Cite as: Bircan, T. and Özbilgin, M. F. (2024) Unmasking Inequalities of the Code: Disentangling the Nexus of AI and Inequality, *Technological Forecasting and Social Change*

Unmasking Inequalities of the Code: Disentangling the Nexus of AI and Inequality

Keywords: Artificial Intelligence, social realism, inequality, digital divide, algorithmic bias, privacy and surveillance

Abstract

This article provides an interdisciplinary exploration of the complex dynamics between artificial intelligence (AI) and inequality, drawing upon social sciences and technology studies. It scrutinises the power dynamics that shape the development, deployment, and utilisation of AI technologies, and how these dynamics influence access to and control over AI resources. To do so, we employ Margaret Archer's social realism framework to illuminate the ways in which AI systems can reinforce various forms of inequalities. This theoretical perspective underscores the dynamic interplay between social context, individual agency, and the processes of morphostasis and morphogenesis, offering a nuanced understanding of how inequalities are reproduced and potentially transformed within the AI context. We further discuss the challenges posed by the access and opportunity divide, privacy and surveillance concerns, and the digital divide in the context of AI. We propose co-ownership as a potential solution to economic inequalities induced by AI, suggesting that stakeholders contributing to AI development should have significant claims of ownership. We also advocate for the recognition of AI systems as legal entities, which could provide a mechanism for accountability and compensation in cases of privacy breaches. Finally, we conclude by emphasising the need for robust data governance frameworks, global governance, and a commitment to social justice in navigating the complex landscape of AI and inequality.

1. Introduction

In the throes of the Fourth Industrial Revolution, the relentless march of Artificial Intelligence (AI) has ignited a transformative epoch, reshaping the contours of myriad sectors with its disruptive prowess (Dellerman et al., 2019). From revolutionising healthcare through automation to tailoring education through personalised learning algorithms, the pervasive influence of AI accentuates the pressing need for a rigorous examination of its foundational principles and the consequential ramifications it engenders on societal disparities (Bostrom, 2014). This paper embarks on an intellectual odyssey to decipher the complex nexus between AI and inequality, posing the research question: How does the interplay between AI technologies and social structures contribute to perpetuating and potentially exacerbating societal inequalities?

The study aims to illuminate the societal chasms that AI systems may inadvertently widen, building on the premise that biases, power dynamics, and ethical quandaries are inextricably knitted into the fabric of AI technologies (Crawford, 2016). This conceptual paper built on illustrative cases has two objectives: first, to elucidate the mechanisms through which AI both perpetuates and is moulded by inequality; and second, to explore transformative solutions that can inform the design, deployment and ownership of AI technologies that embody the principles of fairness, inclusivity, and justice (O'Neil, 2016).

Informed by recent scholarly contributions (Zarsky, 2016; Eubanks, 2018; Birhane & van Dijk, 2020), we extend our conceptual framework to include a critical analysis of case studies drawn from existing regulatory environments and policy frameworks. The intent is to assess their efficacy in mitigating the unintended negative effects of AI on social inequality. By identifying areas of strength and weakness within these regulatory approaches, we aim to propose novel and effective strategies that foreground ethical considerations, prioritise fairness, and optimise social impact.

To achieve these aims, we employ a mixed-methods approach that integrates qualitative and quantitative analyses, allowing for a comprehensive exploration of the broader context and societal milieu within which AI operates. This scrutiny encompasses the convoluted interplay of power, governance, and resource allocation.

The development and deployment of AI technologies are inherently interdisciplinary endeavours, requiring the integration of insights from computer science, data science, social sciences, ethics, law, and other fields. However, the lack of interdisciplinarity in AI development has often led to a narrow focus on technical aspects, overlooking the broader societal implications and potential inequalities perpetuated by these technologies (Broussard, 2018). AI systems are not merely technical artefacts; they are deeply embedded within social, economic, and cultural structures, reflecting and reinforcing societal norms, values, and power dynamics. Their impacts are felt across diverse sectors and communities (Ozbilgin, 2024), and without an interdisciplinary approach, there is a risk of developing AI technologies that are disconnected from the societal contexts in which they operate, leading to unintended consequences and exacerbating existing inequalities (Baum, 2021, Dafoe et al., 2021).

Sociological frameworks provide a lens through which we can understand the societal implications of AI technologies, including their potential to perpetuate and exacerbate inequalities. By deploying late Margaret Archer's conceptual universe, we explore the duality of continuity and change in the interplay of powerful agents and socio-cultural structures. This framework allows us to mobilise the concepts of morphogenesis (signifying change), morphostatis (signifying continuity), along with reflexive monitoring and reflexive engagement (framing agency) to elucidate how the concentration of AI development and ownership within a small number of powerful corporations and investors leads to polarisation of wealth across the value chain. in the hands of a few owners and investors, and a lack of transparency and accountability, with decisions about the design and deployment of AI technologies often made behind closed doors (Pasquale, 2015). This can result in AI systems that serve the interests of the powerful at the expense of disadvantaged communities, reinforcing existing power dynamics and inequalities (O'Neil, 2016).

Hence, fostering interdisciplinarity in AI development and ensuring a fair and accountable ownership structure are crucial steps towards mitigating the potential inequalities perpetuated by AI technologies. This requires a commitment to open, collaborative, and inclusive practices, as well as ongoing research and dialogue across diverse fields and sectors. This requires a commitment to open, collaborative, and inclusive practices, as well as ongoing research and dialogue across diverse fields and sectors. In pursuing these aims, we hope to contribute to a future where AI serves as a powerful ally in the pursuit of societal equality and justice.

2. Conceptual Framework

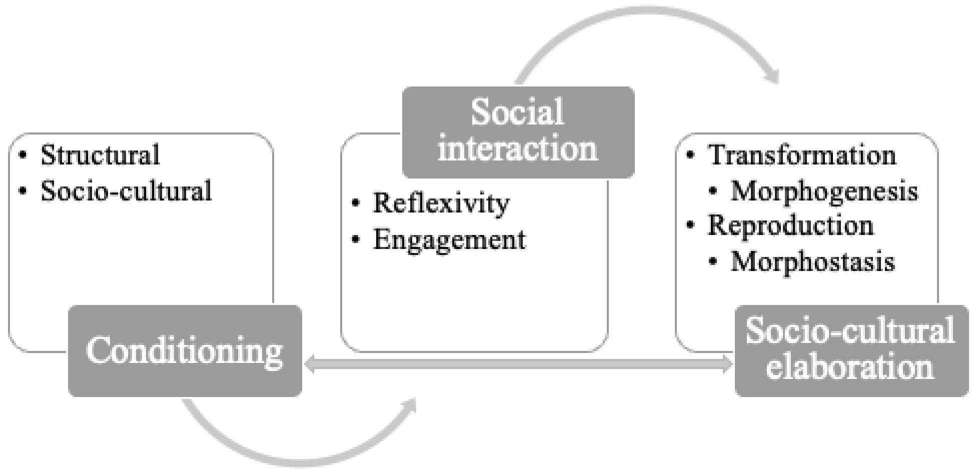
In our conceptual exploration of the interplay between AI and inequality, we draw extensively upon the theoretical underpinnings of social realism, as articulated by Margaret Archer (1995, 2003, 2007, 2016). This conceptual framework offers a robust scaffold to dissect the complex dynamics between social structures and AI systems. It illuminates how AI, as a socio-technical entity, becomes entwined within broader societal frameworks, perpetuating and potentially exacerbating inequalities. Social realism, with its emphasis on the impact of institutions and social dynamics on human behaviour, provides a lens through which we can understand the embeddedness of AI technologies within societal structures. This perspective allows us to move beyond a surface-level analysis and investigate the structural and institutional forces that shape the development, deployment, and societal impact of AI (Archer, 2003, 2008).

Archer's theoretical framework is pivotal for examining the nexus of AI and inequality due to its emphasis on agency and the complex interplay between individual actions and social structures. Her conceptual universe enables us to critically evaluate how the design and implementation of AI systems reflexively change and reinforce societal biases and power dynamics, highlighting the significant role of individuals and social groups in shaping and contesting these technologies. Through this lens, we are offered a pathway for transformative change.

Central to Archer's theory is the duality of change and continuity within socio-cultural systems, encapsulated in the concepts of morphogenesis and morphostasis. Morphogenesis refers to the emergence of changes emanating from the active engagement of actors, altering the status quo of societal conditions. Morphostatis, on the other hand, refers to the enforcement or preservation of existing sociocultural norms and structures, maintaining continuity. These concepts are crucial for understanding the dynamic interplay of stability and transformation within the sphere of AI and societal inequality.

Figure 1 visualises Archer's vision of change and reproduction of inequalities, conditioned by sociocultural context and reflexive agents. This model is instrumental in our analysis of inequalities in artificial intelligence, providing a visual interpretation of the theoretical discussion.

Figure 1: Mechanism of change and reproduction of inequalities conditioned by sociocultural context and reflexive agents.



Note: Authors' illustration based on Archer's social realism framework.

Building upon the recent scholarly contributions (Zarsky, 2016; Eubanks, 2018; Birhane & van Dijk, 2020), we extend our conceptual framework to include a critical analysis of illustrative cases drawn from existing regulatory environments and policy frameworks. This conceptual analysis with case examples assesses their effectiveness in mitigating the unintended negative effects of AI on social inequality. By identifying areas of strength and weakness within these regulatory approaches, we aim to propose innovative and effective strategies that foreground ethical considerations, prioritise fairness, and optimise social impact. Our enriched framework, grounded in social realism and augmented by interdisciplinary insights, provides a comprehensive understanding of the dynamics that shape the impact of AI on societal inequalities. This approach enables us to articulate a nuanced understanding of the interdependencies that influence the relationship between AI and societal inequalities, thereby paving the way for informed dialogues and evidence-based interventions.

Furthermore, Archer's conceptual ideas illuminate the potential for individual agency to disrupt entrenched processes and initiate a process of morphogenesis, leading to transformative change. Through reflexive monitoring, individuals and communities can become aware of the biases and inequalities perpetuated by AI systems and take active steps to challenge and change these systems. Advocating for greater transparency in AI algorithms, pushing for the use of more diverse and representative training data, and developing new AI technologies that explicitly prioritise fairness and equity are examples of how his active engagement can foster morphogenesis, reshaping AI practises to promote greater equality (Eubanks, 2018, Dwivedi, 2021). However, it is crucial to acknowledge that such transformative change is bounded and contingent on broader societal structures and power dynamics. While individual agency can catalyse change, the transformation of deeply entrenched inequalities requires collective action and systemic change (Archer, 2012). Recognising the power of morphostatis, the resistance to progress by entrenched traditional forces, is crucial as it can manifest as backlash against demands for social justice and equality.

All in all, by highlighting the dynamic processes of morphostasis and morphogenesis, Margaret Archer's social realism framework does not only emphasise the potential for both the reproduction and transformation of inequalities within the realm of AI, but also underscores the need for ongoing vigilance and active engagement. This ensures that AI technologies are developed and deployed in ways that promote fairness, equity, and justice. Archer envisions dual mechanisms of change and reproduction which she termed as morphogenesis and morphostatis both of which are conditioned by socio-cultural dynamics and agentic processes of reflexive monitoring, engagement and reproduction.

3. Methodology

This study adopts an interdisciplinary approach, integrating insights from computer science, sociology, ethics, and law to investigate the relationship between AI technologies and societal inequalities. Each discipline contributes uniquely: computer science offers technical insights into AI mechanisms, sociology explores societal impacts and inequalities, ethics addresses normative implications, and law evaluates regulatory frameworks. This synthesis ensures a nuanced examination of the complex interplay between AI technologies and societal structures (Broussard, 2018; Baum, 2021).
The case studies for this research were selected based on their capacity to exemplify the multifaceted dynamics of AI-induced inequalities across diverse socio-economic, geographical, and technological contexts. This selection process followed a systematic and deliberate approach to ensure relevance, diversity, and representativeness.

The first step involved identifying relevant cases through a comprehensive review of academic literature, policy reports, and publicly available data to ensure a balanced perspective (Dafoe et al., 2021). Searches were conducted using databases such as Scopus, Web of Science, and Google Scholar, with keywords including but not limited to “AI and inequality,” “bias in AI,” “algorithmic bias,” “AI in education,” “predictive policing,” and “AI in economic systems.” This process resulted in an initial pool of cases that highlight the relationship between AI technologies and societal inequalities.

To refine the selection, specific evaluation criteria were applied. Geographical diversity was a primary consideration, with cases chosen to represent a variety of regions, including the Netherlands, Australia, the United States, Turkey, and the United Kingdom. This diversity allows the research to explore the global implications of AI technologies while considering distinct cultural, economic, and regulatory environments. Sectoral representation was also crucial, with cases drawn from welfare systems, financial services, education, law enforcement, and the gig economy. These sectors were selected because of their significant societal impact and their potential to reflect the systemic patterns of inequality influenced by AI. Furthermore, priority was given to cases with well-documented societal consequences, such as public controversies, policy failures, or measurable outcomes of inequality. This focus ensures that each case provides rich empirical insights and practical relevance. This diversity ensures a comprehensive understanding of the implications of AI across different settings. Additionally, each case provides sufficient publicly available data to facilitate robust analysis, enabling the study to draw meaningful conclusions (Obermeyer et al., 2019).

The selected cases were subsequently grouped into thematic categories based on the dimensions of inequality they illustrate. For instance, economic inequalities are highlighted through the examination of AI-driven financial services and gig economy platforms. Racial and cultural biases are evident in predictive policing algorithms, while access and opportunity divides emerge starkly in education systems. These thematic groupings serve to organise the analysis and provide a structured lens through which to understand the broader implications of AI technologies.

The final selection includes:

Welfare Systems: AI-driven fraud detection systems in the Netherlands and automated debt recovery systems in Australia illustrate how biases in algorithmic design and deployment can disproportionately impact vulnerable populations, exacerbating socio-economic inequalities.

Financial Services: In the United States, AI mortgage lending systems reveal the persistence of racial disparities in access to financial opportunities, reflecting systemic biases embedded in historical data.

Education: The use of AI algorithms for grading in the United Kingdom during the COVID-19 pandemic highlights how algorithmic decision-making can reinforce socio-economic disadvantages in educational outcomes.

Law Enforcement: Predictive policing systems in the United States demonstrate how AI can perpetuate racial biases, leading to discriminatory outcomes in minority communities.

Gig Economy: AI-driven platforms, such as those operating in Turkey's food delivery sector, showcase the dual role of AI in creating opportunities and deepening inequalities for gig workers.

These cases, while illustrative rather than exhaustive, provide critical insights into the ways AI technologies interact with societal structures to reinforce, perpetuate, or transform existing inequalities.

Margaret Archer's social realism framework provides the theoretical lens for analysing the interplay between powerful agents and socio-cultural structures. The concepts of morphogenesis (signifying change) and morphostasis (signifying continuity) within the context of AI technologies (Pasquale, 2015). This theoretical framework assists in understanding how AI systems can both perpetuate and be shaped by existing societal structures.

Margaret Archer's social realism, this research underscores the importance of examining both the continuity and change in social systems brought about by AI technologies.

The analysis maintained a reflexive stance, evaluating the researchers' positionality and its influence on interpretations (Crawford, 2016). Furthermore, collaboration with scholars from various disciplines aided in identifying potential biases and enriched the analysis (Eubanks, 2018).

Each case directly informs the research question, facilitating a focused examination of the issues at hand. The cases represent a variety of geographical and sectoral contexts, including welfare systems in the Netherlands and Australia, mortgage lending practices in the United States, and educational grading systems in the UK (Eubanks, 2018; Zarsky, 2016).

To further enhance clarity, each discipline's perspective—sociology, law, ethics, and computer science—contributes distinct insights that collectively address AI's societal impacts on inequality. We grouped case studies thematically based on dimensions of inequality, such as economic and racial biases, which facilitated a structured analysis of AI's varied societal implications. This thematic approach aligns with our goal of examining AI's role across diverse social contexts.

In the following sections, we present illustrative case studies that elucidate the complex interplay between AI technologies and societal inequalities. These case studies demonstrate how AI systems can reinforce existing disparities, drawing upon Margaret Archer's social realism framework to analyse the dynamics of power, governance, and resource allocation in each context.

4. AI Induced Scandals of Inequality: When Governance and Policy Fails

AI systems have the potential to reproduce, perpetuate and entrench structures of inequality that are already evident in their operational context. The rapid advancement of AI technologies in environments characterised by unaddressed social inequalities has brought about a series of scandals that underscore the urgent need for a comprehensive examination of the relationship between AI and societal inequalities. These scandals serve as a stark reminder of the potential for AI to perpetuate and exacerbate existing disparities if not properly regulated and scrutinised. A common thread across these cases is the lack of accountability and governance vacuum, which has prevented AI systems from undergoing the necessary legal and societal scrutiny.

4.1 Case Study 1: AI in Welfare Systems:

In the Netherlands, an AI system designed to detect fraud in welfare benefits led to a significant scandal when it was found to disproportionately target low-income families and immigrants, resulting in wrongful accusations of fraud and severe financial hardship for thousands of families (Taylor, 2020; Broussard, 2018). Similarly, the Dutch childcare benefit scandal highlighted the discriminatory impact of AI systems, with migrant families disproportionately denied access to benefits due to biased algorithmic decision-making (Politico, 2023). These cases exemplify a state of morphostatis, wherein the techno-cultural system reproduces and exacerbates existing social inequalities. While the departments that utilised these AI systems faced accountability, the AI systems themselves evaded critical public and legal examination, highlighting the urgent need for reflexivity in the development and deployment of AI systems.

In Australia, the infamous "robodebt" scandal saw an automated debt recovery system erroneously issue debt notices to thousands of welfare recipients, leading to widespread distress and financial hardship (Reuters, 2023). Although designed with the intent to streamline welfare debt recovery, the lack of adequate human oversight resulted in erroneous debt notices disproportionately impacting vulnerable populations. This case further emphasises the significance of reflexive monitoring in design and deployment of AI systems.

4.2 Case Study 2: AI in Mortgage Lending Practices

In the United States, several sectors have witnessed AI-related scandals that reflect broader sociocultural conditioning. In the mortgage industry, an AI system was found to exhibit racial bias, leading to 80% of Black mortgage applicants being denied loans, thereby exacerbating existing racial disparities in homeownership and wealth (Forbes, 2021). Likewise, in the healthcare, an AI algorithm used to allocate resource was similarly found to be racially biased, leading to Black patients being less likely to receive referrals for care compared to their white counterparts with similar health conditions (Obermeyer et al., 2019; ACLU, 2022). These biases highlight how unchecked AI systems may deepen social inequalities when they are left unregulated.

The media industry has also faced controversies related to AI, such as CNET pausing the publication of AI-generated stories following a controversy regarding the lack of transparency in their AI tools (The Verge, 2023). This incident highlights the ethical dilemmas surrounding the use of AI in content creation where lack of transparency and accountability can lead to significant reputational harm.

4.3 Case Study 3: AI in Educational Grading Systems

The impact of AI on education became glaringly evident during the Covid-19 pandemic, when all secondary education examinations were cancelled in 2020. In response, an algorithm was produced by the regulator to determine grades. However, due to its disproportionately negative impact on students from lower socioeconomic backgrounds, the regulators ultimately withdrew this algorithm. The A-level scandal in the UK exemplifies the potential for AI to reinforce existing inequalities, functioning as a form of socio-cultural conditioning in Archer’s framework. The algorithm’s design was influenced by long-lasting societal structures and cultural norms, which led to widespread public outcry – a manifestation of morphogenesis (social interaction). The goal of determining students' grades based on teachers’ predicted grades aimed to maintain qualification standards while ensuring distribution mirrored previous years (The Guardian, 19 August 2020). Consequently, the government's decision to abandon the algorithm represents an instance of morphostasis (structural elaboration), as it resulted in a shift in structural conditions that could have long-term implications for how technology is used in education.

These scandals illustrate the profound societal implications of AI technologies and the potential for these systems to perpetuate and exacerbate existing inequalities, underscoring the dual interplay of morphostasis and morphogenesis They emphasise the duality and interplay of morphostasis and morphogenesis, highlighting the urgent need for rigorous and reflexive scrutiny of AI systems, comprehensive regulatory frameworks, and a steadfast commitment to fairness, inclusivity, and justice in the design and deployment of these technologies.

5. Algorithmic Bias

Algorithmic bias has emerged as a critical concern in the AI sphere, where biased decision-making processes can reinforce existing societal inequalities. To contextualise this issue within Archer's social realism framework, we examine how algorithmic bias serves as a catalyst for both morphogenesis and morphostasis in societal structures. Morphogenesis is evident as AI systems introduce new forms of interaction and decision-making that alter traditional biases, potentially creating new societal norms. Conversely, morphostasis is observed when these systems perpetuate and solidify existing biases, reflecting entrenched societal inequalities.

AI systems, renowned for their ability to process vast amounts of data and make autonomous decisions, are not immune to the biases inherent in the data used to train them. If the training data contains biases related to race, gender, or other protected attributes, the AI system can endure and amplify those biases, leading to discriminatory outcomes (Zou & Schiebinger, 2018; Joyce et al., 2021). The nexus between AI and inequality, which is already a complex and multifaceted issue, deeply rooted in the biases inherent in the design, training, and usage of AI systems. These biases, which can be traced back to the data used to train AI models, the modelling techniques employed, and the interpretation of AI outputs, play a significant role in shaping the impact of AI on societal inequalities.

By inquiring into the complex relationship between algorithmic bias and AI systems, we shed light on how biases present in training data can permeate and amplify within the decision-making processes of AI systems. Specifically, we explore the profound implications of biases associated with race, gender, and other protected attributes, which have the potential to amplify discrimination within AI outcomes.

Through the lens of Archer’s theory, the replication of societal biases in AI systems underscores the concept of morphostasis, where the pre-existing social conditions, including prejudices and disparities, are embedded within the technological processes. This embedding process often goes unchecked due to the opaque nature of algorithmic decision-making, which obscures the biases from stakeholders and limits opportunities for reflexive monitoring.

5.1. Case Study 1: Predictive Policing Algorithms

In the sphere of law enforcement, predictive policing algorithms exemplify can perpetuate biases and lead to discriminatory outcomes. For instance, a notable case in a U.S. city demonstrated that an AI system used to predict crime hotspots disproportionately targeted minority neighbourhoods. This over-policing not only reinforced existing systemic biases within law enforcement (Joh, 2016; Asaro, 2019) but also raised serious ethical and legal implications regarding the deployment of AI technologies. Such outcomes highlight the urgent need for rigorous scrutiny of AI systems and comprehensive regulatory frameworks to promote fairness, inclusivity, and justice.

The amplification of biases predictive policing algorithms is concerning, as these systems are often trained on historical arrest data that reflects racial biases in policing practices. When an AI system is trained on such data, it may learn to associate certain neighbourhoods or demographic groups with higher crime rates, leading to biased policing practices that disproportionately affect marginalized communities (Meijer & Wessels, 2019). This not only perpetuates existing inequalities but also raises questions about the accountability and transparency of AI decision-making processes.

5.2. Case Study 2: AI in Hiring Practices

AI-based hiring systems illustrate the potential for algorithmic bias to sustain discriminatory practices. When trained on historical employment data, if this data reflects biases favouring certain gender or racial groups, the AI system may learn to associate specific jobs with these groups, thereby prioritizing candidates from these demographics. For example, if an AI system is trained on data that exhibits biases in favour of certain genders or races, it might perpetuate discriminatory hiring practices, leading to a lack of diversity in the workforce (Dastin, 2018; Raghavan et al., 2020).

Moreover, the modelling techniques used in AI can further contribute to bias amplification. If an AI model is designed to prioritise certain features over others, it may inadvertently reinforce societal biases, even if the training data itself is unbiased. The complexity and opacity of many AI models make it difficult to identify and address these biases, complicating efforts to ensure fairness and equity in AI systems (Ntoutsi et al., 2020). While the technical opacity of AI should not be used as a blanket excuse to evade scrutiny, it often obscures the biases from stakeholders and limits opportunities for reflexive monitoring.

Even in cases where AI systems are trained on unbiased data and designed with fairness in mind, biases can still emerge in the usage and interpretation of AI outputs. Additionally, if an AI system is used in a context where societal inequalities exist, it may inadvertently contribute to these inequalities, even if the system itself is unbiased (Holzinger et al., 2019). Moreover, AI based decision systems could suffer from scientism, discouraging users from demonstrating reflexive monitoring behaviours to identify biases due to the perception of scientific methods employed in designing these systems (Vassilopoulou et al. 2022).

5.3. Case Study 3: Content Recommendation Algorithms

AI algorithms used for content recommendation can amplify existing prejudices by creating echo chambers. These algorithms often prioritise content that aligns with users' existing views and interests, which can lead to the reinforcement of biases and stereotypes. For instance, if a user frequently engages with content that reflects certain prejudices, the AI system might learn to recommend similar content, thereby reinforcing and amplifying the user's existing biases (Akter et al., 2021). This phenomenon, known as the amplification of biases, poses a significant challenge in ensuring fairness, equity, and non-discrimination in automated decision-making processes. Furthermore, in the context of digital media, algorithmic bias can lead to the spread of misinformation and the polarization of public opinion. As content recommendation algorithms favour engagement over accuracy, they may perpetuate harmful stereotypes and reinforce societal divisions. This amplification of biases in content delivery highlights the necessity for ongoing research and dialogue to understand the complex interplay between AI and societal inequalities, and to develop effective strategies for mitigating the impact of AI on these disparities (Khakurel et al., 2018).

The amplification of biases through AI systems can be conceptualised as a form of morphogenesis, where new patterns of inequality emerge, reshaping societal landscapes in profound ways. The challenge lies in transforming these patterns through informed policy interventions and technological redesign that prioritize equity and justice, moving beyond mere recognition of biases to actively mitigating their impacts in societal applications.

In conclusion, addressing algorithmic bias requires a concerted effort that encompasses both morphogenesis and morphostasis. Strategies must consider the dual processes highlighted by Archer's framework: disrupting the continuity of entrenched biases (morphostasis) while fostering the emergence of more equitable AI practices (morphogenesis). Effective strategies will depend on robust reflexive monitoring mechanisms that not only detect biases but also enable the dynamic adaptation of AI systems in alignment with evolving societal values and norms.

6. Digital Divide: Access and Opportunity

In this analysis, we interpret the digital divide through the lens of Archer's social realism, focusing on how structural, cultural, and agentic dynamics interplay to shape access and opportunities. This divide is not merely a reflection of current technological gaps but also a manifestation of morphostasis, wherein existing societal inequalities are entrenched and perpetuated through new technological formats. The concepts of access and opportunity assumes a great significance also for the AI phenomenon. As AI technologies flourish, demanding access to data, computational resources, and technical expertise, a profound divide emerges between those poised to harness the benefits of AI advancements and those left at its periphery. This divide is not solely a technological issue; it is fundamentally a sociological one, reflecting the broader disparities present in our society. Such a divide is also deepened with the socio-technical division between technophile and technophobe segments of the society (Archer, 2021). Thus, it is essential to embark on an illuminating expedition into the conceptual underpinnings of AI and inequality, with a specific focus on the profound implications of the access and opportunity divide.

AI technologies necessitate access to robust digital infrastructure, reliable internet connectivity, and advanced technological devices. However, the reality of unequal access to these resources, often dictated by socio-economic factors, geographical location, and educational opportunities, creates a digital divide that mirrors and amplifies societal disparities (Robinson et al., 2015;

Lutz, 2019). This divide manifests as a chasm between those who can fully engage with and benefit from AI technologies and those who are left on the periphery. This divide is not merely about access to technology but extends to digital literacy – the skills and knowledge required to use digital technologies effectively and safely. As AI systems become more complex and integrated into daily life, the lack of digital literacy can further marginalise disadvantaged communities, limiting their ability to leverage the benefits of AI and participate in the digital economy (Heeks, 2022). This necessity highlights a morphogenetic shift, where the demands of new technology reshape social and economic landscapes, potentially transforming access paradigms. Simultaneously, the persistence of access disparities underscores morphostasis, as the pre-existing socio-economic stratifications continue to dictate the distribution of technological benefits.

For example, in India, the introduction of the Digital India initiative aimed to increase access to digital technologies. However, significant disparities remain, particularly in rural areas, where infrastructure and digital literacy are often inadequate. According to a report by the Internet and Mobile Association of India (IAMAI, 2022), the digital gender gap is pronounced, with women in rural areas having significantly less access to the internet and digital skills compared to their male counterparts. This situation highlights the necessity of targeted interventions to bridge these gaps and foster more equitable access to AI technologies.

Employing reflexive monitoring, we observe how individuals and communities actively engage with and respond to the digital divide. Through reflexive actions, such as advocating for policy changes or creating community-driven digital literacy programs, stakeholders strive to mitigate the perpetuation of this divide and foster more inclusive technological futures. The digital divide, illuminated by Archer's social realism framework, is not a static phenomenon but is subject to the interplay of structure, culture, and agency. Structural factors such as socio-economic status, education, and geographical location can limit access to digital resources and opportunities (structural constraint). Cultural norms and beliefs can influence attitudes towards technology and its use (cultural constraint). However, individuals and communities are not passive recipients of these constraints. They can exercise agency, engaging with and shaping their digital environments (agency effect).

For instance, Robinson et al. (2015) highlight how low-income parents in the United States navigate structural and cultural constraints to access and use digital technologies for their children's education, demonstrating agency in the face of digital inequality. Similarly, Lutz (2019) emphasises the role of digital skills training in empowering disadvantaged youth in South Africa, illustrating how targeted interventions can foster agency and bridge the digital divide. In South Africa, initiatives such as the Code for South Africa programme exemplify community-driven efforts to enhance access to technology and improve employment opportunities in the tech sector through targeted coding and digital literacy training for underprivileged youth.

Nevertheless, the potential of AI to increase the digital divide raises critical questions about the equitable distribution of AI benefits and the need for inclusive AI development. As AI systems become more pervasive, there is a growing need for policies and initiatives that ensure equal access to AI technologies and promote digital literacy. This includes efforts to democratise AI education, invest in digital infrastructure in underserved areas, and promote the co-creation of AI technologies with diverse communities to ensure that AI systems are inclusive and beneficial for all (Broussard, 2018; Eubanks, 2018).

In summary, the digital divide represents a significant aspect of AI-induced inequality, shaped by the interplay of structural and cultural constraints and individual agency. Addressing this divide necessitates a multifaceted approach that acknowledges the complexity of access and opportunity in the age of AI, ensuring that all individuals and communities can engage meaningfully with technological advancements.

6.1. The Prerequisite of Access

At the heart of the digital divide lies the fundamental prerequisite of access. The utilisation of AI technologies necessitates a sturdy digital backbone, comprising reliable internet connectivity and access to technological devices. However, disparities in access to these crucial resources generate a chasm that limits opportunities and widens the gap between those able to partake in AI's transformative potential and those left on the periphery. Socio-economic constraints, geographical remoteness, and limited digital infrastructure act as formidable barriers, further deepening the divide and hindering the ability of disadvantaged communities to harness the benefits of AI.

AI technologies thrive on a trifecta of essential ingredients such as data, computational prowess, and technical acumen. However, the glaring reality of unequal access to these fundamental resources casts a shadow of disparity upon the AI landscape (Lutz, 2019). The implications are profound, as those with limited access find themselves constrained by the barriers that prevent their meaningful engagement with AI advancements. Whether due to financial limitations, infrastructural gaps, or educational disparities, unequal access widens the chasm between those who can seize the transformative potential of AI and those who are relegated to the sidelines.

The Moral Machine experiment, a study conducted by Awad et al. (2018), highlights the importance of access to AI technologies. The experiment, which collected 40 million decisions from over 2.3 million people in 233 countries, dependencies, or territories, aimed to gauge social expectations about the way autonomous vehicles should solve moral dilemmas. The study emphasises the importance of access to AI technologies, not just in terms of data and computational resources, but also in terms of participation in shaping the ethical landscape and principles that guide AI behaviour.

6.2. Deepening Access Inequalities

In the absence of equitable access, the access and opportunity divide stand as a catalyst for deepening pre-existing inequalities. Disadvantaged communities, already burdened by socio-economic constraints and limited resources, bear the brunt of this divide. Inequitable distribution of access to AI technologies perpetuates systemic disparities, hindering the ability of these communities to fully participate in the benefits offered by AI advancements (Caradaica, 2020). The opportunity divide becomes pronounced in developing countries, where limited infrastructure, educational disparities, and socio-economic barriers hinder access to AI's transformative potential, further entrenching existing inequalities (Vinuesa et al., 2020).

As the digital divide widens, so too do existing inequalities. Disadvantaged communities find themselves disproportionately burdened by the repercussions of limited access to technology and digital literacy (Tinmaz et al., 2022). The digital divide serves as a trigger that intensifies societal disparities and perpetuates systemic inequality. Those without adequate access to technology face profound obstacles in educational opportunities (Meneses & Mominó, 2010), economic empowerment (Tang, 2022), and participation in the digital economy (Curran, 2018). As AI technologies continue to advance, failure to bridge the digital divide risks further marginalising these communities, exacerbating existing inequalities, and stifling progress toward a more inclusive and equitable society.

The repercussions of the access and opportunity divide within the realm of AI extend far beyond mere technological disparities. It amplifies existing socio-economic and educational inequalities, perpetuating a cycle of disadvantage and impeding progress towards a more equitable society (Pedro et al., 2019). Limited access to AI technologies hampers the capacity of disadvantaged communities and developing nations to leverage AI's power for social transformation, economic empowerment, and knowledge advancement. As the world witnesses the rapid evolution of AI, failing to bridge the access and opportunity divide risks leaving significant segments of society behind, exacerbating inequality and stifling progress.

The widening of this divide can be viewed through the framework of morphostasis, where traditional barriers in education, economic status, and geographic location continue to reinforce themselves in the digital age, limiting the transformative potential of AI technologies. This scenario demands a morphogenetic approach, where systemic changes are introduced to break these cycles of disparity.

To effectively address the challenges posed by the digital divide, a concerted effort embracing both morphogenesis and morphostasis is required. Policies and initiatives must not only aim to introduce new technologies in underserved areas but also need to transform the underlying social structures that govern access and equity. This includes fostering an environment where reflexive monitoring by all stakeholders - policymakers, technology developers, and community advocates - is a continuous process, ensuring that the evolution of AI technology is inclusive and equitable.

7. Economic Inequality: Case Studies in Transformation and Continuity

The integration of AI into various sectors of the economy has brought about significant shifts in the labour market and income distribution, exemplifying morphogenesis. Concurrently, it demonstrates morphostasis, as AI-driven growth benefits are unevenly distributed, reinforcing pre-existing economic disparities.

AI-driven automation disrupts industries and job markets, leading to job displacement and increasing income inequality (Goos et al., 2014; Goyal & Aneja, 2020). Low-skilled workers face greater challenges adapting (Schwabe & Castellacci, 2020), while those with AI-related skills benefit from increased job opportunities and higher wages (Acemoglu & Restrepo, 2019; Alekseeva et al., 2021). In regions with minimal or ineffective regulatory protections, AI's impact on technologies on blue collar workers can be severe. For example, the rapid growth of food delivery companies in Turkey, as highlighted by Erbil and Ozbilgin (2021) has resulted in companies - at the expense of worker health and safety- exemplifies how AI can exacerbate inequalities.

These disruptions can be seen as both a challenge and an opportunity for reflexive monitoring, requiring stakeholders to assess AI's impacts and adapt their strategies to mitigate negative outcomes. This adaptive process is essential to prevent the entrenchment of inequalities and to ensure that the benefits of AI are more broadly shared across the economy.

AI's integration into the gig economy has undeniably reshaped the landscape of work, creating new opportunities for income generation and economic growth. Particularly during the Covid-19 pandemic, it has enabled some disadvantaged communities, such as delivery staff, to secure income (Kamasak et al., 2019). However, this transformation has not been without its challenges and controversies; the benefits of this new economy are unevenly distributed, and the lack of stable employment associated with traditional jobs can exacerbate income inequality (Wood et al., 2019). The gig economy, facilitated by AI, represents a morphogenetic shift in the nature of work, offering unprecedented flexibility but highlighting the need for robust reflexive mechanisms to address the morphostasis evident in the increasing precariousness and vulnerability of gig workers.

The AI-powered gig economy has catalysed the creation of a new class of millionaires, particularly among those leveraging these technologies. Founders and early investors of gig economy giants like Uber, Lyft, and DoorDash have amassed significant wealth, highlighting the immense profitability of AI-driven platforms (Kenney & Zysman, 2016). However, gig workers, who are often classified as independent contractors, lack traditional employee protections and face income instability and inequality (Rosenblat & Stark, 2016).

Moreover, the concentration of power within big tech companies and their resistance to unionisation efforts further exacerbate these issues. Workers' attempts to organise and advocate for their rights are often met with resistance, and in some cases, retaliation. This dynamic hinders workers' ability to negotiate better working conditions, fair pay, and other protections (Duggan et al., 2020). This dynamic vividly illustrates the ongoing struggle between morphogenesis, which brings about new economic configurations and opportunities, and morphostasis, which sees the persistence of exploitative practices and concentrated power within a few tech giants. The challenge lies in harnessing the transformative potential of AI to foster economic equity, rather than merely replicating and amplifying existing inequalities.

In the prevailing discourse, the notion of co-ownership surfaces as a promising antidote to the challenges posed by the gig economy and AI-driven platforms. The involvement of disadvantaged communities in the development of AI technologies is critical, however, while these individuals from these communities are often solicited to co-design new AI technologies to encapsulate their needs, they are rarely invited to partake in the ownership structures of AI-led technological innovations. This creates a form of economic exclusion, as co-design does not necessarily translate into co-ownership. Individuals from non-traditional backgrounds are often instrumentalised in the co-design of AI-led products, which could potentially exacerbate their exclusion, otherness, and disadvantages (Ozbilgin, 2023). Promoting co-ownership and participatory development models in AI applications is an example of fostering morphogenesis by actively redesigning economic structures to be more inclusive and equitable. Such initiatives could significantly shift the balance of power, redistributing the economic gains from AI more fairly among those who contribute to its development and deployment. For instance, co-ownership models in the gig-economy, where workers have a stake in the companies they serve, could catalyse a redistribution of wealth and power. Empowering workers with a say in decision-making processes could address the pressing issues of worker exploitation and income disparity. However, the implementation of such models within the gig economy and AI-driven platforms is fraught with significant challenges and would likely necessitate regulatory backing and innovative business practices (Pendleton et al., 2018).

In this context, the gig economy and AI-driven platforms present a paradoxical landscape. On one hand, they democratise access to opportunities, reducing costs and barriers for producers and consumers alike. On the other hand, they create new privately controlled market choke points, with a "winner-takes-all" concentration of market share at the level of platform providers, often fostered by powerful interlocking increasing returns. This contradiction of centralisation through democratisation also manifests geographically. The "superstar effect" prevalent in knowledge-driven sectors frequently intensifies regional disparities instead of establishing an equitable competitive landscape facilitated by ICT (Haberly et al, 2019). The financial sector, characterised by its intrinsic reliance on information and its regulatory and organisational adaptability, is progressively adopting the digital platform model. However, the distinctive attributes of the financial industry have led to a divergent path compared to other sectors. The emergence of Digital Asset Management Platforms (DAMPs) and automation of derivatives exemplify this progression, fundamentally reshaping the market structure, yielding significant cost reductions for investors, and drastically altering established business models. However, the yields that investors make through AI led disruptive changes are adversely impacted by developers and users who contribute to innovation and design of these technologies.

These phenomena exemplify morphostasis, where existing social structures and inequalities are reproduced and reinforced through AI technologies. However, Archer's work also posits the potential for morphogenesis, or the active transformation of these structures through reflexive agency. This suggests the urge for a more inclusive and equitable approach to AI development and ownership, where disadvantaged communities are not only involved in the design process but also equitable share in the ownership and benefits of innovations.

As we explore the economic inequalities exacerbated by AI, it becomes evident that both morphogenesis and morphostasis are at play, reflecting the complex interplay of innovation and continuity within capitalist systems. By integrating Archer's insights into our analysis, we can more effectively identify opportunities for transformative change that not only address the symptoms but also the underlying causes of economic inequality. This entails a continuous commitment to reflexive monitoring and adaptation of AI technologies to serve broader societal goals, ensuring that progress in AI does not come at the cost of widening economic divides

8. Privacy and Surveillance

AI technologies, fuelled by the acquisition and examination of copious amounts of personal data, harbour the potential to disrupt the delicate balance between privacy rights and societal well-being. This potential disruption represents a clear instance of morphogenesis, where the innovative capabilities of AI redefine traditional boundaries and expectations around privacy and surveillance. Simultaneously, the societal impact of these technologies often exemplifies morphostasis, as existing inequalities are deepened through differential impacts on privacy across various communities. Hence, it is not merely a theoretical concern but a tangible reality that disproportionately impacts disadvantaged communities and individuals who are vulnerable to privacy breaches and the augmented reach of surveillance mechanisms. Inadequate safeguards and regulations can lead to privacy breaches and surveillance, disproportionately impacting disadvantaged communities and individuals who may already face discrimination and over-policing.

The COVID-19 pandemic has underscored the potential of AI technologies to address a broad range of biomedical, epidemiological, and socioeconomic challenges. However, it has also highlighted the ethical challenges that these technologies pose, particularly in terms of privacy and surveillance. As Leslie (2020) notes, the need for rapid and global action to combat the pandemic has necessitated unprecedented practices of open research and responsible data sharing. Yet, this urgency has also raised fears of 'surveillance creep' and challenged widely held commitments to privacy, autonomy, and civil liberties. The pandemic has accelerated the morphogenetic processes where AI technologies have rapidly integrated into public health strategies, reshaping norms around privacy and public health surveillance. This rapid integration challenges us to engage in reflexive monitoring to balance innovation with ethical considerations, ensuring that the rights to privacy and autonomy are maintained even as we combat global health crises.

In the rush to harness the power of AI technologies, there is a risk of reinforcing entrenched dynamics of societal inequity. These risks highlight the need for a reflexive approach to understanding and mitigating the impacts of AI on privacy, specifically in disadvantaged communities. Such an approach involves continuous assessment and adaptation of AI technologies to prevent morphostasis - where existing social inequalities are perpetuated and solidified through new technological means. For instance, digital contact tracing technologies, while potentially effective in controlling the spread of the virus, can also lead to privacy breaches and over-policing, particularly in disadvantaged communities, which are already affected by systematic discrimination (Leslie et al., 2021).

Moreover, the opacity of AI technologies can exacerbate these issues. Burrell (2016) identifies three forms of opacity in AI: opacity as intentional corporate or state secrecy, opacity as technical illiteracy, and opacity that arises from the characteristics of machine learning algorithms and the scale required to apply them usefully. This opacity can make it difficult for individuals and communities to understand how AI technologies are being used and to challenge or seek redress for any harm caused.

In the light of these challenges, it is crucial to develop and implement safeguards and regulations that protect individuals' privacy rights while also harnessing the potential of AI technologies to address societal challenges. As Leslie (2020) suggests, this requires a practice-based path to responsible AI design and discovery centred on open, accountable, equitable, and democratically governed processes and products. This development is a critical component of fostering morphogenesis in societal governance of technology, aiming to transform the landscape of AI development and deployment towards more equitable outcomes. It underscores the importance of incorporating democratic, transparent, and accountable processes in the design and implementation of AI systems to ensure they serve the broader societal good without compromising individual freedoms.

As we examine the intersection of AI, privacy, and surveillance through a scientific lens, it becomes imperative to elucidate the implications faced by disadvantaged communities and individuals who are vulnerable to privacy breaches. One reflexive engagement with the challenges of surveillance could involve recognising AI as a legal entity, akin to the legal status of a corporation. This legal responsibilisation of AI could allow contesting privacy breaches induced by AI systems.

8.1. The Data Dilemma

At the core of AI's prowess lies its insatiable appetite for personal data, a resource that fuels the learning and decision-making capabilities of AI systems. However, the acquisition and utilisation of such data can inadvertently compromise individual privacy, paving the way for potential breaches and encroachments. In the absence of robust safeguards and regulations, the collection, storage, and analysis of personal data can become a double-edged sword, increasing inequalities and amplifying existing power imbalances.

Inequities within AI systems become glaringly apparent when examining the disproportionate impact on marginalised communities and individuals who already grapple with discrimination and over-policing. Privacy breaches and the omnipresence of surveillance can reinforce existing inequalities, further marginalising those who are already disadvantaged. Vulnerable populations, often subject to systemic biases and social injustices, face heightened risks as their personal data becomes susceptible to exploitation and discriminatory practices. For instance, in South Africa, the roll-out of facial recognition technology has raised significant concerns regarding privacy and discrimination against black individuals, who are disproportionately affected by inaccuracies in these systems (Dlamini, 2021).

The implications of AI's data practices exemplify how morphostasis can be manifested in the digital sphere, with pre-existing vulnerabilities exploited and exacerbated by advanced technologies. Addressing this requires not only technical solutions but also a robust societal dialogue about the values we wish to uphold in the age of AI, promoting a morphogenetic shift towards greater justice and equity in digital privacy.

The amalgamation of AI, privacy, and surveillance has far-reaching implications for society. The unchecked proliferation of surveillance technologies, coupled with inadequate safeguards, can result in a chilling effect on individual freedoms, curtailment of civil liberties, and the erosion of privacy rights. Moreover, the impacts are not evenly distributed, amplifying the disparities that already permeate our social fabric. Disadvantaged communities, including racial and ethnic minorities, socioeconomically disadvantaged individuals, and vulnerable groups, are bear the brunt of the encroachment of surveillance mechanisms, reinforcing social inequalities and power imbalances.

As we pay more attention to the intersection of AI, privacy, and surveillance, it becomes clear that both morphogenesis and morphostasis are at play, reflecting the transformative potential and the risks of reinforcing existing disparities through new technologies. By employing Archer’s insights, we can better navigate these complex dynamics, advocating for an approach that ensures AI technologies foster societal advancement without compromising fundamental human rights. The challenge lies in maintaining a vigilant and adaptive stance - through reflexive monitoring - to ensure that as AI reshapes our societal landscape, it does so in a way that is inclusive, equitable, and respectful of all individuals' privacy and dignity.

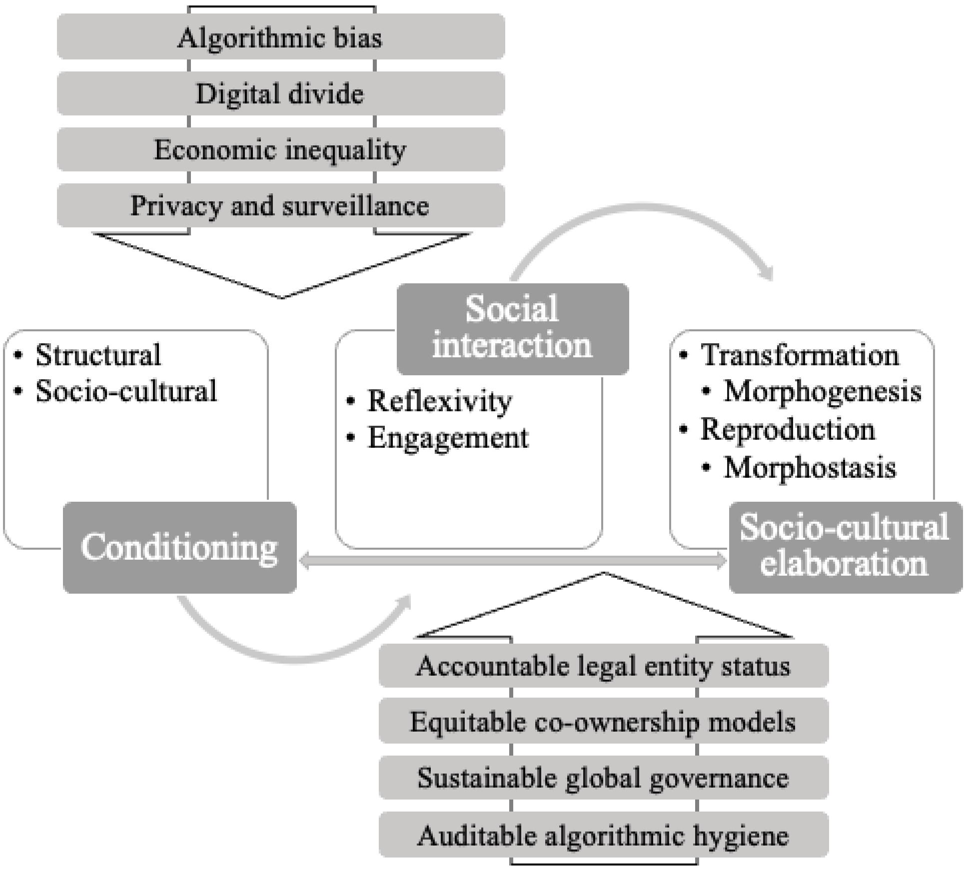
9. Critical Reflections and Discussion

In this study, we have rigorously applied Archer’s comprehensive concepts of sociocultural conditioning, morphogenesis, morphostatis, reflexive monitoring and reflexive action to elucidate the complex interconnections and transformative potential of AI and social inequalities. Through detailed examination, we have illuminated how both individual and collective reflexive engagements, driven by sociocultural conditioning, shape the outcomes of AI technologies.

Our research has identified four key social conditioning mechanisms that significantly lead to the entrenchment of existing inequalities and emergence of new forms of disparity within AI-led systems. in Figure 2 illustrates the interaction between these conditioning mechanisms and the structural and socio-cultural factors underpinning AI's societal impact.

We delineate four primary dimensions of intervention—accountable legal entity status, equitable co-ownership models, sustainable global governance, and auditable algorithmic hygiene—each designed to address specific challenges posed by AI, proposing robust strategies for mitigating inequality and enhancing societal well-being.

Figure 2: Major determinants of AI and inequality nexus and mitigation strategies



Note: Authors' illustration based on Archer's social realism framework.

1. Accountable Legal Entity Status: This intervention, aligning with morphostasis, aims to institutionalise responsibility within AI systems by creating a legal framework that holds AI accountable, akin to corporate entities (Ozbilgin, 2024). This legal grounding adapts static societal structures - such as laws and regulations - to the dynamic and evolving nature of AI, thereby enhancing reflexive monitoring and providing clear pathways for legal recourse and accountability. It directly addresses the issues raised by privacy and surveillance concerns, where AI's role in data handling and personal privacy necessitates clear legal boundaries and responsibilities.

2. Equitable Co-ownership Models: Representing Archer's morphogenesis, this approach proposes a transformation in the ownership and profit-sharing structures of AI enterprises. It challenges traditional hierarchies and power imbalances in technology development, fostering socio-economic changes that enable disadvantaged groups to benefit from AI advancements

(Ozbilgin and Erbil 2023). This model is particularly relevant to addressing economic inequality, as it promotes a more equitable distribution of the economic benefits derived from AI.

3. Sustainable Global Governance: Reflecting both morphogenetic and morphostatic elements, this strategy involves creating robust, adaptable frameworks that not only respond to the evolving challenges posed by AI but also solidify global standards that protect human rights and ensure ethical usage of AI across different societies. This governance is crucial in managing the digital divide, ensuring that AI technologies are accessible and beneficial across diverse global communities.

4. Auditable Algorithmic Hygiene: Directly linked to reflexive monitoring, this practice involves the continuous scrutiny and revision of AI algorithms to prevent and address biases, ensuring that AI operations are transparent and align with ethical standards. This approach operationalises socio-cultural conditioning by influencing the technical and operational aspects of AI systems, ensuring they evolve in ways that do not perpetuate existing social prejudices or inequalities. It is pertinent to algorithmic bias, as it seeks to ensure fairness and non-discrimination in AI applications.

These interventions serve as critical pathways for addressing the complexities of AI-induced inequalities and suggesting mechanisms for their active mitigation. Our findings reveal that AI systems, often trained on inherently biased data, have the potential to reproduce and amplify these biases across various domains, such as hiring processes (Raghavan et al., 2020), criminal justice systems (Završnik, 2020), and social media algorithms (Akter et al., 2021), resulting in discriminatory practices and unequal treatment. This propagation of bias highlights the critical importance of addressing foundational issues in the data used to train AI systems and ensuring that these systems are designed and deployed in a manner that promotes fairness and equity. It also emphasises the need for ongoing monitoring and evaluation of AI systems - morphogenesis- to identify and mitigate potential discriminatory outcomes.

Unmasking algorithmic bias within AI systems is of paramount importance in our quest for ethical and fair AI deployment. Understanding how biases embedded in training data can amplify discrimination is essential for developing strategies to mitigate and eliminate such biases. By addressing algorithmic bias head-on, we can work towards building AI systems that are more equitable, inclusive, and aligned with societal values. It is imperative for researchers, policymakers, and stakeholders to collaboratively explore methods and frameworks to reduce bias and ensure *auditable algorithmic hygiene* that will contribute to a just and unbiased future. In line with Vassilolopulou et al. (2022), we propose that algorithmic hygiene could be integrated in the design of AI systems to bias-proof AI.

10. Conclusions

As we embark upon the dynamic landscape of AI and social inequalities, a sustained commitment to interventions outlined is essential. Future research should focus on developing innovative methods to enhance algorithmic transparency, construct equitable AI governance frameworks, and expand access to AI technologies across diverse communities. Bridging this divide requires concerted efforts to address the barriers hindering access to data, computational resources, and technical expertise.

Sustainable global governance of big data, AI systems and their deployments should foster inclusivity, promote digital literacy, and develop supportive infrastructures. Only then, we can endeavour to narrow the access and opportunity gap, fostering a future where AI technologies are harnessed as a force for equitable progress. Through interdisciplinary collaboration (Greenhalgh et al., 2022) and a commitment to social justice, we can pave the way toward a society that ensures equal access, equal opportunities, and equal benefits from AI advancements.

Emerging Impacts of AI: Deepfakes and Amplification of Inequalities

The rapid evolution of AI technologies has given rise to tools such as deepfakes, which exemplify the dual-edged nature of these advancements. Deepfakes, or AI-generated synthetic media, pose significant threats to public trust, political stability, and personal security. By convincingly mimicking real individuals, deepfakes have been weaponised to spread misinformation and disinformation, exacerbating societal divisions and undermining democratic processes (Kharvi, 2024). Vulnerable groups, including women and marginalised communities, face disproportionate harm, particularly through non-consensual synthetic media (Vaccari & Chadwick, 2020).

These developments emphasise the structural power imbalances inherent in AI's design and deployment, where certain demographics face heightened risks. Deepfakes serve as a compelling example of morphostasis, perpetuating existing inequalities while simultaneously creating new societal disruptions.

To mitigate these risks, robust legal frameworks are necessary. Policies addressing content authenticity, platform accountability, and criminal liability for malicious actors are critical (Acemoglu, 2021; Kanzola et al., 2024). These measures must prioritise transparency and ethical governance to ensure AI technologies do not undermine democratic values.

Broader Implications for AI Governance

The increasing influence of AI technologies demands robust regulatory measures to address their societal risks while maximising their benefits. Transparency and accountability should underpin AI governance. Developers must disclose algorithms, training data, and decision-making processes to enable oversight and mitigate harm. Furthermore, collaborative efforts among governments, international organisations, and industry stakeholders are necessary to establish ethical AI standards that prioritise fairness, inclusivity, and non-discrimination. Unified global regulations are essential for bridging regulatory gaps and upholding ethical principles across jurisdictions.

Addressing these challenges requires a dual approach. First, robust data governance frameworks must safeguard privacy and ensure the ethical use of personal information. Second, public awareness campaigns should mitigate the misuse of AI tools such as deepfakes and promote digital literacy.

Economic Inequality and the Co-Ownership Paradigm

The AI led disruptive technologies and AI led platform, e-commerce and gig economies have been the dominant force of economic growth in industrialised countries. While economic growth of AI led commerce has revolutionised many sectors of work, its impact on different communities has been ambivalent, in terms of job creation, and job destruction. Further, the wealth generated through AI led commerce has been hoarded by owners of AI technologies and investors in these businesses. Polarisation of wealth in the hands of owners and investors has deepened economic inequalities as a result of AI led disruption in economic systems.

To counteract these inequities, we propose a shift from the co-design paradigm to a *co-ownership* paradigm. Under this model, stakeholders contributing to AI development would have significant claims of ownership, redistributing wealth more equitably. This approach aligns with the pursuit of societal equality and justice, ensuring that AI serves as a tool for inclusive economic growth.

In conclusion, our application of Archer's social realism framework has not only provided significant insights into the nexus of AI and inequality but has also offered both theoretical and practical contributions to the field. Recognising that AI systems are not neutral or detached entities but are enmeshed within social systems and structures, we examine the ways in which power dynamics influence the development, deployment, and utilisation of AI technologies. We scrutinise the distribution of power and the hierarchies that shape the access to and control over AI resources, acknowledging that disparities in access can perpetuate existing inequalities and hinder the potential of AI to contribute to societal well-being. By addressing these critical issues, we contribute to a more equitable future where AI serves as an ally in the pursuit of social justice.

References

Acemoglu, D. (2021). Harms of AI (No. w29247). National Bureau of Economic Research. https://doi.org/10.3386/w29247

Acemoglu, D., & Restrepo, P. (2019). Artificial Intelligence, Automation, and Work. In *The Economics of Artificial Intelligence* (pp. 197-236). University of Chicago Press.

ACLU (2022). Algorithms in Health Care May Worsen Medical Racism. Retrieved from https://www.aclu.org/news/privacy-technology/algorithms-in-health-care-may-worsen-medical-racism

Akter, S., McCarthy, G., Sajib, S., Michael, K., Dwivedi, Y. K., D'Ambra, J., & Shen, K. N. (2021). Algorithmic bias in data-driven innovation in the age of AI. *International Journal of Information Management*, *60*, 102387.

Albert, E. T. (2019). AI in talent acquisition: a review of AI-applications used in recruitment and selection. *Strategic HR Review*, *18*(5), 215-221.

Alekseeva, L., Azar, J., Gine, M., Samila, S., & Taska, B. (2021). The demand for AI skills in the labor market. *Labour economics*, *71*, 102002.

Archer, M.S. (1995) *Realist social theory: The morphogenetic approach*. Cambridge, UK: Cambridge University Press.

Archer, M.S. (2003). *Structure, agency and the Internal conversation*. Cambridge, UK: Cambridge University Press.

Archer, M.S. (2007) *Making our way through the world: Human reflexivity and social mobility.* Cambridge, UK: Cambridge University Press.

Archer, M.S. (2008) *The internal conversation: Mediating between structure and agency: Full research repor*t. Swindon: ESRC.

Archer, M.S. (2012) *The reflexive imperative*. Cambridge University Press, Cambridge

Archer, M.S. (2016) 'Reconstructing sociology: The critical realist approach', J*ournal of Critical Realism*, 15(4), pp.425-431.

Archer, M. S. (2021). Friendship Between Human Beings and AI Robots?. *Robotics, AI, and Humanity: Science, Ethics, and Policy*, 177-189.

Asaro, P. M. (2019). AI ethics in predictive policing: From models of threat to an ethics of care. *IEEE Technology and Society Magazine*, *38*(2), 40-53.

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., ... & Rahwan, I. (2018). The moral machine experiment. *Nature*, 563(7729), 59-64.

Baum, S. D. (2021). Artificial interdisciplinarity: Artificial intelligence for research on complex societal problems. *Philosophy & Technology*, *34*(Suppl 1), 45-63.

Bircan, T., & Korkmaz, E. E. (2021). Big data for whose sake? Governing migration through artificial intelligence. *Humanities and Social Sciences Communications*, 8(1), 1-5.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Buhmann, A., & Fieseler, C. (2021). Towards a deliberative framework for responsible innovation in artificial intelligence. *Technology in Society*, *64*, 101475.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big data & Society*, 3(1), 2053951715622512.

Crawford, K. (2016). Can an Algorithm be Agonistic? Ten Scenes from Life in Calculated Publics. *Science, Technology, & Human Values*, 41(1), 77-92.

Caradaica, M. (2020). Artificial Intelligence and Inequality in European Union. *Europolity-Continuity and Change in European Governance*, *14*(1), 5-31.

Curran, D. (2018). Risk, innovation, and democracy in the digital economy. *European journal of social theory*, *21*(2), 207-226.

Dafoe, A., Bachrach, Y., Hadfield, G., Horvitz, E., Larson, K., & Graepel, T. (2021). Cooperative AI: machines must learn to find common ground. *Nature*, *593*(7857), 33-36.

Dastin, J. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. Reuters.

Dlamini, N. J. (2021). Gender-based violence, twin pandemic to COVID-19. *Critical Sociology*, *47*(4-5), 583-590.

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., ... & Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57, 101994.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Fleming, P. J. (2019). Robots and Organization Studies: Why Robots Might Not Want to Steal Your Job. *Organization Studies*, 40(8), 1149-1161

Forbes (2021). AI Bias Caused 80% of Black Mortgage Applicants to be Denied. Retrieved from https://www.forbes.com/sites/korihale/2021/09/02/ai-bias-caused-80-of-black-mortgage-applicants-to-be-denied/

Fountain, J. E. (2022). The moon, the ghetto and artificial intelligence: Reducing systemic racism in computational algorithms. *Government Information Quarterly*, *39*(2), 101645.

Goos, M., Manning, A., & Salomons, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring. American Economic Review, 104(8), 2509-26.

Goyal, A., & Aneja, R. (2020). Artificial intelligence and income inequality: Do technological changes and worker's position matter?. *Journal of Public Affairs*, *20*(4), e2326.

Greenhalgh, T., Ozbilgin, M., & Tomlinson, D. (2022). How COVID-19 spreads: narratives, counter narratives, and social dramas. *bmj*, 378.

Haberly, D., MacDonald-Korth, D., Urban, M., & Wójcik, D. (2019). Asset management as a digital platform industry: A global financial network perspective. *Geoforum*, 106, 167-181.

Heeks, R. (2022). Digital inequality beyond the digital divide: conceptualizing adverse digital incorporation in the global South. *Information Technology for Development*, *28*(4), 688-704.

Holzinger, A., Plass, M., Kickmeier-Rust, M. D., Holzinger, K., Crisan, G. C., Pintea, C. M., & Palade, V. (2019). Interactive machine learning: experimental evidence for the human in the algorithmic loop. *Applied Intelligence*, 49(7), 2401-2414.

IAMAI. (2022). India Internet 2022. Internet and mobile association of India. Retrieved from https://core.ac.uk/download/pdf/382871343.pdf.

Joh, E. E. (2016). The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing. *Harvard Law & Policy Review*, 10, 15.

Joyce, K., Smith-Doerr, L., Alegria, S., Bell, S., Cruz, T., Hoffman, S. G., ... & Shestakofsky, B. (2021). Toward a sociology of artificial intelligence: A call for research on inequalities and structural change. *Socius*, *7*, 2378023121999581.

Kadiresan, A., Baweja, Y., & Ogbanufe, O. (2022). Bias in AI-based decision-making. In *Bridging Human Intelligence and Artificial Intelligence* (pp. 275-285). Cham: Springer International Publishing.

Kamasak, R., Özbilgin, M. F., Yavuz, M., & Akalin, C. (2019). Race discrimination at work in the United Kingdom. In *Race discrimination and management of ethnic diversity and migration at work: European countries' perspectives* (pp. 107-127). Emerald Publishing Limited.

Kanzola, A. M., Papaioannou, K., & Petrakis, P. (2024). Unlocking society's standings in artificial intelligence. *Technological Forecasting and Social Change*, *200*, 123106.

Khakurel, J., Penzenstadler, B., Porras, J., Knutas, A., & Zhang, W. (2018). The rise of artificial intelligence under the lens of sustainability. *Technologies*, 6(4), 100.

Kharvi, P. L. (2024). Understanding the Impact of AI-Generated Deepfakes on Public Opinion, Political Discourse, and Personal Security in Social Media. *IEEE Security & Privacy*, 22(4), pp. 115-122, July-Aug. 2024, doi: 10.1109/MSEC.2024.3405963.

Khogali, H. O., & Mekid, S. (2023). The blended future of automation and AI: Examining some long-term societal and ethical impact features. *Technology in Society*, 73, 102232.

Leslie, D. (2020). Tackling COVID-19 through responsible AI innovation: Five steps in the right direction. *Harvard Data Science Review*, 10.

Leslie, D., Mazumder, A., Peppin, A., Wolters, M. K., & Hagerty, A. (2021). Does "AI" stand for augmenting inequality in the era of covid-19 healthcare?. *BMJ*, 372.

Lutz, C. (2019). Digital inequalities in the age of artificial intelligence and big data. *Human Behavior and Emerging Technologies*, *1*(2), 141-148.

Meijer, A., & Wessels, M. (2019). Predictive policing: Review of benefits and drawbacks. *International Journal of Public Administration*, *42*(12), 1031-1039.

Meneses, J., & Mominó, J. M. (2010). Putting digital literacy in practice: How schools contribute to digital inclusion in the network society. *The Information Society*, *26*(3), 197-208.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 2053951716679679.

Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M. E., ... & Staab, S. (2020). Bias in data-driven artificial intelligence systems—An introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *10*(3), e1356.

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447-453.

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.

Ozbilgin, M. F. (2024). Diversity: A key idea for business and society. Taylor & Francis.

Ozbilgin, M., & Erbil, C. (2021). Post-hümanist İnovasyon: Gig Ekonomi Özelinde Moto Kuryeli Teslimat Sektörü Örneklemi. *Sosyal Mucit Academic Review*, *2*(1), 22-41.

Özbilgin, M. F., & Erbil, C. (2023). Insights into Equality, Diversity, and Inclusion. In Contemporary Approaches in Equality, Diversity and Inclusion: Strategic and Technological Perspectives (pp. 1-18). Emerald Publishing Limited.

Pedro, F., Subosa, M., Rivas, A., & Valverde, P. (2019). *Artificial intelligence in education: Challenges and opportunities for sustainable developmen*t.

Pendleton, A., Poutsma, E., Van Ommeren, J., & Brewster, C. (2003). The incidence and determinants of employee share ownership and profit sharing in Europe. In *Advances in the Economic Analysis of Participatory & Labor-Managed Firms* (Vol. 7, pp. 141-172). Emerald Group Publishing Limited.

Politico (2023). A Dutch Algorithm Scandal Serves a Warning to Europe. Retrieved from https://www.politico.eu/newsletter/ai-decoded/a-dutch-algorithm-scandal-serves-a-warning-to-europe-the-ai-act-wont-save-us-2/

Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020, January). Mitigating bias in algorithmic hiring: Evaluating claims and practices. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 469-481).

Reuters (2023). Australia's 'Robodebt' Scandal Highlights Dangers of Automated Welfare. Retrieved from https://www.reuters.com/article/australia-tech-ai-idUKL8N3340SN

Robinson, L., Cotten, S. R., Ono, H., Quan-Haase, A., Mesch, G., Chen, W., ... & Stern, M. J. (2015). Digital inequalities and why they matter. *Information, communication & society*, *18*(5), 569-582.

Rosenblat, A., & Stark, L. (2016). Algorithmic labor and information asymmetries: A case study of Uber’s drivers. *International journal of communication*, *10*, 27.

Schwabe, H., & Castellacci, F. (2020). Automation, workers’ skills and job satisfaction. *Plos one*, *15*(11), e0242929.

Tang, C. S. (2022). Innovative technology and operations for alleviating poverty through women's economic empowerment. *Production and Operations Management*, *31*(1), 32-45.

The Guardian (2020). Ofqual Ignored Exams Warning a Month Ago Amid Ministers' Pressure. Retrieved from https://www.theguardian.com/politics/2020/aug/19/ofqual-was-warned-a-month-ago-that-exams-algorithm-was-volatile

The Verge (2023). CNET Pauses AI Articles After Controversy Over Lack of Transparency. Retrieved from https://www.theverge.com/2023/1/20/23564311/cnet-pausing-ai-articles-bot-red-ventures

Tinmaz, H., Lee, Y. T., Fanea-Ivanovici, M., & Baber, H. (2022). A systematic review on digital literacy. *Smart Learning Environments*, *9*(1), 1-18.

Vaccari, C., & Chadwick, A. (2020). Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Social Media + Society*, 6(1).

Vassilopoulou, J., Kyriakidou, O., Özbilgin, M. F., & Groutsis, D. (2022). Scientism as illusio in HR algorithms: Towards a framework for algorithmic hygiene for bias proofing. *Human Resource Management Journal*.

Vice (2020). How a Discriminatory Algorithm Wrongly Accused Thousands of Families of Fraud. Retrieved from https://www.vice.com/en/article/jgq35d/how-a-discriminatory-algorithm-wrongly-accused-thousands-of-families-of-fraud

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., ... & Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature communications*, *11*(1), 233.

Wired UK (2020). Welfare Algorithms Could Soon be Labelled Discriminatory. Retrieved from https://www.wired.co.uk/article/welfare-algorithms-discrimination

Wood, A. J., Graham, M., Lehdonvirta, V., & Hjorth, I. (2019). Networked but commodified: The (dis)embeddedness of digital labour in the gig economy. *Sociology*, 53(5), 931-950.

Zarsky, T. (2016). The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making. *Science, Technology, & Human Values*, 41(1), 118-132.

Završnik, A. (2020, March). Criminal justice, artificial intelligence systems, and human rights. In *ERA Forum* (Vol. 20, No. 4, pp. 567-583). Berlin/Heidelberg: Springer Berlin Heidelberg.

Zou, J., & Schiebinger, L. (2018). AI can be sexist and racist—it's time to make it fair. *Nature*, 559(7714), 324-326.