Regulating Algorithmic Assemblages: Exploring Beyond Corporate AI Ethics

Md.Mafiqul Islam

Department of Information Science and Library Management, University of Rajshahi, Bangladesh

Abstract

The rapid advancement of artificial intelligence (AI) systems, fueled by extensive research and development investments, has ushered in a new era where AI permeates decision-making processes across various sectors. This proliferation is largely attributed to the availability of vast digital datasets, particularly in machine learning, enabling AI systems to discern intricate correlations and furnish valuable insights from data on human behavior and other phenomena. However, the widespread integration of AI into private and public domains has raised concerns regarding the neutrality and objectivity of automated decision-making processes. Such systems, despite their technological sophistication, are not immune to biases and ethical dilemmas inherent in human judgments. Consequently, there is a growing call for regulatory oversight to ensure transparency and accountability in AI deployment, akin to traditional regulatory frameworks governing analogous processes. This paper critically examines the implications and ripple effects of incorporating AI into existing social systems from an 'AI ethics' standpoint. It questions the adequacy of self-policing mechanisms advocated by corporate entities, highlighting inherent limitations in corporate social responsibility paradigms. Additionally, it scrutinizes well-intentioned regulatory initiatives, such as the EU AI ethics initiative, which may overlook broader societal impacts while prioritizing the desirability of AI applications. The discussion underscores the necessity of adopting a holistic approach that transcends individual and group rights considerations to address the profound societal implications of AI, encapsulated by the concept of 'algorithmic assemblage'.

Keywords: Artificial Intelligence, Machine Learning, Ethical Dilemmas, Regulatory Oversight, Corporate Social Responsibility, Algorithmic Assemblage, Societal Implications, EU AI Ethics Initiative

Article Information:
Article history: Received: 20-Jan-24 Accepted: 27-Mar-24 Online: 08-May-24 Published: 09-May-24
DOI: https://doi.org/10.60087/jklst.vol3.n3.p28
Correspondences'author: Md.Mafiqul Islam Email: techbook1994@gmail.com

Introduction

The surge in investment and integration of artificial intelligence (AI) systems into various
decision-making processes across commercial, political, and public organizations is largely attributed to advancements in machine learning, facilitated by the availability of vast digital datasets. These datasets, extracted from diverse sources ranging from government records to corporate databases, fuel AI systems with the capability to discern correlations and provide valuable insights into human behavior and other phenomena. However, the pervasive use of AI in decision-making has raised concerns about the neutrality and objectivity of automated processes, as they often reflect embedded values rather than impartial judgments. Consequently, calls for regulatory oversight have emerged to ensure transparency and accountability in AI deployment, analogous to regulatory frameworks governing traditional decision-making processes.

Moreover, the personalized nature of AI-driven services, while offering efficiencies and economic gains, also introduces unintended negative effects that warrant public debate. For instance, highly personalized insurance or credit scoring systems, while advantageous to privileged customers, can exacerbate existing social disparities and undermine notions of distributive justice. Similarly, personalized news and entertainment consumption can foster societal fragmentation by reinforcing filter bubbles. Despite media and academic focus on individual impacts, there's been inadequate consideration of broader societal consequences stemming from the incorporation of AI into existing social systems. For example, while technologies like automated facial recognition or predictive policing promise enhanced law enforcement, they may divert resources from proven community policing methods, further isolating law enforcement from the communities they serve.

This paper examines these consequences and ripple effects through an 'AI ethics' lens, critiquing the prevailing discourse's reliance on corporate self-policing. It highlights the limitations of corporate social responsibility paradigms and critiques well-intentioned regulatory initiatives that prioritize AI's desirability over broader societal impacts. Emphasizing the concept of 'algorithmic assemblage,' the paper advocates for a holistic analysis that transcends individual rights considerations and engages with AI's broader societal impacts. It underscores the necessity of shifting AI ethics discourse away from corporate interests towards a more comprehensive understanding of societal welfare.
A New Ethical Awakening or the Practice of 'Ethics-washing'

The Emergence of an Ethics Industry

The growing public concern surrounding the ethical implications of artificial intelligence (AI) has not escaped the attention of governments, commercial entities, and researchers involved in AI development or implementation. The European Commission, in its AI development strategy, explicitly emphasizes the need for an ethical and legal framework aligned with the Union's values and the Charter of Fundamental Rights of the EU. Similarly, the UK Government's Office for Artificial Intelligence, in its policy paper "The AI Sector Deal" (2018), sets as a policy objective of its AI strategy to lead the world in the safe and ethical use of data through initiatives like the Center for Data Ethics and Innovation. Major tech companies such as Google, Microsoft, Facebook, Amazon, IBM, and Salesforce have also taken steps to address AI ethics concerns by establishing AI ethics councils, launching ethics frameworks, and hiring ethicists to contribute to their AI strategies or fund AI initiatives.

However, while the increasing interest in AI ethics from government and business may initially seem like a positive response to public concerns, it brings its own set of challenges and considerations. The establishment of ethics boards, oversight committees, and codes of practice for AI by corporate entities follows a familiar regulatory pattern seen in the technology sector. This pattern involves industries preemptively addressing concerns to avoid formal governmental regulation by implementing self-regulatory mechanisms. While such mechanisms are sometimes effective in achieving positive regulatory outcomes in specific contexts, they are often criticized for prioritizing commercial interests over public welfare, lacking transparency and accountability, and failing to engage meaningfully with civil society.

It's important to note that corporate adoption of ethical frameworks for AI isn't necessarily a cynical attempt to evade regulation through 'ethics-washing.' It may genuinely reflect efforts to align with values prevalent in contemporary public discourse and to legitimize technology practices or business opportunities. However, whether this alignment leads to meaningful resolutions of social problems caused by AI depends on whether it genuinely internalizes public
values or merely deflects criticism.

**Reinventing Corporate Social Responsibility?**

Examining corporate initiatives to establish internal or industry-wide external frameworks for AI ethics inevitably draws parallels to the broader discourse on Corporate Social Responsibility (CSR). While CSR, broadly defined as the articulated policies and practices of corporations reflecting responsibility for societal well-being, has gained global recognition, its origins trace back to the United States. Observers have noted distinct differences between US and European approaches to CSR, with US CSR emphasizing a clearer role for corporations in independently assuming responsibility for societal interests, while European CSR envisions a partnership involving representative social and economic actors led by government.

Given that many corporate giants leading the AI sector are based in the US, it's not surprising that the initial response to ethical concerns surrounding AI was rooted in corporations developing solutions internally or through trusted external advisors. This early 'Americanization' of the AI ethics debate has strongly influenced national and supranational policy developments, especially when considering the establishment of bodies like the EU High-Level Expert Group on Artificial Intelligence by the European Commission in 2018. The discourse surrounding this group predominantly framed ethical risks and mitigation strategies from a corporate-centric perspective.

Critiques of CSR's effectiveness in achieving long-term societal benefits are longstanding. A key concern is that while CSR practices may benefit primary stakeholders, they often fail to address broader social issues, functioning more as direct influence tactics. An illustrative example in the AI domain is Google's response to employee concerns regarding its AI research and its military collaborations. Google publicly distanced itself from controversial contracts, established AI principles, and formed an AI ethics committee in response to employee protests. However, Google continues to provide support to startups supplying AI technology to military and law enforcement agencies through its venture capital arm. This indirect support undermines the impact of Google's seemingly responsible actions and avoids rather than addresses ethical considerations. Moreover, distancing ethically questionable practices into smaller, ostensibly independent startups makes it challenging to identify, scrutinize, and regulate commercial
practices that may breach socially desirable ethical principles.

**Structural Flaws in Corporate Self-Regulation**

The challenges inherent in determining the appropriateness of government sanctioning corporate self-regulation, coupled with criticisms of US CSR approaches, raise concerns about the consultation processes shaping guidance for national and supranational policymakers.

A critical evaluation of the governmental and corporate approach to developing ethical AI must consider the independence and transparency of bodies drafting AI ethics guidelines on behalf of governments or advising on policy through corporate ethics boards. For instance, in the EU High-Level Expert Group on Artificial Intelligence, questions arise regarding the proportion of members employed by or funded by corporate entities. The Expert Group's report "Ethics Guidelines for Trustworthy AI" is silent on potential conflicts of interest among its members, indicating an ethical oversight. Similar issues plague corporate AI ethics boards, where companies may withhold membership details, meeting participants, decision-making processes, and actions taken based on board suggestions.

These concerns extend to issues of accountability. If companies influence the outcome of 'independent' governmental evaluations and set the parameters of the ethical landscape through their own ethics boards, critical questions may go undebated. Discussions may overlook societal concerns not easily addressed by checklists or limited to impacts on individual rights or specific groups. This narrow focus ignores broader societal implications that may transcend individual companies or industries.

The self-regulatory nature of ethics boards offers weak accountability, even for addressed issues. Establishing an ethics board alone does not ensure ethical behavior without transparent frameworks defining its operations, membership, recommendations' implementation, and accountability mechanisms for breaches. Similarly, codes of practice lack value without publicly accessible consequences for breaches. Without such accountability mechanisms, ethics boards risk being mere regulatory facades, deflecting public concern and state regulation with little
impact on corporate practices. There's also the concern about whether small groups of ethicists and experts, often drawn from a narrow range of disciplines or interest groups, can adequately represent the diverse concerns of wider civil society. When individuals serve on multiple ethics boards or expert groups, there's an increased risk of 'groupthink,' which may hinder the exchange of context-specific information, discourage exploration of alternative viewpoints, and result in the adoption of incomplete or inflexible outputs. Notably, much of the ethical debate surrounding AI-supported personalization, both corporate-sponsored and governmental, has been disconnected from the general public and civil society groups.

For instance, the EU High-Level Expert Group on Artificial Intelligence, comprising predominantly corporate representatives and AI researchers from select academic disciplines, lacks broader civil society representation. The literature cited in their Ethics Guidelines primarily includes theoretical ethical material developed by group members, with little reference to empirical work on the social impacts of AI or civil society critiques. While the group emphasizes the importance of open discussion and stakeholder involvement, their pilot Assessment List appears to have been compiled without significant public input.

These concerns shouldn't be viewed as criticisms of specific individuals but should prompt reflection on the biases and preferences that drafting ethical guidelines under such conditions may foster. Similar to the internet's filter bubble phenomenon, expert groups may inadvertently operate within their own 'groupthink' bubbles, wherein individuals are exposed only to viewpoints aligned with their own. This can lead to the creation of rules and frameworks that seem acceptable within the group but face resistance when applied in diverse social, cultural, or business contexts.

As subsequent sections will discuss, even with regulatory compulsion, such rules may not effectively ensure transparency, oversight, audit, and accountability.

**Institutionalized Ethics**
Serving as Chair of an academic Research Ethics Review Committee offers a unique insight into how researchers in UK Higher Education engage with both the ethical guidelines pertinent to their discipline and the procedural requirements established to uphold those guidelines. There's often a noticeable gap between the perceived ethical principles and guidelines and the actual willingness of researchers to adhere to them in practice.

Researchers often express frustration with the bureaucratic hurdles, administrative burdens, constraints on academic freedom, and methodological limitations imposed by ethical oversight. They may resort to recycling past responses to ethics review questions without adequately considering the specific variables and risks associated with their proposed research. Minimalist responses to inquiries about risks to research subjects and their data are common, and ethics review applications may be left until the last moment before grant submissions or time-sensitive fieldwork commences.

Their communication with research subjects may be laden with technical jargon, consent forms may be vague and confusing, and risk assessments may be hastily conducted. Furthermore, once the research begins, many of the ethical commitments made in the initial applications may be overlooked when time is short. Ethical corners may be cut in pursuit of perceived research opportunities. However, the repercussions of breaches of ethical guidelines, both perceived and actual, for individuals or institutions are rarely considered.

Similarly, academic ethical review processes often fall short of expectations. Criticisms include the perception of merely going through the motions, institutions covering their backs, excessive formalities, a lack of critical self-reflection, and the imposition of inappropriate discipline-specific requirements. Often, the primary institutional motive for incorporating research ethics review processes across all academic disciplines is not necessarily a concern for the fair treatment of research subjects, the welfare of researchers, or the avoidance of negative impacts on wider society. Instead, it revolves around maintaining access to grants and avoiding potential legal issues or negative publicity, essentially catering to primary stakeholders. From this perspective, there are greater similarities between the objectives of institutional research ethics policies and Corporate Social Responsibility (CSR) policies in the broader commercial sector.
than one might anticipate.

This 'institutional protection' may also manifest in the nature of ethical oversight in academia, typically involving significant front-end scrutiny by committees responsible for ethical review at various levels of the institution. While some fields, such as biomedical research, may undergo formal oversight by external bodies, formal audits are unlikely outside specific disciplinary domains, except in cases of severe breaches of guidelines. Various factors contribute to this, including a lack of resources, authority, access, and willingness. Oversight often relies on self-reporting of ethical breaches, reports from those managing the researchers, or reports from third parties, including research subjects, to ensure ongoing monitoring.

It is argued that the establishment of academic ethical standards and processes, alongside broader legal requirements like data protection laws, has created a research environment where academics engaging with human research subjects are generally aware of the overarching ethical principles governing their work, albeit sometimes with a vague understanding of the specifics. However, in practice, these principles are often perceived, consciously or subconsciously, as applying more to others than to oneself. Researchers may believe that their own practices are inherently ethical, any deviations from these principles are minor and forgivable given the circumstances, and it's other researchers who are more likely to significantly deviate and thus warrant scrutiny. This perception fuels some of the resistance towards formal oversight.

A crucial question arises: have ethics guidelines and review processes genuinely improved ethical practices in academic research, or have they merely cultivated a facade of ethical conduct, creating the illusion of an effective and reflective process for considering and mitigating individual, group, and societal risks that it cannot truly deliver? Examining the outcomes of such an assessment might shed light on the challenges likely to emerge when attempting to apply a general ethical framework to a phenomenon as pervasive as the integration of artificial intelligence into decision-making processes.

If it proves challenging to instill and integrate effective ethical practices within a community of researchers that has established numerous codes of ethical conduct, supports specialized journals
dedicated to research ethics, and is subject to formal ethical review requirements from institutions, funders, and increasingly publishers, then the efficacy of ethical guidelines alone in contexts with significant opposing forces, such as governance and commerce, must be seriously questioned.

**Dissecting AI Ethics Guidelines**

A Narrow Toolbox: Those who fail to learn from history are doomed to repeat it. It's often remarked about military commanders that they tend to prepare for future conflicts by studying past ones, rather than devising strategies for the unknown. Yet, this tendency to resort to familiar methods when faced with new challenges, even if those methods have proved less than optimal, is not exclusive to the military domain. One critical concern surrounding AI-driven personalization revolves around privacy, whether in the context of governmental surveillance and predictive profiling or corporate profiling and manipulation of individuals and groups. The issues of privacy and data protection have been central in the relationships between citizens and the State, and consumers and the private sector for over six decades. Despite this lengthy history, existing regulatory strategies aimed at safeguarding privacy and personal data have struggled to prevent their erosion by technology, even in jurisdictions with constitutionally protected privacy rights and comprehensive data privacy laws. Similarly, suggestions that technological solutions (such as privacy-enhancing technologies) could bolster or replace non-technical methods have not gained widespread traction in the private sector - to paraphrase a common saying, it seems that the best way to make a small fortune investing in privacy-enhancing technologies is to start with a large one.

With these challenges in mind, let's examine the Ethics Guidelines. The technical measures it advocates include: embedding ethical rules within AI system architecture (ethics by design), implementing secondary systems to monitor compliance with these rules (Trustworthy AI), explaining system behavior to assess reliability and inform users (explicability of output), testing and validating AI systems throughout their lifecycle (oversight), and developing quality of service indicators to ensure that AI systems prioritize security and safety considerations. While these are commendable objectives, many of them echo suggestions made from a privacy-
enhancing perspective for similar AI systems. For instance, the idea of ethics by design draws directly from the GDPR and related data privacy research and practices. However, the real-world application of such technical measures in the privacy domain suggests, at best, a modest level of adoption, and even where adopted, the practical outcomes often fall short of the expectations of regulators and theorists.

Following these technical measures are recommendations for non-technical approaches, including: regulation, with references to product safety legislation and liability frameworks; soft law and co-regulatory approaches, involving organizations signing up to guidelines, codes of conduct, accreditation systems, professional codes of ethics, or certification regimes run by trusted intermediaries; internal governance frameworks, such as establishing ethics boards; fostering education and awareness among stakeholders, including the general public, through stakeholder panels and social dialogue; and ensuring diversity and inclusivity among AI makers and users. The notion that ethical AI can be advanced with the support of product safety legislation and liability frameworks is intriguing, especially given their limited application to computer software in the UK and the software industry's efforts to maintain this status quo. Beyond this, most of these measures have either been mooted or utilized in the data privacy regulatory sphere. With the possible exception of data privacy legislation, most of these measures, while admirable in principle, have had limited effectiveness in the data privacy realm. It remains unclear why the Guidelines assume they will be more effective in fostering 'Trustworthy AI', rather than simply spawning a secondary industry ecosystem, akin to what we've seen in contemporary data privacy, comprised of consultants, certification bodies, e-training specialists, and administrators of varying quality, reliability, and longevity.

Diversity and inclusivity, from design to procurement and deployment, are undoubtedly desirable, especially considering the technology industry's historical shortcomings in these areas, both in terms of its workforce and organizational leadership. However, even if strides are made to enhance diversity and inclusivity, it's important to recognize that the design phase of technology often operates separately from the procurement and deployment processes. An essential aspect missing from the equation is the aspect of 'selling' - what are the key features upon which AI marketing relies? Do purchasers grasp the ethical implications of its usage in
specific contexts? And do vendors themselves understand and endeavor to communicate to purchasers the ethical issues that might arise from the utilization of their products?

It's entirely possible to have an inclusive and diverse team involved in the design, development, testing, and maintenance of an ethical AI system, yet still encounter problems when it comes to its marketing, implementation, and direction. The concerns regarding 'diversity and inclusivity' are aptly illustrated by various manifestations of data-driven personalization, which, as evidenced in this collection, tend to exacerbate existing socio-economic disparities and vulnerabilities. AI makers or their governmental and corporate users often do not share these vulnerabilities or express significant concern about them. In fact, their personal or institutional interests may directly benefit from ethically questionable personalized applications, such as offering cheaper prices or credit (price or credit discrimination), enhancing public safety (through predictive policing or sentencing), or reducing public health expenditures (via precision medicine) - regardless of the detrimental impact on the already disadvantaged, as detailed in several chapters of this book.

**Limited Perspectives on Affected Interests: An Atomistic Approach**

The language of the Ethics Guidelines predominantly revolves around fundamental rights and freedoms aimed at individuals or specific groups perceived to be vulnerable due to particular characteristics. AI-driven personalization undoubtedly impacts individual and group rights, and while the Ethics Guidelines offer an extensive Assessment List to consider these issues, it often overlooks the broader societal implications of AI-driven personalization. Although the report acknowledges wider social issues to some extent, only one out of 131 questions in the Assessment List pertains to "Society and Democracy," hinting at the potential interests of affected stakeholders beyond the end user. This atomistic approach prioritizes discrete individual or group rights that can be neatly categorized, risk-assessed, and checked off.

Hoffman identifies several shortcomings with this approach. The focus on avoiding breaches of legally protected rights tends to center attention on preventing 'bad actors' from embedding discriminatory biases into AI systems. Alternatively, due to the opaque nature of many AI
systems, blame for unexpected outcomes is increasingly shifted from 'bad actors' to 'bad algorithms,' namely the model or training data. While solutions are often sought in diversifying human teams across the AI lifecycle or implementing technical patches, these fixes for individual or systemic behavior often sidestep broader attempts to address the underlying social and cultural processes that foster discrimination.

Moreover, there is a tendency to address specific disadvantages faced by legally pre-categorized groups without fully understanding how institutional or social contexts may influence the impact of AI decision-making on sub-groupings within and across those categories. Solely focusing on avoiding specific types of disadvantage may result in achieving 'fair AI' in terms of avoiding obvious disparities in treatment among different groups. However, this approach may fail to question the discriminatory effects of systemic advantages enjoyed by particular groups, whether internal or external to the AI system. For instance, an AI decision's real-world consequences may depend entirely on external factors such as wealth or social capital, as described by Eubanks, highlighting the disparate experiences within the digital data regime.

Failing to adopt a holistic perspective on potential structural inequalities means that the injustices stemming from AI outcomes often remain unchallenged and unaddressed. Legal and political discussions surrounding fair treatment and anti-discrimination tend to focus narrowly on the distribution of rights, opportunities, and resources. However, Hoffman argues that this approach is problematic for two main reasons. Firstly, merely redistributing rights, opportunities, and resources is insufficient to uphold human dignity without concurrent changes in social structures and attitudes that prevent harms not remedied by distribution alone. Secondly, framing the central issue as one of distribution overlooks the intrinsic role of AI in shaping ongoing social and cultural dynamics within societal systems.

This 'atomistic' approach poses a significant challenge for the authors of the Ethics Guidelines. On one hand, they are tasked with considering the ethical implications of AI comprehensively, while on the other hand, they aim to provide an Assessment List to assist organizations in integrating ethical considerations into their processes and procedures. While the former objective should entail a holistic assessment of broad social issues and risks beyond individual and group
rights, the latter demands a narrower focus on individual or group rights through micro-risk assessment. However, the risk with the latter approach is that it implies that broader social questions are beyond the responsibility or capacity of individual organizations or industries.

If governments, influenced by corporate interests, delegate the legal and ethical regulation of AI to self-regulatory processes, it becomes even more problematic. Such a scenario may lead to a debate focused on implementing regulation that is business-friendly and individual rights-focused, neglecting engagement with civil society representatives to understand the societal trade-offs citizens are willing to accept for the potential benefits of AI and personalization. Important societal values such as communal solidarity and recognition of structural disadvantages may be overshadowed by an intensified 'user-pay' approach, which clashes with concerns of redistributive justice. Additionally, assuming the legitimacy of predictive sentencing for public safety benefits undermines fundamental values such as the presumption of innocence or principles of justice.

The Ethics Guidelines notably lack a sustained effort to reevaluate or challenge the prevailing framing of the potential ethical risks and the strategies to address them. Recent literature in the field of Science and Technology Studies suggests that much of the discourse in this area has been narrowly focused on technical aspects of algorithms. To achieve broader social objectives such as fairness, justice, and due process, there is a need for a more holistic approach to framing these discussions.

Central to this argument is the recognition that AI algorithms are not devoid of values; rather, they possess sociological and normative features that influence their interactions with humans. These algorithms shape associations, similarities, and actions based on their engagement with data, end users, or other systems. They play a role in structuring how information is produced, interpreted, and attributed public significance. Therefore, creating guidelines for the ethical use of algorithms is challenging because it involves not applying ethical values to a blank slate but attempting to alter values already embedded within specific algorithmic instances.

Moreover, ethical guidelines aimed at influencing the behavior of organizations or industries may not adequately address the impact of "algorithmic assemblages" that span wide socio-
technical networks. The effectiveness of such assemblages lies not only in their ability to process and identify patterns in data but also in their capacity to influence adjacent computational routines, material infrastructures, and human behaviors. Depending on their domain of application, functioning algorithms require integration across hardware, digital flows, organizational structures, analog infrastructure, and socio-economic processes.

The prevailing discourse on AI ethics has largely adopted a US corporate-centric viewpoint regarding the potential ethical risks and the strategies to mitigate them. This perspective reflects a narrow corporate social responsibility (CSR) approach to crafting ethical guidelines, which prioritizes the interests of primary stakeholders and relies heavily on existing legal commitments to protect individual and group rights. However, this limited framing of the ethics debate is insufficient for developing effective ethical approaches to AI in general, and "algorithmic assemblages" in particular.

A fundamental reassessment of how policymakers address the challenges posed by AI and algorithmic assemblages is imperative. Relying on ethical frameworks and regulatory models that fail to engage with the actual and potential structural inequalities perpetuated by AI and algorithmic assemblages is akin to treating symptoms rather than root causes. For instance, consider Amazon's market dominance built on its personalized recommender system, which is supported by a complex assemblage of algorithms that extend beyond analyzing consumer behavior to optimizing various aspects of its operations, including workforce management and supply chain logistics.

This intricate assemblage of algorithms poses significant challenges to the applicability of conventional ethical guidelines. It represents just a fraction of a broader socioeconomic ecosystem centered around Amazon, which transcends the company's direct control and continues to evolve dynamically. Addressing the structural inequalities embedded within such vast assemblages requires an ethical framework capable of dismantling them explicitly. However, this is unlikely to emerge from corporate ethics committees or expert panels focused on producing superficial outputs like checklists and questionnaires for private and public sector organizations.
Towards a Comprehensive Framework for AI Ethics

The current focus of both corporate and governmental entities on developing AI ethics frameworks primarily revolves around shaping future regulations concerning AI usage, particularly in personalization contexts. These frameworks, whether developed by the EU, nation states, or corporate bodies, alongside the establishment of expert groups and institutional ethics boards, aim to advocate for self-regulation as an effective means of addressing the diverse array of risks and challenges associated with AI. As a result, governmental regulation, whether through legislative measures or regulatory agencies, may be deemed unnecessary.

However, this approach is flawed in several critical aspects. Firstly, it tends to evaluate the ethical concerns surrounding AI and algorithmic assemblages through a narrow corporate social responsibility lens and relies on a limited range of conventional regulatory techniques familiar to organizations. While this approach may facilitate easier adoption by organizations, it raises questions about the suitability of these techniques for addressing the unique challenges posed by AI. Additionally, it overlooks alternative technical and non-technical approaches and fails to leverage insights from contemporary science and technology studies literature, which emphasizes the need for a holistic approach to framing the discussion on AI ethics.

Secondly, proponents of ethical frameworks often underestimate the practical challenges of implementing and enforcing them within organizations. Real-world practices of ethical review and oversight within organizations involved in designing, developing, and deploying AI systems often exhibit a lack of transparency and accountability.

Thirdly, when national governments or supranational entities like the EU do not take a leading role in evaluating the ethical risks of AI, regulatory initiatives are left in the hands of the very entities targeted for regulation. This raises concerns about the independence and effectiveness of expert groups dominated by representatives with vested interests in AI production or utilization. Moreover, the lack of input from the public or civil society groups may result in a narrow focus on business-friendly, individual rights-focused regulation, neglecting broader social impacts.
While there has been a critical response to various aspects of "ethical AI," including analysis of fair machine learning methodologies and shortcomings of anti-discrimination laws in complex socio-technical systems, the social sciences and humanities offer valuable evidence and analysis to support more sophisticated regulatory interventions into algorithmic assemblages. Consequently, current ethics guidelines may serve as an initial milestone in the ongoing development of a comprehensive and reflexive regulatory practice in the field of AI ethics.

**Conclusion**

In conclusion, the current discourse on AI ethics, driven primarily by corporate and governmental entities, tends to adopt a narrow perspective influenced by corporate social responsibility principles. This approach, while aiming to promote self-regulation, often overlooks the complex societal implications of AI and algorithmic assemblages. Instead, it relies on conventional regulatory techniques and lacks consideration of alternative approaches suggested by contemporary science and technology studies literature.

Moreover, the practical challenges of implementing ethical frameworks within organizations, coupled with the limited involvement of national governments or supranational entities, raise concerns about the effectiveness and independence of regulatory initiatives. The dominance of stakeholders with vested interests in AI production or utilization further complicates efforts to address broader social impacts.

However, there is growing critical discourse and evidence from the social sciences and humanities that offers insights into more sophisticated regulatory interventions for algorithmic assemblages. While current ethics guidelines may represent an initial step, they underscore the need for a more holistic and reflexive approach to AI ethics regulation. Moving forward, it is essential to incorporate diverse perspectives and consider the broader societal implications to ensure that AI technologies are developed and deployed in a manner that aligns with ethical principles and societal values.
References List:


Journal of Multidisciplinary Research, 2(2).


[45]. Sarioguz, O., & Miser, E. ARTIFICIAL INTELLIGENCE AND PARTICIPATORY LEADERSHIP: THE ROLE OF TECHNOLOGICAL TRANSFORMATION IN BUSINESS MANAGEMENT AND ITS IMPACT ON EMPLOYEE PARTICIPATION.