



Democratizing AI from a Sociotechnical Perspective

Merel Noorman¹ · Tsjalling Swierstra²

Received: 15 February 2023 / Accepted: 30 October 2023 / Published online: 18 November 2023
© The Author(s) 2023

Abstract

Artificial Intelligence (AI) technologies offer new ways of conducting decision-making tasks that influence the daily lives of citizens, such as coordinating traffic, energy distributions, and crowd flows. They can sort, rank, and prioritize the distribution of fines or public funds and resources. Many of the changes that AI technologies promise to bring to such tasks pertain to decisions that are collectively binding. When these technologies become part of critical infrastructures, such as energy networks, citizens are affected by these decisions whether they like it or not, and they usually do not have much say in them. The democratic challenge for those working on AI technologies with collectively binding effects is both to *develop* and *deploy* technologies in such a way that the democratic legitimacy of the relevant decisions is safeguarded. In this paper, we develop a conceptual framework to help policymakers, project managers, innovators, and technologists to assess and develop approaches to democratize AI. This framework embraces a broad sociotechnical perspective that highlights the interactions between technology and the complexities and contingencies of the context in which these technologies are embedded. We start from the problem-based and practice-oriented approach to democracy theory as developed by political theorist Mark Warren. We build on this approach to describe practices that can enhance or challenge democracy in political systems and extend it to integrate a sociotechnical perspective and make the role of technology explicit. We then examine how AI technologies can play a role in these practices to improve or inhibit the democratic nature of political systems. We focus in particular on AI-supported political systems in the energy domain.

Keywords Democracy · AI · Sociotechnical systems · Practices · Critical infrastructures · Energy

✉ Merel Noorman
m.e.noorman@tilburguniversity.edu
Tsjalling Swierstra
t.swierstra@maastrichtuniversity.nl

¹ TILT, Tilburg Law School, Tilburg University, Tilburg, The Netherlands

² Philosophy Department, Faculty of Arts and Social Sciences, Maastricht University, Maastricht, The Netherlands

1 Introduction

Democratizing Artificial Intelligence (AI) is like building a black-box, off-the-shelf, easy to use, affordable, and accessible microwave. At least this is what Google's Chief Decision Intelligence Engineer, Cassie Kozyrkov suggested.¹ Democratizing, as Kozyrkov used the term, means making complex machines more widely available for the larger public and easier to use in daily practices. In doing so, Google is not alone.² The EU-funded I-nergy project, for instance, similarly aims to democratize AI through their efforts “to evolve, scale up and demonstrate innovative AI as a service energy analytics applications” (Barrientos et al., 2023). Here too democratization of AI roughly equals “developing a tool that can help inexperienced users to implement their own AI models”.³

These proposals to democratize AI refer to a new generation of AI technologies driven by recent developments in machine learning. Such technologies can create models from real-world phenomena by detecting patterns in large volumes of heterogeneous data, allowing them to solve problems or perform tasks with a significant level of autonomy.⁴ These new kinds of AI technologies promise to help solve “some of the world's biggest challenges” including fighting climate change and treating chronic diseases (European Commission, 2018).

However, not everyone is convinced that simply making these technologies more widely available will indeed benefit democratic processes. Their skepticism is based on the fact that many of the decisions delegated to these technologies are ‘collectively binding’: citizens are affected by them whether they like it or not. In particular, when AI technologies are used in public-facing data-driven systems, such as traffic control systems, content monitoring systems on social media platforms, or electricity systems, they have direct and difficult to avoid implications for the daily lives of citizens. If we define ‘politics’ with David Easton (1965) as “the authoritative allocation of value”, then it is clear that in these cases *political* decisions get delegated to AI technologies. Concerns about the democratic legitimacy of AI-based decisions have been voiced in the context of infrastructures related to speech, deliberation, and media (see for example Helberger, 2019; Nemitz, 2018). Furthermore, increasing dependency on corporate-controlled opaque, and unaccountable AI-based services and products can increase power asymmetries and weaken the position of democratic states and citizens (Montes & Goertzel, 2019; Sadowski & Levenda, 2020; Taylor, 2021). In light of these concerns, it is clear that ‘democratizing AI’ requires more than users getting access to technology.

How to democratize AI? That is the central question that we will take up in this paper. A growing body of literature explores how to exert some form of democratic control over AI, to ensure that this technology will not exacerbate existing power

¹ <https://www.youtube.com/watch?v=iLu9XyZ55oI>

² Microsoft made similar claims (see for example Boyd, 2017).

³ <https://i-nergy.eu/democratising-ai-transfer-learning-wizard-for-energy-machine-learning-applications>

⁴ AI is a rather big and malleable umbrella term referring to both a field of science as well as a collection of technologies (HLEG, 2018).

asymmetries or bring new ones into existence (Buhmann & Fieseler, 2023; Duberry, 2022; Nemitz, 2018; Sætra et al., 2022). Some believe that this task should be left to existing political institutions, be it on the supranational, national, and/or municipal level (Himmelreich, 2022). Others formulated more radical proposals, arguing that we additionally need to bring democracy *into* the companies that design, develop, and implement AI technologies (Cuéllar & Huq, 2020; Maas & Durán, 2022).

In this paper, we argue that these proposals are too limited. Calls for external democratic control, we contend, underestimate the extent to which democratic processes themselves are influenced by AI. They fail to take a ‘sociotechnical’ perspective, as developed in Science and Technology Studies (STS) (Sovacool et al., 2020). Such a perspective highlights that human values and interests shape the design and implementation of AI technologies *and* that, at the same time, these technologies shape democratic practices—in good or bad ways. The more radical proposals to democratize the design and development of AI technologies do better on the socio-technical point. However, we will show that their pleas to democratize AI tend to draw on conceptions of democracy that are both too restricted and too rigid. To remedy this flaw, we propose a ‘system approach to democracy’, such as developed by Canadian political theorist Mark Warren, that foregrounds the interplay of different political practices within democratic systems. By drawing on recent critical studies of AI, science and technology studies, and political philosophy, we develop a conceptual framework that provides a radical, rich, and flexible understanding of democratizing AI. More concretely: subjecting AI to democratic control involves developing, implementing, and using these technologies in such a way that they foster democratic practices, which requires an analytical focus that encompasses both the social as well as technical.

To illustrate our argument, we chose an application domain that has been relatively underexplored in discussions about democratizing AI: energy (Cuéllar & Huq, 2020; Judson et al., 2022). AI technologies are, for example, being developed to facilitate the integration of renewable energy sources in electricity grids (Noorman et al., 2023). How an AI technology distributes locally produced solar energy in a neighborhood will influence residents regardless of whether they have solar panels or not, because their energy bills may be affected. Decisions affecting this distribution are therefore political. To what extent and how should citizens be involved in these decisions regarding the energy system?

In Sect. 2, we first discuss how a sociotechnical perspective enables us to see that AI technologies are inextricable parts of political systems. In Sect. 3, we show how the existence of such politico-technical systems provides a strong argument to include technologies in discussions about democracy. In Sect. 4, we introduce Mark Warren’s approach to democracy theory. In Sect. 5 we demonstrate how Warren’s framework helps to critically assess previous calls to democratize AI. Finally, using the example of the energy domain, we examine in Sect. 6 how Warren’s framework extended with a sociotechnical perspective can provide a heuristic in the design and deployment of democratic AI technologies. In this way, we offer a rough conceptual framework to assess and develop approaches to democratize AI in the double sense of being both democratically controlled and fostering democratic practices.

2 A sociotechnical perspective: AI as an actor in a political system

AI, like any technology, cannot be separated from political decision-making practices. It constitutes a form of political power, as it can impose the norms inscribed in the technology (Lessig, 2009). The design of technologies *affords* particular behaviors and designers make choices about what *affordances* to design into the technology (Norman, 2013). The decision to equip a smart meter with an on–off switch that can be remotely operated, for example, is a decision with a political dimension. Not including such a switch protects citizens from hackers being able to shut off entire neighborhoods while including the switch makes it easier for companies to offer services that may benefit certain groups in society, such as prepaid electricity (Van Aubel & Poll, 2019). As research in STS has shown, technological development is characterized by value trade-offs, conflicting interests regarding scarce goods, and the unequal distribution of power (Cuppen, 2018; Marres, 2016). Technologies, thus, have values and political dispositions inscribed in their design (Feenberg, 2017; Winner, 1980).

AI technologies bring new affordances that complement, augment, or displace human behaviors and actions. Their ability to find patterns in large amounts of heterogeneous data, to categorize, classify, forecast, and predict trends, enables the creation of, for example, synthetic texts and images as well as the formulation of risk scores, profiles, or decision trees. These in turn afford the creation of new kinds of relationships, of knowing, and of decision-making. But what they do is shaped by the interests of multiple actors involved in their design, development, and use. And possibly the outcomes go against the interests of stakeholders who had less influence on the design, development, and implementation of AI technologies.

AI technologies should, therefore, always be approached as part of broader configurations in which they are shaped by existing practices, consisting of values, norms, institutions, relationships, multiple different actors, and technologies. And in turn, they can change, shift, or disrupt these practices (Schatzki et al., 2001). Such configurations in which social and technical elements *mutually shape* each other are called *sociotechnical systems*. Such systems extend beyond the technical artifact, its developers, and users: tracing the interconnections between the components of such a system quickly expands the focus to encompass organizations, companies, institutions, other technologies, governments, and much more. The attempt to always keep the two sides of such systems in view we call the *sociotechnical perspective* (Sovacool et al., 2020).

The political relevance of AI technologies in sociotechnical systems can be illustrated by renewable energy communities (RECs). RECs are communities that organize “collective energy actions around open democratic participation and governance and to benefit its partakers” (Di Lorenzo et al., 2022). These communities use various kinds of renewable energy sources, such as wind and solar, to locally produce, consume, and trade energy between members of the community as well as between the community and the grid. They are a key component of governmental strategies to facilitate the energy transition through digitization and

to empower citizens in the energy sector. The EU, for instance, has made these communities an integral part of its ambition to comply with the Paris Agreement by boosting renewable energy and incentivizing consumers to become producers (prosumers) of the energy they consume (Hanke & Lowitzsch, 2020). From a sociotechnical perspective, these RECs are focal points in which many different human actors, technologies, norms, interests, relationships, policies, and institutions come together.

AI technologies are increasingly pervasive components of sociotechnical systems centered on RECs and are delegated decision-making tasks with collectively binding effects. Decentralized and volatile renewable energy sources, central to RECs, lead to growing complexity in balancing supply and demand in electricity systems. AI technologies, such as neural networks, reinforcement learning, and deep learning, are employed to manage this complexity in various ways (Hernandez-Matheus et al., 2022). For RECs, applications include forecasting generation and consumption, demand response, and storage (Di Lorenzo et al., 2022). Demand response technologies help to distribute demand for energy to match the supply of energy. For example, a system may use predictive algorithms to produce price incentives to encourage the use of washing machines or coolers during sunny afternoons when the energy supply from solar panels is high (Vázquez-Canteli & Nagy, 2019). Storage can entail scheduling the charging and discharging of local batteries or electric vehicles (EVs), to balance demand and supply.

The design of the AI technologies used in REC shapes which decisions are taken, how they are taken, and by whom (Noorman et al., 2023). For example, certain machine learning models afford efficient optimization of EV charging but leave little room for drivers to charge their EV according to their preferred time schedule, while other models may be less efficient but allow for flexible time schedules using pricing incentives (Al-Ogaili et al., 2019). Decisions delegated to these technologies, such as who gets to charge their EV at a certain hour, are political. Individuals affected might not agree on how these decisions should be made, but they will still be bound by them in most cases. This will open these 'technical' decisions up to public scrutiny, both regarding their content and the way these decisions were taken. What is a fair way of sharing excess energy or scheduling EV charging and discharging? The design and development of these technologies are shaped by broader negotiations about the energy transition, notions of fairness, priorities in public funding, etc. Within RECs this presents questions about how such negotiations should take place, who should be involved, and who should make the decisions and enforce them. Can citizens be participants in these negotiations? Should decision-making be left to local governments and distribution system operators, or to commercial parties through market mechanisms? The answers to such questions are deeply political as they will have implications for what the technology is required to do, and in turn, for how the technology will affect the choices that can be made.

In the following section, we argue that a conception of democratizing AI should adopt a sociotechnical perspective because of this mutual shaping between AI technologies and political systems.

3 The importance of the sociotechnical perspective on democratizing AI

The increasing use of AI technologies in public-facing infrastructures, such as energy networks, presents challenges for existing democratic governance structures. In the energy domain, a variety of institutions, laws, protocols, (market) mechanisms, norms, and agreements currently ensure that multiple voices can be heard, decision-makers can be held to account, and collectively binding decisions are arrived at in a way that confers some degree of democratic legitimacy on the outcomes. These governance structures are already under strain because of the shift to highly distributed renewable energy sources, increasing electrification, and external developments such as war and climate change (Judson et al., 2022; Marinakis et al., 2021). Introducing new AI technologies adds further challenges as they automate decision-making, complicate accountability, present new choices, and introduce new powerful actors (e.g. global tech companies) (Cuéllar & Huq, 2020). They destabilize existing relationships of power and of legal and democratic control. AI technologies also present risks for democratic decision-making processes, including biased algorithms, opaque and complex systems, loss of control, unclear responsibilities, and growing power asymmetries (Malik, 2020; Mittelstadt et al., 2016; Pasquale, 2015; Whittaker et al., 2018). Moreover, big commercial AI companies are becoming increasingly involved in public governance because governments rely on their products and services (Taylor, 2021). This makes it hard to opt out of engaging with AI firms.

In response to the potentially disruptive effects of new AI technologies, scholars have called for democratizing AI (Sætra et al., 2022). They emphasize a range of different aspects of the democratic process. They have called for more accountability and transparency, for explicit articulation of values, or for more stakeholder participation (Buhmann & Fieseler, 2023; Rahwan, 2018; Sudmann, 2019).

However, not everyone agrees. The political philosopher Himmelreich (2022) contends that calls to democratize AI are misplaced. Such calls wrongly assume that democracy provides an answer to the concerns about fairness, freedom, and equality that AI technologies raise. Democratization, says Himmelreich, should be seen as a defensive response when systems trigger legitimization requirements, i.e. when a system leads to coercion, when its decisions have a pervasive impact, or when it involves schemes of social cooperation. States have these three triggers. In all such cases, citizens can be expected to pause and ask: why should I accept this? Why is this legitimate? But according to Himmelreich, AI technologies do not coerce, do not have a pervasive impact, and do not involve schemes of social cooperation. Thus, they do not trigger demands for democratization. At least, not as long as they operate on their own. Of course, they can become part of broader social or political systems—like insurance or taxes. Interestingly, at this point in his argument, he seems to mobilize the sociotechnical perspective. However, he then uses this perspective to defend a vision that in practice neatly *separates* technology and politics. He justifies this separation of technology and

politics by distinguishing coarse-grained decisions, such as regulation issues about standards of performance or norms of practice, from fine-grained, concrete decisions about the development and deployment of AI, for example about what data a model should be trained on. It is sufficient, he argues, to democratically control only the social and political institutions that take coarse-grained decisions. And for that, developed countries already have laws and institutions. Fine-grained decisions should be left to the experts, developers, and entrepreneurs. For those kinds of decisions, democracy is not the right instrument: democracy is costly, and these costs are not justified in the case of fine-grained decisions.

We partially agree and disagree with this seemingly pragmatic approach. The case of REC is illustrative here. The use of AI technologies for energy management, forecasting, or demand response in REC can facilitate the decentralization of control in energy networks and, as such, can contribute to changes in how decisions are made and by whom. They do so as part of broader sociotechnical systems, in which changes such as decentralization already occur regardless of the introduction of AI technologies. Thus, a machine learning technology that is part of a system that automates the coordination of the charging of EVs affords decentralized decision-making about when energy can be generated or consumed and by whom, but it is part of a broader regulatory push to do so. Moreover, AI technologies can only have a collectively binding effect and perform social coordination activities as part of that broader sociotechnical system. Without a physical infrastructure, renewable energy sources, residents of a neighborhood, distribution system operators (DSO), and many other components, the technology cannot operate appropriately. It is thus the broader sociotechnical system that triggers legitimation requirements. So far, we can agree with Himmelreich.

However, Himmelreich does not carry the sociotechnical perspective all the way through. Although AI technologies may not trigger legitimation requirements on their own, they actively co-shape the social systems that do trigger such requirements. Therefore, the technology component cannot be ignored. AI technologies, as Himmelreich also notes, can scale up coercion or shift relations of domination; they may be part of systems with pervasive impacts, such as electricity grids, and they can be embedded in systems of social coordination, such as REC. Their mediating role in these practices should therefore be a key element of any conceptualization of democratization. A case in point is the mediating role in shaping decision-making processes. Himmelreich suggests it is appropriate to ignore the role of technology in a sociotechnical network because it is the domain of fine-grained decisions best left to technical experts. Yet, the distinction between politically relevant course-grained decisions and technical fine-grained decisions is in practice not so clear cut. Decisions about a just distribution of resources, such as electricity, can shift and take on different forms in the translation from coarse-grained decisions to finer-grained decisions about development and deployment (Lorenz et al., 2021; Smith et al., 2010). Such translations happen, for example, in the gray area where the public turns into private (Sharon & Gellert, 2023; Taylor, 2021). Commercial technology firms have increasingly become involved in public governance, through their supply of digital infrastructure for the state's operations and through people's mass engagement with commercial platforms and services, including in the domain of energy (Niet et al.,

2022). Firms in the energy domain offer, for instance, platforms that enable energy trading among peers or allow distributed users to participate in energy markets, challenging existing governance structures (Kloppenborg & Boekelo, 2019). AI technologies are implicated in this broader trend and through their affordances co-shape it. For example, decisions about fairness in the distribution of electricity made by democratic institutions risk being reduced to concrete technical issues about price optimization addressed by private corporations, as these corporations provide the required infrastructures through cloud-based AI services (Niet et al., 2022; Sadowski & Levenda, 2020). In sum, a strict distinction between coarse-grained and fine-grained decisions to delimit the scope of democracy is problematic, because it leads us to ignore the role of technology in sociotechnical networks.

Democratizing sociotechnical systems is not complete without also democratizing the AI technologies that are part of these systems. We propose that *democratizing AI technologies* is, therefore, about developing, deploying, and using these technologies to help make political decisions in a democratic way. A call to democratize AI is specifically targeted at the technology and how it is situated within the broader sociotechnical system. This additional focus on AI technologies is required because, on the one hand, conflicting interests and values connect in the design of these technologies and they constitute a form of power. Especially when these technologies are developed and designed to change the lives of citizens, all those impacted by these technologies should have a say over the values and political dispositions built into these technologies. On the other hand, because of the entanglement of technology and politics, democratic processes should also, self-reflectively, be directed at diagnosing and guiding the ways technologies affect democratic governance. Through their affordances and constraints, technologies shape how democracy is done. Democratization of AI therefore also requires that we explore how AI affects how we do democracy. That is what the sociotechnical perspective is about.

This leaves open the question of what it means to make something more democratic. In the following section, we take a further step in developing the concept of democratizing AI. We take our cue from the ‘problem-based approach to democracy theory’ as developed by Mark Warren (2017). After that, we extend this conceptual framework with a sociotechnical perspective to show how such a framework can also be used as a heuristic to guide the ways technologies affect democratic governance.

4 Warren’s problem-based approach

Normative conceptions of democracy are usually framed as rivaling models, e.g. direct democracy, representative democracy, deliberative democracy, agonistic democracy, etc. Warren suggests that such models typically absolutize one particular feature of democracy at the cost of others (2017). Instead he, like several others (Mansbridge et al., 2012), defends a system perspective that understands democracy as a set of interconnected, interacting parts that realize a set of functions (Dean et al., 2019). Rather than asking after the essence of true democracy, he takes a pragmatic route by postulating that a political system is democratic if it manages three core tasks: “if a political system *empowers inclusion, forms collective*

agendas and wills, and *organizes collective decision capacity*, it will count as ‘democratic’” (p. 39). Democracies solve these three problems by means of a limited set of generic, institutionalized, political practices: *recognizing, resisting, deliberating, voting, representing, joining, and exiting* (see Table 1). All these practices have normative weaknesses and strengths relative to the three core functions of a democratic system. For example, ‘resistance’ enhances democracy if it increases the inclusion of marginalized voices; but it weakens democracy if a powerful elite ‘resists’ democratic decisions. The task of ‘designing’ a democratic political system is, therefore, to “combine these practices, usually into institutions, in ways that maximize their strengths and minimize their weaknesses” (p.39). There is not one way to do this optimally: there will always be trade-offs, and what works in one environment does not necessarily work in another. In this way, Warren outlines an approach to democratic theory and practice that is not essentialist, but flexible and context-sensitive.

There have been critiques of Warren’s approach, such as that it is unclear why he focuses on the three functions that he identifies instead of others, or that he unduly limits the range of democratic practices (Dean et al., 2019; Felicetti, 2021). Nevertheless, we think his approach is a promising way of thinking about democratizing AI. His approach stimulates us to ask questions about what makes political *systems* more democratic rather than simply assume that democracy means more deliberation, more participation, more inclusion, or more resistance.

So, what does Warren mean by empowered inclusion, collective agenda and will formation, and organizing collective decision capacity? *Empowered inclusion* is about individuals having the power to co-shape collective decisions that affect them (Warren, 2017, p. 44). It captures the democratic ideal that all those affected by decisions should have an equal opportunity to influence the decision-making. Democratic inclusion involves more than merely formal rights, like the right to free speech. Other institutional or practical guarantees should endow individuals with real power to exercise and, if necessary, enforce their inclusion. Empowered inclusion is related to *input legitimacy*: a decision has (some) democratic legitimacy if citizens feel they have had real opportunities to influence the decision-making process.

Collective agenda and will formation refers to the political processes through which individual interests, values, perspectives, and preferences are transformed into shared agendas and collective judgments (p. 44). In a democracy, the move from individual to collective self-government typically entails cooperative deliberations, negotiations, and compromises, but also antagonist debates. Collective will formation can be undertaken by the citizens themselves or by their representatives, and it comes with its own set of democratic norms, e.g. that individuals and groups express and listen to the ideas of others in an atmosphere of mutual respect and reciprocity, without coercion or oppression. Collective will formation aligns with *throughput legitimacy*: a decision is accepted because those affected feel that they had a fair chance to shape the decision.

Finally, *collective decision-making* is about “getting things done” (p. 44). It pertains to collective empowerment, where a collective has the capacity to make and impose binding decisions on itself. It requires qualities, like loyalty, as well as legally controlled coercion to ensure that decisions are indeed binding for all

Table 1 Warren's framework

Democratic systems						
Core functions						
Empowered inclusion		Collective agenda and will formation			Collective decision capacity	
Political practices						
Recognizing	Resisting	Deliberating	Representing	Voting	Joining	Exiting

involved. Collective decision-making corresponds to *output legitimacy*. The acceptance of decisions and the use of coercion increases when people see that a system or organization is able to translate a shared agenda into decisions that effectively solve problems that people care about.

The key point now is that Warren allows for various, competing, solutions to how a democratic system can execute these three functions. There is not one right 'model of democracy' as long as the three functions are realized as best as the situation allows. However, as an empirical observation, one can say that so far all more or less successful democratic systems involve combinations of the seven generic, institutionalized, practices we mentioned above. 'Practice' refers to routinized social actions, and encompasses formal and informal activities and mechanisms, norms, duties, obligations, roles, etc. (p. 43). These political practices can be organized, incentivized, or protected by *institutions*, defined as "rule-based, incentivized, and sociologically stable combinations of social actions" that assign roles, duties, obligations, and responsibilities to individuals (*ibid.*).

Now for the seven practices. First, any democracy presupposes a community of members who *recognize* each other's membership as equals: the *demos*. Citizens "establish mutual connections to shared circumstance, affected interests, fate, concern, common injuries, or common aspirations [and] put into place moral relationships" (p. 47). Mutual recognition as fellow citizens is at the core of empowered inclusion. But it is also key to collective will-formation, as it motivates a willingness to see the opponent as motivated by reasons too and to take everyone's perspective into account. Recognition helps to transform raw political conflict into shared deliberation, a search for win-win solutions, a willingness to negotiate fair compromises. Recognition is typically strengthened by formal citizen rights, but it is essentially a moral act, depending on people's voluntary motivations. Moral recognition cannot be coerced. Furthermore, recognition does not *automatically* enhance the normative values at the heart of empowered inclusion and collective will-formation. In the same move that some people get recognized as belonging to the in group, others can get excluded and marginalized.

Second, Warren points out that democracies are usually born out of *resisting* domination: strikes, revolts, demonstrations, picket lines, and civil disobedience. "Resisting is typically combined with moral demands for recognition, often focused by injury or injustice, and usually combined with some kind of power resource" (p.

47). It plays a role in empowered inclusion, but also during agenda and will formation. A mature democracy organizes its own spaces for resistance. Examples are Madison's 'checks and balances' or Montesquieu's *Trias Politica*. The core idea is that power should be distributed over institutions that to some extent resist one another. Another way a democracy organizes resistance is by giving citizens the right to appeal, organize, protest, go to court, etcetera. Resistance enables the democratic system to self-correct (see also: Mouffe, 2011). But resistance is not automatically positive. It can also undermine democracy, e.g. when powerful groups are able to resist democratic procedures or outcomes.

The third practice is *deliberation*, defined as "mediating conflict through the give and take of (cognitively compelling) reasons" about matters of common concern (p. 40). Deliberation is key in forming a collective will, as common and overlapping preferences emerge through the pooling of information—the wisdom of the crowd—and the reasoned revision of preferences. But deliberation too comes with its democratic weaknesses. First of all, in reality it is marred by cognitive and interactional biases, e.g. differences in status and power. Furthermore, it is time-consuming and there is no guarantee of consensus. Often deliberation causes differences of opinion to deepen and harden.

Fourth, the shortcomings of deliberation are why it gets typically combined with *voting*, defined as the aggregation of preferences. Deliberation is essentially qualitative; voting is essentially quantitative. Important strengths of voting are that it affirms equality, that it is relatively easy and fast to organize, that the process is quite transparent, and that the outcomes are usually clear. However, it typically remains opaque *why* people voted for or against a certain outcome. It does not facilitate negotiations and it can lead to winner-takes-all politics, or what John Stuart Mill called 'the tyranny of the majority', which is incompatible with recognizing each other as fellow members of a demos. As the strengths and weaknesses of deliberation and voting mirror each other, most democracies combine the two: first deliberate, then vote.

Fifth, even the most direct and small-scale democracies organize collective agenda and will formation through some form of *representation*: some members of the demos think, speak, vote, and act on behalf of others. Representation helps to overcome the limitations of time, space, and issue complexity, thus facilitating the inclusion of all—as the requirement of direct, personal participation favors those with the most time, those who are closest to the places where decisions are taken, and those who are the most vocal, knowledgeable, and/or skilled. Representative bodies are small enough to focus on an issue, deliberate, and bargain. Representation too comes with its unavoidable democratic weaknesses. The representee must trust the representative sufficiently to delegate parts of the will formation and decision making, and to accept the outcomes of both processes. If the mandate for the representatives is too strict, they lack the space for deliberation and collective decision-making. But if the mandate is too loose, the constituency may not recognize the outcome of the negotiations as legitimate.

Sixth is *joining*. By organizing themselves into groups individuals ensure that they are included in the democratic process. Through joining citizens form 'publics' around shared matters of concern. This practice is both relevant for

empowered inclusion and for the agenda and will formation. But joining can also undermine democracy when powerful associations exclude and overpower other citizen groups, dominate deliberation and voting, and frustrate the effective execution of majority decisions.

Finally, when citizens have little or no leverage to influence decision-making, *exiting*—‘voting with one’s feet’—is an option. Exiting incentivizes organizations to engage with their members. “[O]rganizations faced with loss of members, income, or votes have inducements to reach out to individuals proactively” (2017, p. 50). Like votes, exits are low information signals: it can remain unclear for an organization or party *why* citizens exit. More importantly, exiting is mainly a negative signal and does not contribute to collective will-formation or collective action.

As we have seen, all these practices have their democratic strengths and weaknesses. Furthermore, there can be tension between the practices, e.g. between deliberation and voting, or between joining and exiting. There also could be more than these seven practices (Dean et al., 2019). For example, according to us ‘doing’ or ‘making’ should be added to the list, as experimenting with local forms of Do-It-Yourself democracy—like hacker communities—is also part of a vital democratic culture. So, there is not one ‘right’ model for democracy. Warren’s approach allows for value trade-offs, tensions, and imperfections: democracy is always a work in progress. Citizens may and do disagree over democratic procedures, and there are always democratic values that are less expressed in a system than some citizens deem desirable.

Warren provides us with a checklist with a coherent set of questions that can be posed to any proposal to democratize AI: does the proposal take into account all three core functions of a democratic system (inclusion; will formation; decision capacity)? Does it provide us with an institutional design to combine the generic political practices in such a way as to make the most of their democratic strengths and diminish their democratic weaknesses? In the next section, we will briefly discuss a few proposals to demonstrate the value of Warren’s approach to democracy.

5 Assessing calls for democratizing AI

Earlier critics have already pointed out that calls for democratizing AI often insufficiently draw on democratic theory. Skaug Sætra et al (2022) for example, warn such calls often dilute the concept of democracy by reducing it to majority rule. This “might both undermine the prospects of using AI to foster democracy and the very idea that democracy is something worth defending” (p. 805). But they only point in the direction of what a richer concept of democratizing AI might entail, highlighting leadership, elections, organizations, deliberations, and pluralism. The following section aims to give a more systematic and comprehensive overview of the core ingredients of a vital democracy, as it is to be applied to AI technologies.

Calls for democratizing AI tend to focus on only some of the three core functions of democracy that Warren distinguished, on at most two or three of the practices he identified as helpful (or harmful) to solve those problems, and on only some of

the institutional safeguards that are needed to ensure that these practices enhance rather than weaken democracy. Typically, these calls for democratizing AI give no thought to how to make the most of the democratic strengths of the practices they propose, nor how to avoid their weaknesses. Most authors simply plead for more citizen inclusion in the form of more direct participation. For instance, McQuillan calls for people's councils to "collectively question and challenge decisions made by machines" (2018, p.7). These are "bottom-up confederate structures that act as direct democratic assemblies, based on the face-to-face democracy of the Athenian *ekklesia*" (p. 7). In this plea, we can detect the notions of inclusion and will-formation, realized by practices like deliberation and joining. However, other practices identified by Warren are absent. For example: How are people to be empowered so they can deliberate in the councils; who will populate these assemblies; how are the problems of direct democracy tackled without resorting to a practice like representing; what procedures are in place to deal with dissensus; and how do these councils make sure that their opinions indeed 'count' when faced with opposition?

A richer plea for inserting deliberative democracy in AI can be found in Buhmann and Fieseler (2023): "responsible AI governance needs to be enacted through a deliberative control process." (p. 10) They focus on how to overcome two problems: the opacity of AI and the gap between experts and laypersons. These problems make the direct participation of stakeholders problematic. The authors think these problems can be overcome by "distributed deliberation": in society there exist many venues where people with different types of (experiential) expertise deliberate: "What is most important about the judgments or outputs of deliberative venues is (...) whether that venue's particular discourse leads to a useful output that can be further 'processed' by other venues" (p. 15). A central role in bridging the gap between AI experts and lay audience is to be played by 'mini-publics' like citizen panels, AI think tanks, and interest groups. Rich as Buhmann and Fieseler's description of distributed deliberation may be, from the perspective of Warren's theoretical framework, it leaves many relevant topics untouched. First, it almost exclusively focuses on the stage of collective will formation, largely ignoring problems of empowered inclusion and building collective decision capacity. Who will be included in these mini-publics and who will not? If these mini-publics are informal representative bodies, nothing is said about their legitimacy. On the other end of the spectrum, there is little reflection on the fact that mini-publics typically have limited power to enforce decisions. Although the authors acknowledge the possibility of 'ethics washing', they do not discuss the need for political practices like resistance, joining, and exiting. With regard to deliberation, they only see its strengths, and remain blind to its weaknesses, i.e. that deliberation rarely results in consensus. Voting is absent from their proposal.

Koster et al. (2022) also plead for democratic control and decision-making, but rather than extolling deliberation they rely completely on voting. Their idea is that a machine learning algorithm can offer alternative solutions to problems of distributive justice, one of which is then selected "with an age-old technology for arbitrating among conflicting views—majoritarian democracy among human voters" (ibid.). They call this "human-in-the-loop" approach 'Democratic AI.' But in terms of Warren's framework, again relevant elements are missing. Not only does it remain

opaque who will be recognized as ‘voting citizen’, or how these voters will be empowered to vote, but the authors also fail to acknowledge the weaknesses of voting, i.e. that it can only register existing preferences, rather than help reflect on and revise these initial preferences through a reasoned exchange of arguments. Deliberation plays no role, nor any of the other democratic practices identified by Warren.

Another plea for democratizing AI is Iyad Rahwan’s plea for ‘Society in the loop’ (2018). This is a radicalization of the well-known demand for a human overview of AI technologies: “SITL is about embedding the values of society, as a whole, in the algorithmic governance of societal outcomes [... This] raises a fundamentally different problem: how to balance the competing interests of different stakeholders [...]? This is, traditionally, a problem of defining a social contract.”(p. 7). This contract is supposed to resolve tradeoffs between the different values and to distribute costs and benefits fairly among the members of the community. Rahwan gives a limited summing-up of ways citizens influence the government: “voting, opinion polls, civil society institutions, the media” (p. 9). He mostly elaborates on the practice of voting. The question then how AI technologies should be democratized, boils down to informing engineers in charge of designing AI of societal decisions, e.g. by value-sensitive design, or through algorithms that for instance explore “the aggregation of societal preferences and fair allocation of resources [...] that rational actors would be willing to vote for” (p. 11), or through professional ‘algorithm auditors’ (ibid.). Almost everything that makes democracies tick is absent from this vision, especially recognition (who is considered to be a stakeholder), possibilities for resistance, and deliberation, but also the key question: how to make profit-driven companies obey political demands? There are no reflections on ‘organizing collective action capabilities’.

Rather than deliberation and voting, Cuéllar and Huq (2020) highlight inclusion and the practice of resistance: “Rather than searching for exit routes for some, we would ask, how to educate and empower individuals and to invite mobilizations within and around AI systems” (p. 18). Examples of resistance and exiting are “walk-outs, secondary boycotts, and go-slows” (p. 19). Empowerment also requires ‘joining’: we need “platforms to coordinate responses, so as to influence and even change the policies and values embedded in those systems” (p. 19). The result of such forms of collective action is to enhance ‘recognition’: “Where local institutions are under pressure from engaged parts of the public, they are more likely to make inclusive and ethnically defensible choices about the scope and operation of AI systems” (p. 19). The authors also rightly point out that this democratic inclusion will not be automatic but will require institutionalization in the form of both legal rights for citizens and legal requirements for AI technologies. However, if one wants to know how and where democratic will formation is to take place, that is less clear. Nor do they have an eye for organizing collective action capacity.

Of course, there are more proposals for democratizing AI technologies, and some of them will be more comprehensive than the five we discussed here. The point is not to show that Warren is ‘better’ than everyone who has written about democracy and AI, but that he does provide a conceptual framework that can be fruitfully applied when thinking about that relationship. But there is one thing missing in his framework: a sociotechnical perspective. Warren has little to say about the role of

technology in politics. Yet, as we argued above, political practices and institutions depend on technologies and are co-shaped by them. Technologies may enhance and strengthen but also disrupt or destabilize political practices (Liu et al., 2020). We should study how technologies facilitate, enhance, transform, or threaten empowered inclusion, collective will-formation, and collective decision capacity. We need to extend Warren's framework with the question of how the affordances of technology can best mediate political practices to serve the three democratic core functions. This is what we do in the next section.

6 Democratizing AI in practice

To see how political practices can be improved to strengthen empowered inclusion, collective will formation, and collective decision capacity, we should not only look at institutions and social relationships but also at the technologies that support and shape them. What then can we understand as a political system supported by AI? To what extent, for example, can we consider AI-supported REC as a political system? Like other system approaches, Warren's framework is not tied to the nation-state as the unit of analysis. Rather, it allows for an analysis of democracy as a multi-level governance system (Dean et al., 2019). Warren points out that the boundaries of social systems are fluid. In the context of democratic theory, "we should view systems as comprised of those features of social relationships that are relevant to the ways individuals (or classes of individuals) are enabled, supported, empowered, constrained, dominated, marginalized (etc.) by the social relationships in which they are entangled, or upon which they depend" (Warren, 2017, p. 42). From this perspective, a REC is a social system in which collectively binding decisions are made. It is thus also a political system; a political system within a broader political system that includes national and local governments, energy networks, civil society organizations, and commercial companies. Warren argues that we should avoid "equating "democracy with specific institutions, such as constitutional states, no matter how essential. Instead, we should ask how practices and institutions function within political systems, so that we can identify and assess democratic possibilities within new or novel contexts" (p. 45). Thus, although constitutional states may be necessary for democracy, Warren's framework encourages us to look at practices within political (sub-)systems, such as a REC, to assess how these practices can best serve the three functions for the political system to be democratic. This means that for AI-supported RECs, we also have to consider what role these technologies can play to achieve this goal.

Assessing AI-supported practices in political systems requires that we look at the lifecycle of these technologies. That is, we should look at all the different stages that these technologies go through to see how these technologies can be shaped to contribute to more democratic practices, from the early design stages to the operating stage as well as to changes to or the decommissioning of the technology. We can, for example, push for the democratization of the development of AI technologies by emphasizing the need for deliberation and resistance

in participatory design methods to determine what the technologies in development should do and mean, as some of the described approaches have done. However, we should also pay attention to how these technologies can be (re)designed to improve democratic functions once in operation. Deliberation between and resistance from stakeholders as part of participatory design processes may facilitate collective will formation around issues such as fairness, but the collective will might change during the life cycle of an AI system. For example, let's suppose that developers of a charging algorithm for EVs invite members of a REC to deliberate on what a fair principle for prioritization of charging is. They collectively decide that charging should be done on a first-come-first-served basis. However, over time the prevailing ideas about what is a fair charging principle change. For instance, a majority of REC members now prefer priority be given to people with certain critical jobs, like doctors. What kinds of AI technologies should the energy management systems employ to allow for such changes in conceptions of values, such as fairness? How can such a technology support decision-making processes about how and when to change the prioritization principle? To address such questions, we have to think about how an AI-based system can serve to enhance empowered inclusion, collective will formation, and collective decision capacity throughout the different phases of its life cycle.

We can combine Warren's framework with a sociotechnical perspective to think through what the democratization of AI technologies would entail in the context of a REC. An exhaustive analysis would reflect on how particular AI technologies mediate the democratic strengths and weaknesses of the seven practices within the political system centered on the REC. It would also contain a description of how the REC relates to the larger energy system, the national state, and other relevant systems. However, as our more modest aim in this paper is to illustrate the potential of Warren's framework, we will only briefly highlight how his framework can work as a heuristic in the design and deployment of democratic AI technologies.

Let's first look at how AI technologies can hamper or facilitate empowered inclusion. On the one hand, they may empower prosumers to locally share energy in smart grids and to have more say in collective decisions about energy distribution by ensuring that adding prosumer-generated energy does not overextend the capacity of the grid (Korkas et al., 2018; Sousa et al., 2019). However, empowered inclusion in a REC is also often restricted by technological barriers. It is crucial for inclusion to be recognized as a member of the demos. But to be allowed to participate in collective decision-making, members should not only be formally recognized, for instance by the law, but also by the technologies in question. In practice, this can present a challenge for some individuals and groups. Residents, for example, are only 'visible' for data-driven AI technologies if they have the right hard- and software. These are needed to allow data about, for example, energy consumption to be collected for processing and model training purposes. This poses a barrier for those who cannot afford or do not have access to specialized equipment. Moreover, as discussions on bias and fairness in AI developments have shown, AI technologies, such as machine learning, have a hard time being sufficiently responsive to different and changing cultural, ethnic, and minority groups (Taylor, 2017). This constitutes another technological barrier for these groups to be recognized and included in the REC.

Similarly, AI technologies can enhance or weaken collective agenda and will formation in RECs. They can enable prosumers to ‘join’ a community and facilitate communications and energy sharing within that community (Kloppenburger & Boekelo, 2019; Koirala et al., 2021). However, automating part of the interactions between community members, such as peer-to-peer trading of energy, may also stand in the way of deliberation, which is key to collective will formation. For example, peer-to-peer trading systems generally do not allow explicit negotiations between REC members (Deconinck, 2021). Deliberation takes time and is costly, which is problematic for the real-time management of a smart grid. Therefore, system designers prefer to present choices about how to optimally distribute renewable energy in local grids in terms of selecting the most efficient or cost-effective coordinating mechanism. This approach suggests that there is nothing to deliberate on by prosumers or their representatives. But that is not the case. Members of the community might disagree, for instance, about what to optimize for (e.g. reducing costs or CO₂ omissions) or what fair market rules are. And as the community changes over time this disagreement might take on different forms. This again raises the question of how AI technologies can support, anticipate, and be responsive to continuous deliberations among the members of a REC and other relevant stakeholders to ensure the democratic quality of the collective will formation.

Finally, AI technologies provide effective means of enforcing decisions. This can enhance a community’s collective decision capacity, as regulation can be embedded in the technology (van den Berg & Leenes, 2013). Once a REC, for example, has reached a decision on how to share energy among its members, an AI system can implement the chosen coordination mechanism. However, the downside is that the complexity and opaqueness of AI technologies easily obscure the political relevance of some design choices and can hinder accountability. For example, in some suggested solutions for peer-to-peer electricity trading control over energy distribution is delegated to third parties such as community platforms that offer various kinds of AI-based services. Kloppenburger and Boekelo (2019) note that such community platforms use household energy data to build AI-based algorithms to predict market prices and steer energy flows from and to batteries. However, they point out that the workings of the underlying algorithms are often opaque to users of platforms which makes it hard for these users and other relevant stakeholders, such as a local government, to understand what happens to the energy these platforms buy and sell. This can make practices like resistance harder rather than easier. How can one resist or protest the categorizations, classification, or optimizations that these algorithms enact, especially when these categorizations might not be common sense or conventional? AI technologies can, thus, both strengthen and undermine the collective decision capacity of a REC.

Researchers have recognized some of the problematic aspects of AI described above and have explored alternative approaches to address them. Such approaches could contribute to the development of practices that benefit rather than hamper the three core functions. One example is the growing field of research on explainable AI, which aims to make these systems more transparent, interpretable, and accountable (Barredo Arrieta et al., 2020; Buhmann & Fieseler, 2023; Selbst & Barocas, 2018). Explainable AI techniques are intended to provide some insight into the logic

underlying automated decision-making. Machine learning technologies used for scheduling the charging of electrical vehicles, for example, can be complex and give little insight into how decisions are made. Explainable AI techniques could be used to help various stakeholders understand how the system operates and, thus, support different practices such as recognition, resistance, and deliberation. Other scholars have developed the idea of *contestability by design* “to protect against fallible, unaccountable, illegitimate, and unjust automated decision-making, by ensuring the possibility of human intervention as part of a procedural relationship between decision subjects and human controllers” (Alfrink et al., 2022). This goes beyond making technologies explainable; it also involves building safeguards, allowing for human review and intervention requests, as well as fostering practices such as agonistic approaches to machine learning development, quality assurance during development and after deployment, and third-party oversight. Such initiatives could be directed towards serving the strengths of resistance practices in REC to enhance collective agenda and will formation and collective decision-making in RECs. These are just two examples of initiatives that have explored alternative technological approaches to developing AI. Warren’s framework can be used to inform other such initiatives.

Again, these are only snippets of a more comprehensive and systematic analysis, based on Warren’s framework, of assessing the democratic opportunities and risks of AI technologies in the context of larger sociotechnical systems. Table 2 provides a starting point for such a comprehensive analysis, with the risks and opportunities that we mentioned above filled in. The empty boxes indicate the work to do. As mentioned, a comprehensive approach should also broaden the focus to the larger sociotechnical system.

As these few examples show, Warren’s framework combined with a sociotechnical perspective provides a heuristic that can help to systematically think through how AI technologies can serve or hamper political practices in support of democratic governance. These technologies can affect changes for the better or the worse. We, therefore, need to understand how they mediate political practices and where interventions can be made to strengthen empowered inclusion, collective will formation, and collective decision capacity. This may involve pushing for more accountability, transparency, accessibility, or citizen participation, but such demands should be seen in light of the different political practices.

7 Conclusion

Recent calls to democratize AI present a patchwork of approaches, where each approach highlights one or a few features of democracy. Such calls tend to respond to concerns about the risks accompanying the growing pervasiveness of AI technologies in the daily lives of citizens, such as potential injustices resulting from their use or the growing reliance on opaque, unaccountable global superpowers. Democratization is proposed as a way of addressing such concerns, but a coherent understanding of what democratizing AI entails is still lacking. On the one hand, democratization is often

narrowly understood in terms of one or two political practices, for instance, deliberation and citizen engagement in design and development practices, or voting and organizing resistance. Valuable as these contributions may be in exploring alternative trajectories for the widespread adoption of AI in society, they provide only partial solutions to the problem of making AI more democratic. On the other hand, few have looked at the mutual shaping between AI technologies and political practices around issues of collective concern. Current approaches do not seem to be sufficiently aware that the political consequences of ‘finegrained’ decisions, to borrow Himmelreich’s concept, reverberate through the whole complex sociotechnical system AI technologies are increasingly part of. This means that the seemingly technical or business decisions made by experts have significant consequences for what appears on the political agendas, who is allowed to participate in the making of collectively binding decisions, how that decision-making is organized, and finally, how decisions are executed and put into practice.

As we have argued in this paper, a richer conception of democratizing AI is required, one that can account for the different elements of democracy as well as for the role of technology in the practices that constitute democratic systems. We propose that the democratization of AI technologies should be about finding ways of developing, deploying, and using these technologies in such a way that they are conducive to more democratic ways of making decisions with collectively binding effects. This requires looking at both how these technologies can be democratically controlled and how they can foster democratic practices.

A system approach to democracy, such as the one developed by Warren, can provide the basis for a richer understanding of democracy to further elaborate the idea of democratization. Rather than arguing for a single model of democracy, it highlights three problems that a democratic political system needs to solve: empowered inclusion, collective agenda and will formation, and organizing collective decision capacity. These three functions of democracy, Warren argues, are fulfilled by different political practices, such as deliberation, representation as well as voting and exit. Although Warren’s framework is just one example of a system approach to democracy and has its own limitations, identifying such practices and the functions they serve, as we have shown, can help to assess the different proposals for democratization of AI but can also serve as heuristics in guiding the development of AI technologies to be conducive to democratic governance.

One limitation of Warren’s framework is that it does not take the role of technology in political practices into account; it therefore needs to be complemented with a sociotechnical perspective. An analysis of political systems cannot do without acknowledging the role of technology, nor can an understanding of technology and its political dimensions do without a view of the broader political system. AI technologies are a part of political systems, and they can contribute to improving the democratic quality of these systems or inhibit it. To analyze, assess, and intervene in these processes of mutual shaping a view of the sociotechnical systems is required, including the multiple actors, technologies, and the relationships between them as well as the context in which these systems are situated.

Table 2 Overview of risks and opportunities of AI technology in REC for political practices in relation to core functions

	Empowered inclusion		Collective agenda and will formation		Collective decision capacity	
	Risk	Opportunity	Risk	Opportunity	Risk	Opportunity
Recognizing	Specialized soft- and hardware can exclude groups and individuals	AI can enable access to energy markets for prosumers				
Resisting					Opaque and complex algorithm make it hard to protest	Explainable AI, Contestable by design
Deliberating			AI can lead to complex and opaque systems and technocratic solutions	AI can facilitate peer-to-peer communication within REC		
Representing	Misrepresentation and underrepresentation in data				Opaque proprietary algorithms stand in the way of accountability	AI can enforce collective decisions
Voting						
Joining				AI can facilitate community-based peer-to-peer energy sharing		
Exiting						

Funding The research leading to these results received funding from Dutch Research Council (NWO) under Grant Agreement Nos P19-25 and 313-99-308.

Declarations

Competing interest The authors have no relevant financial or non-financial interests to disclose. The funders took part in the decision to publish the results.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alfrink, K., Keller, I., Kortuem, G., & Doorn, N. (2022). Contestable AI by design: Towards a framework. *Minds and Machines*. <https://doi.org/10.1007/s11023-022-09611-z>
- Al-Ogaili, A. S., Hashim, T. J. T., Rahmat, N. A., Ramasamy, A. K., Marsadek, M. B., Faisal, M., & Hannan, M. A. (2019). Review on scheduling, clustering, and forecasting strategies for controlling electric vehicle charging: Challenges and recommendations. *IEEE Access*, 7, 128353–128371.
- Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- Barrientos, F., Mouzakitis, S., Lombardo, L., Hernandez-Moral, G., Corral-Vicente, A. E., & Rodriguez-Alonso, C. (2023). *AI-based Energy Services for Reinforcing the AI-on-Demand Platform*.
- Boyd, E. (2017). Microsoft and Facebook create open ecosystem for AI model interoperability. *Microsoft Com*.
- Buhmann, A., & Fieseler, C. (2023). Deep learning meets deep democracy: Deliberative governance and responsible innovation in artificial intelligence. *Business Ethics Quarterly*, 33(1), 146–179. <https://doi.org/10.1017/beq.2021.42>
- European Commission. (2018). *Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Artificial Intelligence for Europe* (COM/2018/237). <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0237&from=EN>
- Cuéllar, M.-F., & Huq, A. Z. (2020). Toward the democratic regulation of AI systems: A prolegomenon. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3671011>
- Cuppen, E. (2018). The value of social conflicts. Critiquing invited participation in energy projects. *Energy Research & Social Science*, 38, 28–32. <https://doi.org/10.1016/j.erss.2018.01.016>
- Dean, R., Rinne, J., & Geissel, B. (2019). Systematizing democratic systems approaches: Seven conceptual building blocks. *Democratic Theory*, 6(2), 41–57.
- Deconinck, G. (2021). Decentralised control and peer-to-peer cooperation in smart energy systems. In M. P. C. Weijnen, Z. Lukszo, & S. Farahani (Eds.), *Shaping an inclusive energy transition* (pp. 121–138). Springer. https://doi.org/10.1007/978-3-030-74586-8_6
- Di Lorenzo, G., Stracqualursi, E., Micheli, L., Martirano, L., & Araneo, R. (2022). Challenges in energy communities: State of the art and future perspectives. *Energies*, 15(19), 7384. <https://doi.org/10.3390/en15197384>
- Duberry, J. (2022). Artificial intelligence and democracy: Risks and promises of AI-mediated citizen-government relations. *Edward Elgar Publishing*. <https://doi.org/10.4337/9781788977319>

- Easton, D. (1965). *A framework for political analysis*. Prentice-Hall.
- Feenberg, A. (2017). Critical theory of technology and STS. *Thesis Eleven*, 138(1), 3–12. <https://doi.org/10.1177/0725513616689388>
- Felicetti, A. (2021). Learning from democratic practices: New perspectives in institutional design. *The Journal of Politics*, 83(4), 1589–1601.
- Hanke, F., & Lowitzsch, J. (2020). Empowering vulnerable consumers to join renewable energy communities—Towards an inclusive design of the clean energy package. *Energies*, 13(7), 1615. <https://doi.org/10.3390/en13071615>
- Helberger, N. (2019). On the democratic role of news recommenders. *Digital Journalism*, 7(8), 993–1012. <https://doi.org/10.1080/21670811.2019.1623700>
- Hernandez-Matheus, A., Löschenbrand, M., Berg, K., Fuchs, I., Aragués-Peñalba, M., Bullich-Massagué, E., & Sumper, A. (2022). A systematic review of machine learning techniques related to local energy communities. *Renewable and Sustainable Energy Reviews*, 170, 112651. <https://doi.org/10.1016/j.rser.2022.112651>
- Himmelreich, J. (2022). Against “Democratizing AI.” *AI & SOCIETY*. <https://doi.org/10.1007/s00146-021-01357-z>
- HLEG. (2018). *A definition of AI: Main capabilities and scientific disciplines*. European Commission. https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_december_1.pdf
- Judson, E., Fitch-Roy, O., & Soutar, I. (2022). Energy democracy: A digital future? *Energy Research & Social Science*, 91, 102732. <https://doi.org/10.1016/j.erss.2022.102732>
- Kloppenburg, S., & Boekelo, M. (2019). Digital platforms and the future of energy provisioning: Promises and perils for the next phase of the energy transition. *Energy Research & Social Science*, 49, 68–73.
- Koirala, B., de Koning, N., Kort, J., Iannone, A., Bisconti, P., Claessens, B., Bellesini, F., Mancinelli, E., Tribbolati, G., & Boijn, E. (2021). Deliverable D3. 1 Overview of barriers and drivers for consumer engagement in demand response. In *Bright-boosting DR through increased community-level consumer engagement by combining data-driven and blockchain technology tools with social science approaches and multi-value service design*.
- Korkas, C. D., Baldi, S., & Kosmatopoulos, E. B. (2018). Grid-connected microgrids: Demand management via distributed control and human-in-the-loop optimization. In *Advances in renewable energies and power technologies* (pp. 315–344). Elsevier.
- Koster, R., Balaguer, J., Tacchetti, A., Weinstein, A., Zhu, T., Hauser, O., Williams, D., Campbell-Gillingham, L., Thacker, P., Botvinick, M., & Summerfield, C. (2022). Human-centred mechanism design with Democratic AI. *Nature Human Behaviour*, 6(10), 1398–1407. <https://doi.org/10.1038/s41562-022-01383-x>
- Lessig, L. (2009). *Code: And other laws of cyberspace*. ReadHowYouWant. com.
- Liu, H.-Y., Maas, M., Danaher, J., Scarcella, L., Lexer, M., & Van Rompaey, L. (2020). Artificial intelligence and legal disruption: A new model for analysis. *Law, Innovation and Technology*, 12(2), 205–258. <https://doi.org/10.1080/17579961.2020.1815402>
- Lorenz, L., Meijer, A., & Schuppan, T. (2021). The algocracy as a new ideal type for government organizations: Predictive policing in Berlin as an empirical case. *Information Polity*, 26(1), 71–86. <https://doi.org/10.3233/IP-200279>
- Maas, J., & Durán, J. M. (2022). *Legitimacy of what?: A call for democratic AI design*.
- Malik, M. M. (2020). *A Hierarchy of Limitations in Machine Learning* (arXiv:2002.05193). arXiv. <http://arxiv.org/abs/2002.05193>
- Mansbridge, J., Bohman, J., Chambers, S., Christiano, T., Fung, A., Parkinson, J., Thompson, D. F., & Warren, M. E. (2012). A systemic approach to deliberative democracy. *Deliberative Systems: Deliberative Democracy at the Large Scale*, 1–26.
- Marinakos, V., Koutsellis, T., Nikas, A., & Doukas, H. (2021). AI and data democratisation for intelligent energy management. *Energies*, 14(14), 4341. <https://doi.org/10.3390/en14144341>
- Marres, N. (2016). *Material participation: Technology, the environment and everyday publics*. Springer.
- McQuillan, D. (2018). People’s councils for ethical machine learning. *Social Media + Society*, 4(2), 205630511876830. <https://doi.org/10.1177/2056305118768303>
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 205395171667967. <https://doi.org/10.1177/2053951716679679>

- Montes, G. A., & Goertzel, B. (2019). Distributed, decentralized, and democratized artificial intelligence. *Technological Forecasting and Social Change*, *141*, 354–358. <https://doi.org/10.1016/j.techfore.2018.11.010>
- Mouffe, C. (2011). *On the political*. Routledge.
- Nemitz, P. (2018). Constitutional democracy and technology in the age of artificial intelligence. *Philosophical Transactions of the Royal Society a: Mathematical, Physical and Engineering Sciences*, *376*(2133), 20180089.
- Niet, I. A., Dekker, R., & van Est, R. (2022). Seeking public values of digital energy platforms. *Science, Technology, & Human Values*, *47*(3), 380–403. <https://doi.org/10.1177/01622439211054430>
- Noorman, M., Espinosa Apráez, B., & Lavrijssen, S. (2023). AI and energy justice. *Energies*, *16*(5), 2110.
- Norman, D. A. (2013). *The design of everyday things*. MIT Press.
- Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.
- Rahwan, I. (2018). Society-in-the-loop: Programming the algorithmic social contract. *Ethics and Information Technology*, *20*(1), 5–14. <https://doi.org/10.1007/s10676-017-9430-8>
- Sadowski, J., & Levenda, A. M. (2020). The anti-politics of smart energy regimes. *Political Geography*, *81*, 102202. <https://doi.org/10.1016/j.polgeo.2020.102202>
- Sætra, H. S., Borgebund, H., & Coeckelbergh, M. (2022). Avoid diluting democracy by algorithms. *Nature Machine Intelligence*, *4*(10), 804–806.
- Schatzki, T. R., Knorr-Cetina, K., Von Savigny, E., et al. (2001). *The practice turn in contemporary theory* (Vol. 44). London: Routledge.
- Selbst, A. D., & Barocas, S. (2018). The intuitive appeal of explainable machines. *Fordham Law Review*, *87*, 1085.
- Sharon, T., & Gellert, R. (2023). Regulating Big Tech expansionism? Sphere transgressions and the limits of Europe's digital regulatory strategy. *Information, Communication & Society*, 1–18.
- Smith, M. L., Noorman, M. E., & Martin, A. K. (2010). Automating the public sector and organizing accountabilities. *Communications of the Association for Information Systems*. <https://doi.org/10.17705/ICAIS.02601>
- Sousa, T., Soares, T., Pinson, P., Moret, F., Baroche, T., & Sorin, E. (2019). Peer-to-peer and community-based markets: A comprehensive review. *Renewable and Sustainable Energy Reviews*, *104*, 367–378.
- Sovacool, B. K., Hess, D. J., Amir, S., Geels, F. W., Hirsh, R., Medina, L. R., Miller, C., Palavicino, C. A., Phadke, R., Rygshaug, M., et al. (2020). Sociotechnical agendas: Reviewing future directions for energy and climate research. *Energy Research & Social Science*, *70*, 101617.
- Sudmann, A. (Ed.). (2019). *The democratization of artificial intelligence: Net politics in the era of learning algorithms* (1st ed., Vol. 1). transcript Verlag. <https://doi.org/10.14361/9783839447192>
- Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society*, *4*(2), 205395171773633. <https://doi.org/10.1177/2053951717736335>
- Taylor, L. (2021). Public actors without public values: Legitimacy, domination and the regulation of the technology sector. *Philosophy & Technology*, *34*(4), 897–922. <https://doi.org/10.1007/s13347-020-00441-4>
- Van Aubel, P., & Poll, E. (2019). Smart metering in the Netherlands: What, how, and why. *International Journal of Electrical Power & Energy Systems*, *109*, 719–725. <https://doi.org/10.1016/j.ijepes.2019.01.001>
- Van den Berg, B., & Leenes, R. E. (2013). Abort, retry, fail: Scoping techno-regulation and other techno-effects. *Human law and computer law: Comparative perspectives* (pp. 67–87). Springer.
- Vázquez-Canteli, J. R., & Nagy, Z. (2019). Reinforcement learning for demand response: A review of algorithms and modeling techniques. *Applied Energy*, *235*, 1072–1089. <https://doi.org/10.1016/j.apenergy.2018.11.002>
- Warren, M. E. (2017). A problem-based approach to democratic theory. *American Political Science Review*, *111*(1), 39–53. <https://doi.org/10.1017/S0003055416000605>
- Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kazianus, E., Mathur, V., West, S. M., Richardson, R., Schultz, J., & Schwartz, O. (2018). *AI now report 2018*. AI Now Institute at New York University New York.
- Winner, L. (1980). Do artifacts have politics? *Daedalus*, *109*(1), 121–136.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.