

Legitimacy of what?: a call for democratic AI design

Jonne MAAS^{a,1} and Juan M. DURÁN^a
^a*Delft University of Technology*

Abstract. The legitimacy of AI decision-support systems raises several concerns. Especially in the field of public decision-making, the consequences of the use of opaque machine learning systems have been under critical investigation as these systems jeopardize democratic rights like transparency and contestability. To question how such AI systems affect democratic rights is worthwhile, as legitimacy is concerned with justified exercises of *public* (i.e. the state's) power. However, the focus on the system itself (and its consequences for democratic rights) overshadows a distinct type of legitimate decision-making, namely the legitimacy of the design decisions - which we understand as exercises of power - underlying the development process of an AI system. Such decisions can affect the public domain, even when the system is designed privately. Based on two contextual case studies, one in the legal (public) domain and one in the medical (private) domain, we argue that the legitimacy of an AI system predominantly depends on legitimate design decisions, which we claim should be rooted in democratic ideals. First, following political philosophers like Rawls and Cohen, democratic procedures are a requirement for overall fairness. Second, following political philosophers like Pettit, democratic procedures ensure legitimate exercises of power. Hence, democratic design is paramount to a system's overall fairness and ensures control of public power inherent to design decisions.

Keywords. AI systems, legitimacy, decision-making, democratic design

1. Introduction

As AI systems are increasingly used in decision-making processes, the validity of said decisions has been under critical investigation regarding their legitimacy, particularly in the field of public decision-making and governance (e.g., Montreal Declaration for Responsible AI 2018; de Fine Licht & de Fine Licht 2020). In particular, most questions are focused on the extent the output of the system affects political and moral legitimacy in current societies (Danaher 2015). This focus makes sense, given that legitimacy directly corresponds to public authority and the justification of the authority's exercises of power (Peter 2010).

Although the attention to the system's use and consequences for the legitimacy of public decision-making itself is worthwhile, such a narrow focus neglects the distinct worry of the legitimacy regarding decisions made during the development of an AI system. Design decisions include value-judgements, making them a political act (Busuioic 2020) and should be seen as exercises of power (Maas 2022). Some of these

¹ Jonne Maas, Delft University of Technology, Faculty of Technology, Policy and Management, Jaffalaan 5, 2628 BX Delft, The Netherlands; E-mail: j.j.c.maas@tudelft.nl.

decisions can have a significant impact on an end-user's life, regardless whether the system is developed and used in a public or private sphere. "Public" decision-making then becomes a flexible term and requires a re-evaluation of legitimacy in the context of AI systems.

In this paper, we distinguish between two different decision-making processes regarding AI systems that are subject to a legitimacy evaluation: (1) non-trivial design and development decisions² that influence a system's behaviour, and (2) the AI system itself as being part of a decision-making procedure. Based on two contextual AI applications in the public, legal (COMPAS) and medical, private (NarxCare) domain, we argue that it is in fact the legitimacy of non-trivial design and development decisions that predominantly contributes to the overall legitimacy of an AI system. Moreover, the overall legitimacy is improved through democratic design procedures.

Our argument is as follows. The legitimacy of a system depends both on the legitimacy of the development process as well as the legitimacy of the system's output. The legitimacy of the system's output, in turn, depends on its context of use (i.e., public vs. private). Yet *even if* the system is used in a legitimate manner, when the design decisions are made illegitimately the system cannot be fully legitimate. The opposite, however, is possible: when design decisions are legitimate, it is possible for the system to be fully legitimate depending on the context of use. Because of the potentially strong public influence of design choices, it is not enough that the system itself is deployed legitimately, it is essential that these choices are embedded within legitimate decision-making procedures (that we argue should be rooted in democratic ideals). Legitimate design decisions are hence *necessary* but not necessarily *sufficient* for a legitimate AI system.

2. Legitimate decision-making

In political theories, an entity is legitimate when "it is morally justified in wielding political power", with political power understood as "the attempt to make, apply, and enforce rules monopolistically over the broadest class of citizens" (Buchanan 2002, 689-690). Political power is handed down by the state to smaller institutions through laws, allowing individual public agents (e.g., a policeman) and public institutions (e.g., tax affairs, schools) besides the government itself to exercise this power. So, legitimacy and political power are concerned with exercises of power in the public realm.

There are numerous political theories about what makes an authority 'legitimate' and an exercise of power 'justified'. The traditional understanding of legitimacy follows consent theory, with the idea that once one consents to an authority (1) this authority has legitimate power over the consenter and (2) the consenter is both obliged to obey laws, rules and orders. This view, however, has been contested for numerous reasons, including consent being an 'unsatisfiable demand' (we cannot expect each citizen to agree to *all* exercises of political power); an 'unnecessary demand' (we managed to form a state prior to consent); and an 'insufficient demand' (citizens cannot be expected to

² With 'non-trivial design and development decisions' we refer to design choices that have a significant impact on a system's end-user. For instance, how to define a concept like 'fairness' has non-trivial effects on the system's end-user, as one interpretation need not have the same effect on a system's end-user as another interpretation.

obey immoral acts) for political legitimacy and authority (Buchanan 2002, 699-702; see also Simmons 1980).

More recently, scholars converge on the view that democratic procedures are required for legitimate authority and legitimate exercises of power. Some argue that democratic procedures lead to the best outcome and the 'best outcome' makes a decision legitimate (e.g., Estlund 2008) and others claim that democratic procedures ensure equal treatment of persons (e.g., Rawls 1997; Cohen 1996; Cristiano 2004; see also Viehoff 2014 for an overview). These views need not be separate. For instance, Rawls' account of justice as fairness combines procedural and substantive fairness. On his account, a society requires necessary elements (e.g., basic liberties, public reason) that ensure a legitimate democratic state that upholds fair decisions for its citizens. Rawls formulates these elements from the original position. He imagines, using the 'veil of ignorance', what people would reasonably agree on if they would not know their place in society. Based on these elements, or 'procedures', Rawls develops his own theory of what a just society should look like, one rooted in democratic ideals of freedom and equality that are expressed in the basic liberties. So, the idea is that democratic procedures not only contribute to fairness and 'perceived legitimacy' by the public, but actually also improves the overall epistemic output of the decision-making process.

This is possible, Rawls (1997) argues, through the notion of 'public reason', according to which a decision must be based on acceptable reasons by rational human agents (see also Cohen 1996). In line with this ideal of public reason is the moral requirement that those subjected to the decision-making are provided with an acceptable reason, requiring both (1) reason-giving of the decision-maker as well as (2) some degree of transparency to share the reasons. These two requirements together form the 'principle of publicity' (Cristiano 2009).

Other political scholars claim that democratic procedures in the form of contestation are the only way to ensure legitimate exercises of power thanks to public control (e.g., Pettit 2012). Without the possibility to contest decisions, one is subjected to an uncontrolled power. Legitimate authority, then, is an accountable authority. Following this line of thinking, the only possibility to ensure contestability is through involving the public (Pettit 2012). One way or another, public reason or public contestation, in contemporary political philosophy democratic thought is central (though not exclusive) to legitimate authority. This implies that a legitimate authority, such as in the case of a public decision-maker, is itself subjected to democratic norms to ensure that the democratic rights of those subjected to the decisions are respected³.

3. Issues with Legitimacy and AI

Since AI systems are used in the making, applying and enforcing of rules in the public realm, these systems are now subject to questions of political power and legitimacy as well. AI systems, in particular opaque machine learning systems, create an obstacle for

³ Note that we distinguish between 'legitimacy' and 'justice' as two separate concepts. Whereas Rawls conflates these concepts in his theory of justice, Pettit for instance clearly distinguishes between these two. For Pettit, 'justice' is *what* the state imposes (which on his account should be non-domination) and 'legitimacy' is *how* the state imposes this (which on his account should be done in a non-dominating manner). See Pettit (2012a; 2012b, 107, particularly fn 21).

legitimacy due to arising responsibility gaps and the ‘black-box’ phenomenon (Matthias 2004; de Fine Licht & de Fine Licht 2020). The lack of accountability mechanisms and lack in transparency threaten democratic rights (e.g., Maas 2022), the public’s perceived legitimacy of the system (e.g., de Fine Licht & de Fine Licht 2020) and particular democratic ideals such as Cristiano’s principle of publicity (e.g., Jebari et al. forthcoming). Given the impact of AI systems on legitimate authority, it is hence understandable why legitimacy is a big concern in the context of AI systems used for public decision-making.

Nonetheless, we argue that this focus on legitimacy for public decision-making obscures other, more fundamental concerns regarding legitimacy for AI systems. In particular, while these issues regarding one’s conceptualization of legitimacy (perceived vs. moral) and the potential concerns with the use of an AI system in public decision-making (responsibility gaps, lack of accountability mechanisms) are essential to highlight, they also show that the literature lacks conceptual clarity regarding legitimacy in the context of AI systems. Besides this lack of conceptual clarity, these concerns predominantly relate to the system itself, and to what extent its use increases or decreases legitimate exercises of power.

This focus on the system has three main concerns. First, it suggests that the main problem lies with the system. This quickly results in a discussion on whether to use or not use the system. Such a binary focus overshadows potential benefits if we choose to not use the system and risks decreasing the legitimacy of the public authority if we do use it. This, then, creates a dilemma without the potential for a ‘win-win’ solution, where we have both the benefits of the system *and* a legitimate authority (morally as well as perceived).

Second, and more importantly, a focus on the system itself risks overlooking the social aspects inherent to the system’s output. Legitimacy regarding AI systems is more complex than whether the system’s use and output itself is ‘legitimate’. Decisions made by developers and users can have a significant impact on the end-user, for instance when deciding how to interpret a complex concept like fairness. Different interpretations have different effects on end-users, and hence such interpretations and decisions become a “political act” (Busuioc 2020, 7). The point is that when we discuss the legitimacy of an AI system, it is essential that we take these social aspects into account. Legitimacy does not involve merely the AI’s decision. It similarly involves the legitimacy regarding the decision-making processes of the developers and users.

And this leads us to our third concern. Precisely because of the social aspects inherent to the development of an AI system, and hence to the output of an AI system, the line between the public and private becomes increasingly blurred. An AI system used for public decision-making is clearly part of the public - and hence political - domain. Yet what about systems that are used privately but have a public effect? With social media platforms, we increasingly see the interconnectedness between private companies and the public sphere. Decisions whether to filter misinformation or other forms of harmful content rest in the hands of private companies, even though these decisions have public effects (e.g., Myanmar genocide, Cambridge Analytica, Covid-19 misinformation on vaccinations).

Particularly with the second and third concern, we see how it is irrelevant whether non-trivial design and development decisions are made privately or publicly as they can have a public effect nonetheless. For this reason, we argue it is essential to distinguish between two different - though related - types of legitimacy: one relates to the decision-making process regarding the development and deployment of said system - regardless

whether this is in the private or public realm - and the other relates to the *output* of a system – and to what extent the system’s use and particular decision is legitimate. More precisely, the former focuses on the social aspects behind the AI system, the latter focuses on the system itself. In Figure 1, we have highlighted these different types of decision-making, with the first type being referred to as influence #1 and the second type as influence #2.

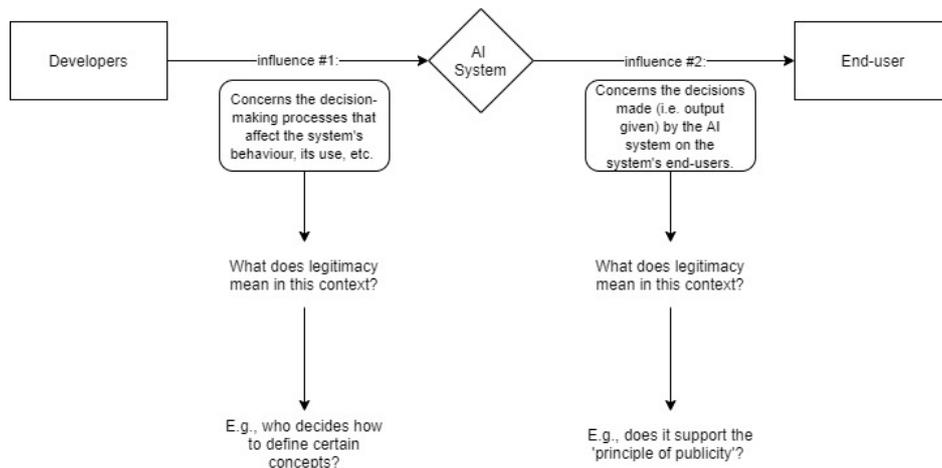


Figure 1. Legitimacy of what? Different types of legitimacy in the context of AI systems

To illustrate these different types of legitimacy, consider the following two AI systems used in different domains. In the legal domain, we have the notorious case of COMPAS, a risk-scoring system to predict a criminal’s likelihood to recidive. Jebari et al. (forthcoming) point out that the use of statistical similarity within judicial courts is democratically illegitimate, as it defeats the principle of publicity formulated by Thomas Cristiano. According to this principle, those subjected to the rule of law, or ‘law-takers’ must be treated as an individual. Yet since COMPAS uses statistical similarity, COMPAS is at odds with this democratic ideal. In this context, Jebari et al. convincingly argue why COMPAS is illegitimate, as its output is considered democratic illegitimate in the framework of Cristiano’s principle of publicity.

Compare COMPAS to the medical domain, where people are often known to be treated based on statistical similarity (e.g., Karmen et al. 2019). Indeed, many assumptions for diagnosing patients or treatment relies on statistical similarity with previous known cases. A key difference is that when subjected to the law, the individual’s role takes place in the political sphere, and thereby the public domain. When subjected to a doctor’s expertise, one finds themselves in the private domain. And as we saw, legitimacy belongs to the public domain.

Yet is this then to say that all ‘private’ AI systems cannot be assessed in terms of legitimacy? We argue that the common understanding of legitimacy and its relation to political power in the public sphere overshadows non-trivial and in fact politically laden decisions *during* the design and development of a system. Indeed, Jebari et al. make an interesting case why COMPAS is illegitimate. However, this case relates to the decision

made by the AI system (i.e., influence#2), and does not include decisions that are made during the design process (i.e., influence#1). Arguably, there is a difference in whether it is democratically legitimate to use a system as decision-maker (or decision-support) and whether the decision-making process during the development of a system is legitimate.

Consider the AI system NarxCare, used to predict a patient's risk for drug overdose (Szalavitz 2021). One assumption implicit in this system is that people who travel far are at higher risk for drug abuse. People who live in a rural area, however, necessarily travel further for their drug prescriptions. The assumption 'travel far implies drug abuse' does not hold for people in rural areas. In fact, the use of either wrong data or problematic assumptions made by the developers contributes to discriminatory bias. The underlying assumption then has political connotations that extend into the public realm, even when used in a private context that is commonly less considered with issues of political legitimacy.

What we can learn from these scenarios is that legitimacy in the decision-making process of AI developers can mean something different than legitimacy in the decision-making of an AI system itself. While for decisions made by judicial AI systems there is a strong need for democratic legitimacy, individual-based reasoning, and transparency, decisions made by medical systems are less concerned with these matters as in this case end-users are 'patients' rather than 'law-takers'. When we discuss legitimacy of an AI system, we can discuss the output in a particular context, such as Jebari et al. do in the context of the legal systems. Yet legitimacy regarding developmental decisions become more complex, as it is unclear to what extent a particular decision affects an end-user in a non-trivial way. When designing AI systems, particularly the notorious opaque ML systems, social factors matter regarding a system's behaviour, and hence how the system affects its end-users.

4. Decision-making, Developers, and Value-judgements

As Busuioc shows with the concept of fairness and as the example of NarxCare illustrates with assumptions underlying complex social phenomena, at least some design decisions have a political dimension due to the value-judgements embedded within these decisions. This then raises the question to what extent such political decisions should be legitimate, and building on this, what such a legitimate decision-making process in the context of AI development must look like. With regards to the first question, we argue that design and development decisions should be legitimate. Design and development decisions entail exercises of power (Maas 2022), more precisely *political* power when these decisions have a significant influence on the public realm. Legitimacy is hence essential to justify this political power.

Regarding the concept of fairness, besides the fact that this concept has multiple interpretations, it is also the case that different societies - or even communities within one society - may not agree on the same interpretation, yet a system can only accommodate one particular interpretation. This becomes an issue when a system is used intersocially, where one interpretation is favored over the others without accounting for cultural differences.

In particular, such favoring is but one facet of the 'fossilization' of concepts in ML. A required practice in ML is to code concepts that work as the basis for its ontology. But whereas some are concrete concepts (e.g., house, human, rectangle, etc.), some are more

abstract (e.g., health, criminality, humanity). Both sets of concepts are treated equally by ML, but whereas the first group does not introduce special considerations, the second group becomes fossilized or naturalized. In other words, once a concept is coded into an ML system, it is universally and indistinguishably applied across large and heterogeneous databases with varying degrees of success. Consider the concept of “health”. One interpretation takes statistical measurements and standards of normal biological measurements of the body as the baseline that indicates when a person is healthy. But the concept of health can also be interpreted as social constructions based on what the relevant individuals or groups consider to be healthy (Richman, 2004). For instance, a group that considers blood transfusion as harmful will treat any of its members who have received a transfusion as unhealthy (Richman, 2000). If the ML implements the former interpretation of health, the concept is fossilized as a measurement of the standard bodily function. As a consequence, all groups of patients will be treated equally, regardless of the interpretation of “health”. It is situations similar to these that are at the root of Watson for Oncology's difficulties in South Korea and Denmark: the medical knowledge, practices, and mores in these places differ, however marginally, from the standard American patient.

In the previous section, we highlighted the nuance between a system's output and use and the design choices made during the development process by distinguishing between the two types of ‘legitimacy’ related to AI systems. In this section, we see how non-trivial design decisions become incorporated within a system. Especially with the global use of certain AI systems and the potential for the fossilization of concepts along with their implicit value-judgements, it is essential that design choices are rooted in legitimate decision-making procedures. And though other artifacts raise similar moral concerns (Winner 1980; Verbeek 2008), the opaque and learning characteristics of an ML system make it difficult to trace back particular decisions and how they ultimately affect the output, increasing potential threats of fossilized concepts. By focusing on the design process of an AI system, we see that when discussing the legitimate use of an AI system we must always consider the value-judgements made during the design process as well. This furthermore implies that legitimacy regarding design decisions relates to *all* AI systems whereas legitimacy of a system's use and output is more context-dependent.

So far, we have argued that legitimate decision-making *during* the development process of an AI system is essential due to the political dimension of design decisions. Moreover, we have seen that a legitimate decision-making procedure during the development process is *necessary* for an AI system to be legitimate (though not sufficient). Whereas the context of use (public/private) determines legitimacy requirements for the system itself (such as with COMPAS) we see that design decisions - even when done privately - can extend into the public realm. So whereas it is more evident in how to set ‘legitimacy assessment criteria’ regarding a system's use, this becomes more of an issue regarding design choices. In particular because the effects of these decisions might not always be clear beforehand (such as with NarxCare), and unwanted biases or other moral concerns arise *after* the system has been already put into use.

5. Democratic design for improving a system's legitimacy

Having argued in favour of a legitimate decision-making procedure during an AI system's development, we now argue that AI design should be democratic to promote legitimate decision-making. First, we have seen that some scholars claim that democratic procedures ensure overall legitimacy and improve the epistemic output of a state. Where for Rawls (and others) democratic procedures ensure a legitimate state, we could argue that an AI system increases in overall legitimacy if the system's output rests on a legitimate decision-making procedure, grounded in democratic norms, during the development process of the system itself. The developers of an AI system are those that find themselves in the position to develop an AI system that conforms to a particular interpretation. A democratic procedure, similarly to a state's democratic procedures, hence enforces the legitimacy and increases the epistemic value of a system's output itself⁴.

Our second reason is rooted in theories developed by neo-republican scholars like Pettit (1997; 2012), who argue that the only means to achieve such legitimate exercises of power in the context of the state is through democratic control, as this provides people the means to control the governmental authority. Thus, if we extend this idea to AI decision-making, we need democratic control regarding the decision-making processes of the developers and users. This reason supports why a legitimate development process is necessary for a system's overall legitimacy: all systems include value-judgements - i.e., exercises of power - and therefore democratic design provides an overarching means to ensure at least *some* legitimacy for AI systems.

Note that democratic design ensures a system's overall legitimacy that other political theories for legitimacy fail to achieve. We, people of a society, are necessarily subjected to AI systems - both in public and private domains. We are not offered to 'consent' to this subjection of power yet even if we were we might wonder whether that would increase legitimacy. Besides current issues with legitimacy and consent (e.g., cookies on the internet that we need to accept in order to use a website), consent to be subjected to an AI system reflects the binary dimension of the use/do not use legitimacy question: either we consent and have legitimacy when subjected to an AI system, or we do not consent and miss out on potential benefits of the AI system. 'Legitimacy' then becomes subjective: if I choose to consent to be subjected by the system, the system then is legitimate for me. If you do not, the system is not. This seems fundamentally morally problematic. Moreover, it does not increase the 'overall' legitimacy of the AI system. It only increases the legitimacy with respect to individual relations with the system. By focusing on legitimacy *during* the development process, we open up the third possibility of legitimacy through democratic design that works for all citizen relations with the system which actually facilitates the system's 'overall' legitimacy.

As an additional benefit, democratic design may diminish problematic consequences. If, for instance in the case of NarxCare, citizens living in rural areas would have been

⁴ Such reasoning is already supported by several responsible design and assessment methodologies, such as Responsible Research and Innovation (RRI) (Owen, Macnaghten & Stilgoe 2020) or Participatory Technology Assessment (pTA) (Joss & Bellucci 2002). We, however, go a step further and argue that democratic design does not just contribute to responsible development (by increasing overall legitimacy). Democratic design is essential to control private exercises of political power.

able to interfere with the system's design, they may have been able to point out that the developers' assumption 'travel far implies drug abuse' could lead to discriminatory output in their situation. This is but one example of how democratic procedures not only treat people as equals and ensure that exercises of power are grounded in public control, thereby creating an accountable authority, but also shows how democratic procedures contribute to a more epistemic reliable output of the system itself.

Finally, a concern with technologies is how to justify their use if they can have unforeseen morally problematic consequences. With development of AI systems, in particular opaque machine learning systems, not only does implementation exceed regulation speed, yet *ex post* responsibility and accountability is especially challenging (Maas 2022). Democratic design and improved legitimacy of an AI system contributes to the justification of using these systems when we know that there are likely to be some negative consequences - even when unknown what they are.

6. Concluding remarks

In this paper, we have attempted to highlight the important nuance regarding legitimacy in the context of AI systems: not just the output and use of the system should be assessed in terms of legitimacy, also the design choices during the development process have a political dimension and hence require moral justification. Moreover, precisely because *all* AI systems (public and private) entail value-judgements that have the potential to fossilize or naturalize standards in society, we argue in favour of a strong focus on a legitimate development process. 'Legitimacy of what?', then, is answered with 'design choices'. In order to achieve such a legitimate process, we have argued that democratic design best ensures a system's overall legitimacy. Unlike other political theories for legitimacy such as consent theory, design rooted in democratic theory ensures (1) better epistemic output, (2) equality between people, (3) controlled exercises of power, and (4) actual *overall* legitimacy. Democratic design can thus be used as a justification for uncertain moral concerns.

References

- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). *Machine bias*. ProPublica. Retrieved March 15, 2022, from <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Buchanan, A. (2002). Political legitimacy and democracy. *Ethics*, 112(4), 689-719.
- Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. *Public Administration Review*, 81(5), 825-836.
- Cohen, J. (1996). Procedure and substance in deliberative democracy. *Democracy and difference: Contesting the boundaries of the political*, 95.
- Christiano, T. (2004). The authority of democracy. *Journal of Political Philosophy*, 12(3), 266-290.
- Danaher, J. (2016). The threat of algocracy: Reality, resistance and accommodation. *Philosophy & Technology*, 29(3), 245-268.

- Estlund, D. (1997). Beyond fairness and deliberation: The epistemic dimension of democratic authority. *Deliberative democracy: Essays on reason and politics*, 173.
- de Fine Licht, K., & de Fine Licht, J. (2020). Artificial intelligence, transparency, and public decision-making. *AI & society*, 35(4), 917-926.
- Jebari, K., Beckman, L., Rosenberg, J.H. (forthcoming). Artificial intelligence and democratic legitimacy: the problem of publicity in public authority.
- Joss, S., & Belucci, S. (2002). European Participatory Technology Assessment, Participatory Methods in Technology Assessment and Technology Decision-Making. *Center for the Study of Democracy, Westminster*.
- Karmen, C. L., Reisfeld, M. A., McIntyre, M. K., Timmermans, R., & Frishman, W. (2019). The clinical value of heart rate monitoring using an apple watch. *Cardiology in review*, 27(2), 60-62.
- Maas, J. (2022). Machine learning and power relations. *AI & SOCIETY*, 1-8.
- Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and information technology*, 6(3), 175-183.
- Montréal Declaration Responsible AI. (2018). Retrieved March 15, 2022, from <https://www.montrealdeclaration-responsibleai.com/reports-of-montreal-declaration>.
- Owen, R., Macnaghten, P., & Stilgoe, J. (2020). Responsible research and innovation: From science in society to science for society, with society. In *Emerging Technologies: Ethics, Law and Governance* (pp. 117-126). Routledge.
- Peter, F. (2010). Political legitimacy. Retrieved from <https://plato.stanford.edu/entries/legitimacy/>
- Pettit, P. (1997). *Republicanism: a theory of freedom and government*. Oxford University Press.
- (2012a). Legitimacy and justice in republican perspective. *Current Legal Problems*, 65(1), 59-82.
- (2012b). *On the people's terms: a republican theory and model of democracy*. Cambridge University Press.
- Rawls, J. (1997). The idea of public reason revisited. *The University of Chicago Law Review*, 64(3), 765-807.
- (2020). *A theory of justice*. Harvard university press.
- Richman, K. A. (2004). *Ethics and the metaphysics of medicine: Reflections on health and beneficence*. MIT Press.
- Richman, K. A., & Budson, A. E. (2000). Health of organisms and health of persons: an embedded instrumentalist approach. *Theoretical medicine and bioethics*, 21(4), 339-354.
- Szalavitz, M. (2021, August 11). *A drug addiction risk algorithm and its grim toll on chronic pain sufferers*. Wired. Retrieved March 15, 2022, from <https://www.wired.com/story/opioid-drug-addiction-algorithm-chronic-pain/>.
- Verbeek, P. P. (2008). Morality in design: Design ethics and the morality of technological artifacts. In *Philosophy and design* (pp. 91-103). Springer, Dordrecht.
- Viehoff, D. (2014). Democratic equality and political authority. *Philosophy & Public Affairs*, 42(4), 337-375.
- Winner, L. (1980). Do Artifacts Have Politics?. *Daedalus*, 109(1), 121-136.