CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

From Poisons to Antidotes: Algorithms as Democracy Boosters

Paolo Cavaliere[1] and Graziella Romeo[2]*

[1]Lecturer in Digital Media and IT Law at University of Edinburgh Law School, Edinburgh, UK and
[2]Associate Professor of Comparative Constitutional Law at Bocconi University, Milan, Italy
*Corresponding author. Email: graziella.romeo@unibocconi.it

Abstract

Under what conditions can artificial intelligence contribute to political processes without undermining their legitimacy? Thanks to the ever-growing availability of data and the increasing power of decision-making algorithms, the future of political institutions is unlikely to be anything similar to what we have known throughout the last century, possibly with parliaments deprived of their traditional authority and public decision-making processes largely unaccountable. This paper discusses and challenges these concerns by suggesting a theoretical framework under which algorithmic decision-making is compatible with democracy and, most relevantly, can offer a viable solution to counter the rise of populist rhetoric in the governance arena. Such a framework is based on three pillars: (1) understanding the civic issues that are subjected to automated decision-making; (2) controlling the issues that are assigned to AI; and (3) evaluating and challenging the outputs of algorithmic decision-making.

Keywords: artificial intelligence; decision-making; democracy; political rights

I. Introduction

Disco Sour is a dystopian novel published in 2017 in which the author imagines a post-nation-state world where elections are replaced by a Tinder-like app called *Plebiscitum*.[1] The book proved successful enough to inspire a cross-party conference at the European Parliament on "Democracy in the Age of Algorithms"[2] to discuss the ethical issues and threats that artificial intelligence (AI) can pose to democracy. The novel (and, more relevantly, the high-level conference that followed) is but one of many voices that have raised alarms in the last few years over the looming end of democracy at the hands of digital technologies and automation.

Thanks to the ever-growing availability of data and the increasing power of decision-making algorithms, the future of political institutions is unlikely to be anything similar to what we have known throughout the last century, possibly with parliaments handing over their traditional authority to largely unaccountable technologies. Critics have described

---

[1] G Porcaro, *Disco Sour* (London, Unbound 2017).

[2] "Democracy in the Age of Algorithms", European Parliament, Brussels, 7 November 2017 <https://www.internetforum.eu/events/657-scaling-the-sharing-economy>.

this as a "toxic cocktail for democracy".[3] This paper is framed as a (provocative but constructive) response to the concerns prevalent among academic circles and public opinion. We respond to these fears with a classic trope of dystopian literature: *what if* all is not lost?

In the same vein, the paper aims to make the case for harnessing the potential of digital technologies to increase the quality of democracy in times of rampant populism. In fact, we suggest that the current debate over the use of AI in the public sphere needs to be reframed against the backdrop of rampant populism, as opposed to an idealised concept of democracy.

The main focus of our discussion is on technology-enabled policymaking mechanisms and their potential to positively affect democratic representation and legitimisation. The theoretical background to our analysis involves democratic theories that reject the simplistic identification of democracy with the manifestation of popular will. This strand of literature characterises democracy as being linked to a certain level of protection of participatory rights that is functional for the realisation of the democratic principle. Such an approach directly challenges the rhetoric of populism that emphasises disproportionately the representative element at the expense of a more articulated notion of democracy as a regime that is supposed to protect individual and collective rights.

The paper suggests that algorithmic decision-making can contribute to an output-oriented democratic process in which fundamental rights can regain centrality – one that can be challenged on substantive grounds by voters who are self-aware and have experience with technology. The key shift that an AI government may have in store for us is the elimination of the need to cater to voters' often irrational and detrimental concerns from the process of policymaking. This shift is likely to be a positive one.

While not discounting existing concerns, the paper discusses the merits of algorithmic decision-making in the public sphere from the perspective of democratic theories, as elaborated by Robert Dahl, Charles Beitz and Fritz Scharpf, and it offers a change of perspective from the currently dominant narratives. Instead of accepting the narrow view of technology as a binary alternative to traditional representative democracy, where participation is mainly channelled through periodic voting, we look at the current dynamic through the prism of the protection of fundamental rights and output legitimacy.

With a view to supporting our claim, the paper proceeds as follows: it will firstly engage with the concept of populism with a view to exposing the incomplete (and partially distorted) concept of democracy it purports to represent (Section II); it goes on to propose an alternative reading of the relationship between algorithms and democracy, which relies on the notion of output legitimacy and its fundamental rights implications (Section III). The article then builds an agenda for an algorithmic decision-making framework based on a methodological framework for the analysis of the relationship between algorithms and democracy (Section IV). The framework is centred on three main pillars: (1) understanding civic issues that are subjected to automated decision-making; (2) controlling issues that are assigned to AI; and (3) evaluating and challenging the outputs of algorithmic decision-making (Sections IV.1, IV.2 and IV.3). Finally, the paper offers some concluding remarks on the way in which algorithmic decision-making can be framed to be (1) compatible with democracy and (2) an effective alternative to populist rhetoric.

II. The narrative of "the people"

The rise of populism has made the headlines and has been a topic of discussion in academic and public contexts in recent years. Yet, despite its renewed popularity, the concept itself

---

[3] C O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York, Crown Books 2016).

is far from new. Among early commentators from the mid-twentieth century, the notion of populism emerged as markedly "characterized by a peculiar negativism [and] great doses of blind hatred"[4] towards multiple facets of society. Isaiah Berlin was among the first to identify a distinct element common to all types of populism, namely the centrality of the notion of the people, invariably characterised as the have-nots, in stark opposition to an enemy – a dominant group said to have caused them damage, usually identified as the capitalists, the bureaucracy or some other group identified on the basis of an ethnic, racial or national otherness.[5] Along the same lines, Hugh Seton-Watson described populism as the "idolization and worship of the people … contrasted with the vices of the elite".[6] Other observers have pointed out the emphasis on the popular will as the ultimate source of legitimisation for political authority. Lloyd Fallers, most influentially, suggested that both populistic and nationalistic ideologies share this same trait.[7] Even with the lack of a generally agreed upon definition, the relevant literature has been in agreement in seeing populism as a pathological feature of democracy – a "syndrome" rather than a specific doctrine.[8]

In keeping with this tradition, recent works have almost universally framed populism in negative terms. Yet identifying precisely its practical impact has proven less than straightforward. Paul Taggart discusses current forms of "new" populism as the "rejection of the political agenda, institutions, and legitimacy of the modern welfare state model of mixed economy capitalism".[9] David Landau has expressed a largely similar concern while also providing an explanation of the progressive unfolding of the detrimental effects of populistic institutions across three consecutive phases, such as "undermining the existing institutional order, constructing a new order built on a critical vision of the old one, and consolidating power in the hands of populist leaders".[10] The dynamic between constitutional democracy and populism has been described as a form of "parasitism" where the latter exploits certain features and fragilities of the former to tip the balance between the rule of majority and the rule of law on which constitutional democracy is traditionally founded.[11] From this perspective, the practical threat that populism poses to democracy is understood as a toxic rhetoric that, over time, manipulates both of these pillars by proposing a procedural vision of democracy that disproportionally prioritises the rule of the majority, in pursuit of aims incompatible with the spirit of democracy and its values, such as political pluralism, transnational solidarity and the protection of minority rights.[12]

Crucially, the prioritisation of the rule of majority is made possible by the vision, intrinsic to populistic rhetoric, of the majority as the sole legitimate ruling entity.[13] Oran Doyle elaborates on this concept by offering a convincing account of populism as a form of constituent power that, while delegitimising non-majoritarian institutions (such as those that typically provide checks and balances), twists the constitutional system towards serving the needs of a specific part of the population.[14]

---

[4] I Berlin et al, "To Define Populism" (1968) 3 Government and Opposition 137, 169.

[5] ibid, 175.

[6] ibid, 156.

[7] L Fallers, "Populism and Nationalism" (1964) 6 Comparative Studies in Society and History 445, 447.

[8] P Wiles, "A Syndrome, Not a Doctrine: Some Elementary Theses on Populism" in G Ionescu and E Gellner (eds), *Populism: Its Meanings and National Characteristics* (London, Weidenfeld and Nicolson 1970) p 166.

[9] P Taggart, *Populism* (New Delhi, Viva Books, 2002) p 75.

[10] D Landau, "Populist Constitutions" (2018) 85 University of Chicago Law Review 521, 523.

[11] T Fournier, "From Rhetoric to Action, a Constitutional Analysis of Populism" (2019) 20 German Law Journal 362, 364.

[12] ibid, 381.

[13] ibid, 380.

[14] O Doyle, "Populist Constitutionalism and Constituent Power" (2019) 20 German Law Journal 161.

The most immediate effect of populist rhetoric is thus the de-legitimisation of certain traditional aspects of democratic procedures, most typically all of those that provide some forms of checks and balances; and this, in turn, unleashes a further consequence on the quality of democratic governance and policymaking. Populism's first victims are usually the procedural mechanisms typical of parliamentary politics, such as multiparty competition and democratic checks and balances that, in the view of populist rhetoric, "illegitimately restrict the will of the people by checking the power of the majority and empowering minorities" in a way that time and again throughout history has led to "the crisis of parliamentary institutions".[15] The proceduralist conception of political legitimacy traditionally places a strong emphasis on parliamentary deliberation and constitutional-level decision-making rules. But rising populist movements have attacked the role of institutions underpinned by those procedures – traditional parties in the first place, as well as all other institutions normally overseeing governments' operations – and, lastly, are undermining procedural legitimacy.

However, by disrupting the legitimate standing of those bodies whose specific function is to counter the view of hegemonic majorities, to contribute the perspectives of diverse groups and to provide mechanisms for the oversight of governments' decisions, populist stances cause practical and direct consequences at the governance level. Empirical studies have confirmed that populist political forces, when they are in government, typically cause a decline in the enjoyment of civil liberties and political rights,[16] lead to significant surges in corruption[17] and tend to perform proportionately less effectively than non-populist governments in domains such as basic welfare, gender equality and local democracy.[18]

The complexity of this dynamic can be tentatively summarised as follows: while populism primarily and most directly affects and undermines the procedural legitimacy of democratic decision-making, its most practical and perceivable impact is, instead, on the quality of policy decisions. Both the inputs (procedures) and outputs (actual policy decisions) of democratic governance are negatively affected by populism; however, solutions proposed so far have struggled to capture this ambivalence, as they largely focus on correcting the quality of outputs alone, disregarding the centrality of procedures and their perceived legitimacy in leading to such outcomes. In fact, a common reaction all across Europe to the surge of populist movements has been a renewed appeal for technocratic governments and their promise of better governance[19]; meanwhile, resorting to algorithmic decision-making is often decried on the grounds of its innate inability to achieve fair and non-discriminatory results. In both cases, the acceptance of technocratic governance and the rejection of AI-assisted solutions are based on a similar inability to reconcile procedural and substantive elements of democracy. We suggest instead that, under certain circumstances, AI can be deployed as a useful tool to increase the quality of democratic decisions – not in and of itself, but by enabling new mechanisms for democratic procedures to regain their legitimacy.

III. The counterclaim: the notions of input and output legitimacy and their implications for algorithmic decision-making

Democratic legitimacy is a potentially elusive locution, which deserves to be clarified for the present purposes. It can be defined as the principles and procedures through which

[15] MP Saffon and N Urbinati, "Procedural Democracy, the Bulwark of Equal Liberty" (2013) 41 Political Theory 441, 454.

[16] J Kyle and Y Mounk, "The Populist Harm to Democracy: An Empirical Assessment" (Tony Blair Institute for Global Change 2018) p 18.

[17] ibid, 19.

[18] A Silva-Leander, "Populist Government and Democracy: An Impact Assessment Using the Global State of Democracy Indices" (2020) Global State of Democracy in Focus no. 9, p 7.

[19] See H Kudnani, "Technocracy and Populism After the Coronavirus", in R Youngs (ed.), *How the Coronavirus Tests European Democracy* (Brussels, Carnegie Europe 2020).

collective decisions are accepted and deemed to be binding by those who have not directly participated in making them.[20] The locution thus refers to that particular kind of justification for the exercise of power under a democratic framework that is normative in the sense that it is objectively based on normative principles and procedures and informed by political theory.

Towards the end of the twentieth century, a cluster of different phenomena led scholars in law and political science to question the meaning of democratic legitimacy.[21] First of all, the progressive particularisation of governmental functions, which went hand in hand with the emergence of new rights, determined the articulation of the administrative apparatus in a series of agencies and bodies taking the responsibility for important policy decisions.[22] Second, the ideological clash between capitalism and socialism led some scholars to explore different models of legitimisation of political institutions by focusing on the interplay between the terms of political participation and the desirability of certain political results.[23] Third, judicial activism in the field of fundamental rights urged intellectuals to address the issue of democratic legitimacy by advancing the need to frame democracy in light of a "strong principle of equality".[24] Finally, the establishment of supranational levels of government, in Europe and globally, forced scholars to find a conceptual frame for the exercise of public powers beyond state borders, which happened via centralised decision-making processes, only marginally involving national democratic circuits.[25]

Thus, the literature on democratic legitimacy has grown mainly with the goal of deepening the understanding of democracy in the context of multi-layered and complex decision-making processes within which democratic deliberation is one step in the articulated procedures of political actions.

The main contribution of those doctrinal studies was the identification of different kinds of democratic legitimisation. Scholars such as Joseph Weiler went as far as proposing a typology within which it is possible to distinguish between: (1) input or process legitimacy; (2) output or results legitimacy; and (3) legitimacy based on *telos*, which means that legitimacy is gained by promising a desirable result.[26] Weiler's classification builds on the work of prominent political philosophers who engaged in defending democracy. Robert Dahl, for example, reflected on democracy by focusing on three components: voting, political participation and understanding of civic issues. According to Dahl, individuals subjected to collective decisions should be able to have their interests equally taken into consideration, as well as to gain control over the matters that reach the decision-making agenda.[27] Therefore, Dahl built his democratic theory upon a strong understanding of the principle of political equality, which is functional for the protection of a complex set of collective and individual social, cultural and economic interests.[28] Within Dahl's democratic theory, the essential core of democratic legitimisation lies in its ability to reflect

---

[20] P Rosanvallon, *Democratic Legitimacy* (Princeton, NJ, Princeton University Press 2011). For a European perspective, see S Bartolini, "The Nature of the EU Legitimacy Crisis and Institutional Constraints: Defining the Conditions for Politicisation and Partisanship" in O Cramme (ed.), *Rescuing the European Project: EU Legitimacy, Governance and Security* (London, Policy Network 2009) p 57.

[21] JHH Weiler, "The Transformation of Europe" (1991) 100 Yale Law Journal 2403.

[22] See S Rose-Ackerman, "American Administrative Law Under Siege: Is Germany a Model?" (1994) 107 Harvard Law Review 1279, 1279, arguing that "democracies need to strike a balance between popular control and expertise"; and RH Pildes and CR Sunstein, "Reinventing the Regulatory State" (1995) 62 University of Chicago Law Review 1, 3. The authors delve into the challenges that the role of expertise poses to political decision-making.

[23] CR Beitz, *Political Equality: An Essay in Democratic Theory* (Princeton, NJ, Princeton University Press 1990).

[24] R Dahl, *Democracy and Its Critics* (New Haven, CT, Yale University Press 1991).

[25] J Lodge, "Transparency and Democratic Legitimacy" (1994) 32 Journal of Common Market Studies 343.

[26] Weiler, supra, note 21, 2405.

[27] Dahl, supra, note 24, 322.

[28] ibid, 92.

the popular will in a meaningful way, such as through means of effective control over (1) the selection of political options and (2) political decisions.

Charles Beitz started by acknowledging that political equality is the main challenge for any model of democratic legitimacy. At the same time, he warned that the output ("result" in his words) of democratic processes should not be underestimated. In his conception of democracy, institutions gain legitimacy when they maximise the expected values of an independently specified social welfare function. In this context, only fair terms of political participation are likely to produce the most desirable results. According to Beitz, though, this understanding of democratic legitimacy is not equivalent to an outcome-oriented theory because the logic of the maximisation of expected values works within the framework of a social function that is concerned about alternative political outcomes, which essentially means that the social function is drawn by political preferences and, therefore, by inputs.[29]

Almost ten years after Dahl and Beitz, Fritz Scharpf identified two frames of input and output legitimacy by conflating the normative and social legitimacy of a given political regime.[30] Input legitimacy refers to the bottom-up process through which the *people* make political choices concerning how they want to be governed. Within the "input frame", political choices are legitimate to the extent that they reflect the will of the people. The latter, in turn, expresses a collective self-determination of preferences that are expected to be addressed by the representatives.

Input legitimacy resorts to a kind of consensus rhetoric whereby the people define preferences on the basis of a minimum agreement on some values. It comes as no surprise, then, that populist movements escalated the input legitimacy rhetoric by drawing a direct connection between the will of the people and the accomplishment of an authentic democratic model of government. In other words, according to populist rhetoric, democracy exists insofar as it consists of the realisation of people's determinations without the need for further political assessment or appreciation.[31]

On the contrary, under an output-oriented model, legitimacy concerns are addressed by focusing on the effective promotion of a constituency's common welfare through a number of political actions designed to solve problems of a collective nature. The output perspective takes into consideration a political environment of articulated needs and preferences in which shared values do not necessarily express a common identity that is translated into a unitary political will. To that extent, an output legitimacy frame suits the reality of pluralist societies where decision-making processes are located at more than one level of government. The output perspective engages with the substantive content of democracy rather than with its procedural meaning.[32]

Against this backdrop, focusing on outputs means considering democracy as a political regime in which a given set of individual and collective needs are recognised in the form of rights with a view to promoting social welfare and ensuring peaceful coexistence. By contrast, the input frame is more concerned with fulfilling democracy from the standpoint

---

[29] Beitz, supra, note 23.

[30] F Scharpf, *Governing Europe: Effective and Democratic* (Oxford, Oxford University Press 1999) pp 11–12.

[31] Fournier, supra, note 11.

[32] Democracy is an ambiguous term: it can refer to majority rule or it can capture a broader meaning whereby a political regime can be identified as democratic only if a set of fundamental rights is guaranteed. The conceptualisation of a substantive meaning of democracy has been offered by TM Franck in "The Emerging Right to Democratic Governance" (1992) 86 American Journal of International Law 46–91. Constitutional theory is also concerned with the substantive meaning of democracy and generally reluctant to identify the latter merely with a procedural rule. Under a constitutional model, democracy is framed by substantive constraints deriving from common shared values enshrined in constitutional texts, which in turn define the boundary of legitimate democratic choices: see B Ackerman, *We the People: Foundations* (Cambridge, MA, Harvard University Press 1991) pp 3–33.

of democratic deliberation (majority rule), while it does not address the effectiveness of democratic values in a given constituency.

Output legitimacy, however, does not necessarily rule out the bottom-up element of democracy. In looking at ways to reconcile the quality of outputs with the requirements of procedural legitimacy, we note that even in the field of proceduralism several authors, while not dismissing the role of procedures as a foundation stone for legitimacy, include as a further condition that outcomes of democratic decision processes have certain qualities. Kenneth Arrow's rational social choice theory, in a most influential manner, emphasised the requirement that political decisions produce rational outcomes.[33] Different authors have discussed the requirement of rational outcomes from different perspectives, some looking at the consistency of preferences and choices, others at the reasoning provided during deliberation. Under a model of legitimacy defined as "rational deliberative proceduralism", requirements include the political equality of all of the individuals comprised in the polity and called to participate in the political process and the rationally justified outcomes of collective decision-making.[34]

Input and output legitimacy tend to coexist in mature democracies. Through an input framework, needs and priorities are identified, legitimising the pursuit of some goals; in parallel, outputs measure the results and legitimise political choices that may not represent the specific content of democratic deliberation but nevertheless effectively achieve those desired goals.

From this perspective, algorithmic decision-making can be seen as an instrument of output legitimacy when looking at algorithms through the lenses of democratic theories stressing the importance of individuals' chances of understanding and assessing political choices. The perspective we suggest for this, however, is different from merely resorting to technocratic governance; in fact, as we discuss below, neither algorithms, which would certainly not be capable of ensuring "better" output per se, nor technocracy as a model of governance – whether AI- or human-led – fully address the issue of the de-legitimisation of democratic procedures that lies at the root of the problem.

IV. Algorithms and democratic legitimisation: a framework for analysis

A compelling interpretation of populism and technocracy suggests that the two tendencies are fundamentally complementary to one another.[35] The complementarity resides in the idea that both populism and technocracy can be understood as reactions to the system of government by political parties and its provision of political mediation and the procedural conception of political legitimacy to democratic governance. In this sense, populism and technocracy share a similar approach in that they both constrain and eventually void political antagonism, whereas procedural democracy offers a system to channel it. From this perspective, both populism and technocracy undermine two key features of parliamentary democracy, which are political mediation among different social groups and a procedural conception of political legitimacy.

If technocratic government seems unfit to offer a valid solution to the issue of delegitimised democratic procedures, so do algorithms if they are expected to reach substantive political decisions independently. In fact, the application of AI and algorithms to the public sector has been framed in slightly different terms by different regional and international institutions in recent years. The European Union (EU) defines AI as "systems that display

---

[33] K Arrow, *Social Choice and Individual Values* (New York, Wiley 1951).

[34] F Peter, "Democratic Legitimacy and Proceduralist Social Epistemology" (2007) 6 Politics, Philosophy & Economics 329, 335–37.

[35] C Bickerton and CI Accetti, "Populism and Technocracy: Opposites or Complements?" (2017) 20 Critical Review of International Social and Political Philosophy 186.

intelligent behaviour by analysing their environment and taking action – with some degree of autonomy – to achieve specific goals"[36]; the Organisation for Economic Co-operation and Development (OECD) defines AI as "machine-based systems that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments";[37] and the United Nations Educational, Scientific and Cultural Organization (UNESCO) defines AI as "technological systems which have the capacity to process information in a way that resembles intelligent behaviour, and typically includes aspects of reasoning, learning, perception, prediction, planning or control".[38] Even with the lack of a generally agreed upon definition, all of these formulations focus on some common features, such as the ability of these technologies to process information and perform tasks intelligently. Interactions between public bodies and AI can take several forms, including financing and developing technologies, harnessing the data to feed into the algorithms, imposing regulatory standards and, finally, utilising the technology to provide a range of services.[39] We are particularly interested in the intersection between these last two functions – more specifically, how algorithms can help boost the democratic legitimisation of the public bodies that utilise them.

In the context of public decision-making, algorithms can function in several ways. Let us attempt to provide a non-exhaustive classification by looking at the steps of policymaking and decision-making procedures. AI can: (1) represent the world by, for example, proxying demographic data; (2) predict or test the desirability of a given course of action in light of the results that the latter may determine, such as immigration detention risk assessment[40]; (3) reach a decision, generally on the assumption of the inherent reliability of the result and/or efficiency of the process, by, for example, selecting individuals who will benefit from an allocation decision that had been made within the traditional political process[41]; and (4) act as an algorithm-manager by supervising and controlling public servants who are required to make complex decisions.[42]

In each of these scenarios, algorithms play different roles. Representation and prediction functions, described under (1) and (2), imply that algorithms can provide decision-makers with accurate information concerning a given political option. Therefore, algorithms do not replace political choices, but rather they create the conditions for a political choice to be confronted with concrete outputs. The selection function

---

[36] EU Commission, Communication From the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee of the Regions on "Coordinated plan on Artificial Intelligence", COM(2018) 795 final, 7.12.2018.

[37] OECD, Recommendation of the Council on Artificial Intelligence, C/MIN(2019)3/FINAL adopted on 22 May 2019.

[38] UNESCO, First draft of the Recommendation on the Ethics of Artificial Intelligence, SHS/BIO/AHEG-AI/2020/4REV.2, 7 September 2020.

[39] B Ubaldi et al, "State of the Art in the Use of Emerging Technologies in the Public Sector", OECD Working Papers on Public Governance No. 31, 2019.

[40] Let us imagine a situation in which an algorithm is used to predict the likelihood that a migrant would face torture or inhuman and degrading treatment if they are sent back to their country of origin. In this particular case, algorithm settings will probably consider the safety of the country of origin on the basis of a number of features whose consistency and accuracy may be open to question. See M Nofferi and R Koulish, "The Immigration Detention Risk Assessment" (2014) 29 Georgetown Immigration Law Journal 45.

[41] This is the case for using an algorithm in the context of welfare allocation decisions. There are a number of such examples: see infra, note 54.

[42] Let us take as an example the use of algorithms in the context of judicial proceedings. The Council of Europe addressed the potentialities of the use of AI in judicial decisions: see "European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment", European Commission for the Efficiency of Justice, Strasbourg, 3–4 December 2018. See also M Bovens and S Zouridis, "From Street-Level to System-Level Bureaucracies: How Information and Communication Technology is Transforming Administrative Discretion and Constitutional Control" (2002) 62 Public Administration Review 174.

mentioned under (3) uses algorithms to speed up procedures that would otherwise require careful and lengthy examination. Algorithms can thus ensure the efficiency of the selective process and the consistency of results. In the scenarios depicted under (3) and (4), the algorithm makes decisions instead of the decision-maker identified via the ordinary political process. Computer science is used to replace political deliberation with possible benefits in terms of the efficiency of a democratic system.

It is worth noting that none of these scenarios envisage algorithms independently defining preferred outcomes of political decisions, which would completely exclude any grounds for democratic procedural legitimacy; instead, we focus on the opportunities they bring for disentangling political legitimisation from the merely formulaic input-centric aspect of the will of the majority. Algorithms offer the opportunity to refocus on output legitimacy by connecting inputs and outputs and making such connections rationally appraisable.

From this perspective, a notable element of the rational deliberative procedural model is that, when the rationality test is applied to the outcomes in order to assess their consistency with the premises of any political decision (defined as "the reasons given during deliberation in favour of or against certain alternatives"),[43] it may be the case that the rational outcome(s) is different from what voters would have immediately chosen if asked directly. In this way, rational deliberative proceduralism separates legitimacy from input-centrism by offering alternative procedural grounds.

In light of this theoretical model, we suggest that algorithms offer a viable way to implement both of these conditions of (rational deliberative procedural) democratic legitimacy by offering the possibility of equal democratic participation and a system by which to rationally assess the consistency between premises and outcomes.

In translating the model from theoretical to more practical perspectives, the impact of algorithmic decision-making on democratic governance and legitimacy can be discussed with respect to its contribution to policymaking, and particularly through the prism of consolidated findings in policy studies where, over the last few decades (largely following seminal works through the 1960s to the 1980s), a direct influence of the content of public policy decisions on the character of democracy has become clear. More specifically, distinct features of the policy cycle – such as the framing of issues, the construction of targets, the structure of implementation and delivery systems and the transparency of the whole process – impact foundational elements of democratic life, such as citizens' trust in the government's ability to solve public problems, public support for the government's action and public accountability. Helen Ingram and Anne Schneider have helpfully suggested a framework through which to understand and evaluate the effectiveness of public policies in contributing to democracy: by building on core literature in the field, they suggested that policies contribute to democratic governance inasmuch as they allow it to expand its franchise, scope and authenticity of democratic governance.[44] These three domains respectively refer to the number of participants, the number of domains of life directly and legitimately affected by political decisions and the possibility for citizens to exercise "substantive, informed and competency-engaged" democratic control. The three domains exist in delicate balance with one another, as, for instance, expanding the scope of democratic governance could come at the cost of superficial deliberation, thus undermining authenticity. In practical terms, the first two domains refer to the processes of the identification of issues that belong in the political discourse and the terms of the debate regarding which actions public authorities would be expected to take (the authors exemplify this dynamic by describing the switch in the discourse over water policy in the

---

[43] Dahl, supra, note 24, 335.
[44] H Ingram and AL Schneider, "Policy Analysis for Democracy" in RE Goodin, M Moran and M Rein (eds), *The Oxford Handbook of Public Policy* (Oxford, Oxford University Press 2008) p 169.

USA from water as a public good to the conceptualisation of water as a privatised commodity, and criminal policies built around the idea of crime as a violation against an individual as opposed to an offence against the society or state). The terms and the results of such definitory processes depend naturally on how open and inclusive these are with respect to their ability to include those who are directly affected as well as those with a specialised knowledge in the field. The process of identifying the issues leads to two relevant effects: by defining any societal issues as a question of political relevance, the process identifies public authorities as competent decision-makers; and at the same time, by "pitching" the debate at a more or less specialised level, the process inevitably defines its boundaries by making it more or less accessible to ordinary citizens.

A crucial aspect of this consists in the creation of open public forums where citizens can discuss policy problems openly and directly. The way a policy is framed and designed is directly responsible for the appearance of such forums. The process of policymaking, in other words, needs to be framed in a way that facilitates the emergence of such opportunities for civic discussion.

The model, as explained in those terms, does not fail to consider a substantive change in how accountability ought to be construed in the current context of the decentralisation of political power and the increasing distances between citizenries and those in government, which, in turn, calls for more direct citizen involvement in holding governments accountable and, crucially, a different approach to assessing the specificities of public governance by evaluating aspects that are not limited to mere effectiveness and efficiency. Ingram and Schneider suggest that government action instead be measured by "its ability to intervene strategically in the complex networks of policy delivery systems to encourage better access to information, to correct for power imbalances and damaging stereotypes and social constructions among stakeholders, and to create arenas and spheres of public discourse".[45]

It is therefore useful to examine algorithmic decision-making by carefully disentangling the cluster of issues surrounding it through the lenses of democratic theories. By unpacking the logic of democratic legitimacy and still adhering to the input/output framing, it is possible to identify at least three different, though related, cruxes: (1) the problem of understanding and selecting civic issues that deserve to be addressed by political institutions; (2) the problem of controlling which issues reach the democratic institutions; and (3) the problem of evaluating and challenging the results of a given course of political action. The inherent complexity of political processes makes these three aspects particularly sensitive with respect to algorithmic decision-making. Algorithms work on the basis of instructions related to data processing; the selection and the organisation of data derive from human input or they can be learnt by the computational machine itself according to a logic that is completely artificial.

Therefore, tackling the three aforementioned problems in the context of algorithmic decision-making involves different consecutive steps. The first issue to address is determining the extent to which the instruction algorithms reflect political preferences that have legitimately reached democratic institutions. This issue can also be framed as a problem of understanding the "democratic soundness" of algorithmic decision-making by clarifying how computers process data or even how they learn to *select and process* data. The understandability of the process and the selection of data are key to ensuring that participation is broad enough to let democracy expand its franchise.

The second issue is the ability of the political community to control the algorithmic decision-making process by deciding which choices can be dealt with by AI. Control over the algorithm here means that the political community has the opportunity to select the issues that are allocated to AI and, therefore, to own the scope of democratic governance.

---

45 ibid, 184.

Finally, the third issue is the opportunity to challenge algorithmic decision-making, understood as the opportunity to assess, question and potentially change the outcome of any given non-human decision.

With a view to applying the output legitimacy frame to algorithmic decision-making and to test its impact on democratic processes, the paper next explores these three different dimensions of the interplay between algorithmic decision-making and democracy.

1. Understanding civic issues under an algorithmic decision-making framework

For a democratic process to be authentically based on political freedom, citizens should be able to understand civic issues; that is to say that the citizenry should, with a reasonable intellectual effort, understand the justification for decisions taken collectively. Dahl identified understanding civic issues as one of the conditions under which democracy can properly function as a reliable means for protecting and promoting the goals of persons that are subjected to collective decisions.[46] To be more precise, according to Dahl, the opportunity to understand matters that reach the decision-making agenda enables people to properly enjoy political equality, which means to have their interests equally addressed by legislators. In democratic theory, this need is commonly expressed by an emphasis on education and participatory rights. In such a context, understanding means being able to have a sufficiently clear idea of the functioning of the institutions as well as (at least) a basic understanding of the language of the public discourse.[47] Understanding is essential to building trust in political institutions; without it, people are left with no clear point of reference regarding the reliability of democratic processes.

When it comes to algorithmic decision-making, understanding civic issues that are processed through algorithms requires a second level of comprehension concerning the peculiar mechanism of AI, especially when it is performed by machine learning and with limited involvement from human intelligence. This is to say that for any algorithmic decision-making process to be compatible with democracy it should be coupled with the guarantee of a minimum level of technological education, as the EU has recently recognised.[48] Algorithmic decision-making therefore needs to be grounded in a "cultural cognition" of what an AI decision-making process is about. By "cultural cognition" we refer here to the shared view or frame of the world and/or of a given portion of reality of a group of people.[49] Such a concept builds on Mireille Hildebrandt's idea of enabling citizens to counter-profile AI. According to Hildebrandt, if citizens are equipped with proper intellectual instruments by which to understand the AI ecosystem, then they may be able to challenge the modelling that they have been represented with or even to question the set of instructions used by algorithms.[50] Counter-profiling requires citizens to engage actively with AI, an attitude that can be stimulated by spreading a cultural cognition of AI decision-making.

By assuming this perspective, we can imagine a situation in which citizens may approach algorithm decisions in order to make sense of the unfolding of the political decision-making process.

---

[46] Dahl, supra, note 24, 322.

[47] See R Dworkin, "What Is Equality – Part 4: Political Equality" (1987) 22 University of San Francisco Law Review 1.

[48] See, for example, the European Commission Strategy on Artificial Intelligence, which states that modernisation of education in light of the impacts of AI on social dynamics should be a priority for governments: see COM(2018) 237 final, at 4 and 12.

[49] The concept of "cultural cognition" is borrowed from JK Sax, "The Problems with Decision-Making" (2020) 56 Tulsa Law Review 39.

[50] M Hildebrandt, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology* (Cheltenham, Edward Elgar Publishing 2015) pp 100–03.

Let us consider the example of immigration. AI can be employed across the whole spectrum of procedures related to migration, from the screening of applicants for asylum protection to the risk assessment decisions related to individual migrants who may pose a threat to national security. Examples include the USA and Canada, where governments considered using algorithms developed through the analysis of wide swaths of data to realise trend studies and to make predictions as to the influx of migrants in a particular context.[51] Such use comes with many concerns, including risks of human rights violations and forms of direct or indirect discrimination deriving from implicit biases in algorithm setting or simply from the absence of reasonable assessments of results.[52] Those risks can be reduced by increasing the sophistication of the instructions given to the machine or through a careful human check on biases. These issues, though important, are not the focus of our attention here. Instead, we are attempting to look at the problem from a different angle. Let us imagine using an algorithm, or even a machine learning algorithm, to predict the likelihood that a migrant would face torture or inhuman and degrading treatment if they are sent back to their country of origin. Let us also imagine that the algorithm's settings consider the safety of the country of origin on the basis of a number of features derived from decisions made by competent authorities in the past. What the algorithm is doing in this situation is predicting how many people, given the instructions on which it operates, will be denied entry, eventually by also classifying different levels of risks of inhuman or degrading treatment.

A decision ordering expulsion when the risk of inhuman or degrading treatment exists can have devastating effects on the right to life of the individual concerned if it is made based on incorrect assumptions. As such, there can be reasonable agreement on the fact that this is a technical issue that impacts the general public and also signals a state's commitment to human rights protection. It comes as no surprise, then, that scholars urge governments to avoid using AI in such a context on the basis of the assumption that relying on the empathetic assessment of human beings can produce fairer results.[53] Even when the emphasis is not strongly on empathetic assessment but on implicit biases, such as automatic decisions on the level of risk that an individual migrant may pose to national security, scholars insist that AI does not tailor decisions on well-assessed grounds.[54]

In such a context, it is possible to imagine legal solutions ranging from the right to challenge the automatic decision, as codified in Article 22 of the GDPR 8,[55] or the duty of double-checking on the part of public authorities. In both cases, a second step involving human intelligence can correct the unfair, unreasonable results of an algorithm or simply place a second deliberation before the final decision so that the decision-making process is still not determined by an algorithm.

---

[51] In Canada, this is the result of a broader strategy to implement ethically aware artificial intelligence decision-making processes: see the Pan Canadian Artificial Intelligence Strategy, available at <https://policyoptions.irpp.org/magazines/august-2018/responsibly-deploying-ai-in-the-immigration-process/>.

[52] See D Robinson and K Vold, "Immigration Decisions Are Complex with High Stakes for the People Involved. The Government Must Tread Carefully on Using AI in the Screening Process", available at <https://policyoptions.irpp.org/magazines/august-2018/responsibly-deploying-ai-in-the-immigration-process/>.

[53] Nofferi and Koulish, supra, note 40.

[54] When looking closely at both arguments, though, a fundamental logical flaw emerges: any decision based on some kind of generally applicable criteria, as decisions based on laws generally are, faces the risk of being insufficiently tailored to a peculiar case and thus proving unfair with respect to an individual situation while still pursuing an objectively fair result on the whole. Therefore, the problem is not a problem of AI, but rather one of the general or blind application of non-empathetic intelligence, which may lack reasonableness or be fundamentally biased. See J Dickinson, "Legal Rules: Their Function in the Process of Decision" (1931) 79 University of Pennsylvania Law Review 835.

[55] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

From an alternative viewpoint, algorithms such as the one used in the example above can perform an informative function by making people aware of the actual numbers of individuals risking death or torture because of expulsion decisions made under a certain set of criteria reflecting a given political choice. Moreover, people may be able to frame those pieces of information against the backdrop of broader national or supranational policies in order to better understand their implications (and eventually to better understand their preferences or priorities).

Along these same lines, algorithms can enhance the understanding of the real dimensions of a problem and expose populist rhetoric by confronting it with facts and data. This is a kind of "use" of algorithms that, when properly accessible and transparent, equips voters with fact-checking instruments concerning policies and their effectiveness.

The relationship between data processing and participation can be further explored by addressing the ways in which data processed by algorithms are fed into the political decision-making process to generate informed political choices.

A key contribution in the field of data science is that extremely fragmented, individual data collected by technologies are often insignificant if taken by themselves. At a later stage, technologies re-aggregate data in the process of generating outputs and "fold [the data] back into the experience of everyday life". This portion of the process offers an opportunity to harness data within governance procedures in a way that can effectively contribute to a more meaningful and thorough understanding of civic issues by decision-makers.[56]

Examples of such processes have been identified and discussed with respect to specific policy sectors, such as environmental policies. It has been discussed how citizens' participation in the data-led delivery of public services within the context of smart cities allows local governments to tune in to the everyday experiences of citizens and their use of space and urban infrastructure. Citizens' sharing of individual data performs a function that is substantively equivalent, for its contribution to democratic engagement, to participatory media, to the point of eventually becoming a "constitutive [practice] of citizenship".[57] Even beyond the specific cases of smart cities and environmental policies, the positive impact of harnessing data in public decision-making processes has been identified as twofold; it enables public action to become better informed and more responsive to citizens' needs, which, in turn, places a specific responsibility on public decision-makers to develop legal and policy frameworks for the responsible sharing and processing of data to feed democratic processes.[58]

From these doctrinal perspectives, data-sharing emerges as a practice that enables citizens to connect with public decision-making processes in a way that continues and complements the more traditional role of the news media to build and inform political agendas. From this premise, a further reflection can follow regarding the opportunity that such new modalities of citizenship offer to feeding the policymaking process with qualitative and reliable information that neither mainstream media[59] nor social media[60] seem capable of contributing in the current climate due to the policy and legal frameworks in which they operate.[61] By contrast, citizen data could break the cycle of biased, sensationalistic narratives cycled into the political agenda.

---

[56] N Couldry and A Powell, "Big Data from the Bottom Up" (2014) 1 Big Data & Society 1–5.

[57] J Gabrys, "Programming Environments: Environmentality and Citizen Sensing in the Smart City" (2014) 32(1) Environment and Planning D: Society and Space 30–48.

[58] WJ Mitchell and F Casalegno, *Connected Sustainable Cities* (Cambridge, MA, MIT Mobile Experience Lab Publishing 2008).

[59] G Mazzoleni, "Populism and the Media" in D Albertazzi and D McDonnell (eds), *Twenty-First Century Populism: The Spectre of Western European Democracy* (New York, Palgrave Macmillan 2008) pp 49–64.

[60] T Flew and P Iosifidis, "Populism, Globalisation and Social Media" (2020) 82(1) International Communication Gazette 7–25.

[61] D Freedman, "Populism and Media Policy Failure" (2018) 33(6) European Journal of Communication 604–18.

Whereas concerns regarding the impacts of the mass collection of individualised data on privacy and data protection rights have (quite rightfully) been at the centre of the academic literature, this alternative perspective speaks to the current rise of populism in a way that has often been overshadowed by the more prominent conversation surrounding the risks connected to such practices. Without dismissing the valid concerns that the academic literature has unveiled and continues to discuss to date, re-centring the conversation in a way that acknowledges this potential opportunity can instead be key to making progress in countering the rise of populism.

For data to feed into the political process, they need to be discernible from those in charge of shaping public policies. This idea is based on the observation that social analytics provide information and can perform a role in the policy cycle akin to the role traditionally played by content in the media sphere. But whereas media content can be understood and interpreted semantically, data convey information that can be made sense of in different ways.[62] Algorithms can provide decision-makers with several alternative outputs that depend on how a given set of instructions elaborates data. From a democratic standpoint, then, concerns are raised by the reliability of the particular elaboration or modelling of the available dataset. The literature has already addressed this problem in the context of auto-mated data processing, with a view to controlling and directing machine learning processes. In particular, Hildebrandt advocates for "agonistic machine learning", which she identifies as a requirement that "companies or governments that base decisions on machine learning must explore and enable alternative ways of datafying and modelling the same event, person or action".[63] Such an agonistic frame would provide decision-makers with several accounts of processed data, which, in turn, would enable them to detect biases or incorrect assump-tions in machine learning processes. An agonistic approach, however, benefits algorithmic decision-making more generally. In fact, it helps public institutions make sense of data by confronting multiple models or alternative readings of the reality. Public institutions can then couple this piece of information with citizens' inputs to make informed choices that are based on a thorough understanding of civic issues.

2. Controlling and selecting civic issues that are assigned to algorithmic decision-making

It may well be pointed out that no decision-maker would like to expose their choices to a fact-checking system that is the result of a non-empathetic assessment not connected to a genuine political appreciation of their choices. This is where the relevance of the second element of the algorithmic decision-making frame comes in: the opportunity to exercise control over the issues that can be delegated to algorithmic decision-making processes.

Algorithmic decision-making is often praised for being efficient, while critical voices have been raised concerning the discriminatory biases that algorithms often incorporate and reproduce.[64] Therefore, allocating public choices to AI can result in a blank check, even when precise instructions have been set up. This is especially true in the context of

---

[62] Couldry and Powell, supra, note 56, 5.

[63] M Hildebrandt, "Privacy as Protection of the Incomputable Self: From Agnostic to Agonistic Machine Learning" (2019) 83 Theoretical Inquiries in Law 83–121.

[64] F Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Cambridge, MA, Harvard University Press 2015) pp 1–17 (arguing that data analyses about consumers are often not disclosed and even hidden from legal process). See also A Chander, "The Racist Algorithm" (2017) 115 Michigan Law Review 1023. Studies have been specifically conducted on the credit sector to demonstrate biases on the allocation of loans: L Rice and D Swesnik, "Discriminatory Effects of Credit Scoring on Communities of Colour" (2013) 46 Suffolk University Law Review 935, showing that credit-scoring systems in the USA systematically discriminate against communities of colour vis-à-vis White Americans. See also B Birnbaum, *Credit Scoring and Insurance: Costing Consumers Billions and Perpetuating the Racial Divide* (Boston, MA, National Consumer Law Center 2007), maintaining

machine learning when human intelligence's contribution is limited and the machine is capable of setting autonomous premises for logical deductions.[65] An uncontrolled use of AI increases people's scepticism regarding algorithmic decision-making, thus leading them to question its legitimacy in a liberal democracy grounded on moral values that have legal significance. As Tom Tyler has argued, people's perceptions of legitimate authority are more likely to be met if authorities ensure the existence of procedural safeguards as well as the opportunity to be heard for individuals.[66] To avoid the blank check effect, then, the processes through which AI reaches decisions need to be complemented by procedural safeguards that should, at least, enable people, through their representatives, to exercise full control over determining the issues that are delegated to an AI process. In this respect, the role of parliaments is crucial. Algorithmic decision-making can contribute to democratic legitimisation to the extent that parliaments are enabled to control the issues that are delegated to AI, the procedures and conditions under which the issues are evaluated and the checks on the fairness of the overall decision-making process. Such an approach requires that parliamentary procedures are in place for (1) selecting issues that can be assigned to AI decision-making and (2) proceduralising this particular kind of decision-making process in a way that enables individuals to evaluate and challenge its outcomes, as will be further explained in Section IV.3.

As for the selection of issues, this essentially means that parliaments should assess whether certain issues can be dealt with by AI and the extent to which it can do so. The principle that statutory law is required to regulate AI decision-making is not included in existing international legal instruments, and such a principle is uncertain or weak in national jurisdictions as well. Neither is Article 22 GDPR diriment – it has not been interpreted in such a way as to demand that the law should explicitly authorise resorting to automated decision-making. For example, the UK's Information Commissioner Office (ICO) has stated that when either legislation or common law authorises the exercise of power, the choice of means to achieve purposes consistent with legitimate powers includes resorting to automated decisions, with no need for an *ad hoc* authorisation by law.[67]

Let us consider the case of welfare benefits allocation choices, which in recent years have been delegated to AI to process the identification of beneficiaries as well as the sums they were entitled to receive.[68] There are hardly any cases of formalised delegation from democratically legitimised public authorities to automated processes of such substantive decisions.

Most often, automated decisions pertain to technical tasks to be performed downline of a political decision (eg whether to establish a certain welfare benefit or not) made elsewhere, on the assumption of the efficiency of AI.

Concerns regarding the impacts of automated decision-making on socioeconomic rights have been expressed by academics and civil society organisations. Among the most high-

---

that discrimination for residents in minority communities has been demonstrated by studies on insurance scores calculated by algorithms that had eliminated factors such as income, education or unemployment status.

[65] C Riegler, "The Moral Decision-Making Capacity of Self-Driving Cars: Socially Responsible Technological Development, Algorithm-Driven Sensing Devices, and Autonomous Vehicle Ethics" (2019) 11(1) Contemporary Readings in Law and Social Justice 15.

[66] T Tyler, *Why People Obey the Law* (Princeton, NJ, Princeton University Press 1990) pp 96, 137–38.

[67] See <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/when-can-we-carry-out-this-type-of-processing/#id1>.

[68] Examples include resorting to automated decisions in relation to eligibility to receive welfare benefits such as Medicaid and food stamps at the state level in the USA: see V Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor* (New York, St Martin's Press 2018). Local authorities also use automated decision-making procedures to assign welfare benefits in the UK: see L Dencik, J Redden, A Hintz and H Warne, "The 'Golden View': Data-Driven Governance in the Scoring Society" (2019) 8(2) Internet Policy Review <https://doi.org/10.14763/2019.2.1413>.

profile and influential critical voices is the report on "Extreme Poverty and Human Rights" released by the UN Special Rapporteur on extreme poverty and human rights in 2019.[69] The report expresses concerns regarding the application of algorithmic decision-making to welfare services for at least three separate ranges of reasons: (1) lack of privacy and data protection (especially in the phase of identity verification), even veering towards systemic surveillance (exploited on the pretext of preventing and detecting fraud); (2) system failures (while assessing receivers' eligibility, benefit calculations and payments); and (3) unfairness (from basing decisions affecting individual rights on group-based predictions), lack of transparency and the risk of reinforcing existing inequalities and discrimination, particularly in relation to risk scoring and need classification. Such concerns question the appropriateness of AI decision-making on socioeconomic rights, irrespective of the constitutional or legislative status they enjoy in a particular country. In fact, the danger of discrimination and inequality predates the development and possible uptake of AI. Where the relevant rights are granted at the legislative level, the risk is one of violating due process guarantees; by contrast, where they enjoy constitutional status, the problem is even greater in terms of the limitation or curtailment of fundamental rights. The choice of assigning a given issue to automated decision-making should therefore be carefully assessed within the political process and not made on the simplistic assumption of the efficiency that AI may bring into the public institutions.

The Commission White Paper on Artificial Intelligence engages with similar issues.[70] Indeed, the Commission states the need for the EU to implement the use of AI in public sectors to the advantage of citizens, while at the same time it remarks on the relevance of a regulatory framework ensuring the protection of European values and fundamental rights. The White Paper's concern for fundamental rights, however, is not exclusively devoted to protecting privacy and avoiding discrimination. In fact, the Commission is preoccupied with envisioning the use of AI in a context in which citizens trust institutions because they perceive that automated decisions are encapsulated in a system that takes into consideration human dignity and all of the fundamental rights protected at the EU level.[71] The relevance of the principle of trust is also reflected in the Commission's approach to the assessment of risks connected to AI. In particular, the White Paper stresses the importance of evaluating the impacts of AI not only from an individual perspective, but also from a societal one. In doing so, the Commission favours the idea of AI as a functional tool to enhance equality of opportunity, socioeconomic welfare and democracy. The Commission has advanced a similar approach in the Communication to the European Parliament titled "Fostering a European Approach to Artificial Intelligence", in which it has identified a strategy built on several requirements for an AI ecosystem that also tackles public responsibility, including societal and environmental wellbeing.[72]

Neither the White Paper nor the AI strategy, however, addresses openly how the use of AI can become entrenched in the democratic process. Rather, at its present stage of evolution, EU law conceives of trust in the AI ecosystem as a principle that is fulfilled by substantive rules protecting fundamental rights and ensuring oversight and accountability mechanisms. Such an approach is also confirmed by the Commission's Proposal for a Regulation on Artificial Intelligence.[73] Taking stock of the White Paper, the Proposal focuses on addressing risks associated with the use of algorithm technology by advancing "a legal framework for trustworthy AI". To that effect, the proposed regulation identifies

---

[69] United Nations General Assembly, "Extreme Poverty and Human Rights", Report of the Special Rapporteur on extreme poverty and human rights, 11 October 2019.

[70] On Artificial Intelligence – A European Approach to Excellence and Trust, COM(2020) 65 final.

[71] See also Communication on Building Trust in Human-Centric AI, COM(2019) 168.

[72] COM(2021) 205 final.

[73] COM(2021) 206 final.

different levels of risks connected to algorithmic technology and establishes transparency obligations for those systems that interact with humans or with their emotions, as well as for those generating or manipulating contents. The legal framework is completed by setting up a governance system for AI that establishes a system based on the interplay between the EU and national authorities. At the EU level, the Commission proposes the creation of a European Artificial Intelligence Board with the task of guaranteeing the effective implementation of the regulation. At the national level, Member States will be required to designate competent authorities to apply the regulation, including a national supervisory authority. Those authorities will be designed to supervise AI systems providers. In particular, those bodies will set out to monitor and report obligations regarding use, incidents or malfunctioning of AI technology.[74] While the Proposal is mainly concerned with guaranteeing fundamental rights and the functioning of the internal market, nothing in the Proposal suggests that EU institutions in fact are specifically concerned with the broader approach involving the political decisions of resorting to AI in decision-making process in the first place.

While the proposal put forward by the EU authorities focuses on a (legitimate and timely) concern for the concrete impacts of AI on fundamental rights, it appears to disenfranchise representative political processes from the governance of AI. However effective it may be to set up dedicated agencies to monitor the impacts of AI on fundamental rights, proceduralisation of AI is not a guarantee per se,[75] and the current European framework disregards the political value of the choice of authorising the use of AI in decision-making processes.

3. Evaluating and challenging algorithmic decision-making

Following on from this analytical framework, we now look at the potential for algorithms to increase the accountability of public policies by making relevant information more easily accessible, providing information more accurately and facilitating public debate among citizens.

Looking at the impact of algorithms from this specific angle raises apparent challenges, even at first glance. The academic debate on the topic has traditionally discussed the two issues of algorithmic transparency and accountability as being closely intertwined. In the academic literature, the issue of the chronic lack of meaningful ways to subject algorithm-led decision processes to thorough scrutiny has been connected to characteristics such as complexity and secrecy, leading to the now popular expression "algorithmic black box".[76]

It may be that, due to the very nature of such technologies, black-box issues cannot find satisfactory solutions. It has been emphasised that the level of sophistication of current AI techniques and deep learning oftentimes makes the details of decisional rules and patterns impossible to explain, even for the developers of the technology, and when decision-making processes in the public sector involve a certain degree of discretion it may be necessary to always require human intervention.[77] More optimistically, other academic authors have already suggested that algorithmic decision-making processes could, both in theory and in practice, lead to "fairer and more objective decisions, grounded in data that are representative of the community where the decisions apply"[78] thanks to the

---

[74] COM(2021) 206 final, Titles VI, VII and VIII.

[75] R Koulu, "Proceduralizing Control and Discretion: Human Oversight in Artificial Intelligence Policy" (2020) 27 Maastricht Journal of European and Comparative Law 720–35.

[76] Pasquale, supra, note 64.

[77] H-W Liu et al, "Beyond *State v Loomis*: Artificial Intelligence, Government Algorithmization and Accountability" (2019) 27 International Journal of Law and Information Technology 122, 139.

[78] B Lepri et al, "Fair, Transparent, and Accountable Algorithmic Decision-Making Processes. The Premise, the Proposed Solutions, and the Open Challenges" (2018) 31 Philosophy & Technology 611, 622.

implementation of a range of possible technical solutions (eg deploying participatory processes to include diverse and local voices in co-designing algorithms and vetting procedures to allow for the sharing of data and algorithm templates).[79] Other authors have expressed optimistic views that research is underway to promote more transparency, ideally to the point of providing an explanation of decision models sufficient to allow a human to understand how inputs relate to predicted results.[80]

However, one question that remains is how to make algorithms transparent; another related yet different question is how to make sure that algorithms can contribute to the accountability of policy decisions and democratic legitimacy. While technology and law hopefully both progress on parallel tracks to make "algorithmic boxes" less "black", we turn to the question of whether and how algorithmic decision-making can be subjected to thorough evaluation and, when necessary, challenged in a way that boosts democratic legitimacy.

Several regional and international organisations have considered this issue and recommended policy frameworks that expressly connect requirements such as transparency and explicability with accountability. UNESCO's draft recommendation is particularly straightforward in making the connection between explicability (defined as "making intelligible and providing insight into the outcome of AI systems", with specific regard for the "input, output and behaviour of each algorithmic building block and how it contributes to the outcome of the systems") and transparency, as well as the connection between these two principles and accountability and trustworthiness, since making processes and outcomes transparent and traceable is conducive, in turn, to keeping AI actors responsible and liable for their decisions. Appropriate oversight and due diligence mechanisms throughout the life cycle of AI systems are thus recommended.[81] Similarly, the EU's Ethics Guidelines for Trustworthy AI require that automated processes be "transparent" and any decisions "explained to those directly and indirectly affected" "to the extent possible" so that they can be "duly contested". When black-box or other technical difficulties make full explicability impossible, then other back-up measures should be provided (eg "traceability, auditability and transparent communication on system capabilities"), proportionately to the context and severity of the consequences of a possible erroneous policy outcome.[82] The OECD recommends "transparency and responsible disclosure regarding AI systems" through the provision of "meaningful information, appropriate to context and consistent with the state of art" to "enable those affected by an AI system to understand the outcome" and "enable those adversely affected by an AI system to challenge its outcome based on plain and easy-to-understand information of the factors, and the logic that served as the basis for the prediction, recommendation or decision".[83]

Transparency and explicability are also the core principles in Articles 15 and 22 GDPR, which have made the rights to an explanation and to contest decisions parts of a fundamental right to data protection. Any references to Articles 15 and 22 GDPR need to be interpreted, in this context, only *lato sensu*, as in fact in many cases public policy decisions may be based on anonymised data, which in any form do not involve the right to personal data protection in a direct way. These provisions ultimately appeal to the idea that the opacity of "black-box" decisions could be countered by a right to an explanation and to "open" the box.

[79] ibid, 623.

[80] B Goodman and S Flaxman, "European Union Regulations on Algorithmic Decision-Making and a 'Right to Explanation'" (2016) 38(3) AI Magazine DOI: 10.1609/aimag.v38i3.2741.

[81] UNESCO Recommendation, paras 40–41.

[82] High Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI" (2019) p 13.

[83] OECD Recommendation, Principle 1.3.

Other suggestions have been advanced in the literature to make algorithms explicable and accountable. Doshi-Velez and Kortz, for instance, considered the possibility of legally mandating that governments explain the ways in which inputs affect outcomes, concluding that such an obligation would most likely be technically feasible but onerous.[84] The requirement of a "human in the loop" has also proven to be popular; however, concerns have been raised over this possibly being counterproductive, for it would exclude from the scope of application of GDPR's safeguards a large number of decisions in high-risk areas.[85]

The right to an explanation is thus not enough in and of itself to enable a meaningful right to challenge algorithmic decisions. Yet some recently developed legal frameworks evidently focus on the former but not the latter; an example of this is France's Digital Republic Law.[86] This law requires large public-sector bodies to make publicly available, in an easily accessible format, the algorithmic processes that were used to reach decisions affecting individuals. This provision takes a step forward compared to the GDPR in that it applies to partially automated decisions as well; it does not, however, provide for a specific procedure to challenge such decisions. While the original bill was debated, however, the advisory commission had recommended introducing a new individual right to oppose profiling, to require human intervention and to oppose discriminatory decisions.[87]

Another suggestion for extending the scope of this newly established right was put forward by researchers working in the French Prime Ministerial task force for open data and open government, Etalab. The researchers identified a set of discrete (yet interconnected) obligations that public authorities should meet when making decisions based on algorithms, such as acknowledging the (entire or partially) automated nature of the process, explaining the functioning of the algorithm, offering a justification for choosing that particular algorithm, making the relevant source code and documentation public and, finally, providing mechanisms for challenging the outcomes.[88]

The idea of turning the principles of explicability and accountability into actual legal rights certainly marks a step in the right direction. However, Edwards and Veale noted that when automated processes affect large groups, individuals are usually unlikely to successfully mount challenges, and representative bodies may be in a better position to exercise this function, provided they are equipped with certain technical requirements, such as access to data used for training and modelling purposes, amongst others.[89]

It may seem far more appropriate for a representative body to accept and mount challenges than each individual user when it comes to public policy decisions. Following on from Edwards and Veale's suggestion that representative bodies step up, it is useful to recall here two principles included in the Council of Europe Recommendation on the human rights impacts of algorithmic systems,[90] namely the principle of democratic participation and awareness (urging state authorities to "foster general public awareness of the capacity, power and consequential impacts of algorithmic systems") and the expectation

---

[84] F Doshi-Velez and M Kortz, "Accountability of AI Under the Law: The Role of Explanation" (2017) Berkman Klein Center Working Group on Explanation and the Law, Berkman Klein Center for Internet & Society working paper.

[85] S Wachter, B Mittelstadt and L Floridi, "Transparent, Explainable, and Accountable AI for Robotics" (2017) 2(6) Science Robotics eaan6080.

[86] Law no. 2016-1321 du 7 octobre 2016 pour une République numérique.

[87] <https://www.assemblee-nationale.fr/14/pdf/rapports/r3119.pdf>.

[88] S Chignard and S Penicaud, "'With Great Power Comes Great Responsibility': Keeping Public Sector Algorithms Accountable" Etalab Working paper on algorithmic accountability 2019.

[89] L Edwards and M Veale, "Enslaving the Algorithm: From a 'Right to an Explanation' to a 'Right to Better Decisions'?" (2018) 16 IEEE Security & Privacy 46.

[90] Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems, adopted by the Committee of Ministers on 8 April 2020.

that adequate institutional frameworks are set up providing for "general or sector-specific benchmarks and safeguards".

In the context of public policies, the best-suited bodies seem to be – once again – parliaments rather than data protection authorities as a way of reprising what is ultimately a traditional role of parliaments in keeping governments accountable, though possibly in a different form. As much as populist rhetoric aims to cut parliaments out of their traditional intermediary role between the people and decision-making circles, implementing algorithms – in a way that is mindful of the principles included in the policy and legal frameworks illustrated above and the findings from the literature – offers a way to reinstate them in the policymaking cycle. If parliaments – along with individuals – were granted access to specific pieces of information, as described above (the functioning and justification of the algorithm, including the source code), and to a dedicated mechanism for challenging the outcomes, this could reinforce the democratic accountability of algorithmic decision-making and even fill in a longstanding gap in the field of policy evaluation.

The debate surrounding the assessment of the outcomes of policy decisions and the kind of metrics that would be best suited for this purpose in fact predates the rise of algorithmic decision-making. Especially after the advent of "new public management", emphasis has progressively grown on finding ways to quantify and measure government performances. Yet the relationship between performance and its quantification has proven a difficult one: different alternative metrics (such as targets, rankings and intelligence) have all been found to present shortcomings and only work effectively under specific circumstances.[91] The advent of digital technologies, and big data in particular, has offered new solutions and raised several new concerns in turn. Academic commentators have questioned the possibility of completing thorough evaluations by meaningfully connecting data collected through all of the phases of the policy cycle[92] and warned about the risk of an excessive focus on data in evaluation processes distracting from underlying substantive problems.[93] There are thus inherent difficulties in choosing the method, or even the relevant data, to enable a thorough assessment of government performance.

It is important to consider that for parliamentary scrutiny to contribute to the procedural legitimacy of policymaking decisions scrutiny itself ought to focus on the effects and impacts of such decisions rather than on their processes. This is only apparently a contradiction; substantive scrutiny on the outcomes can well be a fundamental aspect of procedural legitimacy. In fact, if black-box issues can challenge parliaments' ability to be effectively capable of performing thorough scrutiny of the technicalities of the process, it may then seem paradoxical that algorithms could contribute to the legitimacy of democratic governance while their own internal legitimacy and transparency are disputed. The two grounds, however, are to be kept separate.

In his strong critique of algorithmic decision-making and its underlying neoliberal mind-set, Ari Ezra Waldman has contended that there is a lack of effectiveness of any such mechanisms aimed at correcting the process of algorithmic decision-making based on the traditional equation of fair processes with fair results.[94] However, in the author's view, boosting transparency and other procedural safeguards cannot overcome the underlying characteristic of algorithmic decision-making as "agnostic about its sociopolitical and

---

[91] C Hood, "Public Management by Numbers as a Performance-Enhancing Drug: Two Hypotheses" (2012) 72 Public Administration Review S85–S92.

[92] DF Kettl, "Making Data Speak: Lessons for Using Numbers for Solving Public Policy Puzzles" (2016) 29(4) Governance 573–79.

[93] P White and RS Breckenridge, "Trade-Offs, Limitations, and Promises of Big Data in Social Science Research" (2014) 31(4) Review of Policy Research 331–38.

[94] AE Waldman, "Algorithmic Legitimacy" in W Barfield (ed.), *Cambridge Handbook of the Law of Algorithms* (Cambridge, Cambridge University Press 2020) p 107.

economic implications".[95] In the context of decisions made by AI, the "emphasis on efficiency ... undermines the effectiveness of procedures".[96] In turn, Waldman suggests that scrutiny of algorithmic decisions should focus on substantive outcomes and their impacts on fundamental rights rather than on their mere procedural fairness.

In our framework, however, the emphasis on results is a means to an end; focusing on the results helps to create a place for public discussion where discursive legitimacy is built by offering the opportunity for rational debate. A set of legal provisions granting parliaments (as well as individuals) access to specific pieces of information and the opportunity to challenge the outcomes of decisions against specific benchmarks, such as their impacts on fundamental rights, can streamline the process and offer a more robust underpinning to it (focusing on outcomes rather than data per se).

In line with the rationalist deliberative proceduralist approach illustrated above[97] and the focus on output legitimacy, empowering parliaments to scrutinise the outcomes of automated decisions can offer an opportunity for institutional scrutiny in order to minimise the potential negative impacts of technology. At the same time, algorithmic decision-making can offer a ground for parliamentary scrutiny and enable the public debate to refocus on rational discourses and their ability to produce rationally justified outcomes through fair procedures.

V. Concluding remarks

Algorithmic decision-making has started to be used in many policy areas in recent years, and it is now on trial in many others, such as welfare benefits and immigration management. Although critics point out the discriminatory effects, the human rights violations and the distorting effects on political communities as a whole of such systems, algorithmic decision-making seems to be "here to stay".[98] As is often the case with instruments of scientific progress, their mere existence makes the case for their use. Without dismissing those relevant and salient concerns, we focused on a complementary perspective to discuss the opportunities and the potentialities of algorithmic decision-making.

The model suggested in this article tested the hypothesis that algorithms can contribute to increasing democratic legitimacy at times of rampant populism, provided that their use takes place within a framework that maximises political equality and rational decision-making by enabling wider participation, consideration of diverse social issues and oversight of the decisions made.

We suggest that these conditions are met when citizens are able to: (1) understand the civic issues assigned to AI; (2) control the agenda of AI decision-making; and (3) evaluate and challenge the outcomes. For these conditions to be met, we suggest that frameworks such as a right to explanation, parliamentary oversight and opportunities to challenge the elements of an algorithm's processes, both judicially and in public debate, are put in place.

The efficiency as well as the representation of reality that an algorithmic process may be able to produce expose the fallacies or the "easy truth" of populist rhetoric. From such a viewpoint, algorithmic decision-making is not good because the output is intrinsically

---

[95] ibid, 116.

[96] ibid.

[97] F Peter, "Democratic Legitimacy and Proceduralist Social Epistemology" (2007) 6(3) Politics, Philosophy & Economics 329–53.

[98] As has been shown by the recent study "Understanding Algorithmic Decision-Making: Opportunities and Challenges" of the European Parliamentary Research Service EPRS | Scientific Foresight Unit (STOA), PE 624.261, March 2019, available at <https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU(2019)624261_EN.pdf>.

trustworthy. It is good as long as it is embedded in a democratic frame that enables both the represented and representatives to exercise choices and control over the decision-making process. In this way, algorithms can expose the populist rhetoric by being an instrument of knowledge and therefore a tool to read the reality and solve its problems.

Competing interests. The authors declare none.