# ALGOCRACY AND THE CRISIS OF POLITICAL LEGITIMACY: A PHILOSOPHICAL ANALYSIS OF HUMAN AUTONOMY IN ALGORITHMIC GOVERNANCE

**Johnpaul Enemuo, PhD**
Department of Philosophy
Nnamdi Azikiwe University, Awka
cj.enemuo@unizik.edu.ng
&
**Anthony Ugochukwu Nwokoye**
Department of Philosophy
Nnamdi Azikiwe University, Awka
au.nwokoye@unizik.edu.ng

**Abstract**
The rise of algocracy, which referred to "governance by algorithms" poses challenges to traditional notions of political legitimacy and human autonomy in democratic societies. While algorithmic systems promise efficiency, objectivity, and data-driven decision-making, their opacity and lack of accountability raise critical philosophical questions about their compatibility with democratic principles. This research examines the tension between algocratic governance and democratic legitimacy through a philosophical lens, drawing on John Danaher's foundational critique of algocracy, as well as frameworks from political philosophy such as Habermasian deliberative democracy, Rawlsian public reason, and Frankfurtian autonomy theories. Central to the analysis are two ethical concerns identified by Danaher: the hiddenness concern, which highlights the covert collection and use of personal data without informed consent, and the opacity concern, which underscores the inscrutability of algorithmic decision-making processes. These concerns create a legitimacy deficit, as algorithmic systems operate as "black boxes," undermining transparency, accountability, and public deliberation, which are the cornerstones of democratic governance. The research further explores how algocracy erodes human autonomy by displacing moral reasoning and participatory decision-making, reducing citizens to passive subjects of automated governance. The debate extends to technocratic paternalism, where efficiency-driven algorithmic governance risks sidelining democratic participation and moral agency. Ultimately, the research contends that the greatest challenge lies in balancing efficiency with democratic accountability. Without meaningful transparency and mechanisms for contestation, algocracy risks becoming a post-democratic order. Yet, if carefully designed, algorithmic governance could enhance democracy by addressing human cognitive limitations and systemic inefficiencies. The study concludes that the future of legitimate governance hinges on integrating algorithmic means with democratic ends, ensuring that technological advancements serve rather than sabotage the principles of autonomy, justice, and collective self-determination.
**Keywords:** Algocracy, political legitimacy, human autonomy, algorithmic governance, democracy, opacity, efficiency, hybrid governance.

## Introduction

The rapid integration of algorithmic decision-making into governance what philosopher Danaher (2016) terms "algocracy" has sparked a philosophical debate about the future of democracy, autonomy, and political legitimacy. As artificial intelligence and automated systems increasingly shape public policy, criminal justice, social welfare, and administration, fundamental questions arise on whether rule by algorithms ever be truly legitimate? And if so, under what conditions can algocracy coexist with democratic self-governance? Algorithmic governance promises efficiency, objectivity, and data-driven precision, which helps to mitigate human biases. However, Burrell (2016) argues that this shift introduces a legitimacy deficit, as algorithmic systems often operate as "black boxes," obscuring the reasoning behind their decisions. This challenges core democratic principles such as transparency, accountability, and public deliberation that underpin political legitimacy (Habermas, 1996; Floridi, 2022).

Danaher's (2016) seminal work, "The Threat of Algocracy," sees this tension as a conflict between epistemic efficiency, that is, algorithms' superior problem-solving capabilities and procedural legitimacy, that is, the democratic requirement that citizens understand and contest decisions. When courts, legislatures, and administrative agencies delegate authority to AI, they risk creating a system where power is exercised without meaningful human oversight;

a scenario Danaher describes as "governance without comprehension" (Danaher, 2016, p. 12). A central concern in this debate is the erosion of human autonomy, the capacity for self-determination and participation in the norms that govern society (Kant, 1785; Frankfurt, 1971). Algorithmic systems, even when well-intentioned, may reduce citizens to passive subjects of automated decision-making, undermining their role as active participants in democratic processes (Susskind, 2022).

As argued by Pasquale (2015) in The Black Box Society, the lack of transparency in algorithmic governance creates a power asymmetry, where individuals cannot meaningfully challenge decisions that affect their lives, which is a violation of what Forst (2012) calls the "right to justification." However, fundamental issues remain. Can algorithms ever replicate the deliberative legitimacy of human democratic processes (Habermas, 2015)? Or does algocracy inevitably lead to a post-democratic order, where governance is outsourced to unelected systems (Zuboff, 2023)?

**The Concept of Algocracy**

The concept of algocracy is derived from the words "algorithm" and "government," and emerged in response to the increasing integration of automated decision-making systems into governance. While algorithmic governance has roots in early cybernetic theory (Wiener, 1948) and computational public administration (Simon, 1947), the term itself was formally introduced by philosopher John Danaher in his foundational 2016 paper, "The Threat of Algocracy: Reality, Resistance, and Accommodation." According to Danaher (2016, p. 248)

> To be more precise, I use it to describe a system in which algorithms are used to collect, collate and organise the data upon which decisions are typically made and to assist in how that data is processed and communicated through the relevant governance system. In doing so, the algorithms structure and constrain the ways in which humans within those systems interact one another, the relevant data and the broader community affected by those systems.

This explains that algogracy is a system of governance where algorithms play a central role in collecting, organizing, and processing data used for decision-making. These algorithms not only manage the flow of information but also influence how it is communicated within the system. In doing so, they shape and limit how individuals within the system interact with each other, the data itself, and the wider community. Essentially, algorithms become active participants in governance by structuring the ways decisions are made and relationships are formed. However, Danaher's work crystallized growing concerns about the displacement of human judgment by opaque, data-driven systems, framing algocracy as a distinct mode of governance that challenges democratic legitimacy. He argues that algocratic systems raises two moral and political concerns.

> Hiddenness concern: This is the concern about the manner in which our data is collected and used by these systems. People are concerned that this is done in a covert and hidden manner, without the consent of those whose data it is.
> Opacity concern: This is a concern about the intellectual and rational basis for these algocratic systems. There is a worry that these systems work in ways that are inaccessible or opaque to human reason and understanding (Danaher, 2016, p. 249).

The two critical ethical concerns raised by Danaher are central to the question of algorithmic legitimacy. The hiddenness concern reflects anxieties about how personal data is often collected and utilized without individuals' informed consent or awareness. Zuboff (2023), in The Age of Surveillance Capitalism, argue that data is frequently extracted surreptitiously, creating a power imbalance between those who generate data and those who control it. This covert data collection undermines principles of autonomy and transparency. The opacity concern, on the other hand, addresses the unintelligibility of algorithmic decision-making processes. As Pasquale (2015) discusses in The Black Box Society, many algorithms operate in ways that are unreadable to the average person and sometimes even to experts making it difficult to assess their fairness, logic, or potential biases.

Prior to Danaher's formulation, scholars had examined the role of technology in governance, but without a unified theoretical framework. Early discussions focused on technocracy, which is referred to as the rule by experts, and later cybernetics, which viewed governance as an information-processing system (Wiener, 2019). However, the rise of machine learning and big data in the 21st century shifted the discourse toward automated decision-making, particularly in welfare allocation (Eubanks, 2018), and bureaucratic administration (Zuboff, 2019). Danaher's intervention was important because it identified algocracy not merely as a tool of governance but as a structural transformation of

political authority; one that risks undermining human autonomy and democratic accountability (Danaher, 2016, p. 248).

Danaher's analysis built upon earlier critiques of algorithmic opacity (Burrell, 2016) and surveillance capitalism (Zuboff, 2015), synthesizing them into a coherent critique of governance-by-algorithm. He argued that algocracy differs from traditional technocracy because it does not rely on human experts but on "epistemically superior but incomprehensible systems" (Danaher, 2016, p. 250). This shift, he contended, creates a legitimacy crisis: if citizens cannot understand or contest algorithmic decisions, the very basis of democratic consent erodes (Forst, 2012).

**The Legitimacy Deficit of Algocracy**

The rise of algocratic systems in public administration and policy-making has introduced a fundamental paradox in modern governance, while these systems frequently demonstrate superior efficiency compared to human decision-makers, they simultaneously suffer from a profound legitimacy deficit rooted in their lack of procedural accountability (Danaher, 2016). This crisis between technical efficacy and democratic validity forms the core of contemporary debates about the role of artificial intelligence in governance structures. Traditional democratic systems derive their legitimacy from well-established mechanisms of accountability including periodic elections, public deliberation, transparent reasoning processes, and institutional checks and balances (Habermas, 1996). Algocracy, by contrast, operates through decision-making processes that are frequently opaque, technically complex, and removed from conventional channels of public scrutiny and contestation (Burrell, 2016).

Danaher's (2016) epistemic opacity argument provides an important framework for understanding this legitimacy crisis. He contends that the very features that make algorithmic decision-making effective - its ability to process vast datasets and identify complex patterns - simultaneously render it incomprehensible to ordinary citizens and even to the experts who design these systems. This creates what might be termed a "democratic black box," where consequential decisions affecting citizens' lives are made through processes that resist meaningful examination or challenge (Pasquale, 2015). The implications of this opacity extend beyond technical concerns to touch upon fundamental questions of political philosophy, particularly what Forst (2012) identifies as the "right to justification" - the basic democratic principle that those affected by decisions must be able to understand and contest the reasons behind them.

Furthermore, Volkov (2025) argues that legitimately wield power, a decision-maker must uphold moral responsibility by avoiding violations of fundamental rights and opaque or unaccountable practices. This presumes that the decision-maker is capable of being held morally accountable, that is, that it possesses moral status, which in turn requires autonomy and agency. Despite advances in AI, the dominant view in the literature is that current and foreseeable algorithmic systems do not meet the threshold for moral agency(Volkov, 2025). While some scholars, such as Sullins (2011) entertain the possibility of moral AI, these views depend on speculative forms of autonomy not present in today's systems. Volkov (2025) writing to show that machine learning algorithms are human-curated data and code cited Mele's (2005) distinction which is akin to person A who is entirely dependent on person B for all their knowledge; while A acts, they do so without independent reasoning, and thus are only partially accountable for the outcomes. Algorithms, like A, are heavily determined by exogenous human inputs and lack the capacity for self-reflection or moral reasoning.

List (2023) argues that human decisions are also influenced by external and evolutionary forces, this ignores the critical distinction: humans, despite those influences, possess the capacity for reflective endorsement of beliefs and the ability to act contrary to deterministic inputs. Thus, the ability to do otherwise is foundational to ascriptions of moral agency. Algorithms, by contrast, cannot meaningfully revise or resist their programming. Thus, the opacity of algocratic systems is not just technical or epistemic; it is moral—they are not and cannot be accountable moral agents. Therefore, when such systems make decisions with serious consequences, they introduce a form of opacity that is delegitimizing, since the power exercised lacks the moral grounding necessary for legitimacy (Volkov, 2025).

This crisis of legitimacy becomes particularly serious when Habermasian (1996) theories of deliberative democracy is put into consideration, which emphasize the importance of communicative reasoning in legitimate governance. Algocratic systems fundamentally alter the structure of political discourse by removing or obscuring the grounds for decision-making, making genuine deliberation impossible. Where democratic legitimacy traditionally requires that

decisions be justified through public reason (Rawls, 1993), algorithmic governance substitutes technical optimization for reasoned debate.

**Autonomy Under Algorithmic Rule**

The advent of algorithmic governance presents profound challenges to traditional conceptions of individual autonomy, particularly when understood through the rich philosophical frameworks developed by Kant, Habermas, and Frankfurt. The formulation of autonomy by Kant is accepted, where he defined autonomy as the capacity for self-governance through rational moral law, combined with Habermas's (1996) emphasis on discursive participation in norm-creation, then algocracy threatens to fundamentally alter the conditions for autonomous agency in modern societies. This transformation goes beyond simple questions of coercion or freedom from interference but touching instead on the more subtle but equally the vital capacity to engage meaningfully with the normative structures that shape collective life.

At the heart of this concern lies what Frankfurt (2018) identified as the distinction between freedom of action and freedom of the will. Even when algorithmic systems produce outcomes that are technically fair or even superior to human decisions in terms of reducing bias as some claim occurs in predictive policing or sentencing algorithms (Angwin, Larson, Mattu, & Kirchner, 2022), they may nonetheless undermine autonomy by displacing the practice of moral reasoning itself. When judges, for instance, defer to algorithmic risk assessments in sentencing decisions (Starr, 2014), they are not just using a tool but outsourcing a fundamentally moral judgment; one that requires the exercise of practical reason and accountability that Kant saw as essential to human dignity. The problem is not merely that the algorithms might be wrong or biased; though they frequently are, but that their use changes the nature of the decision-making process in ways that diminish human moral agency (Starr, 2014).

Habermas's (1996) theory of communicative action helps illuminate why this matters for autonomy. For Habermas, legitimate norms emerge through processes of discursive justification where all affected parties can participate. Algorithmic governance, by contrast, tends to replace such discursive processes with technical optimization. Where democratic autonomy requires that citizens can understand and challenge the reasons for decisions that affect them (Forst, 2012), algocracy offers only the "black box" of algorithmic output (Burrell, 2016). This creates what Danaher (2016) calls the "problem of incomprehensibility" even if an algorithmic decision is technically correct, its inscrutability makes meaningful engagement impossible, reducing citizens to passive recipients rather than active participants in governance.

The implications extend beyond political philosophy into practical psychology. Research on automation bias (Skitka, Mosier, & Burdick, 1999) shows how humans tend to over-trust algorithmic systems, suspending their own judgment even when they have reason to question the system's output. This suggests that algorithmic governance may create what Vallor (2016) terms "technosocial moral atrophy" the gradual weakening of moral reasoning capacities through disuse. If moral judgment is like a muscle that strengthens with exercise as Aristotle's virtue ethics would suggest, then widespread delegation to algorithmic systems risks leaving that muscle dangerously underdeveloped.

**Technocratic Paternalism: An Objection to Algocracy**

The fear of technocratic paternalism in algorithmic governance presents a fundamental philosophical challenge to proponents of algocracy. This critique which is rooted in democratic theory and moral philosophy, questions whether efficient governance should ever supersede individual autonomy as a political value. As Sandel (2020) argues in The Tyranny of Merit, systems that prioritize technical efficiency over democratic participation risk creating a new form of oligarchy; one where decisions are made by algorithmic elites rather than through collective self-determination. This concern echoes Habermas's (1996) theory of communicative action, which positions democratic deliberation as essential for legitimate norm-creation, not merely as an inefficient means to effective governance.

The paternalism critique gains particular force when examining how algorithmic systems reshape moral agency. Frankfurt's (1971) concept of second-order desires reminds us that autonomy involves not just freedom from interference, but the capacity to reflectively endorse one's choices, a capacity potentially undermined when algorithms make decisions on our behalf. This aligns with Pettit (2012) republican theory of freedom as non-domination, which would view algocracy as problematic even if benevolent, because it places citizens in a position of dependence on systems they cannot understand or contest.

Philosophical objections to technocratic paternalism also draw on Kantian ethics. O'neill (2002) emphasizes that treating individuals as ends rather than means requires respecting their capacity for practical reasoning, and this principle is violated when algorithmic systems make opaque decisions affecting life chances without providing justifications accessible to those affected. This connects to Forst's (2012) "right to justification" framework, which holds that legitimate governance requires giving citizens adequate reasons for decisions that constrain their freedom. The incomprehensibility of many machine learning systems makes this fundamentally impossible in current implementations of algocracy.

Even proponents of algorithmic governance acknowledge these objections. Danaher (2016), while advocating for algocracy's potential benefits, concedes that epistemic opacity creates a "legitimacy deficit" that purely efficiency-based arguments cannot resolve. Floridi (2023) work on digital ethics similarly emphasizes that technical competence cannot substitute for democratic accountability mechanisms. These philosophical perspectives suggest that the trade-off between autonomy and efficiency presents not just a practical challenge, but a fundamental normative dilemma for algorithmic governance; one that efficiency arguments alone cannot adequately address.

**Efficiency as the New Political Imperative**

The dominant critique of algocracy rests on the normative assumption that democratic participation and human moral agency constitute inviolable political values. This perspective, however, fails to engage with the evolutionary necessity of governance efficiency in an era defined by unprecedented complexity, global interconnectedness, and existential risks ranging from climate collapse to artificial superintelligence (Mulgan, 2016). If we accept that the primary function of the state is not merely to uphold abstract ideals of autonomy but to solve problems ensuring public safety, economic stability, and ecological sustainability, then algorithmic governance, with its superior speed, scalability, and data-driven precision, should not be artificially constrained by outdated human-centric models of legitimacy (Danaher, 2016).

Democratic systems, while committed to collective self-rule, are in practice riddled with inefficiencies that undermine their problem-solving capacity. Political gridlock, short-term electoral incentives, and cognitive biases frequently result in substandard policy outcomes (Gilens & Page, 2014). Human decision-makers are also susceptible to corruption, ideological capture, and emotional reasoning, all of which distort rational governance (Acemoglu & Robinson, 2019). On the other hand, Danaher (2016) argues that algocracy lacks "intrinsic legitimacy" because it bypasses democratic deliberation. However, this critique assumes that existing democratic systems truly deliver meaningful self-rule. As argued by Gilens and Page (2014) that policy outcomes in so-called democracies are overwhelmingly shaped by elite interests, not public. Most citizens have negligible influence on legislation, and electoral participation is often reduced to a ritualistic choice between pre-selected elites.

In this light, performance-based legitimacy where governance is judged by results rather than process may be a superior framework. Citizens may not understand how an AI determines tax policy or infrastructure investments, but they also do not comprehend the opaque deal-making of legislative bodies. If algorithmic governance delivers higher living standards, better public services, and fairer resource allocation, its legitimacy derives from efficacy, not participation.

The persistent failures of human-led governance systems demand radical rethinking of political structures. High-performance algocracy where optimization algorithms replace electoral politics in technical governance domains represents not just an improvement, but a necessary evolution for societies facing existential challenges. This transition acknowledges three fundamental realities: human cognitive limitations in complex systems, the demonstrated superiority of algorithmic decision-making in technical domains, and the urgent need for governance systems capable of responding to 21st-century challenges with precision and speed.

Traditional democratic systems suffer from well-documented challenges that make them ill-suited for technical governance. As Brennan (2016) demonstrates in Against Democracy, the average voter lacks both the expertise and motivation to make informed decisions about complex policy issues; a phenomenon he terms "political ignorance." To understand the position of Brennam, lets imagine this scenario where people are picking what to order for a big office lunch. Normally, they would vote democratically; where everyone gets a say. But here's the problem: most people don't know which foods are actually good for a work meal too greasy? makes people sleepy?, don't care enough

to research, and just vote for their personal favorite without thinking about what's best for the whole group. Some might even vote randomly just to get it over with. That's "political ignorance" - when voters don't know and often don't bother to learn what they're really voting about. On the other hand, imagine when there is an app that knows Everyone's dietary needs; What foods keep people energized; What fits the budget; Which restaurants deliver fastest. The app would likely choose a better lunch than the flawed group vote. That's the core idea for complex decisions, sometimes data and algorithms might make better choices than uninformed voters picking what "sounds good" without really understanding the issues. Thus, democracy assumes voters will be informed and thoughtful, but in reality, most of the people are busy with their lives and don't have time to deeply research every policy; so they vote based on emotions, headlines, or habit rather than real understanding.

As Landemore (2020) argues, democracy's justification has always been primarily epistemic; its ability to produce good decisions and not procedural. When superior decision-making mechanisms emerge, clinging to outdated forms becomes irrational. This is not to suggest eliminating all democratic participation, but rather to confine it to value-setting while delegating technical implementation to optimized systems. Thus, efficiency is not the enemy of democracy, but it is its prerequisite, such that the adoption of a hybrid model where democratic ends are achieved with algorithmic means remain the best hope for governance that is both legitimate and effective in an increasingly complex world.

**Democratic Ends; Algorithmic Means**

The conventional view pits efficiency against democracy, framing them as opposing values where efficiency prioritizes speed and optimal outcomes, and the Democracy emphasizes participation and deliberation. However, this dichotomy is misleading. In reality, efficiency is not the enemy of democracy but its prerequisite without effective governance, democratic ideals become meaningless, as citizens lose faith in systems that cannot deliver basic services, economic stability, or security (Fukuyama, 2013). The solution to this dilemma lies in a hybrid governance model, where democratic ends such as justice, equity, and collective self-determination are pursued through algorithmic means that enhance, rather than replace, human decision-making.

As rightly observed by Burrell (2016), traditional democratic systems struggle with slow deliberation, cognitive biases, and political gridlock, making them ill-equipped to handle modern challenges like climate change, pandemics, and economic instability. For example, legislative bodies often delay critical policies due to partisan conflicts, while voters lacking expertise make choices based on emotions rather than evidence (Achen & Bartels, 2017). The COVID-19 pandemic exposed these weaknesses better as countries with fragmented decision-making suffered higher death rates, while those that used data-driven responses e.g., South Korea's AI-powered contact tracing, mitigated outbreaks more effectively. If democracy cannot produce timely, rational policies, its legitimacy erodes not because people reject participation, but because they demand governance that works.

Efficiency is often mischaracterized as a technocratic imposition that sidelines public input. However, when properly structured, algorithmic governance can enhance democracy by improving policy accuracy. AI can analyze vast datasets to predict the outcomes of welfare programs, infrastructure projects, or environmental regulations thereby reducing wasteful spending and unintended harms (Helbing et al., 2019). On the other hand, human bureaucracies are vulnerable to favoritism and discrimination. But, algorithms, if transparently designed, can enforce rules impartially. These examples show that efficiency does not suppress democracy—it makes it more responsive. When governments fail to deliver services, citizens grow disillusioned; when they succeed, trust in institutions strengthens (Rothstein, 2011).

The hybrid governance system harmonizes human oversight with algorithmic precision, creating a synergistic model where democratic values guide AI-powered execution while maintaining essential safeguards for transparency and accountability. At its core, this hybrid approach begins with democratic deliberation—where citizens and elected representatives collectively define societal priorities and overarching goals, such as reducing inequality or achieving social justice, through inclusive public debate. Once these values are established, sophisticated algorithms take over the complex task of policy implementation, leveraging real-time data analytics to optimize resource allocation, adjust economic levers, and monitor outcomes with a level of precision and responsiveness impossible for human bureaucracies alone (Floridi, 2023).

Crucially, this system maintains robust democratic checks through two key mechanisms: first, by ensuring complete transparency in algorithmic decision-making processes, and second, by preserving citizens' fundamental right to challenge, audit, and when necessary, override automated decisions (Danaher, 2016). Climate policy provides a compelling illustration of this model in action, legislatures would democratically set binding emissions targets after thorough public deliberation, while AI systems would dynamically manage carbon budgets across sectors, optimize energy pricing structures, and ensure compliance without being distorted by special interest lobbying. An independent oversight body, composed of both technical experts and civil society representatives, would continuously verify that these algorithmic executions align with societal values and constitutional principles (Mittelstadt, 2019). This carefully balanced framework elegantly navigates between the shortcomings of pure democracy such as chronic inefficiency, policy stagnation, and vulnerability to populist distortions and the dangers of pure algocracy such as opaque technocratic rule that erodes human agency. By distributing roles according to their inherent strengths; human judgment for value determination and ethical oversight, algorithmic precision for complex optimization and execution; this hybrid model represents the most viable path forward for governance systems that must simultaneously address urgent global challenges while maintaining their democratic legitimacy in an increasingly complex world.

The greatest threat to democracy is not efficiency but irrelevance; if governments cannot solve problems, citizens will seek alternatives, whether authoritarian strongmen or corporate technocracies (Zuboff, 2019). A hybrid system that marries democratic legitimacy with algorithmic effectiveness offers a viable path forward. By letting humans define what they want and algorithms determine how to achieve it best, humans can build governance that is both fair and functional thereby fulfilling democracy's promise rather than abandoning it.

**References**


Acemoglu, D., & Robinson, J. A. (2019). *The Narrow Corridor: States, Societies, and the Fate of Liberty: Winners of the 2024 Nobel Prize in Economics*: Penguin UK.

Achen, C. H., & Bartels, L. M. (2017). Democracy for realists: Why elections do not produce responsive government.

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2022). Machine bias. In *Ethics of data and analytics* (pp. 254-264): Auerbach Publications.

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

Danaher, J. (2016). The threat of algocracy: Reality, resistance and accommodation. *Philosophy & technology, 29*(3), 245-268.

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

Floridi, L. (2023). The ethics of artificial intelligence: Principles, challenges, and opportunities.

Forst, R. (2012). *The right to justification: Elements of a constructivist theory of justice* (Vol. 46): Columbia University Press.

Frankfurt, H. (2018). Freedom of the Will and the Concept of a Person. In *Agency And Responsiblity* (pp. 77-91): Routledge.

Fukuyama, F. (2013). What is governance? *Governance, 26*(3), 347-368.

Gilens, M., & Page, B. I. (2014). Testing theories of American politics: Elites, interest groups, and average citizens. *Perspectives on politics, 12*(3), 564-581.

Habermas, J. (2015). *Between facts and norms: Contributions to a discourse theory of law and democracy*: John Wiley & Sons.

Helbing, D., Frey, B. S., Gigerenzer, G., Hafen, E., Hagner, M., Hofstetter, Y., . . . Zwitter, A. (2019). Will democracy survive big data and artificial intelligence? *Towards digital enlightenment: Essays on the dark and light sides of the digital revolution*, 73-98.

Landemore, H. (2020). Open democracy: Reinventing popular rule for the twenty-first century.

List, C. (2023). Agential possibilities. *Possibility Studies & Society, 1*(4), 461-470.

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature machine intelligence, 1*(11), 501-507.

Mulgan, T. (2016). Superintelligence: Paths, dangers, strategies. In: Oxford University Press.

O'neill, O. (2002). *Autonomy and trust in bioethics*: Cambridge University Press.

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*: Harvard University Press.

Pettit, P. (2012). *On the people's terms: a republican theory and model of democracy*: Cambridge University Press.

Rawls, J. (1993). Political Liberalism Columbia University Press. *New York, 1993*, 162.
Rothstein, B. (2011). *The quality of government: Corruption, social trust, and inequality in international perspective*: University of Chicago Press.
Sandel, M. J. (2020). *The tyranny of merit: What's become of the common good?* : Penguin UK.
Skitka, L. J., Mosier, K. L., & Burdick, M. (1999). Does automation bias decision-making? *International Journal of Human-Computer Studies, 51*(5), 991-1006.
Starr, S. B. (2014). Evidence-based sentencing and the scientific rationalization of discrimination. *Stan. L. Rev., 66*, 803.
Sullins, J. P. (2011). When is a robot a moral agent. *Machine ethics, 6*(2001), 151-161.
Vallor, S. (2016). *Technology and the virtues: A philosophical guide to a future worth wanting*: Oxford University Press.
Volkov, M. (2025). The Root of Algocratic Illegitimacy. *Philosophy & technology, 38*(2), 1-15.
Wiener, N. (2019). *Cybernetics or Control and Communication in the Animal and the Machine*: MIT press.
Zuboff, S. (2023). The age of surveillance capitalism. In *Social theory re-wired* (pp. 203-213): Routledge.