AI Policy for Whom? Reclaiming Governance from Capitalist Capture

Petter Ericson[1], Rachele Carli[1], Jason Tucker[2], Virginia Dignum[1]

[1]Department of Computing Science, Umeå University, Umeå, Sweden
[2]Institute for Future Studies, Stockholm, Sweden
pettter@cs.umu.se, rachele.carli@umu.se, jason.tucker@iffs.se, virginia@cs.umu.se

Abstract

Contemporary AI policy is dominated by hegemonic neoliberal ideology, embedding assumptions of individualism, rationality, and market fundamentalism into its regulatory frameworks. This is evident in major policy efforts (e.g., the EU AI Act or the OECD principles) which prioritize economic growth and innovation over justice, equity, and collective welfare, and in the current policy landscape that favors market incentives and private sector leadership while sidelining democratic control and structural critique. This paper questions these prevailing paradigms and exposes how they reflect and reinforce capitalist power structures through corporate lobbying, the pursuit of specific kinds of AI models motivated primarily by usefulness to capital, and the externalization of social and environmental costs. We argue that effective AI governance must confront, rather than accommodate, capitalist interests. Drawing on legal and political theory, we propose an explicitly anti-capitalist approach to AI policy, that centers on social well-being, redistributive justice, and democratic control over technological infrastructures. In doing so, we outline essential counter-balancing policy approaches to reclaim AI governance from capitalistic capture and advance just and sustainable technology futures.

Introduction

AI governance today is shaped less by democratic deliberation than by the imperatives of capital, embedding neoliberal assumptions into the very fabric of regulatory frameworks. These hegemonic world views pose serious challenges to justice, equity, and democratic accountability. They simultaneously fail to recognize and exacerbate the reality that capital (i.e. specific large corporations) have material interests and inherent dynamics that run counter to societal and individual benefits. This is true even if one's goal is the perpetuation of capitalism as an underlying system. Within AI this impact is particularly pernicious, being a result of the targeted framing of the danger of AI governance, under the mantra of "regulation stifles innovation" (Schaake 2024) .

More specifically, despite a growing body of AI ethics and governance initiatives (Klein and Patrick 2024), most contemporary AI policy remains rooted in neoliberal ideology, prioritizing market efficiency, shareholder profits, unchecked innovation, and self-regulation, often at the expense of structural justice, democratic control, and collective well-being. This ideological foundation shapes how rights, risks and responsibilities are implemented in AI systems, reinforcing the influence of capital and sidelining the public interest (Schaake 2024). Moreover, current policy discourse lacks a fundamental interrogation of these ideological assumptions and rarely considers anti-capitalist, or even non-neoliberal alternatives to technology governance. We argue that taking a broader view is essential and a critical element in the design of public policy. Its absence is highly problematic and speaks to the need to reassess the worldviews shaping theses policies.

This paper interrogates these prevailing paradigms and exposes how they reflect and reinforce capitalist power structures through corporate lobbying, the pursuit of specific kinds of AI models motivated primarily by usefulness to capital, and the externalization of social and environmental costs. We argue that effective AI governance must confront, rather than accommodate, capitalist interests. As a means to provide counter hegemonic narratives, we draw on legal and political theory, to propose an explicitly anti-capitalist approach to AI policy. This approach takes social well-being, redistributive justice, and democratic control over technological infrastructures as its ideological bedrock. In doing so, we outline an essential counter balance to support regulators and their academic advisors to reclaim AI governance from capitalistic capture and advance just and sustainable technology futures.

This forms the motivation to our research questions:

- How does neoliberal ideology manifest in current AI policy frameworks, and what assumptions about individuals, rationality, and the role of the market underlying these approaches?
- In what ways do existing AI policies reflect and reinforce capitalist power structures, particularly through mechanisms like corporate responsibility, private-sector leadership, and technological design?
- What would an explicitly anti-capitalist approach to AI policy look like?

The State of Global AI Governance and the Regulatory Turn

Unlike other historical phases in AI hype, the current one has spurred what can be termed a "regulatory turn" in AI. In the last decade we have seen a massive growth in high level principles, guidelines, horizontal soft law, industry self-regulation and standards in the AI governance ecosystem (World Bank Group 2024; Walter 2024). At the same time, we are witnessing the development of a more layered and complex regulatory landscape, comprising both horizontal laws (broad rules that apply across sectors) and vertical laws (sector-specific regulations), being enacted at local, national, and international levels. The EU AI Act, South Korea's national AI legislation, China's vertical AI governance, and even local examples such as the state of Colorado in the USA point to this trend (Shimpo 2025; Sheehan 2024; Siegal and Garcia 2024; World Bank Group 2024). This is despite recent efforts to stymie this development, e.g. through a 10-year moratorium on further state-level AI regulations proposed by the Trump administration in the USA or the strong corporate resistance to the EU AI ACT and related legislation (Politico 2025).

The development of this multi-layered web of AI governance may seem positive to those who see regulation as vital to innovation and to realizing the social good of AI. Indeed, there is a growing body of work highlighting the dangers and of unregulated AI development and deployment, and how the costs of underregulated AI development and deployments are often shouldered by already marginalized and disenfranchised members of society (Noble 2018; Crawford 2021; Schaake 2024).

Yet, the dominant discourses guiding these governance efforts are very narrow. They frame the purpose of AI governance as being to facilitate the realization of the benefits of AI adoption, while keeping the associated risks limited to an acceptable level. Rarely are red lines around AI use drawn, and for the cases where they are, exceptions are often made, e.g. for scientific research, or for national security. This is even the case in very high risk areas, such as healthcare (Tucker 2023).

This approach has been challenged for its perceived perpetuation of uncritical techno-solutionism (Lindgren 2023), i.e. the belief that technology alone can solve complex social, political, or environmental problems, without considering deeper issues or potential negative impacts. Building on this work, we argue that we are at a critical juncture where there is a pressing need to interrogate the underlying assumptions of AI governance from an anti-capitalist perspective. Failure to do so will mean that the prevailing paradigms in AI governance will continue to reflect and reinforce capitalist power structures to the detriment of society at large. To critically engage with this tension we focus on how the hegemonic neoliberal ideology embeds assumptions of individualism, rationality, and market fundamentalism into its regulatory frameworks. By doing so we are able to identify and propose counter-hegemonic policy tools, yielding AI governance that is fit for purpose.

Hegemonic Neoliberal Models

The current regulatory landscape for AI is largely based on legal traditions derived from a colonialist West, which implies a very specific underlying world-view. Successive economic and social movements have modified this tradition, further entrenching certain parts of it in global treaties and institutions. A neoliberal cultural paradigm is reflected in certain principles and models that characterize legal architecture, making the ideology both descriptive and prescriptive. In the following, we discuss a number of such underlying concepts, which figure in particular in EU and US legislation and regulatory frameworks on AI, though they are also prevalent globally (OECD, UN, G7, WTO, South Korea, China, etc.).

It is crucial to underscore that the neoliberal economic tradition, which advocates for the autonomy of the market from external regulation, inherently incorporates a regulatory approach, that ends up favouring the already dominant actors. States intending to adopt, support, or enhance such a model are required to implement specific legislative policy. In fact, at the policy-making level, even the decision not to regulate is a legal choice. Therefore, even more so is the choice of how to regulate, and thus to prepare a set of norms, but also of legal categories, that make it possible to realise the socio-political end to be pursued. Indeed, law as well is both descriptive and prescriptive (Braybrooke 1990). Descriptive laws reflect the state of society in which they operate and are produced, including the values and priorities perceived or selected. Prescriptive laws use a set of rules to steer the *status quo* in some desired direction. In particular, a cross-sectional analysis of Western lawmaking highlights how its central pillars are the concepts of individualism and rationality.

The Individual

Though later reinforced and expanded under neoliberalism, with its prioritization of free markets, deregulation, and minimal state intervention, individualism has been central to liberal thought since John Locke, who claimed that human beings were endowed with innate rights by nature (Locke 1990). This conception was taken up and reinforced following World War II with the drafting of the UN Declaration of Human Rights, which reaffirms in Article 1 that all human beings are born equal in dignity and rights (United Nations General Assembly 1949). While the programmatic scope of these norms is certainly of central importance, it is easy to see that this is more a goal yet to be achieved than a real ontological state of humanity. As the term implies, fundamental rights are defined in the legal scholarship as those that should not be susceptible to renunciation even by their possessor (Henkin 1989; Kaufmann 2009). In fact, the principle that, with the consent of the holder, the right can be waived or negotiated with another does not apply to them. Nevertheless, it is easy to note that the very means of claiming their respect or fighting against their abuse are often difficult to implement if one does not already possess the starting conditions that facilitate access to justice, such as the right to appear in court, or even just the cultural substratum

sufficient to even conceive of the need for such access (Sen 2005). In other words, the real capacity to enjoy fundamental rights even today still too often depends on the specific characteristics of the individual whose rights are being violated. This has led some legal scholars to point out that liberal rights are in fact conceived, applied and protected much more in their individual dimension than in the universal and collective dimension in which they were theorised (Fineman 2013). Thus, we are faced with the rights of *one* enforced potentially against *all others*, rather than universal rights enjoyed by everyone.

This conception has undoubtedly nurtured the liberal-capitalist culture that makes competition — both between actors and in the market itself — the main drive towards self-regulation of economic exchange and growth. Within this scenario, therefore, reinforcing the idea of a subject who is endowed by nature with all fundamental rights, who therefore only has to protect themselves from external threats, also means outlining a model of a perfectly self-sufficient person. The latter, relying on their autonomy and focusing on their freedom of action as free from conditioning as possible, would be perfectly capable of promoting the maximisation of their own utility and benefits (Constantin 2013) .

This view of the human being mirrors the liberal market ideal: just as markets are seen to self-regulate through the 'invisible hand,' individuals, left to autonomous choice, are presumed capable of fully realising their innate rights (Michelman 1994). Such a vision ignores the profound impact of historical, social, political, and economic contexts on the actual enjoyment and enforcement of fundamental rights, the legacy of colonialism being a poignant example of this (Badawi 2024).

The Rational

A second foundational concept of hegemonic neoliberalism is that of a knowable, quantifiable world, which is open to rational study by an objective and impartial observer through empiricism, experiment, and modelling. Like individualism and liberalism, this conception has deep roots in Enlightenment thought, and gained further definition as Modernity developed further. The capacity for understanding and gaining knowledge about the world in this way is assumed to be available to all citizens. There is a tendency, however, to further assume or impose reversals of these ideas. Taking the basic position that the world is rationally knowable by citizens, certain groups can be excluded from citizenship (or even humanity) by assigning them an (assumed) *incapability* to gain certain knowledge or a (real) disagreement about established 'facts' ('women are too emotional and irrational to be given the vote'). In a similar reversal, the assigning of categories to natural or social phenomena has lead to attempts at enforcing those categories on the world, e.g. through political attempts to enforce a binary and/or unchanging view of sex or gender, or to a rationalising view of nature that takes optimizing 'productivity' according to some specific metric to be the goal of stewardship, e.g. through monoculture agriculture and clear-cut forestry (Scott 1998).

This rationalist view of the world is also expressed in how neoliberal regimes tend to seek *objective metrics* for governance purposes. Instead of directly referring to 'subjective' judgments and moral arguments, politics is treated, at best, like a multi-dimensional optimization problem, with every value and constraint ultimately being expressed in money, with any political goal ultimately being expressible by way of redistribution of money, or by way of defining and pricing commodities and externalities properly. In this way, the market is further idealized as the 'most rational' arena for economic planning and distribution of resources. This has further effects in wider public management, with paradigms such as New Public Management (NPM) promising substantial efficiency gains (Hood 1991). These are to be accomplished by on the one hand opening up previously public industries to private exploitation and 'free enterprise' through procurement and privatized public services, and on the other through internal billing and budgetary constraints on the parts of the public sector that still remain. This is driven by both the ideology that efficiency is the preserve of the private sectors, and a positive thing, and forced upon public actors through reductions in funding and resource.

The Rational Individual

Together, the neoliberal and techno-solutionist assumptions, outlined in the previous sections, produce an idealised subject: a rational, self-sufficient individual presumed capable of independently understanding, enjoying, and enforcing their innate rights. Such a rational and autonomous individual also embodies the ideal model of what has been defined as the liberal subject of law (Michelman 1994).

In fact, like any other science, law also makes use of abstractions to represent reality through models and categories that allow for a regulation of society and its dynamics in both descriptive and prescriptive terms. In this case, the juridical category of the subject of law represents the model of both the human being with rights and obligations towards the community, and the human being whom the law itself should protect. Such a model has been delineated on the image of the liberal individual, strengthening the role of decision-making and behavioural autonomy (Dagan and Heller 2019). If a subject is endowed with rights due to the mere fact of belonging to the human species, this means that no special effort must be made from the outside to ascribe them to such an individual. Furthermore, if this same subject is also perfectly rational, they have an objective view of both their internal world – i.e. of their own interests, how to prioritise them, and how to realise them in the most efficient way – and of the external world – they are able to recognise opportunities and to recognise and exploit the resources necessary to maximise the satisfaction of their needs (Schlag 1998).

This can result in an inversion of the cause-effect order in the act of abstracting the model from reality (Read 2009). Instead of starting from the observation of actual human dynamics and constructing on the basis of it a category that will work in defining and achieving the intended policy objectives, we start from the objective in order to shape from it the type of agent that would have been able to operate in the scenario that was intended to be realised through the policy.

Consequently, in scenarios where the objective is to maxi-

mize the exchange of goods and services, it becomes imperative for economic actors to articulate their consent to transactions of various natures in a clear, definitive, and legally binding manner (Schlag 1998). In order to achieve this objective, it is necessary to assert distinct positions, and establish one's own individuality and rights as an individual in contrast to those of other citizens.

Accordingly, the liberal subject of law that derives from it was modelled with these characteristics, not because it reflects a real abstraction from the real human being, but because it possessed ideal characteristics that alone were conducive to the development and success of a liberal market and of a society guided by the principles of liberalism first and neoliberalism later. Once the legal category was established, it functioned as the reference prototype for the drafting of rules and regulations, as well as for their interpretation in application.

Rational Individuals and Posed Vulnerabilities

In a socio-economic reality that is crystallized through legal categories, in which the default is the valorization of individuality and the exaltation of choices made on the basis of objective reason, those who do not conform to this model must, by necessity, represent an exception. The latter must necessarily be also covered by the law, which thus makes them a non-negligible subject of policy choices (Collins 2013). The result is what has been defined as 'posed vulnerabilities'. The adjective 'posed' is due to the fact that, in order to be recognized as not perfectly rational and to be considered as bearers of interests and rights not only individual, but shared with other subjects, these people must be officially recognized as vulnerable by law. Consequently, vulnerability emerges as a label that can be assigned by law, contingent upon the possession of specific criteria, which are meticulously delineated in established norms (Rossi et al. 2024). The result is of a dual nature. On the one hand, vulnerability confers entitlement to exceptional forms of protection, which are, however, subject to precise restrictions. On the other hand, individuals who do not align with these categories are held accountable for the consequences of decisions that are presumed to be within their purview. This even if, in practice, they often lack the specific competencies necessary to execute these very decisions.

Hegemonic Neoliberalism in AI and AI Policy

Although AI policy sometimes refers to collective rights such as inclusion or access, its dominant framing is individualistic. The current policy frameworks, such as the EU AI Act and various national data protection laws, reproduce the idea of a liberal subject in both design and justification, which significantly constrains the potential for addressing collective harms or structural inequalities (Edwards 2022).

In AI policy, this liberal conception manifests most explicitly in mechanisms such as informed consent and data subject rights. Under the GDPR and similar regimes, the burden of protection is placed on the individual: once informed, their consent legitimises data collection and usage (Breen, Ouazzane, and Patel 2020). Yet these assumptions are unworkable in practice. The complexity of AI systems, combined with opaque data flows and power asymmetries, renders meaningful consent nearly impossible (Carli and Najjar 2023). Moreover, it neglects the broader implications of data aggregation, where individual choices contribute to collective profiling and structural discrimination (Gillis and Spiess 2019).

This individualistic framing also informs the EU AI Act's risk-based approach, which primarily evaluates harms in terms of their impact on specific users or data subjects. The absence of a framework to assess communal, group-based, or systemic risks reflects a continuation of the liberal legal subject: rational, isolated, and self-sufficient. Even when the policy acknowledges group harms—as in Article 5(1)(b), which references the exploitation of vulnerable populations—these are treated as exceptions (European Parliament and the Council 2024). The law thus reinforces the idea that only those who deviate from the rational subject require special protection, relegating structural vulnerability to an auxiliary consideration.

These choices are not politically neutral. They legitimize a regulatory model where corporations self-regulate and responsibility is individualized. Citizens are treated as isolated users rather than members of a political community. This normalizes the externalisation of social and environmental harm, provided formal compliance mechanisms are in place and individual consent has been obtained.

Furthermore, the insistence on rationality as a defining trait of the legal subject aligns AI development with a technocratic and optimization-driven ethos (OECD 2019). AI systems are designed to emulate or augment rational decision-making, reinforcing the ideal of objectivity and calculability (Chien and Danks 2024; Pataranutaporn et al. 2023). This dovetails with neoliberal governance models that prioritize efficiency and measurability, where values are translated into metrics and political problems are treated as optimization challenges (Hintz, Dencik, and Wahl-Jorgensen 2019). In this schema, the liberal subject becomes both the model for AI design and the presumed beneficiary of its deployment.

However, this alignment obscures the deeply situated, embodied, and relational nature of human agency (Fineman 2020). Cognitive science, feminist theory, and postcolonial critiques have long challenged the universality of the rational subject, emphasizing how decision-making is shaped by emotion, context, and social embeddedness. By ignoring these insights, AI policy continues to reproduce exclusionary standards that marginalize those who do not conform to the rational, self-determining norm.

Legal recognition of vulnerability offers a case in point. To receive protection under existing AI policies, individuals must often be categorized as vulnerable—a label that is both stigmatizing and contingent on narrowly defined criteria. This reinforces a binary between the rational majority and the irrational minority, where deviation from autonomy and self-sufficiency becomes the basis for exceptional treatment. Such a model not only fails to recognize the universality of interdependence but also delegitimizes claims to justice that are framed in collective or relational terms.

It follows that AI, as a technology intended to allow individuals to use their reason to its fullest potential, by and large seeks to reproduce very high levels of rationality. Thus, such a technology can also serve as a tool to correct and overcome the possible limitations of the average user.

AI as an Ultimate Expression of Rationality

The *rational agent* was for a long time a prime model of intelligence in AI research. Indeed, Russell (1997) accounts for four different definitions of 'intelligence' of systems (artificial or not), none of which assume anything but an underlying goal of achieving 'rational' action. That is, the intelligence of a system is defined as its ability to take the optimal actions to achieve some well-defined goal. This presumes a quantifiable and predictable world, one that can be optimized. The COMPAS (Brackey 2019) risk assessment system is emblematic: marketed as an objective tool to improve justice outcomes, it in fact replicated and concealed existing racial biases (Washington 2018). In today's context it is interesting to note that a central part of intelligence is taken to be *agency*, i.e. the ability to *take independent action*, something which is conspicuously absent in most of the high-profile systems marketed as AI: Chatbots and classification/prediction systems, both of which will only respond to input. Recent development of so-called 'agentic' AI appears to be a much closer analogue to the descriptions of Russel, as is the long-standing research fields of multi-agent systems and agent-based modelling, which both share an underlying assumption of agents being, in some sense, rational (Dignum 2017).

In addition to these basic assumptions in the field, the integration and use of AI systems is often also motivated with rationalist language. The COMPAS system, for example, was marketed by its seller Northpointe as being able to reduce bias and increase accuracy in the criminal justice system. Similar claims of reducing bias and producing more objective metrics, judgments, and outcomes, are legion both in marketing materials, policy documents, and research on AI.

Sanctions and Policy Pressures in AI Regulation

Current AI regulatory frameworks primarily rely on monetary penalties to enforce compliance, grounded in the neoliberal belief that market incentives can guide corporate behavior. However, for large firms, such fines often represent minor operational costs and do little to shift institutional priorities or redress harm. This limitation is compounded by how the neoliberal state recognizes corporations not as communities of people, but as abstract, autonomous entities. Within this model, individual executives, directors, and workers are often structurally distanced from both decision-making power and responsibility, diluting accountability across the organization. The global power relations in setting and enforcing sanction and policy pressure also remain highly unequal. The core actors (USA, China and the EU) very much set the agenda for peripheral and semi-peripheral actors (Klein and Patrick 2024).

To move beyond this limited paradigm, and the impact of global power asymmetries, regulatory regimes must adopt non-monetary sanctions that emphasize public accountability and structural change. In the section "Non-Monetary Sanction Regimes", we provide concrete examples and further elaboration. These illustrate how a more democratic and substantive form of oversight can be realized in practice.

From Liberal Autonomy to Collective Responsibility in AI

As discussed in the previous sections, contemporary AI governance is fundamentally shaped by a liberal tradition that centers the autonomous, rational individual as the default subject of law and policy. This conception, embedded in legal, political, and economic theory, underpins how rights, risks, and responsibilities are structured in AI regulation. The current policy frameworks, such as the EU AI Act and various national data protection laws, reproduce this liberal subject in both design and justification, which significantly constrains the potential for addressing collective harms or structural inequalities.

The foundation of this model lies in the Enlightenment-era notion that individuals are endowed with innate rights and possess the rational capacity to understand and pursue their own interests. This view was institutionalised through legal constructs, notably in the aftermath of World War II with instruments like the Universal Declaration of Human Rights. While these documents assert universal dignity and rights, their operationalisation in law tends to privilege individual autonomy over collective well-being (Freeman 1995; Jones 1999). As a result, the protection and exercise of rights are often contingent on an individual's capacity to access legal remedies, understand procedural norms, and navigate complex regulatory landscapes.

This legal architecture has profound implications. It legitimises a regulatory framework where responsibility is privatised and diffused, with corporations tasked primarily with self-regulation and risk mitigation through transparency and documentation (e.g. as in the EU AI Act). The public is constructed as a collection of individual users, rather than a polity with shared interests and vulnerabilities. Consequently, the externalisation of harm—whether environmental, social, or epistemic—is normalised, as long as individual consent has been secured and procedural safeguards are in place.

The abstraction of the liberal subject also distorts our understanding of rights. It constructs rights as static possessions rather than dynamic relations shaped by context and power. The subject is presumed to be equally positioned to enjoy and assert these rights, yet in practice, access to justice, representation, and redress is profoundly unequal. The framework fails to account for how socio-economic conditions, cultural capital, and historical marginalization affect one's capacity to participate in AI-mediated systems on equal footing.

A reorientation of AI governance demands a post-liberal conception of subjectivity—one that acknowledges the interdependence, situatedness, and material embeddedness of human beings. Rather than treating autonomy as a default, policy must recognize it as a function of social, economic,

and infrastructural conditions. Rights should be understood relationally, as entitlements that arise not just from individual status but from membership in communities and the collective impacts of technological systems.

This shift would entail a rethinking of regulatory priorities: moving beyond individual data rights to collective data governance; expanding the notion of harm to include epistemic, ecological, and social damages; and centering democratic participation over procedural compliance. Only by abandoning the fiction of the liberal subject can AI policy begin to address the real and unequal ways in which technology mediates power, shapes subjectivity, and structures society.

Failures of Neoliberal Models

There are multiple circumstances in which the limitations of the neoliberal cultural and political paradigm in the development of AI systems may be highlighted.

One of these is surely the claim to compartmentalise human existence and behaviour into a measurable and predictable order. The prevailing consensus in the scientific community acknowledges the inherent presence of biases in data, social and cultural structures (Olteanu et al. 2019), and the manner in which individuals process information and behave in their daily lives (Lotto 2017). Consequently, algorithms that are developed, trained, and programmed to act on the basis of such data, or to interact in a social fabric or with human beings who themselves possess biases, cannot be immune to them, regardless of the technical mitigation systems that may be implemented. It follows that there is a substantial risk of projecting biases from the past onto the future, as happened in the COMPAS case discussed above (Mayson 2019).

Furthermore, in numerous instances, the advent of novel technologies has been employed to reinforce pre-existing power dynamics, obviating the necessity for an initial evaluation of their fairness and suitability. The AI Act is one example of this phenomenon, as it established an exemption to the established limit on the use of biometric identification systems, in addition to substantial exceptions to the transparency requirements for AI utilized by law enforcement, migration control, and national security authorities in the performance of their duties. This has the potential to contravene the fundamental tenets of egalitarianism and anti-discrimination policies, particularly with regard to groups that are already subject to marginalization, such as migrants. Moreover, an imbalance in the distribution of power is perpetuated, thereby exacerbating the evident lack of instruments available to the individual to enforce the protection of their fundamental rights.

This brief mention of concrete examples of the failure of the neoliberal model in AI highlights the core of the analysis carried out in this study. Designing and regulating new technologies on the basis of categories instrumental to neoliberal criteria implies a misalignment from the reality of the actual subjects who will use them and the actual society in which they will be embedded. This results in inevitably inefficient and misaligned solutions, which are incapable of guaranteeing effective protection for both individuals and for the community.

No, We Are Not All Individuals

By emphasizing the notion that fundamental rights are exclusively individual rights, endowed to every individual by virtue of their shared humanness, the collective value of humankind has become obscured. If it is indeed the case that each individual is endowed with their own unique qualities, which is also reflected in the ways in which they enjoy rights and are called upon to respond to duties, it is also, however, accurate to assert that ‘”no man is an island”, quoting John Donne (1987).

Indeed, the notion of individualism itself is inherently relational. From the moment children begin to develop a perception of themselves and to identify themselves as single individuals — with respect first to the mother’s body and then to the family unit (Hatfield, Ferguson, and Alpert 1967)— they do so by differentiation. In essence, they initiate the process by examining the perception of the *Other*, seeking to identify both commonalities and divergences (Dyk and Witkin 1965). Without another subject to act as a mirror, it is problematic for people to develop self-knowledge, a sense of right and wrong, and thus also a conception of an organised society (Gammelgaard 2003; Lacan 1953; Ramaioli 2005; Cimino 2019).

Similarly, on a collective scale, each person is embedded within a political-economic fabric that is influenced by how the individual socio-economic actors operate within it, but which in turn influences them. Therefore, human beings are interdependent, requiring one another to satisfy fundamental needs, including a sense of belonging and emotional connection, as well as more sophisticated economic and political requirements (Maslow and Frager 1987). Moreover, the population as a whole is dependent on the institutions that organise the social fabric (Fineman 2010). These institutions are also dependent on each other, as well as collectively on the policy decisions made by the State to which they belong, and on the international balances that States establish and regulate among themselves (Fineman 2010).

This very brief excursus demonstrates that the presumed autonomy and independence of individuals’ choices is, in practice, significantly influenced by collective dynamics at various levels much more than it is possible to realise in everyday activities.

For this reason, it is opportune to highlight the pivotal role that policy choices play in shaping the distribution of resources and privileges among citizens. These choices also influence the manner in which rights are exercised and enforced, both at the individual and collective level. In this sense, exacerbating the value of individual choices and rights signifies reinforcing the idea that those who hold more privileges will have greater opportunities to collect more resources and thus maximise their right to – even legal – protection. The result is the perpetuation of inequalities, potentially hidden by the myth that everything depends solely and exclusively on the individual’s decisions. Conversely, it should be crucial to understand, analyse, and correct these

inequalities at a collective level, taking into account the social and political factors that influence them.

The development and governance of AI is not immune to similar observations. Each instance of interaction with an AI system is influenced by the cultural and social context in which it is initiated, the political decisions that have enabled its development and dissemination, and the market environment into which it has been placed prior to its acquisition by the particular user. The manner in which the subject utilizes the technology, the extent to which it exerts influence, and the degree of trust placed in it are contingent not only on the development and dissemination of the AI system but also on the characteristics of the societal context in which the subject is situated. The alternatives offered to the user, along with the resources made available for addressing or developing needs that are satisfied through interaction with the technology, play a crucial role in shaping these outcomes.

No, We Are Not Fully Rational

In contrast to the prevailing prototype of the liberal subject and the liberal subject of law, human beings are not, by nature, perfectly rational and objective. Scientific studies have demonstrated that individuals' perceptions of reality and its causal connections are shaped by neuro-cognitive structures that are significantly influenced by prior experiences, the psycho-emotional sphere, the cultural context, and the socio-historical period in which the cognitive subject is situated (Lakoff and Wehling 2016). These elements are instrumental in the process of evaluating reality and making decisions. Evidently, the cultivation of particular competencies and the attainment of specialized knowledge in specific disciplines and domains can empower individuals to make more informed decisions. However, even in such instances, it is implausible to assert the absence of pre-existing biases and conditioning. Moreover, the idea of infinitely looking for the maximisation of profit or benefit is a market ideal rather than a real inclination of the human being. Rather, the latter has a tendency to seek satisfaction, which means accepting a good outcome while fearing the possibility of losing it by seeking better (Dignum 2017).

Likewise, the best scientific models remain maps of an impossibly complex territory, which will inevitably fail to accurately represent reality, be it through lack of context, too strict categories and boundaries, or simply through irreducible errors in the initial conditions.

Alternatives to Neoliberalism

Against the dominant neoliberal framing of AI as a market-driven tool for efficiency and innovation detached from social struggle, an alternative vision is needed, one grounded in democratic engagement, contextual understanding, and social justice. Rather than abandoning technology, this perspective calls for reclaiming it from capitalist imperatives and technocratic control.

Alternative Techno-Optimisms

Technological solutionism, as critiqued by Morozov (2014), rests on the assumption that social and political problems are reducible to puzzles with optimal technical solutions. It invites interventions before the problem is even fully defined, framing efficiency and automation as inherently desirable. This ideology underpins many AI policy agendas, which presume that more data, more automation, and more optimization will inevitably lead to social improvement (Narayanan and Kapoor 2024) . But such assumptions mask the reality that technological systems are always socially embedded and politically charged (Hare 2022; Schaake 2024).

This techno-optimist and techno-centric mode of thinking frames complex issues like inequality, climate breakdown, or systemic discrimination as technical puzzles. AI becomes the presumed remedy, its limitations addressed through tweaks in accuracy, fairness metrics, or transparency features. As in the critique offered in (Lindgren and Dignum 2023), such approaches obscure the political nature of these problems and constrain how we imagine responses.

What is needed is not a wholesale rejection of technology but a reorientation of its purpose. A socially grounded optimism begins with political questions: what should AI be used for, who decides, and how are collective interests represented? This reframes technology not as an autonomous force but as a tool embedded in power relations, shaped by struggle and choice.

Technological change, particularly in the context of AI, must be treated as conditional—not inevitable. Democratic control, not elite expertise or corporate ethics guidelines, is the basis of legitimate governance. Rather than assuming AI will deliver justice, we must design institutions that make justice possible, even if that means refusing or undoing technological systems that entrench harm.

A different kind of optimism is possible—one that places solidarity, care, and redistribution at the centre of technological futures. Not technology that solves injustice, but collective governance that reshapes what we build, and why.

Power-Distributive AI? As noted by Winner (1986), while a deterministic view of technology as controlling society is overly simplified, artifacts do have, and imply, politics and political structures. The vast majority of resources currently being used for AI research is going towards what could be described as 'centralising' varieties, and arguments have been made (e.g. by McQuillan (2022)) that AI as currently constituted is inherently hierarchical and even fascistic in nature. We can imagine other types of AI technologies which acknowledge difference, specificity, and context, however, and which instead of conserving and strengthening existing hierarchical structures, could act to weaken and break them. We could use local and specific AI systems to solve problems in a way that is specific to their context, in a way that is significantly less wasteful in terms of resources, and apply pattern-finding and -matching systems to identify and counter instances of bias, rather than perpetuating them.

Worker and Citizen Empowerment

Empowering workers and citizens in AI governance means moving beyond narrow consumer rights or transparency checklists. It requires institutional mechanisms through

which those affected by AI systems—workers surveilled by algorithmic management, communities targeted by predictive policing, or patients subject to automated triage—can meaningfully influence how these systems are designed, implemented, and challenged.

Such empowerment demands both collective voice and structural leverage. This includes rights to contest and halt harmful systems, but also participation in the design and oversight of AI technologies. Models such as data cooperatives, algorithmic impact assessments with community input, and union-driven AI bargaining provide promising pathways (Veale and Edwards 2018). Just as labour protections were institutionalized to balance industrial power, AI governance must embed democratic rights in technological systems.

Non-Monetary Sanction Regimes

Current AI regulation heavily relies on financial penalties to enforce compliance—assuming, the underlying argument being that, in the neoliberal tradition, market incentives can realign behaviour. However, monetary fines alone are often insufficient deterrents for large firms and do little to redress harm or shift institutional priorities.

Non-monetary sanctions offer a complementary or alternative regulatory logic. These include mandatory system suspension, public transparency obligations (e.g., registries of high-risk systems), exclusion from public procurement, and enforced algorithmic audits or redesign. They also help politicize compliance by treating governance as a matter of public responsibility rather than private cost management (Yeung 2018). By shifting focus from liability to accountability, such regimes can foster a more substantive and democratic form of oversight.

Planning and State Interventions

Advocating for a reconceptualisation of the subject of law, and for their rights to be guaranteed and protected in a less individualistic way, does not mean denying or flattening the inevitable peculiarities of individual subjective legal situations. Indeed, the characteristics inherent to each individual, and the manner in which they influence an individual's capacity to enjoy or enforce the rights bestowed upon them, should not be disregarded. Nevertheless, in our paper we stress the need to claim for the promotion of an augmented and more clearly delineated state responsibility.

In fact, the state is the entity that possesses the capacity to redistribute power and benefits, not only among citizens but also through the institutions it establishes and which act on its behalf in specific domains of citizens' lives (Fineman 2010; Fineman and Spitz 2023). Furthermore, the state is able to exercise its authority through international agreements or by providing support for certain supranational political-economic currents (Fineman 2021).

In this sense, responsibility here does not necessarily imply mere accountability for the damage or failure of a given policy or product introduced into the market, as it is currently often conceived. In order to comprehend the significance of the state's role in shaping the influence of AI on various facets of society — including employment, socio-economic and socio-political dynamics, as well as the psychological profiles of groups and masses — it is necessary to adopt a comprehensive, systemic perspective. It is imperative to acknowledge the web of micro-decisions that are concurrently executed by policymakers which lay the foundations for the impact that AI will have on individuals and the communities. This underscores the necessity for comprehensive planning of social policies, accompanied by meticulous evaluation of the domains in which the responsible utilization of artificial intelligence systems may be both pertinent and safe. For instance, if the primary strategy of State A to address the shortage of healthcare personnel involved the delegation of an escalating number of activities to intelligent agents, this would result in an augmented exposure of the population to the potential risks associated with the implementation of AI in healthcare. Such an exposure would be difficult to counterbalance by the personal resources of citizens, and would obviously benefit those who have the possibility of paying privately for a second or better consultation, thus exacerbating social inequalities. This same manoeuvre would also affect occupation, leading to a different distribution of investments and political attention among the different sectors that make up the state welfare. From a technological standpoint, this would determine a reorientation of market focus and corporate priorities, resulting in the systematic promotion of specific technological solutions over others. We are currently seeing this play out in states with even the most established and entrenched norms around public healthcare provision (Strange and Tucker 2024) .

In light of the prevailing neoliberal imprint on market dynamics and the approach to technological development, a scenario analogous to the one depicted above could also occur through the simple lack of state intervention. To illustrate this point, consider the case of State B, which does not directly incentivize the adoption of a specific technology, yet neglects to address the lack of resources and services to meet a specific need of the population. This could be sufficient to prompt the production of a technology that claims the ability to meet this need. This is due to the fact that the neoliberal cultural paradigm also encompasses the notion that, when unencumbered by external constraints, the market can not only regulate itself, but also intervene in areas of greatest interest. Such an outcome can have a number of consequences for the human beings potentially affected which do not foster their protection. Firstly, it is important to note that this ideology reinforces the techno-solutionist paradigm, in which technology is regarded as the ultimate solution to human needs, particularly those that are regarded as imperfections or failings inherent to the human condition – included lack of full rationality or emotion-driven decision making. Secondly, it perpetuates the view that regulation and policymaking have a reactionary role with respect to technological development. That is to say, it is believed that they can only intervene *ex post* and only when technology alone has not been sufficient or has caused damage. In this way, justice acquires a mere reparatory function, which debases its primary programmatic aim to provide *a priori* protection and prevention from harms for citizens. Furthermore, it has

been demonstrated how intervening when a given technology has already been placed on the market and is widespread among the population, makes intervention far less effective and, at times, fallacious from the outset due to the speed of innovation and its effects, as opposed to the predefined set of procedures from which the law and governance cannot escape.

These examples show how even collective choices not strictly related to the governance of AI can actually influence the technological sphere and have repercussions on society as a whole and also on individual experiences. It follows that advocating for state responsibility means raising awareness among institutions and policymakers about the role that the entirety of their decisions plays in shaping not only the form but also the direction of technological development. Therefore, a prescriptive intervention should be planned that considers the interests and protection of the community and the social fabric, along with but also beyond those of single individuals.

Government Investment Choices

While regulation shapes the boundaries of AI development, past and current public investment decisions determine its direction. Yet current AI policy debates often overlook how government funding actively steers technological priorities, by deciding what is worth researching, which sectors receive subsidies, and which values are embedded in state-sponsored innovation (Jasanoff 2005).

Crucially, governments are not neutral actors merely reacting to market dynamics. Through public Research and Development (R&D) funding, procurement policies, and infrastructure development, they co-create the AI landscape (Mazzucato 2013). For example, investing in military AI, predictive policing tools, or automated welfare screening embeds authoritarian logics, even in ostensibly democratic contexts. Conversely, funding AI for climate adaptation, public health diagnostics, or multilingual education systems could align technological progress with social needs, if governed carefully and conditionality established from the outset (Maggor and Tucker 2025).

Moreover, these choices are not merely about end-use applications, they determine who gets to participate in AI development. Investment in public universities, open-source platforms, and cooperative tech models can redistribute power and knowledge production, challenging corporate concentration. Conversely, channeling public money through public-private partnerships without accountability mechanisms often results in socializing risk and privatizing reward . To be consistent with a post-capitalist or justice-oriented AI policy, governments should:

- Make investment decisions transparent, participatory, and accountable to democratic oversight.
- Prioritize funding toward non-extractive, community-driven technologies.
- Redirect incentives away from speculative AI ventures and toward public infrastructure and resilience.
- Assess the long-term social impact of investments, not just short-term returns or innovation metrics.

Strategic public investment is not just economic policy, it is governance. Without deliberate alignment with justice and collective well-being, it risks reinforcing the same extractive systems AI policy seeks to challenge.

Building Public AI Infrastructure

To move beyond critique and toward transformative alternatives, AI policy needs to focus on the creation of public technological infrastructure which can support mission-oriented innovation (Kattel and Mazzucato 2018). Just as public roads, electricity, and libraries enabled broad-based development in previous eras, democratizing access to data, compute, and models is essential to counter corporate concentration and unlock collective innovation. Public AI infrastructure is a test case for this principle, demonstrating how investment can align technological development with democratic values and social goals. Examples of public AI infrastructure include:

- Public data commons: Secure, transparent data pools governed by communities, not harvested by corporations. Examples include Barcelona's DECODE project or France's HAL repository, as well as Te Hiku Media's project to preserve and safeguard Te Reo Māori.
- Public compute infrastructure: Government-funded compute clusters, like those supporting academic research, could be expanded and made accessible to civil society and cooperatives.
- Open models and public AI labs: Analogous to public broadcasters, states could fund independent AI labs tasked with developing socially beneficial systems outside commercial imperatives.

Without such infrastructure, even the best-intentioned regulatory reforms risk failure, as private firms retain control over core AI capabilities. Public investment here is not just a corrective but a constitutive act: it enables pluralism, participation, and power redistribution in the AI ecosystem.

Toward Anti-Capitalist AI Policy

To meaningfully confront the harms perpetuated by current AI governance, policy must explicitly challenge capitalist imperatives rather than accommodate them. This requires both a clear confrontation of how capitalist dynamics shape AI policy and the articulation of concrete alternatives that redistribute power and prioritize collective well-being.

Confronting Power Means Confronting Capital

One does not have to be a communist or socialist to recognise the deleterious effects of unfettered capitalism, and to argue for the responsibility of the state and other major organisations to act in a way that is explicitly counter to the interests of capital. Mainstream economic theorists from Adam Smith to John Maynard Keynes to Paul Krugman have all called for the regulatory state to rein in the worst excesses of capitalism. However, when it comes to AI policy, policy proposals seldom make this an explicit conflict, and many 'AI Roadmaps' and similar documents caution against

the risk of stifling 'innovation' and 'growth' through heavy-handed regulation.

With this in mind, we propose a number of different pathways and practices for regulators and their academic advisors to properly consider how capitalist pressures and dynamics can be recognized, and how they can either be properly channeled or directly countered, depending on the situation. It should be stated that the policies proposed below are common in public policy to limit the impacts of a range of public and corporate actors and advance the social good. Where there seem to be in significant tension with current state of play in AI governance, this speaks more the overbearing role of certain neoliberal discourses than the policies themselves.

Concrete Policies

Developing anti-capitalist AI policy requires more than critique—it demands actionable frameworks that counteract corporate capture, democratize AI governance, and reorient technological development toward collective, just, and sustainable outcomes. In the following, we outline specific measures to challenge corporate power, prioritize the public interest, and enable democratic oversight:

- Center human wellbeing, labour, and societal sustainability. Pragmatically, international human rights law provides a robust framework to guide this.
- Be explicitly and consciously sceptical, even antagonistic, towards capital and capitalism
- Aim at material outcomes that tangibly improves the life of people, rather than succumbing to the ideology of 'innovation at all costs'
- Be cognizant of the capitalist drive towards externalization, misdirection, and overhype of technology
- Read and apply relevant research on legitimate and illegitimate uses of AI technology, e.g. Wang et al. (2024)
- Openly and clearly question the inflated valuations of AI and tech firms as a product of overfinancialization of the economy rather than a reflection of actual material value for society
- Adopt and or enforce policies to limit lobbying and ensure transparency for when it occurs
- Conflicts of interest (COI) should be made explicit for all parties, and transparency in decision-making should allow external actors to assess any undisclosed COI.
- Adopt and or enforce legislation to limit revolving-door dynamics between public and private organizations.

Conclusion

In this paper, we have argued that dominant AI policy frameworks are deeply shaped by neoliberal assumptions that prioritize individualism, rationality, and market efficiency. These paradigms often obscure structural inequalities and power imbalances that AI technologies both reflect and reinforce. The current regulatory emphasis on innovation and risk management fails to address the broader political-economic conditions under which AI is developed and deployed.

To advance a more equitable and accountable approach to AI governance, we propose a policy orientation grounded in anti-capitalist principles. This is not in opposition to technological development, but a deliberate shift in governance focus: from economic growth and corporate interests to collective well-being, democratic participation and redistributive justice. This reorientation calls for regulatory frameworks that recognize collective rights and harms, not just individual entitlements. It demands mechanisms of consent and accountability that acknowledge and address structural imbalances in power and knowledge. Central to this approach is the meaningful inclusion of marginalized communities in shaping how AI systems are designed, deployed, and governed. Ultimately, governance must be reclaimed as a democratic and participatory endeavour, rather than a technocratic process shaped by market logics.

Anti-capitalist AI policy does not reject technology; it rejects the commodification of technology under exploitative systems. It asks: Who benefits, who decides, and on what terms? A truly democratic AI future is one where these questions are answered collectively, not dictated by capital. Or, as Ted Chiang was quoted in AI Snake Oil (p.278), "Fears about technology are fears about capitalism" (Narayanan and Kapoor 2024), and by properly regulating the latter, we will have a better chance to make full use of the former.

Acknowledgments

We would like to extend our gratitude to the anonymous reviewers for their thoughtful comments. This work was partly supported by the Wallenberg AI, Autonomous Systems and Software Program – Humanity and Society (WASP-HS) funded by the Marianne and Marcus Wallenberg Foundation and the Marcus and Amalia Wallenberg Foundation.

References

Badawi, Z. 2024. *An African History of Africa: From the Dawn of Humanity to Independence*. London: W.H. Allen. First published 2024; paperback edition 2025.

Brackey, A. 2019. *Analysis of racial bias in Northpointe's COMPAS algorithm*. Master's thesis, Tulane University.

Braybrooke, D. 1990. Natural Law Theory: The Link Between Its Descriptive Strength and Its Prescriptive Strength. *Canadian Journal of Philosophy Supplementary Volume*, 16: 389–418.

Breen, S.; Ouazzane, K.; and Patel, P. 2020. GDPR: Is your consent valid? *Business Information Review*, 37(1): 19–24.

Carli, R.; and Najjar, A. 2023. A Vulnerability-oriented Impact Assessment for the Development of Human-Centred and Fundamental Rights-Empowering Social Robots. In *Proceedings of the 11th International Conference on Human-Agent Interaction*, 404–406.

Chien, J.; and Danks, D. 2024. Beyond behaviorist representational harms: A plan for measurement and mitigation. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, 933–946.

Cimino, C. 2019. Le origini del riconoscimento nella costruzione del soggetto. Lacan con Winnicott. *Hegel & sons. Filosofie del riconoscimento*, 281–290.

Collins, J. 2013. The contours of 'vulnerability'. In *Vulnerabilities, Care and Family Law*, 22–42. Routledge.

Constantin, H. 2013. Homo Oeconomicus: Conceptual aspects between dilemmas and paradoxes. *European Scientific Journal*, 9(19).

Crawford, K. 2021. *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.

Dagan, H.; and Heller, M. 2019. Why Autonomy Must Be Contract's Ultimate Value. *Jerusalem Review of Legal Studies*, 20(1): 148–171.

Dignum, V. 2017. Social Agents: Bridging Simulation and Engineering. *Communications of the ACM*, 60(11): 38–40.

Donne, J. 1987. Meditation XVII: no man is an island. *Devotions upon Emergent Occasions*.

Dyk, R. B.; and Witkin, H. A. 1965. Family experiences related to the development of differentiation in children. *Child Development*, 21–55.

Edwards, L. 2022. Regulating AI in Europe: Four Problems and Four Solutions. Technical report, Ada Lovelace Institute, United Kingdom. Expert opinion commissioned by the Ada Lovelace Institute.

European Parliament and the Council. 2024. Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828, (Artificial Intelligence Act).

Fineman, M. 2010. The Vulnerable Subject and the Responsive State. *Emory Law Journal*, 60(2): 251.

Fineman, M. A. 2013. *Vulnerability: reflections on a new ethical foundation for law and politics*. Ashgate Publishing, Ltd.

Fineman, M. A. 2020. Reasoning from the body: Universal vulnerability and social justice. *A Jurisprudence of the Body*, 17–34.

Fineman, M. A. 2021. Rights, resilience, and responsibility. *Emory Law Journal*, 71: 1435.

Fineman, M. A.; and Spitz, L. 2023. *Law, vulnerability, and the responsive state: beyond equality and liberty*. Taylor & Francis.

Freeman, M. 1995. Are there collective human rights? *Political studies*, 43(1): 25–40.

Gammelgaard, J. 2003. Ego, self and otherness. *The Scandinavian Psychoanalytic Review*, 26(2): 96–108.

Gillis, T. B.; and Spiess, J. L. 2019. Big data and discrimination. *The University of Chicago Law Review*, 86(2): 459–488.

Hare, S. 2022. *Technology is not neutral: A short guide to technology ethics*. London Publishing Partnership.

Hatfield, J. S.; Ferguson, L. R.; and Alpert, R. 1967. Mother-child interaction and the socialization process. *Child Development*, 365–414.

Henkin, L. 1989. The universality of the concept of human rights. *The Annals of the American Academy of Political and Social Science*, 506(1): 10–16.

Hintz, A.; Dencik, L.; and Wahl-Jorgensen, K. 2019. *Digital citizenship in a datafied society*. Cambridge, UK ; Medford, MA: Polity Press. ISBN 978-1-5095-2715-1 978-1-5095-2716-8.

Hood, C. 1991. A Public Management for All Seasons? *Public Administration*, 69(1): 3–19.

Jasanoff, S. 2005. *Designs on Nature: Science and Democracy in Europe and the United States*. Princeton University Press.

Jones, P. 1999. Human rights, group rights, and peoples' rights. *Human rights quarterly*, 21(1): 80–107.

Kattel, R.; and Mazzucato, M. 2018. Mission-oriented innovation policy and dynamic capabilities in the public sector. *Industrial and Corporate Change*, 27(5): 787–801.

Kaufmann, M. 2009. *Diritti umani*, volume 20. Guida Editori.

Klein, E.; and Patrick, S. 2024. Envisioning a Global Regime Complex to Govern Artificial Intelligence. Technical report, Carnegie Endowment for International Peace. Accessed: 2025-05-19.

Lacan, J. 1953. Some Reflections on the Ego. *The International Journal of Psycho-Analysis*, 34: 11.

Lakoff, G.; and Wehling, E. 2016. *Your brain's politics: How the science of mind explains the political divide*, volume 59. Andrews UK Limited.

Lindgren, S. 2023. *Critical theory of AI*. John Wiley & Sons.

Lindgren, S.; and Dignum, V. 2023. Beyond AI solutionism: toward a multi-disciplinary approach to artificial intelligence in society. In *Handbook of critical studies of artificial intelligence*, 163–172. Edward Elgar Publishing.

Locke, J. 1990. Natural rights. *Moral Reasoning: A Philosophic Approach to Applied Ethics, Dryden Press, London*, 133–5.

Lotto, B. 2017. *Deviate: The science of seeing differently*. Hachette UK.

Maggor, E.; and Tucker, J. 2025. Towards successful industrial policy on AI in healthcare: establishing the conditions for future public benefit. In *AI Policy Exchange Forum, AIPEX*.

Maslow, A. H.; and Frager, R. 1987. *Motivation and personality*. New York: Harper and Row, 3rd ed edition.

Mayson, S. G. 2019. Bias In, Bias Out. *The Yale Law Journal*, 128(8): 2218–2300.

Mazzucato, M. 2013. *The Entrepreneurial State: Debunking Public vs. Private Sector Myths*. Anthem Press.

McQuillan, D. 2022. *Resisting AI: an anti-fascist approach to artificial intelligence*. Bristol: Bristol University Press.

Michelman, F. I. 1994. The Subject of Liberalism. *Stanford Law Review*, 46(6): 1807–1833.

Morozov, E. 2014. To save everything, click here: the folly of technological solutionism. *Journal of Information Policy*, 4(2014): 173–175.

Narayanan, A.; and Kapoor, S. 2024. *AI snake oil: What artificial intelligence can do, what it can't, and how to tell the difference*. Princeton University Press.

Noble, S. U. 2018. *Algorithms of oppression: How search engines reinforce racism*. New York university press.

OECD. 2019. Artificial Intelligence in Society. *OECD Digital Economy Papers*, 1 – 159.

Olteanu, A.; Castillo, C.; Diaz, F.; and Kıcıman, E. 2019. Social data: Biases, methodological pitfalls, and ethical boundaries. *Frontiers in big data*, 2: 13.

Pataranutaporn, P.; Liu, R.; Finn, E.; and Maes, P. 2023. Influencing human–AI interaction by priming beliefs about AI can increase perceived trustworthiness, empathy and effectiveness. *Nature Machine Intelligence*, 5(10): 1076–1086.

Politico. 2025. Ranked: The 10 most intensely lobbied EU laws. *Politico Europe*. Accessed August 1, 2025.

Ramaioli, I. 2005. Note sull'identificazione: da Freud a Lacan. *Costruzioni psicoanalitiche. Fascicolo 1, 2005*, 1000–1014.

Read, J. 2009. A genealogy of homo-economicus: Neoliberalism and the production of subjectivity. *A Foucault for the 21st Century: governmentality, biopolitics and discipline in the new millennium*, 2.

Rossi, A.; Carli, R.; Botes, M. W.; Fernandez, A.; Sergeeva, A.; and Chamorro, L. S. 2024. Who is vulnerable to deceptive design patterns? A transdisciplinary perspective on the multi-dimensional nature of digital vulnerability. *Computer Law & Security Review*, 55: 106031.

Russell, S. J. 1997. Rationality and intelligence. *Artificial Intelligence*, 94(1-2): 57–77.

Schaake, M. 2024. *The Tech Coup: How to Save Democracy from Silicon Valley*. Princeton University Press.

Schlag, P. 1998. *The enchantment of reason*. Duke University Press.

Scott, J. C. 1998. *Seeing like a state: how certain schemes to improve the human condition have failed*. Yale agrarian studies. New Haven: Yale University Press.

Sen, A. 2005. Human rights and the limits of law. *Cardozo Law Review*, 27: 2913.

Sheehan, M. 2024. Reverse Engineering China's AI Governance: Tracing the Roots of China's AI Regulations. https://carnegie-production-assets.s3.amazonaws.com/static/files/Sheehan_Reverse_Engineering_AI_Gov-UPDATED-1.pdf. Accessed: 2025-05-07.

Shimpo, F. 2025. AI Governance and the Future of the Law: International Trends in Developing AI Regulatory Legislation. *Japanese Society and Culture*, 7: 4.

Siegal, A.; and Garcia, I. 2024. A Deep Dive into Colorado's Artificial Intelligence Act. https://www.naag.org/attorney-general-journal/a-deep-dive-into-colorados-artificial-intelligence-act/. Accessed: 2025-05-07.

Strange, M.; and Tucker, J. 2024. A Paradigm Shift in Plain Sight? AI and the Future of Healthcare in the Nordic States. *Nordisk välfärdsforskning*, 9(2): 168–179.

Tucker, J. 2023. The future vision (s) of AI health in the Nordics: Comparing the national AI strategies. *Futures*, 149: 103154.

United Nations General Assembly. 1949. *Universal declaration of human rights*, volume 3381. Department of State, United States of America.

Veale, M.; and Edwards, L. 2018. Clarity, surprises, and further questions in the Article 29 Working Party draft guidance on automated decision-making and profiling. *Computer Law & Security Review*, 34(2): 398–404.

Walter, Y. 2024. Managing the race to the moon: Global policy and governance in artificial intelligence regulation—A contemporary overview and an analysis of socioeconomic consequences. *Discover Artificial Intelligence*, 4(1): 14.

Wang, A.; Kapoor, S.; Barocas, S.; and Narayanan, A. 2024. Against Predictive Optimization: On the Legitimacy of Decision-making Algorithms That Optimize Predictive Accuracy. *ACM Journal on Responsible Computing*, 1(1): 9:1–9:45.

Washington, A. L. 2018. How to argue with an algorithm: Lessons from the COMPAS-ProPublica debate. *Colo. Tech. LJ*, 17: 131.

Winner, L. 1986. *The whale and the reactor: a search for limits in an age of high technology*. Chicago: University of Chicago Press. ISBN 978-0-226-90210-4.

World Bank Group. 2024. Global Trends in AI Governance: Evolving Country Approaches. Accessed: 2025-05-07.

Yeung, K. 2018. Algorithmic regulation: A critical interrogation. *Regulation & governance*, 12(4): 505–523.