

An Evaluation of the Impact of Ethical Dilemmas for Artificial Intelligence (AI) Integrated in Business Strategies

Dr. Aaron M. Wester

Alliant International University, 10455 Pomerado Rd.
San Diego, CA 92131, USA.

Dr. Rachna Kumar

Alliant International University, 10455 Pomerado Rd.
San Diego, CA 92131, USA.

Dr. Shams Al Ajrawi

Alliant International University, 10455 Pomerado Rd.
San Diego, CA 92131, USA.

Dr. Saba Senses-Ozyurt

Alliant International University, 10455 Pomerado Rd.
San Diego, CA 92131, USA.

Dr. Cory Scott

Alliant International University, 10455 Pomerado Rd.
San Diego, CA 92131, USA.

Tushar Madhukar Survase

Alliant International University, 10455 Pomerado Rd.
San Diego, CA 92131, USA.

Abstract

With the rapid evolution of Artificial Intelligence (AI), businesses have included generative AI in their operations and harnessed multiple benefits such as workflow efficiency, reduced cost through task automation, human error reduction, rapid product and service development, and improved data-driven decision-making. Although numerous benefits may be achieved through generative AI inclusion, the problem is AI integration in business strategies across organizational departments includes significant consequential ethical challenges. These points of opposition may disrupt and hinder competitive advantage and damage customer acquisition as well as lead to erosion in loyalty, trust, and brand building. Ethical dilemmas in AI integration include data bias that tend to skew results discriminately, misuse of AI-driven customer targeted data, complex accountability determinacy, as well as reliability and safety risks in rapid deployments. In this paper we evaluate the ethical risks and dilemmas for AI integrated business strategies, such as AI-driven customer targeted data misuse, excessive profiling and surveillance, manipulative advertising, and challenges in assigning accountability, and provide recommendations to manage potential ethical risks associated with AI in business.

Key words: AI, ethical dilemma, business strategies, generative, challenges, consequences, accountability

1. Introduction

Artificial Intelligence (AI) is rapidly evolving across all vertical industries. Through generative AI inclusion, multiple benefits are being harnessed by organizations such as workflow efficiency, reduced cost through task automation, human error reduction, rapid product and service development, and improved data-driven decision-making (Daniels, 2023). These benefits are actualized through “harnessing machine learning (ML) and deep learning (DL) techniques to glean insights from data, recognize patterns, and make decisions with limited human intervention” (Kumar et al., 2020).

With the expectation of remaining relevant and competitive, many organizations are pushing for rapid deployments with AI integrated. The Future of Jobs Report found across 27 industry clusters, nearly 75% of all surveyed companies (n = 803) expect to adopt artificial intelligence (World Economic Forum, 2023). With the rise of AI adoption and the speed at which generative technology is quickly evolving, the risk of ethical friction and negatively impactful consequences rises as well. The pressure on organizations to anticipate and plan for all potential contingencies with AI integration is tremendous, and impossible.

Many businesses have processes in place to limit their exposure to potential liability prior to services or products release (Pazzanese, 2020). With AI, these organizations are in uncharted territory given how rapidly machine

intelligence is evolving. However, Schubmehl, Research Director, Cognitive/Artificial Intelligence Systems at IDC pointed out "... AI will be the disrupting influence changing entire industries over the next decade." (D'Aquila & Shirer, 2019). While there is a substantial lack of government regulation, monitoring and oversight on AI inclusion (Pazzanese, 2020), organizations are willing to confront the risks and potential backlash to achieve competitive advantage and actualized monetary gains based on promising and potentially lucrative projections. The risks may be more significant and consequential than business leaders realize as Sandel warned, "AI not only replicates human biases, it confers on these biases a kind of scientific credibility." (Pazzanese, 2020).

2. Problem statement

Although numerous benefits may be achieved through generative AI inclusion, the problem is AI integration in business strategies across organizational departments includes significant consequential ethical challenges. These points of opposition may disrupt and hinder competitive advantage and damage customer acquisition as well as lead to erosion in loyalty, trust, and brand building. Ethical dilemmas in AI integration include data bias that tend to skew results discriminately, misuse of AI-driven customer targeted data, complex accountability determinacy, as well as reliability and safety risks in rapid deployments.

3. Benefits of business departments integrating AI into inclusive strategies

Across industries and organizations, there are several emerging benefits influencing businesses to eagerly incorporate AI. Types of AI algorithms which may enhance departmental business strategies include "convolutional neural networks (CNNs), generative adversarial networks (GANs), natural language processing (NLP), neural style transfer, motion detection, speech recognition, video summarization, and emotion recognition" (Huang, et al., 2023) among others. AI First and foremost is the opportunity to achieve maximal revenue gains. McKinsey Global Institute found that generative AI "has the potential to generate value equivalent to \$2.6 trillion to \$4.4 trillion in global corporate profits annually." (Chui & Yee, 2023). There are intradepartmental advantages that indirectly tie into revenues through scalable and achievable internal efficiencies as well.

Within customer relationship management (CRM), even after several years of availability, many of the primary tools and features have been underutilized or not used at all. In 2022, a Sales Mastery's Sales Performance Scorecard survey showed 15% of sales managers (n=332) identified their organizations were not utilizing CRM while 42% were only partially using CRM to retain customer or prospect data (Dickie, et al., 2022). These deficiencies are examples of why AI integration may provide needed improvements to the sales cycle. Key stages within CRM that may be optimized with AI tooling include new customer outreach, prospect to customer lead generation, new customer acquisition, goal-based conversions, as well as customer retention and loyalty optimization. Through autonomous generative integration, CRM teams can focus on more strategic initiatives such as expanding competitive advantage through new Statement of Work (SOW) service offerings, developing relationships with key clients, and signing critical service contracts with those key clients.

Marketing department leaders are continuously re-evaluating optimization tactics to improve marketing mix factors ranging from product and price to place and promotion. AI has the capability to improve the customer experience through expanded data collection and machine modeled interpretation methods, distinguish and optimize primary factors that produce enhanced results for targeted marketing campaigns, and generate influential creative content through "text-to-image synthesis" (Quan, et al., 2023) based on extracted user experience data and machine learned customer behaviors. As a result, businesses can maximize revenue through contextual generative sales intelligence and targeted recommendations aggregated across aggregated prospect and sensor data (Dickie, et al., 2022).

In product development, AI frameworks have the procedural capacity to assist in information collection, dimensionalized data classifications, Web API service provisioning optimization, generative ideation, coding optimization, and testing. Products may be developed and deployed to production faster with a more intrinsic focus on customer needs. Further, in operations management, AI may help automate tasks such as inventory management, supply chain optimization, and quality control. Businesses may achieve reduced costs, increased quality, optimized manufacturing, and more granular dimensionalized factors for quality control and performance analysis (Greenfield, 2020).

4. The impact of ethical dilemmas

4.1 Data bias skews results discriminately

As discussed in the above sections, the volume of data has been ever increasing in every context - business operations, corporate regulations and reporting, governmental requirements, schools and social landscapes, even personal everyday management of life. Data is the natural definition and identification of any animate or inanimate existence. The first and most basic challenge which was faced in the data era was data storage and this was addressed easily.

The second challenge was one of analyzing large volumes of data and big data. Advances in computing power, statistical applications, and powerful algorithmic programs allowed data to be analyzed and fed to relevant and responsible data owners. The third wave in the data era was about guiding data owners, generators, and decision makers in utilizing what the data analysis told them and what the analyzed data patterns indicated. Interestingly, this third wave has faced several issues. The inclusion and integration of AI techniques in data analysis has, on the one hand made adoption very critical and on the other hand has made results and analysis processes difficult to understand for most data managers, owners, generators, and decision makers (Chen & Zhang, 2014).

4.1.1 Data analysis is only as good as the data it works with

Data is very often selectively collected, bringing in bias, sometimes intentionally, sometimes due to known or unknown prejudices of the data designer, and often due to accepted norms and perspectives of society. In addition to the data being selectively skewed, the analysis algorithm can unintentionally be programmed to discover patterns and results in the data which ends up missing important irregularities in population segments, thus reinforcing discrimination (Barocas & Selbst, 2016). Critiques of data bias leading to skewed results are most often related to discriminately understanding outcomes for population groups based on race, ethnicity, culture, age, gender, social strata, and sexual orientation. It is also important to note that often algorithms used to analyze data can be developed in ways which lead to systematic bias in results. This is referred to as overfitting or underfitting the algorithm wherein the algorithm is trained to analyze such that it either ignores certain features of the data or tends to overweigh the previously seen data examples.

4.1.2 Negative impacts of data bias skewing results discriminately

UNESCO identified AI systems embed biases based on the data used by the AI analysis systems and algorithms (UNESCO, 2023). UNESCO recognized that AI-based systems have the potential to analyze big data and beneficially impact critical social, global, and individual problems such as healthcare, climate challenges, mental and behavioral well-being, business efficiencies, and workplace economies to name a few. But risks of skewed results due to data bias can have negative impacts on social injustice and inequalities, prejudices to disadvantaged groups, and to already marginalized groups.

A study of Facebook's algorithm which is AI-based and decides which Facebook users will be shown an ad discovered that the algorithm showed ads discriminately based on gender and ethnicity (Hao, 2019). They found that ads for teachers, nurses and secretaries were shown to many more women for example, ads of home sales were shown to more white users, and ads for rentals, taxi drivers and janitors were shown to more minorities. They describe that an AI-based algorithm in Facebook allows advertisers to maximize the number of views and clicks an ad gets (among a few other criteria). These lead to higher business goals and higher customer acquisition; these are not driven by intentionally discriminating based on gender, ethnicity, or race. However, an AI algorithm which has been trained to decide based on the business goal criteria, automatically reinforces Facebook user engagement, and shows more women ads which experiences higher percentage of clicks from this segment and thus systematically excludes the other gender from such employment opportunities.

An algorithm used in a separate mobile device application, FaceApp, utilized a neural network model which exhibited unintended racial discrimination. The app intended to transform a face through AI to make an individual look older or change a facial expression to portray a different emotion than displayed in photos uploaded by users. However, when darker skin toned individuals, such as an African American, used the app, the AI transformed results included significant biased alterations such as changing to substantially lighter skin tones and European influenced facial features such as nose alterations. Snapchat faced similar discriminatory concerns for AI generative facial alterations through its Bob Marley filter which transformed faces to exhibit blackface elements (Morse, 2017). Blackface is a term referring to a historically offensive and derogatory practice where non-black individuals would wear theatrical makeup to portray exaggerated demeaning stereotypes of African Americans. These are examples of unintended consequences of irresponsibly using limited range data in AI training models which in turn falsely normalizes and "define[s] the algorithm's standard..." (Varsha, 2023).

In sum, it will be important to institute policy and better monitoring of data collection designs, data analysis programs, and AI-based algorithms so that they do not skew results and discriminate.

4.2 AI-driven customer targeted data misuse

AI gives rise to numerous data privacy issues. The escalation of privacy abuse has become a significant worry in a time when consumer daily lives are progressively becoming more digitized. The term "privacy abuse" pertains to the inappropriate or unauthorized utilization of an individual's personal information, frequently carried out without their awareness or consent. Further, the potential for such misconduct has grown substantially with technological progress, notably due in part to the integration of AI (Koome, 2023).

One prominent privacy challenge within the scope of data management is data persistence, wherein information endures longer than the human subjects who generated it, facilitated by the increasingly economical nature of data storage (Pearce, 2021). This persistent nature raises questions about the long-term implications for individual privacy and the potential for unauthorized access or misuse without consent. Another facet is data repurposing, wherein data is used for purposes beyond their initial intent, presenting ethical dilemmas and the risk of unintended consequences. Additionally, the phenomenon of data spillovers underscores the inadvertent collection of information on individuals who are not the primary targets of data gathering, leading to unforeseen privacy infringements (Pearce, 2021). When photos or videos are recorded that unintentionally capture other individuals in the background and are published with their consent, data spillover occurs. Addressing these challenges is imperative to ensure the responsible and ethical use of AI technologies while safeguarding individuals' privacy rights.

4.2.1 Deepfakes

Deepfakes represent a concerning factor of AI misuse, where images or videos, accompanied by cloned audio voices, are fabricated to create a deceptive appearance of individuals engaging in actions or making statements that never occurred (Kite-Powell, 2023). Manipulated generative multimedia content is devised using sophisticated artificial intelligence techniques, adding a layer of realism that can easily mislead viewers. Deepfakes exploit the power of AI to create convincing yet entirely fictional scenarios, raising significant ethical and social concerns regarding the potential for misinformation, identity theft, and the erosion of trust in digital content.

The deceptive inauthentic nature of deep fakes underscores the need for heightened awareness and countermeasures to mitigate the adverse impacts of such AI-driven manipulations. Consider a situation where a misleading advisory video, seemingly originating from the head of the Centers for Disease Control and Prevention (CDC), gains widespread attention, advising against vaccination. In this hypothetical scenario, deceptive manipulation, through the utilization of advanced deepfake technologies, poses a significant risk to public health (Reed, 2023).

4.2.2 Autonomous weapons systems

“Autonomous weaponry is the third revolution in warfare, following gunpowder and nuclear arms”. (Lee K.-F., 2021). One of the most alarming misuses of AI involves the creation of autonomous weapons systems. These weapons have the capability to choose and involve targets without requiring human intervention. The concept of such weapons is not novel, but advancements in AI have significantly increased their practicality (Trager & Luca, 2022). There are numerous concerns surrounding autonomous weapons. They have the potential for criminal misuse, including the intentional targeting of civilians. Additionally, once deployed, recalling these weapons becomes challenging, if not impossible, leading to the risk of unintentional escalation in conflicts. Lastly, the question of accountability arises. In situations where a machine autonomously decides to take lethal action, determining responsibility becomes a complex issue. Therefore, a fresh perspective is needed on Lethal Autonomous Weapons Systems (LAWS) (Trager & Luca, 2022).

“Common sense and humanity must prevail—before it is too late” (Sharkey, 2020).

4.2.3 Profiling and surveillance

Issues surface with excessive profiling of individuals, eroding privacy, infringing on civil liberties, and risking data misuse by authorities and businesses alike. Facial recognition and other surveillance technologies also enable more precise discrimination, especially as law enforcement agencies across countries make misinformed and inaccurate predictive decisions around arrest and detainment that disproportionately impact marginalized populations (Lee & Chin-Rothmann, 2022). In 2019, AI Global Surveillance (AIGS) conducted a survey showing 32% of all countries polled (N = 176) used AI for predictive surveillance. A central controversy then is in breach-of-privacy surveillance without consent in “privacy-aware geographics” (Barroca & Antunes, 2023) where individuals are more sensitive to their retained civil rights and liberties. There is a significant difference in the acceptability of predictive policing through AI by constituents in authoritarian countries when compared to democratic countries where citizens are more privacy aware.

4.2.4 Manipulative advertising & human behavior

The misuse of AI in manipulative advertising, wherein algorithms analyze customer behavior to create exploitative and personalized ads, poses risks of misleading consumers, fostering addictive behavior, and causing potential financial or emotional harm. It is not an overstatement to assert that well-established platforms with dedicated user bases, akin to major internet entities, possess a profound understanding of their users surpassing that of even their closest family and friends. Numerous companies amass vast datasets to fuel their artificial intelligence algorithms. For instance, actions such as expressing approval on social media can be leveraged to anticipate a spectrum of characteristics more accurately about users.

Further, actions tied to non-consensual extracted profile data may include encompassing aspects like personal orientation, cultural background, religious and political inclinations, individual traits, cognitive abilities, emotional well-being, engagement in substance use, experiences of parental separation, age, and gender. “Transparency over systems and algorithms, rules and public awareness are needed to address potential danger of manipulation by artificial intelligence”. (Petropoulos, 2022)

4.3 Complex accountability determinacy

The assimilation of Artificial Intelligence (AI) within contemporary business paradigms has precipitated a confluence of ethical quandaries. The advent of AI has not only augmented the efficiency and analytical capacity of business processes but also introduced profound ethical complexities. Inherent biases within algorithmic data processing can distort decision-making frameworks, while the utilization of AI-driven targeted data presents grave ramifications, ranging from privacy breaches to the manipulation of consumer behavior (Kupfer et al., 2023). This evolutionary juncture in business methodologies compels an exigent reevaluation of accountability mechanisms—a reevaluation that is quintessential for recalibrating the paradigms of responsibility and ethical integrity within the corporate domain (CIO, 2023).

4.3.1 The Imperative of Accountability

In the realm of AI-driven business strategies, accountability transcends the conventional scopes of operational error mitigation and compliance to statutory norms. It encapsulates the foundational ethos of corporate integrity, espousing a culture where ethical considerations are ingrained in the DNA of organizational decision-making (Kazim & Koshiyama, 2021). The pursuit of accountability is rendered multifaceted by the intrinsic opacity of AI decision-making processes, the confluence of stakeholder interests, and the relentless pace of technological progression. These factors collectively mandate a governance model that is both dynamic and responsive, capable of adapting to the evolving landscapes of AI applications and their societal implications (Zoldi, 2022).

4.3.2 Challenges in Assigning Accountability

The delineation of accountability within AI-enhanced business ecosystems is fraught with challenges. Traditional models of decision-making, characterized by their linear trajectories and clear attribution to human agents, starkly contrast with the nebulous decision-making apparatus of AI systems. The latter, driven by intricate algorithmic underpinnings, effectively dissipates the focal point of decision-making, engendering a diffusion of responsibility (Telkamp & Anderson, 2022). This diffusion is accentuated within complex organizational structures, where the recommendations of AI systems are enmeshed with strategic corporate imperatives, thus clouding the lines of accountability (Busuioc, 2021).

The accountability conundrum is further complicated by the distributed nature of responsibility among the myriad agents involved in the AI lifecycle—from the developers who architect the systems to the business leaders who deploy them, and the end-users whose interactions are governed by them. Such a distributed model calls for a paradigm of collaborative governance wherein the roles and responsibilities of all stakeholders are not only recognized but are also synergized within the broader narrative of AI integration (Vice & Khan, 2022).

4.3.3 Impact of Opaque AI Processes

The ethical landscape of AI accountability is significantly marred by the opacity inherent in certain AI systems' decision-making processes, often referred to as the "black box" dilemma. These systems execute complex operations that elude straightforward interpretative analyses, obstructing the elucidation of the underpinning rationale (de Fine Licht & de Fine Licht, 2020). Such opacity not only obfuscates the predictability of AI behavior but also amplifies the intricacies of ascribing liability, engendering multifarious ethical concerns (Hunkenschroer & Luetge, 2022). Consequently, the enhancement of transparency and interpretability through the implementation of rigorous standards and meticulous oversight emerges as an imperative (Busuioc, 2021).

4.3.4 Dynamic Nature of AI Systems

AI systems, characterized by their intrinsic potential for continual learning and adaptation, present formidable challenges in the sustenance of consistent accountability measures. The fluidity of their evolving capabilities necessitates vigilant and perpetual oversight to ensure that the systems' transformative trajectories remain aligned with predefined ethical benchmarks (Zoldi, 2022). This dynamic evolution accentuates the necessity for robust version control protocols and meticulous documentation to meticulously track the genealogy of AI decisions.

4.3.5 Legal and Regulatory Perspectives

The prevailing legal and regulatory edifices frequently prove insufficient in subsuming the autonomous operationality of AI systems. This inadequacy beckons the formulation of novel legislative and regulatory architectures that are sufficiently resilient to enforce accountability upon AI, acknowledging its unique attributes and potentially bestowing a distinct legal status upon it (Hamilton & Davison, 2022; Holder et al., 2019).

Such reconceptualized frameworks must be adept at navigating the complex interplay between evolving AI capabilities and the foundational principles of law and ethics.

4.3.6 Ethical Implications and Considerations

Accountability in AI is inextricably linked with profound ethical implications, encompassing the potential for biases and the overarching societal impacts. The imperative to institute mechanisms that ensure equity in AI decision-making cannot be overstated. It demands a dynamic ethical framework, one that is capable of responding to the concurrent technological and societal flux, to steer AI development and application towards pathways that are congruent with the collective societal ethos (Morse et al., 2022; Ryan, 2020).

In summation, this critical inquiry into the realm of accountability within the milieu of AI-infused business strategies has illuminated the multifaceted complexity inherent in the assignment of accountability, the ethical quandaries precipitated by the opaque nature of AI processes, and the challenges introduced by the dynamic nature of AI systems. It underscores an indispensable need for business entities to integrate ethical considerations into their strategic AI implementations and for policymakers to promulgate adaptable regulatory frameworks. Engaging in a concerted discourse on accountability, ethics, and responsibility is essential in charting a course for AI that not only catalyzes innovation but also steadfastly upholds the tenets of societal and moral integrity (Rodgers & Nguyen, 2022).

4.4 Reliability and safety risks in rapid deployments

As organizations scramble to incorporate Artificial Intelligence into their operations, the fast-paced adoption of these technologies presents serious concerns for reliability and safety. The urge for quick implementation often results in neglecting important steps, such as thorough testing and validation, that are imperative for guaranteeing the AI system's durability and security in practical scenarios (Shaw, Rudzicz, Jamieson, & Goldfarb, 2019).

4.4.1 Reliability Concerns

Deploying AI models too quickly often means cutting corners on testing, which can lead to serious problems with reliability. Machine learning models in particular need a large amount of diverse data and scenarios to learn and adjust effectively (Khaliq, Farooq, & Khan, 2022). Skipping out on this crucial phase could mean that AI systems are not fully prepared to handle the complexities of the real world, resulting in unreliable and inaccurate outputs.

4.4.2 Safety Risks

The potential consequences of hastily implementing AI systems are especially concerning in key industries such as healthcare, transportation, and financial services. For example, in healthcare, a hastily implemented AI system could potentially misinterpret patient data, resulting in inaccurate diagnoses or treatment recommendations (Bajwa, Munir, Nori, & Williams, 2021). Likewise, the financial sector is witnessing a tremendous surge in the use of Artificial Intelligence (AI), bringing to light the pressing concerns surrounding poorly tested AI systems and their potential to execute harmful and unsound investment strategies. As evidenced by Cao's (2021) study, this issue demands urgent attention. The rapid growth of AI adoption, coupled with the constant advancements in generative technology, amplifies the risk of ethical conflicts and detrimental consequences. As we forge ahead with AI methodologies such as machine learning and deep learning, it is crucial to acknowledge the formidable challenges they may bring, not just technical and operational, but also in terms of ethical implications.

4.4.3 Mitigating Measures

To effectively address these potential hazards, a careful and thoughtful approach must be taken towards the implementation of AI technology. This necessitates the following measures:

- **Extensive Testing:** comprehensive testing in a variety of scenarios is crucial in enhancing the reliability of AI systems, as proposed by Myllyaho, Raatikainen, Männistö, Mikkonen, & Nurminen (2021).
- **Iterative Deployment:** Adopting a gradual approach to implementing AI systems, starting with less critical functions and gradually scaling up based on performance evaluations, can help reduce risks (Enholm et al., 2022).
- **Safety Measures:** The development of robust safety protocols and contingency plans is crucial, particularly in industries with high potential risks (Neto, Camargo, Almeida, & Cugnasca, 2022).
- **Continuous Monitoring:** Ongoing monitoring of AI systems after deployment can facilitate prompt identification and resolution of any issues, ensuring safety and effectiveness.

Undoubtedly, the benefits of swiftly implementing AI in the business world cannot be ignored. However, it is crucial to consider the possible drawbacks in terms of reliability and safety. Striking a balance between these opposing factors requires a thorough testing process, gradual integration, and constant monitoring. Such a comprehensive approach not only guarantees that AI systems contribute positively to business strategies, but also ensures their safe and dependable operation (Bankins & Formosa, 2023).

5. Conclusion

As we navigate uncharted territories with AI technologies, the pace and scope of AI advancement has been put into question by leading AI researchers and industry leaders, such as the founder and ex-CEO of Open AI, Sam Altman and Sundar Pichai, CEO of Google, who urged lawmakers in Washington to regulate AI. Governments can require audits and validation processes for generative AI systems and independent assessments can evaluate the ethical considerations of AI models to ensure they align with democratic values and human rights (Khalil, 2023). Many AI ethical risks can be addressed under the existing US regulatory bodies. For example, “the FTC has authority over “false and deceptive” practices such as AI deepfakes. The Equal Employment Opportunity Commission is addressing potential bias of AI models in hiring processes. Microsoft proposes “developing a broad legal and regulatory framework based on the technology architecture of AI.” Others oppose creating a new legal regime and think AI is best regulated under existing laws. (Whyman, 2023).

Until clearer regulations are put into place in the US and globally, overcoming ethical risks associated with AI in business will require a thoughtful and comprehensive approach at the organizational level. Organizations should establish ethical guidelines that align with the organization’s values and principles. These guidelines should encompass transparency, fairness, accountability, and privacy. By adopting a consumer-centric approach to AI, organizations can support the ethical development and utilization of AI technologies.

Transparency and Explainability: Ensure that AI systems are transparent and explainable. To build trust with users and stakeholders, communicate how AI decisions are made and used.

Fairness and Bias Mitigation: Regularly assess and mitigate biases in AI algorithms. Implement fairness measures to ensure that the AI systems do not discriminate against specific groups or individuals. This involves diverse and representative training data and ongoing monitoring.

Data Privacy: Prioritize data privacy and adhere to relevant data protection regulations (e.g., GDPR, CCPA). Clearly communicate how data is collected, used, and stored. Obtain informed consent from users before utilizing their data.

Accountability and Responsibility: Clearly define roles and responsibilities for AI development, deployment, and monitoring. Establish accountability for the outcomes of AI systems and ensure there are mechanisms in place to address unintended consequences.

Regular Audits and Assessments: Conduct regular ethical audits and assessments of your AI systems. This includes evaluating the impact of AI on different stakeholders, checking for biases, and ensuring compliance with ethical guidelines.

Human Oversight: Integrate human oversight into AI systems, especially in critical decision-making processes. Humans should be able to intervene if the system behaves inappropriately or makes a decision that could have ethical implications.

Continuous Education and Training: Keep your teams updated on the latest developments in AI ethics and ensure that they are aware of potential ethical risks. This includes providing training on responsible AI development and deployment.

Collaborate with the AI Community: Collaborate with the broader AI community, including researchers, ethicists, and other organizations, to stay informed about best practices and emerging ethical considerations.

References

- Bajwa, J., Munir, U., Nori, A., & Williams, B. (2021). Artificial intelligence in healthcare: Transforming the practice of medicine. *Future Healthcare Journal*, 8(2), e188-e194. Retrieved from: <https://doi.org/10.7861/fhj.2021-0095>
- Bankins, S., & Formosa, P. (2023). The ethical implications of artificial intelligence (AI) for meaningful work. *Journal of Business Ethics*, 185, 1-16. Retrieved from: <https://doi.org/10.1007/s10551-023-05339-7>
- Barocas, S., & Selbst, A. D. (2016). Big data’s disparate impact, *California Law Review*, 104, 671. Retrieved from <https://www.jstor.org/stable/24758720>
- Barroca, J., & Antunes, M. E. (2023). Surveillance and predictive policing through AI. Deloitte. Retrieved from: <https://www2.deloitte.com/fi/fi/pages/public-sector/articles/urban-future-with-a-purpose/surveillance-and-predictive-policing-through-ai.html>
- Busuioc, M. (2021). Accountable Artificial Intelligence: Holding Algorithms to Account. *Public Administration Review*, 81(5), 825–836. Retrieved from: <https://doi.org/10.1111/puar.13293>
- Cao, L. (2021). AI in finance: Challenges, techniques and opportunities. arXiv. Retrieved from: <https://arxiv.org/abs/2107.09051>

- Chen, P., & Zhang, C. (2014). Data-intensive applications, challenges, techniques, and technologies: A survey on big data. *Information Sciences*, 275, 314-347. Retrieved from: <https://www.sciencedirect.com/science/article/abs/pii/S0020025514000346>
- Chui, M., & Yee, L. (2023). AI could increase corporate profits by \$4.4 trillion a year, according to new research. McKinsey Global Institute. Retrieved from: <https://www.mckinsey.com/mgi/overview/in-the-news/ai-could-increase-corporate-profits-by-4-trillion-a-year-according-to-new-research>
- CIO. (2023). How enterprises can navigate ethics and responsibility of generative AI. CIO. Retrieved from <https://www.proquest.com>
- Continuous control monitoring with AI. (n.d.). Medium. Retrieved from: <https://medium.com/@cognitiveview/continuous-control-monitoring-with-ai-b6caee4d4131>
- Daniels, E. (2023, July 18th). Responsible use of AI is not just a nice to have – it’s essential. GDS. Retrieved from: <https://gdsgroup.com/insights/article/responsible-use-of-ai-is-not-just-a-nice-to-have-its-essential/>
- D’Aquila, M., & Shirer, M. (2019). Worldwide spending on artificial intelligence systems will be nearly 98 billion 2023, according to new IDC spending guide. BusinessWire. Retrieved from: <https://www.businesswire.com/news/home/20190904005570/en/Worldwide-Spending-on-Artificial-Intelligence-Systems-Will-Be-Nearly-98-Billion-in-2023-According-to-New-IDC-Spending-Guide>
- de Fine Licht, K., & de Fine Licht, J. (2020). Artificial intelligence, transparency, and public decision-making: Why explanations are key when trying to produce perceived legitimacy. *AI & Society*, 35(4), 917–926. <https://doi.org/10.1007/s00146-020-00960-w>
- Dickie, J., Groysberg, B., Shapiro, B. P., & Trailer, B. (2022, November-December). Can AI really help you sell? Harvard Business Review. Retrieved from: <https://hbr.org/2022/11/can-ai-really-help-you-sell>
- Enholt, I. M., Papagiannidis, E., Mikalef, P., et al. (2022). Artificial intelligence and business value: A literature review. *Information Systems Frontiers*, 24, 1709-1734. Retrieved from: <https://doi.org/10.1007/s10796-021-10186-w>
- Greenfield, D. (2020). How artificial intelligence works in quality control. AutomationWorld. Retrieved from: <https://www.automationworld.com/factory/sensors/article/21198005/how-artificial-intelligence-works-in-quality-control>
- Hamilton, R. H., & Davison, H. K. (2022). Legal and Ethical Challenges for HR in Machine Learning. *Employee Responsibilities and Rights Journal*, 34(1), 19–39. <https://doi.org/10.1007/s10672-021-09377-z>
- Hao, K. (2019). Facebook’s ad-serving algorithm discriminates by gender and race. MIT Technology Review. Retrieved from <https://www.technologyreview.com/2019/04/05/1175/facebook-algorithm-discriminates-ai-bias>
- Holder C, Iglesias, M., Triaille J.-P., & Van Gysegem J.-M. (2019). Legal and regulatory implications of artificial intelligence (AI): the case of autonomous vehicles, m-health and data mining. Publications Office.
- Hunkenschroer, A. L., & Luetge, C. (2022). Ethics of AI-Enabled Recruiting and Selection: A Review and Research Agenda. *Journal of Business Ethics*, 178(4), 977–1007. <https://doi.org/10.1007/s10551-022-05049-6>
- Kazim, E., & Koshiyama, A. S. (2021). A high-level overview of AI ethics. *Patterns* (New York, N.Y.), 2(9), 100314. <https://doi.org/10.1016/j.patter.2021.100314>
- Khalil, Lydia. (2023). How can we regulate AI? Let’s just ask it. The Interpreter. <https://www.lowyinstitute.org/the-interpreter/how-can-we-regulate-ai-let-s-just-ask-it>
- Khalilq, Z., Farooq, S. U., & Khan, D. A. (2022). Artificial intelligence in software testing: Impact, problems, challenges and prospect. arXiv. Retrieved from: <https://arxiv.org/abs/2201.05371>
- Kite-Powell, J. (2023, September 20). Deepfakes are here, can they be stopped? Forbes. Retrieved from: <https://www.forbes.com/sites/jenniferkitepowell/2023/09/20/deepfakes-are-here-can-they-be-stopped/?sh=1bb648ef3a43>
- Koome, B. (2023, November 15). Exclusive Report: What Are Some Real-Life Examples of Privacy Abuses? Cryptopolitan. Retrieved from: <https://www.msn.com/en-us/news/technology/exclusive-report-what-are-some-real-life-examples-of-privacy-abuses/ar-AA1jYz7E>
- Kumar, V., Ramachandran, D., & Kumar, B. (2020). Influence of new-age technologies on marketing: a research agenda. *Journal of Business Research*. Retrieved from: <https://www.sciencedirect.com/science/article/abs/pii/S0148296320300151?via%3Dihub>
- Kupfer, C., Prassl, R., Fleiß, J., Malin, C., Thalmann, S., & Kubicek, B. (2023). Check the box! How to deal with automation bias in AI-based personnel selection. *Frontiers in Psychology*, 14, 1118723. <https://doi.org/10.3389/fpsyg.2023.1118723>
- Lee, K. F. (2021, September 11). The third revolution in warfare. The Atlantic. Retrieved from: <https://www.theatlantic.com/technology/archive/2021/09/i-weapons-are-third-revolution-warfare/620013/>
- Lee, N. T., & Chin-Rothmann, C. (2022, April 12). Police surveillance and facial recognition: Why data privacy is imperative for communities of color. Brookings. Retrieved from: <https://www.brookings.edu/articles/police-surveillance-and-facial-recognition-why-data-privacy-is-an-imperative-for-communities-of-color/>

- Morse, J. (2017). App creator apologizes for ‘racist’ filter that lightens skin tones. Mashable. Retrieved from: <https://mashable.com/article/faceapp-racism-selfie#zeUItoQB5iqI>
- Morse, L., Teodorescu, M. H. M., Awwad, Y., & Kane, G. C. (2022). Do the Ends Justify the Means? Variation in the Distributive and Procedural Fairness of Machine Learning Algorithms. *Journal of Business Ethics*, 181(4), 1083–1095. <https://doi.org/10.1007/s10551-021-04939-5>
- Myllyaho, L., Raatikainen, M., Männistö, T., Mikkonen, T., & Nurminen, J. K. (2021). Systematic literature review of validation methods for AI systems. *Journal of Systems and Software*, 181, 111050. Retrieved from: <https://doi.org/10.1016/j.jss.2021.111050>
- Neto, A. V. S., Camargo, J. B., Almeida, J. R., & Cugnasca, P. S. (2022). Safety assurance of artificial intelligence-based systems: A systematic literature review on the state of the art and guidelines for future work. *IEEE Access*, 10, 130733–130770. Retrieved from: <https://doi.org/10.1109/ACCESS.2022.3229233>
- Pazzanese, C. (2020). Ethical concerns mount as AI takes bigger decision-making role in more industries. *The Harvard Gazette*. Retrieved from: <https://news.harvard.edu/gazette/story/2020/10/ethical-concerns-mount-as-ai-takes-bigger-decision-making-role/>
- Pearce, G. (2021, May 21). Beware the Privacy Violations in Artificial Intelligence Applications. ISACA. Retrieved from: <https://www.isaca.org/resources/news-and-trends/isaca-now-blog/2021/beware-the-privacy-violations-in-artificial-intelligence-applications>
- Petropoulos, G. (2022, February 02). The dark side of artificial intelligence: manipulation of human behaviour. Bruegel. Retrieved from: <https://www.bruegel.org/blog-post/dark-side-artificial-intelligence-manipulation-human-behaviour>
- Quan, H., Li, S., Zeng, C., Wei, H., and Hu, J. (2023). Big data and AI-driven product design: A survey. MDPI. Retrieved from: <https://www.mdpi.com/2076-3417/13/16/9433>
- Reed, T. (2023, November 14). Deepfakes could supercharge health care's misinformation problem. Axios. Retrieved from: <https://www.axios.com/2023/11/14/ai-deepfake-health-misinformation-fake-pictures-videos>
- Rodgers, W., & Nguyen, T. (2022). Advertising Benefits from Ethical Artificial Intelligence Algorithmic Purchase Decision Pathways. *Journal of Business Ethics*, 178(4), 1043–1061. <https://doi.org/10.1007/s10551-022-05048-7>
- Ryan, M. (2020). In AI We Trust: Ethics, Artificial Intelligence, and Reliability. *Science and Engineering Ethics*, 26(5), 2749–2767. <https://doi.org/10.1007/s11948-020-00228-y>
- Sharkey, N. (2020, February 01). Fully Autonomous Weapons Pose Unique Dangers to Humankind. SCIAM. Retrieved from: <https://www.scientificamerican.com/article/fully-autonomous-weapons-pose-unique-dangers-to-humankind/>
- Shaw, J., Rudzicz, F., Jamieson, T., & Goldfarb, A. (2019). Artificial intelligence and the implementation challenge. *Journal of Medical Internet Research*, 21(7), e13659. Retrieved from: <https://doi.org/10.2196/13659>
- Telkamp, J. B., & Anderson, M. H. (2022). The Implications of Diverse Human Moral Foundations for Assessing the Ethicality of Artificial Intelligence. *Journal of Business Ethics*, 178(4), 961–976. <https://doi.org/10.1007/s10551-022-05057-6>
- Trager, R. F., & Luca, L. M. (2022, May 11). Killer Robots Are Here—and We Need to Regulate Them. FP. Retrieved from: <https://foreignpolicy.com/2022/05/11/killer-robots-lethal-autonomous-weapons-systems-ukraine-libya-regulation/>
- UNESCO (2023). Ethics of artificial intelligence. Artificial Intelligence. Retrieved from: <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>
- Varsha, P. S. (2023, April). How can we manage biases in artificial intelligence systems - A systematic literature review. *Information Management Data Insights*, 3(1). Retrieved from: <https://www.sciencedirect.com/science/article/pii/S2667096823000125>
- Vice, J., & Khan, M. M. (2022). Toward Accountable and Explainable Artificial Intelligence Part Two: The Framework Implementation. *IEEE Access*, 10, 36091–36105. <https://doi.org/10.1109/ACCESS.2022.3163523>
- Whyham, Bill. (2023). AI Regulation is Coming- What is the Likely Outcome? Center for Strategic And International Studies (CSIS) <https://www.csis.org/blogs/strategic-technologies-blog/ai-regulation-coming-what-likely-outcome#>
- World Economic Forum (2023). The future of jobs report 2023. Retrieved from: <https://www.weforum.org/publications/the-future-of-jobs-report-2023/>
- Zoldi, S. (2022). Making AI accountable: Blockchain, governance, and auditability. InfoWorld. Retrieved from <https://www.proquest.com>