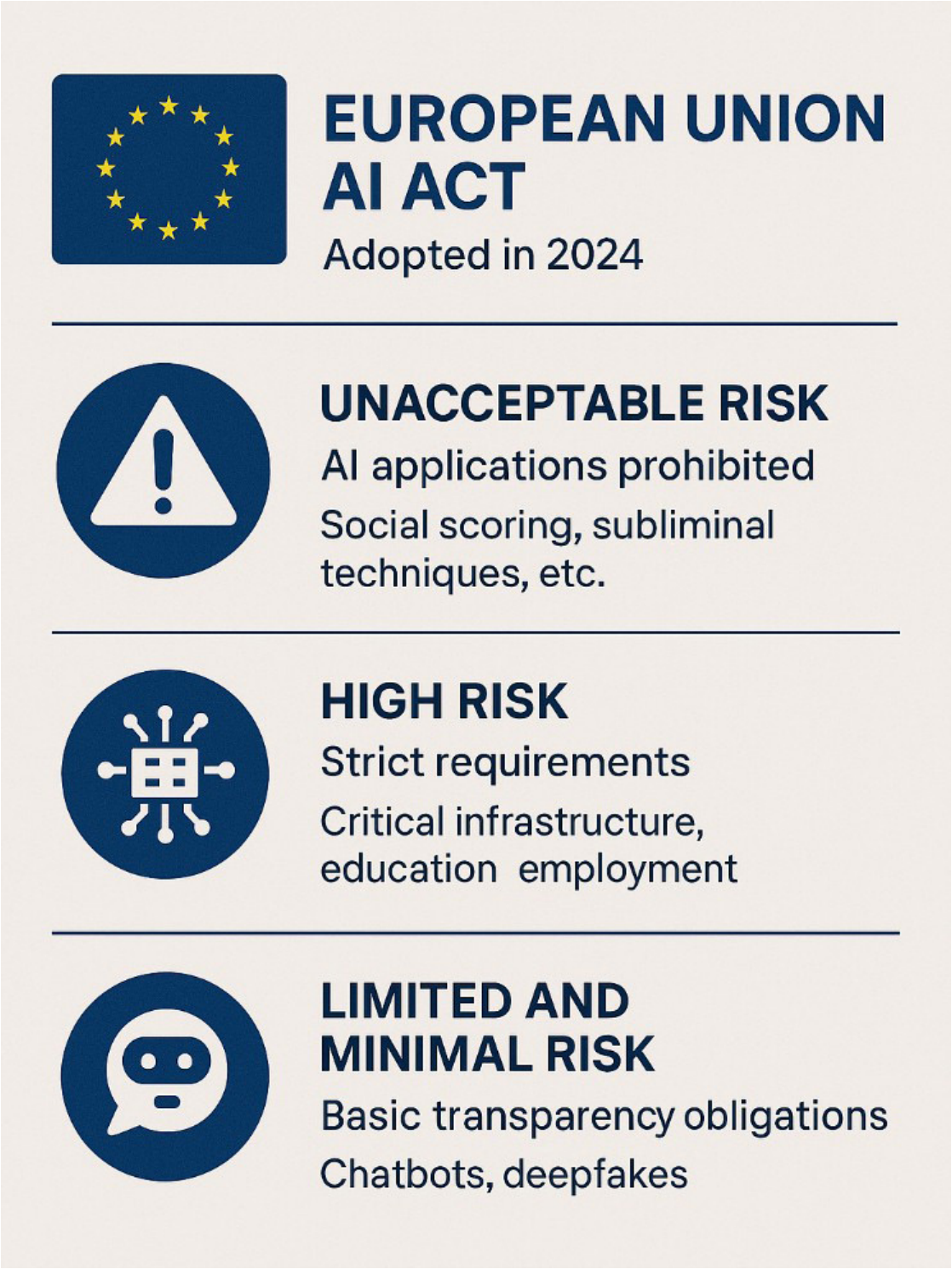

Fig 3 | European Union AI act and risks

bias in AI systems. However, at the same time, in accordance with the GDPR, "data minimization" must be adopted, and hence, certain vital data entities required for debiasing must not be obtained.[34]

Moreover, it must be noted that the GDPR's "right to explanation of automated decision making," as per Article 22, often conflicts with business interests in terms of IP and business secrets.[34] The nuances of the above-mentioned regulations must be addressed through a sophisticated approach in which the interface of regulation and law converges with the technical and design aspects of AI systems. Failure to do so may result in "innovation paralysis," in which the business may not choose to use AI due to the fear of costly regulations and subsequent fines from various regulatory bodies, estimated in millions of euros.[34,41]

DOI: https://doi.org/10.70389/PJBM.100014 | Premier Journal of Business and Management 2026;7:100014

While the EU AI Act provides a risk-based classification, global alignment is found in:

- NIST AI RMF 1.0 (2023): Provides a voluntary framework for managing AI risks in the U.S. context.
- ISO/IEC 42001:2023: The international standard for an AI Management System (AIMS).
- OECD AI Principles: The foundation for international cooperation on trustworthy AI.
- Canada's AIDA & the UK's Pro-Innovation Approach: Representing emerging statutory and sectoral-led models, respectively.

Philosophical Foundations of Governance

The laws of modern-day AI are not completely the offspring of legislative acts but are, in fact, significantly impacted by the classical ethical theories, which are then implemented with some necessary modifications of the modern-day application of these theories. The conflict between deontology and utilitarianism is an integral part of law and legislation.[52,53]

- Bentham and Mill (utilitarianism): The motivation for AI efficiency and the promise of GDP growth of unimaginable proportions can be said to be purely utilitarian, where the aim is to accomplish the "greatest good for the greatest number" of people. However, the AI Act stresses the importance of ensuring that the few are not harmed, which becomes a constraint on the principle of utilitarianism, so that the majority do not celebrate at the cost of the systemic injustices committed against the few.[53]
- Theory of ethics by Kant: The autonomy and transparency principles of Kant also comprise several elements of data transparency, where the person is considered as the end and not as a data point to be used for achieving a particular outcome. There is a significant principle of autonomy and data transparency in providing "informed consent," because of human dignity and not due to "technological efficiency so lost in the process."[54,55]

With the international markets becoming more and more interconnected, the philosophical premises are working to bridge the gap between the different international frameworks, thereby moving toward creating a cohesive framework across the globe on the subject of responsible AI.[55]

Practical Implementation Artifacts

The AI Audit Maturity Model provides specific thresholds for "drift detection lag" and "disparate impact"—moving beyond vague ethical statements to measurable KPIs (Table 2). Second, the Accountability Map aligns with the EU AI Act, assigning specific duties to the C-suite for risk classification and to functional owners for input data quality. Finally, the Fairness Decision Tree offers a compliance-first approach to metric selection, ensuring that technical teams choose statistical measures that satisfy GDPR Article 22 requirements for transparency while balancing the inherent trade-offs between model precision and group equity.

Table 2 | Maturity levels of current AI deployment

Maturity Level	Metric/Threshold	Requirement for Compliance
1. Initial	Documentation	Ad-hoc notes; no centralized model registry
2. Defined	Disparate impact	Tracking the 80% rule (***p***-rate) for protected groups
3. Managed	Drift detection	Automated alerts for "Concept Drift" (Lag < 48 h)
4. Optimized	Audit trail	Full lineage from raw data to decision (GDPR Art. 22 ready)

Artifact A: AI Governance & Audit Maturity Model

To bridge the gap between theoretical principles and corporate execution, a tripartite governance framework is proposed.

This checklist allows managers to score their current deployment:

Artifact B: Provider–Deployer Accountability Map

The EU AI Act distinguishes between those who *build* the AI (Providers) and those who *use* it (Deployers).

- Board of Directors/C-Suite:
  - *Duty:* Ultimate accountability for "High-Risk" system classification and systemic risk mitigation.
- Provider (Internal Dev or Vendor):
  - *Duty:* Technical documentation, logging, and ensuring the "Human-in-the-loop" interface is functional.
- Deployer (Functional Owners—e.g., HR, Finance):
  - *Duty:* Quality of input data, monitoring for "Real-world" bias, and notifying the provider if the model behaves unexpectedly.

Artifact C: Fairness-Metric Selection Decision Tree

This logic flow helps managers choose the right metric based on legal constraints.

1. Is there a "Right to Explanation" requirement (GDPR Art. 22)?
   - *Yes:* Prioritize Local Interpretable Model-agnostic Explanations (LIME) or SHAP values.
2. Is the primary legal risk "Disparate Treatment"?
   - *Yes:* Use Anti-classification (remove protected attributes).
3. Is the primary legal risk "Disparate Impact"?
   - *Yes:* Use Demographic Parity or Equalized Odds to ensure outcomes are balanced across groups.
4. Are there Trade-offs?
   - *Note:* Accuracy vs. Fairness—if the fairness constraint reduces accuracy below [X% threshold], document the "justifiable business necessity" as per EU AI Act recitals.

Conclusion

There is a sense of tension in the literature on AI ethics, as, on one hand, there is a clear indication of idealism, based on theoretical analysis, in terms of the potential of AI ethics in corporations, as presented by Floridi et al.[35] and Dignum,[26,51] who emphasize the need for "value-sensitive design" where ethical principles are the primary constraints for engineers, as AI is not neutral but value-saturated. On the other hand, based on the

empirical analysis of AI adoption in corporations,[3,4,11] there is a sense of applying ethics retrospectively as a way of "ethicswashing" for public relations purposes.

Another important difference among the reviewed studies is related to the "Accountability vs. Efficiency" trade-off. In this context, Char et al.[8] and Kaminski[21] argue that the incorporation of human oversight (HITL) will always result in decision-making delay, which might undermine the speed advantage offered by AI systems. However, Novelli et al.[7] argue that without this "friction," the speed of harm mentioned in the introduction will be uncontrollable. It appears that there is an emerging consensus in the literature[13,27,28] that the most pressing legal issue is indeed the "responsibility gap," and some authors even propose new legal definitions of "algorithmic personhood."

This discussion emphasizes bias with a clear distinction between technical and socio-technical methods. Albaroudi et al.[19] emphasize the mathematical possibility of removing bias (purely technical). Mittelstadt et al.[34] and Martin,[29] on the other hand, emphasize that the idea of objective data is impossible. Rather, these views emphasize that the pursuit of a "perfectly fair" algorithm is perhaps a category mistake in view of its social constructiveness, which has nothing to do with purely mathematical ideas. This, in essence, requires an interdisciplinary partnership between data and social scientists with the aim of identifying what "fairness" actually looks and feels like in a particular business setup.

One of the greatest challenges that organizations face in the 21st century is how to satisfactorily blend AI with ethical business decision-making processes. Although in this paper, it has been proven that AI technology has a tremendous potential to maximize business and economic growth, it has also been proven that AI technology is a destroyer of corporate accountability. For example, a "black box" technology, the use of proxy variables, and a "responsibility gap" in AI technology are not just mere teething problems in the development of technology; they are major obstacles to the creation of an ethical economy.

Going forward, the paradigm needs to shift from "move fast and break things" to a paradigm of "responsible innovation." This needs a threefold approach. The industry needs to move on from voluntary compliance to standardized independent algorithmic auditing, similar to financial auditing. Policymakers need to remove friction between data protection - GDPR and AI quality standards - AI ACT, so that a compliance roadmap is laid out for businesses. Human-centric re-alignment includes a plea to re-invest in human agency so that AI is again perceived as an augmentative approach through which professional judgment is strengthened, and not as an autonomous and independent force that seeks to replace it.

The objective, therefore, of an "ethic" in AI development is not to impede the advancement of technology but to ensure that the advancement of technology occurs in a manner that ensures that the values of humanity are reflected. In creating the "transparency," "justice," and "accountability" aspects of an algorithm as the "building blocks" of AI development, the business world has an opportunity to leverage the power of technology to create not only a "more efficient" world but a "more equitable" world as well. The shift to move from "ethic guidelines" to "legal frameworks" signaled a significant shift in this direction, suggesting that the days of "unguided" AI development are numbered.

References

1 Alzubi J, et al. Artificial intelligence and its context in the Fourth Industrial Revolution. *Front Artif Intell*. 2018;1:12.
2 Schwaeke M, et al. Reshaping human-technology interactions. *Front Hum Dyn*. 2021;3:669041.
3 Ransbotham S, et al. Reshaping business with artificial intelligence. *MIT Sloan Manag Rev*. 2017;59(1):1–10.
4 Bitkina O, et al. Failure rates in AI adoption: a systematic review. *Int J Inf Manage*. 2020;52:102068.
5 Glikson E, Woolley AW. Trust in artificial intelligence: a multiscale feedback loop. *Acad Manag Ann*. 2020;14(2):627–678. https://doi.org/10.5465/annals.2018.0057
6 Cheong BC. Transparency and accountability in AI systems. *Front Hum Dyn*. 2024;6:1421273. https://doi.org/10.3389/fhumd.2024.1421273
7 Novelli C, et al. Accountability in AI governance. *Sci Eng Ethics*. 2023;29(3):18.
8 Char DS, Shah NH, Magnus D. Ethical and legal considerations in healthcare AI. *N Engl J Med*. 2018;378(11):981–983. https://doi.org/10.1056/NEJMp1714229
9 IEEE Global Initiative. Ethically Aligned Design: A Vision for Prioritizing Human Well-being. IEEE; 2019.
10 Chen Y. Artificial intelligence (AI) and ethical concerns: a review. *Cogent Bus Manag*. 2024;11(1):2303001.
11 Grove M, et al. Artificial intelligence in corporate governance. *Corp Gov (Bradford)*. 2021;21(7):1314–1330.
12 Timmons S, et al. The black box problem in corporate decision making. *Bus Ethics Q*. 2023;33(2):145–167.
13 Gunz S, Thorne L. Responsibility gaps in the age of technology. *J Bus Ethics*. 2020;161(2):153–170. https://doi.org/10.1007/s10551-019-04404-4
14 Jarrahi MH. AI and the future of work: a human-centered approach. *Bus Horiz*. 2018;61(4):577–587. https://doi.org/10.1016/j.bushor.2018.03.007
15 Lehner OM, Ittonen K, Silvola H, Ström E, Wührleitner A. Artificial intelligence based decision-making in accounting and auditing. *Account Audit Account J*. 2022;35(9):109–135. https://doi.org/10.1108/AAAJ-09-2020-4934
16 Wang Y, et al. Responsible AI practices in corporate settings. *Proceedings of HICSS*. 2018.
17 Bankins S. The ethical use of artificial intelligence in human resource management. *Ethics Inf Technol*. 2021;23(4):841–854. https://doi.org/10.1007/s10676-021-09619-6
18 Algorithmic Bias. Financial Inclusion, and Gender. *Women's World Banking*; 2021.
19 Albaroudi E, Mansouri T, Alameer A. A comprehensive review of AI techniques for addressing algorithmic bias. *AI*. 2024;5(1):383–404. https://doi.org/10.3390/ai5010019
20 Felzmann H, Fosch-Villaronga E, Lutz C, Tamò-Larrieux A. Towards transparency by design for artificial intelligence. *Sci Eng Ethics*. 2020;26(6):3333–3361. https://doi.org/10.1007/s11948-020-00276-4
21 Kaminski ME. The EU AI act and small and medium-sized enterprises. MDPI. 2024;5(1):40.
22 European Parliament. *Interplay between the AI Act and the EU digital legislative framework*. European Parliamentary Research Service. 2024.
23 The Artificial Intelligence Act. OJEU. 2024;L 2024/1689.
24 Wagner G. Algorithmic regulation. *Oxf Rev Econ Policy*. 2018;34(3):397–421.
25 Floridi L. The global challenges of AI. *Ethics Inf Technol*. 2019;21(4):249–251.
26 Dignum V. Ethics in artificial intelligence. *Ethics Inf Technol*. 2018;20(1):1–3. https://doi.org/10.1007/s10676-018-9450-z
27 Leyer M, Schneider S. Decision-making with AI. *J Bus Res*. 2021;124:441–452.

28 McGregor S. Accountability in AI systems. *Bus Technol Rev*. 2020;32(1):45–60.
29 Martin K. Ethical implications of AI algorithms. *J Bus Ethics*. 2019;155(3):835–850. https://doi.org/10.1007/s10551-018-3921-3
30 Lombardi R, et al. Big data in accounting and auditing. *Int J Account Inf Syst*. 2015;16:1–12.
31 Bedford D, et al. Cognitive states in ethical decision-making. *Manag Account Res*. 2019;45:101062.
32 Orlitzky M. Ethical decision-making in organizations. *J Bus Ethics*. 2016;135(1):1–10.
33 Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. *Nat Mach Intell*. 2019;1(9):389–399. https://doi.org/10.1038/s42256-019-0088-2
34 Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L. The ethics of algorithms: mapping the debate. *Big Data Soc*. 2016;3(2):2053951716679679. https://doi.org/10.1177/2053951716679679
35 Floridi L, Cowls J, Beltrametti M, et al. AI4People—An ethical framework for a good AI society. *Minds Mach*. 2018;28(4):689–707. https://doi.org/10.1007/s11023-018-9482-5
36 King TC, et al. Artificial intelligence in healthcare: review and recommendations. *Health Technol*. 2020;10(1):1–18.
37 Crootof R. The killer robots are here: legal and policy implications. *Cardozo Law Rev*. 2018;39:1837.
38 Bankins S. The ethical use of artificial intelligence in HRM. *Ethics Inf Technol*. 2021;23(4):841–854. https://doi.org/10.1007/s10676-021-09619-6
39 Binns A. Understanding the impact of AI on business ethics. *J Bus Ethics*. 2018;149(4):935–950.
40 Gani A, Darmawan D. Ethics in AI-based managerial decision making. *J Soc Sci Stud*. 2022;2(1):147–152.
41 Verhezen P. AI in corporate governance: boards of directors. *J Bus Strategy*. 2020;41(4):1–12.
42 Wang Y, et al. AI adoption in regulatory compliance. Preprints. 2023.
43 Hazaimeh A, et al. AI-driven Financial Transparency. MDPI; 2021.
44 Wirtz BW, Weyerer JC, Geyer C. Artificial intelligence and the public sector. *Int J Public Adm*. 2019;42(7):596–615. https://doi.org/10.1080/01900692.2018.1498103
45 Cohen J. Statistical Power Analysis for the Behavioral Sciences. Routledge; 1988.
46 Krejcie RV, Morgan DW. Determining sample size for research activities. *Educ Psychol Meas*. 1970;30(3):607–610. https://doi.org/10.1177/001316447003000308
47 Hagendorff T. The ethics of AI ethics: an evaluation of guidelines. *Minds Mach*. 2020;30(1):99–120. https://doi.org/10.1007/s11023-020-09517-8
48 Siau K, Wang W. Ethical aspects of artificial intelligence and robotics. *J Database Manag*. 2020;31(2):74–92. https://doi.org/10.4018/JDM.2020040105
49 Bentham J. An introduction to the principles of morals and legislation. Clarendon Press; 1789. https://doi.org/10.1093/oseo/instance.00077240
50 Mill JS. Utilitarianism. Oxford University Press; 1861.
51 Dignum V. Responsible Artificial Intelligence. Springer Nature; 2019. https://doi.org/10.1007/978-3-030-30371-6
52 McElheran K, et al. AI Adoption and Inequality. NBER; 2024.
53 Kazim E, Koshiyama AS. A high-level overview of AI ethics. *Patterns (NY)*. 2021; 2(9):100314. https://doi.org/10.1016/j.patter.2021.100314
54 Hauer T. Ethical principles for AI explainability. *Front Comput Sci*. 2022;4:845132.
55 Resnik DB, Elliott KC. Ethics in AI-driven scientific research. *Account Res*. 2019;26(2):1–15.