Making Sense of Digital Domination

Jonne Maas

Technology, Policy and Management
Delft University of Technology
Delft, The Netherlands
j.j.c.maas@tudelft.nl

Abstract

This extended abstract discusses the three approaches to epistemic risk in general-purpose AI that characterize my research: measurement, design, and transparency.

Introduction

Digital technologies in general—and platforms such as Google, Facebook, and Uber in particular—are notorious for their concentrated and uncontrolled power. This has led to concerns about algorithmic or 'digital' domination as understood in the republican tradition (Aytac 2022; Muldoon and Raekstad 2023; Susskind 2022; Hoeksema 2023). According to this tradition, an agent is in a dominating position when they have the power to interfere in another agent's choices in an arbitrary or unchecked manner (Pettit 1997; Lovett 2010). As the debate currently stands, much of the literature has been primarily focused on how the economic, social, and political power of online platforms produce various relations of domination (Aytac 2022; Muldoon and Raekstad 2023). What is lacking, however, is a sustained argument for why AI systems in core societal sectors (henceforth 'core AI systems') such as healthcare, finance, judicial, and public governance dominate. Especially given that they are used in decision-making processes to allocate various goods like social benefits and healthcare.

Susskind (Susskind 2022) is one exception, who discusses digital domination in the context of such AI systems. However, Susskind moves between online platforms and core AI systems, which does not provide a coherent account of why these systems count as dominating. Moreover, in his positive proposal to mitigate digital domination, he focuses solely on social media platforms, and doing so offers little guidance on how we can mitigate digital domination beyond online platforms. The result is that it remains unclear whether it is coherent to hold that AI systems in core societal sectors count as digitally dominating at all, despite possessing the relevant hallmarks.

In this paper, I provide an account of digital domination in the context of core societal sectors. The basic idea is that core AI systems have arbitrary power over citizens' basic liberties. Furthermore, I argue that the root cause of this digital domination can be found in broader societal structures, particularly various legal limitations. In what follows, I discuss the concept of domination, outline how this applies to AI in core societal sectors, and argue that digital domination is a concern that applies to all citizens of modern, digital societies. I conclude that the solution is to shift our focus to the broader societal context in which AI is developed and deployed.

In republican theory, domination indicates a relation between two agents in which one agent has the arbitrary power to interfere with the other agent's choices (Pettit 1997). Interference, here, I take to be the imposition of negative social constraints and coercion. With arbitrary power, I mean power that is not under the control of the subordinate agent. Think of a master-slave relation, where the slave has no control over how the master will interfere with him. There are no means to contest the master's power or to hold the master accountable. This is a classic case of domination.

For republicans, a just society is a non-dominating society in which democratic processes force the government to track the public's best interests (Pettit 2012). The 'public's best interests' refer to the necessary features to function in life, as agreed upon by the members of a given society. These features—or basic liberties—must be equally accessible to each citizen. A non-dominating society thus entails that every member of a society can enjoy the basic liberties in an equal manner.

For digital domination to occur beyond online platforms, citizens would have to be subject to an arbitrary power to interfere in their basic liberties. This is why core societal sectors such as medical, financial, judicial, or public governance are relevant. These sectors enable people to act on their basic liberties and must therefore function to promote society's best interests. That is, citizens must have equal access to these sectors and be able to hold the decisions made in these sectors accountable.

Increasingly, these sectors are algorithmically structured (Gabriel 2022). Citizens depend on the AI systems deployed in these sectors, constituting a power relation between AI systems (including their developers and deployers) and the people subjected to the system (Maas 2023). With well-known algorithmic biases, maintaining citizens' equal opportunities is under threat. For example, if someone does

not receive the health care they need, this may affect their basic liberties of having an equal opportunity to function in society by preventing them from finding a job. As a result, basic liberties can be undermined by digital technologies.

One might object that such systems can be held to account. However, AI systems often pose technical limitations that complicate accountability. AI systems are often referred to as a black box, where we know their input and output, but we do not know how the AI system precisely turned the input into the specific output (Burrell 2016; Lipton 2018). This technical opacity contributes to accountability gaps, as it becomes difficult to hold someone accountable for a particular decision if we do not know what decision made by whom caused the system to produce harmful output (Matthias 2004; de Sio and Mecacci 2021). Accountability plausibly requires explainability, yet the opacity of such systems makes explainability impossible.

Even if the explainability issue is resolved, there is a more fundamental limitation with regard to decision-making power. Whereas explainability is arguably a necessary condition for accountability, it is not sufficient. Consider a dictator who provides explanations about his dictator-particularities, such as when, how, and why he will interfere with his subjects. Although his explainability provides a greater insight in his actions, it does not actually force him to track his subjects' best interests. Similarly, despite the best and necessary efforts to increase a system's explainability, this offers no guarantee of tracking the best interests of the citizens subjected to these systems.

What does provide such a guarantee is regulation and legislation enforced by a democratically legitimate body alongside the active input of the citizens. Numerous scholars have urged more public engagement during the design and development of AI systems. They argue that a form of direct or participatory design is essential to better align the system's output with societal values, norms, and desires, and to address the power relations underlying AI development and deployment (Powles and Nissenbaum 2018; Birhane et al. 2022). The problem with this proposal, however, is that 'participation' does not necessarily transform underlying power asymmetries (Birhane et al. 2022). As a result, citizens' inputs can fail to be taken into account, and they are left with no recourse available to them. As a result, participation does not in and of itself provide a meaningful sense of accountability.

Instead, a meaningful approach to participatory design requires formal background structures (i.e., regulation and legislation) that give citizens the relevant 'participation teeth'—or what (Pettit 1996) calls 'antipower'—to balance out the power-dependency relation between AI and the people subjected to it. Such participation teeth comes about through formal regulation of the inclusion of citizens during decision-making procedures underlying AI development and deployment. As the literature currently stands, such decision-making power primarily rests with the developers and deployers, and included citizens have no concrete tools and mechanisms at hand to force developers and deployers to include their input (Delgado et al. 2023). If citizens could fall back on regulation and legislation that would provide them such tools, citizens can bite their way into AI decision-making. This paves the way to non-dominating core AI systems.

References

Aytac, U. 2022. Digital Domination: Social Media and Contestatory Democracy. *Political Studies*, 00323217221096564.

Birhane, A.; Isaac, W.; Prabhakaran, V.; Diaz, M.; Elish, M. C.; Gabriel, I.; and Mohamed, S. 2022. Power to the People? Opportunities and Challenges for Participatory AI. In *Proceedings of the 2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '22)*, Article 6, 1–8. New York, NY, USA: Association for Computing Machinery.

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

de Sio, F. S.; and Mecacci, G. 2021. Four responsibility gaps with artificial intelligence: Why they matter and how to address them. *Philosophy & Technology*, 34: 1057–1084.

Delgado, F.; Yang, S.; Madaio, M.; and Yang, Q. 2023. The Participatory Turn in AI Design: Theoretical Foundations and the Current State of Practice. In *Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '23)*, Article 37, 1–23. New York, NY, USA: Association for Computing Machinery.

Gabriel, I. 2022. Toward a theory of justice for artificial intelligence. *Daedalus*, 151(2): 218–231.

Hoeksema, B. 2023. Digital Domination and the Promise of Radical Republicanism. *Philosophy & Technology*, 36(1): 17.

Lipton, Z. C. 2018. The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, 16(3): 31–57.

Lovett, F. 2010. *A General Theory of Domination and Justice*. OUP Oxford.

Maas, J. 2023. Machine learning and power relations. *AI & Society*, 38(4): 1493–1500.

Matthias, A. 2004. The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*, 6: 175–183.

Muldoon, J.; and Raekstad, P. 2023. Algorithmic domination in the gig economy. *European Journal of Political Theory*, 22(4): 587–607.

Pettit, P. 1996. Freedom as antipower. *Ethics*, 106(3): 576–604.

Pettit, P. 1997. *Republicanism: A Theory of Freedom and Government*. Oxford University Press.

Pettit, P. 2012. *On the People's Terms: A Republican Theory and Model of Democracy*. Cambridge University Press.

Powles, J.; and Nissenbaum, H. 2018. The Seductive Diversion of "Solving" Bias in Artificial Intelligence. https://onezero.medium.com/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53.

Susskind, J. 2022. *The Digital Republic: On Freedom and Democracy in the 21st Century*. Bloomsbury Publishing.