



The epistemic impossibility of an artificial intelligence take-over of democracy

Daniel Innerarity¹

Received: 4 July 2022 / Accepted: 16 January 2023
© The Author(s) 2023

Abstract

Those who claim, whether with fear or with hope, that algorithmic governance can control politics or the whole political process or that artificial intelligence is capable of taking charge of or wrecking democracy, recognize that this is not yet possible with our current technological capabilities but that it could come about in the future if we had better quality data or more powerful computational tools. Those who fear or desire this algorithmic suppression of democracy assume that something similar will be possible someday and that it is only a question of technological progress. If that were the case, no limits would be insurmountable on principle. I want to challenge that conception with a limit that is less normative than epistemological; there are things that artificial intelligence cannot do, because it is *unable* to do them, not because it *should not* do them, and this is particularly apparent in politics, which is a peculiar decision-making realm. Machines and people take decisions in a very different fashion. Human beings are particularly gifted at one type of situation and very clumsy in others. The part of politics that is, strictly speaking, *political* is where this contrast and our greatest aptitude are most apparent. If that is the case, as I believe, then the possibility that democracy will one day be taken over by artificial intelligence is, as a fear or as a desire, manifestly exaggerated. The corresponding counterpart to this is: if the fear that democracy could disappear at the hands of artificial intelligence is not realistic, then we should not expect exorbitant benefits from it either. For epistemic reasons that I will explain, it does not seem likely that artificial intelligence is capable of taking over political logic.

Keywords Democracy · Politics · Algorithms · Decision-making

1 Introduction

The attempt to replace politics with activities that are similar (administration, knowledge, technology) is long-standing. The lure of neighbouring logics becomes more powerful while politics, in the way, it is habitually carried out, disappoints us with its failures. Turning to experts or technology, its valuation with economic categories or as a procurer of social order seem more promising than the old politics, which was ideological and elusive, risky and inexact. The temptation of leaving behind that time of fury and ideological imprecision is now driven by the technologies that accompany artificial intelligence, algorithmic decision-making, data analytics and automatization. This colonization begins with an epistemic confusion by virtue of which ways

of thinking that make perfect sense in one realm and are admired for their precision are extrapolated to other realms where they do nothing but distort reality. Algorithmic exactitude becomes unfairness when it abandons its instrumental nature, and its logic invades the spaces in which questions about the values and goals of society should be decided through political and democratic procedures.

Algorithmic governance attempts to reintroduce in democratic societies the secure, exact and incontrovertible criteria that experts represented in the past and that political pluralism resisted accepting. Then like now, democratic equilibrium is achieved by challenging expert knowledge with public criticism and appraising algorithms with criteria that include a complete vision of the political community in automated decision-making systems. It makes no sense for us to use algorithmic governance to commit the error that we resisted committing when faced with technocratic enticement.

That something is epistemically impossible means that it cannot happen for epistemic reasons; it can happen, but not

✉ Daniel Innerarity
dinner@ikerbasque.org

¹ Chair AI & Democracy, Ikerbasque Foundation for Science, European University Institute, Fiesole, Italy

because it is epistemically better. The announcements of an artificial overcoming of our political practices assume that one day artificial intelligence will be smarter than human intelligence, and this, at least from the point of view of political intelligence, seems to me not only a mistake but a misguided prediction. What can happen is that we may in fact entrust an AI with tasks for which it is not fully competent and that it performs them with less effectiveness or legitimacy. My thesis about impossibility refers not to the fact that such a departure cannot be made but that if it is made it will not be without regrettable losses. This “take-over” would only be possible as an inappropriate substitution. It could happen factually, but it would lead to a series of bad decisions about the central issues of politics, which I have synthesised here into decisions taken in situations of complexity, uncertainty, ambiguity and contingency.

2 Two ways of thinking and deciding

The best way to defend democratic politics is to fully identify its nature. The defence of political logic in the face of its colonization comes about through correctly identifying what makes politics so unusual in the face of those who attempt to replace it. Its first characterization is of an epistemological nature: politics is an activity that does not exercise linear and deductive reasoning; it manages particularly ambiguous situations; it must take decisions in situations of great uncertainty and contingency. This peculiarity characterizes it in the face of algorithmic logic, which demands clarity, objectivity and accuracy. This creates the true limit to treating political matters algorithmically, but it is also the foundation of democracy. The fact that we organize society democratically is not a normative concession but, more than anything, an intelligent consequence of the experience that fundamental issues regarding public life need to be decided through tools that are able to manage a high degree of uncertainty.

The idea that algorithmic procedures cannot yet take over all political decisions does not merely allude to a lack of development of the technology but to an inability of the technology itself, whatever its evolutionary state, to take on functions for which it is by its very nature not suited. It is not an exhortation or a moral prohibition for technology not to do certain things, but a realisation that certain issues cannot be adequately resolved by algorithmic procedures.

Democratic concerns (guaranteeing pluralism, minimizing imposition, impeding the concentration of power, encouraging the review of agreements) are, first and foremost, epistemological strategies. We do all of that to protect ourselves from error and to make correct responses more probable. While the algorithmization of political decisions is a great tool to confront certain forms of complexity (particularly those that require a lot of data or precision when

measuring preferences or impacts), it is inconvenient when it comes to other forms of complexity that stem from the ambiguous and contingent nature of political situations. The question about whether artificial intelligence is capable of taking on political prominence recalls previous reflections about whether the reductionism of calculation or reflexive pluralism is most aligned with politics in a democratic society and how to balance them (Koster 2021). We protect democracy to the same extent that we protect our own space to manage issues that are not resolved with deductive logic, issues that are ambiguous, uncertain or contingent, or when we resist the depoliticization that would arise from treating them with tools that are only pertinent when dealing with the problems of a completely different nature.

The first difference between humans and machines is found in the type of reasoning that leads to political decisions in contrast to the way algorithmic decisions work. Algorithms decide relatively easily when using “if... then” decision trees, where the relationship between an input and an output is clear. The knowledge found in artificial intelligence refers to an objective world that is both reducible to binary categories and calculable. For artificial intelligence, the world is a series of facts that can be logically deduced based on concrete computer-generated rules. From the epistemological point of view, the elaboration of information follows a model of logical calculation. The calculation performed by artificial intelligence does not include the context or specificity of input as a reference point; instead, it focuses on whether the logical operations are correct. In this way, the issues that it manages are separated from their specific contexts and considered as isolated phenomena, such that the logical coherence of the system is more important than the situation’s multiple possible interpretations (Bächle 2016, 22).

If human beings function reasonably well in ambiguous situations, it is because we take the context into consideration, which is something that algorithms or data analyses struggle to capture. Any time something models or automates, there is a simplification that ignores the context. The success of the Fordist mass production industry is due to standardization, but this type of organization is no longer ideal when there are activities that cannot be understood and organized through its classification scheme. In the current digital environment, one example of this is seen in the recommendations that respond to our spontaneous purchases and that do not identify what our long-term preferences might be (O’Neil 2016, 12), or more generally, the data obtained in one context (purchasing preferences) that are utilized in contexts that are quite different (a political decision). Human beings commit many errors, but we are in principle more capable of understanding different contexts than algorithms are.

3 Political ambiguity

Algorithms function with a 0/1 logic that is the complete opposite of ambiguity. When algorithms work with categories of likelihood, they end up resorting to binary responses, because those are the only kind that can be computerized with objective categories. Anything that is indistinct, undefined, not formalizable or imprecise is difficult to handle with binary logic. Algorithms are appropriate to navigate defined or quantifiable circumstances, but they are unable to question whether something makes sense or is valid. The only way to correct this bias is to promote a value table to consider other factors such as common sense, empathy, deviation or quality.

This is the main reason algorithmic mechanisms are unsuccessful when it comes to political decisions. Politics consists of taking decisions when there is no indisputable evidence, where the objectives are often contested, ambiguous or in need of being specified. Machines are of limited utility when dealing with problems that are not well structured, for which there is not much evidence, which are difficult to identify or to quantify, that are not predictable and are repetitive. Machines serve as an instrumental type of rationality, but they are generally unable to resolve complex problems where the issue actually revolves around the definition of the problem, rather than the implementation of a solution.

Algorithms can only resolve predetermined problems and those that can be translated into mathematical codes. They cannot identify a problem outside of their defined realm or determine what is or is not relevant in each situation, in other words, what has been called the “frame problem” (McCarthy/Hayes 1969), which is not a practical problem, but an epistemological one in the area of representation (Dennett 2006, 148). Artificial intelligence is not able to define every possible scenario as relevant or irrelevant. Here, we run into a very important limit, because if something defines politics—as opposed to administrative rationality, for example—it is its disposition to take on unforeseen realities, constellations that do not fully fit all the standardized predictions. Automatization has a logic that does not fit well with the changes that open societies must confront. Democracy is a way of organizing collective life that has an unlimited repertory of responses to new situations, that allows for the questioning of tradition, where there is nothing impertinent, that politicizes everything, that always tolerates criticism. All of that stems not from virtue but from the epistemological principle that no opinion exhausts the possible interpretations of reality.

Algorithmic rationality, on the other hand, does not tolerate the variety and ambiguity of social phenomena. Instead, it reduces them to a number from which it deduces structures

and causalities (Becker/Seubert 2020, 238). The desire to assure political freedom has always coexisted with the attempt to limit the field of the politically questionable. The exactitude that we expect from algorithms is a sophistication of the old project of transforming a continuous reality into a discrete representation, in such a way that, because of that quantification, human affairs become commensurate and calculable, thus increasing the possibilities of control (Mau 2017). Any irreflexive use of indicators and calculations is attempting to provide a version of reality that is more exact than the one provided by the appearances of perception. When this quantitative rationality overcomes its ancillary possibilities and is constructed as the functional equivalent of politics, it not only reduces the task of governing but also questions the democratic value of reflexivity. Politics has a dimension of calculation and measurement, but what most characterizes it in a democratic society is that it takes care of articulating the discussion about the meaning of the reality that we have quantified.

Algorithmic governance that is unconscious of its own limits commits from the beginning the error of thinking that social situations and political solutions can be categorized with a clarity that resolves any ambiguity. While that objective is very reasonable, it does not seem possible, and we run the risk of confusing a desire with reality and declaring the questioning of diagnostics and decisions as inappropriate. Democracy, in contrast, owes a great deal to the ambiguous nature of the realities in which we live. Perhaps, we would not need to organize the discussion, tolerate criticism or allow alternation if reality were incontrovertible. There is a background connection between the basic logic of human decisions and democratic logic. The attempt to govern with the greatest possible exactitude must guard against the temptation to declare the exactitude that is reached unquestionable and, especially, it should allow the intervention of other modalities of knowledge that are not actually governed by criteria of exactitude.

Politics has a lot to do with that human capacity to put diverse ways of knowing in play, but especially with the skills that allow us to manage ourselves in uncertain situations. Compared with machines, human beings reason and take decisions in an incredibly right way in situations of ambiguity, confusion and uncertainty. Let us think about the classic example of a child’s ability to recognize a cat after seeing only two images, compared to the enormous number of images that a machine needs to achieve the same skill. If we are characterized by anything, it is that we are not as inept as one might expect in situations with minimal or poor quality information. Human beings have to take many of our decisions in situations of ambiguity and uncertainty. If this is uncomfortable, we should keep in mind that this is also what leads us to pluralism, so a forced reduction in this

diversity would imply limiting the diversity of opinions and values that characterize a democratic society (Herzog 2021).

It is true that humans have many characteristic biases and there is a large literature on this, but what this literature reveals is that these biases are very different from those of algorithmic procedures and that there is no bias that does not have a positive reverse side. Many biases are often due to a heuristic strategy. That humans decide relatively well with scarce information, for example, is because throughout history, we have had to survive in environments of scarce information. Given the current information overload, this ability does not exempt us from taking advantage of technologies that can process enormous amounts of information. The evolutionary advantage of humans is that we can benefit from data processing technologies without having to sacrifice our special ability to decide when data is scarce because, in fact, no matter how much data we have, there will always—especially in politics—be situations where the decision is not entirely clear.

4 Political contingency

Algorithmic procedures operate with a degree of exactitude that contrasts with the management of uncertainty that is inherent to politics. Much of the relevance or lack thereof of political decisions stems not from technical, logical or moral criteria, but from political specificity that has to do with contextual factors, criteria of opportunity, equilibrium and transactions. Political decisions are rarely based on binary categories (good versus evil, truth versus falsity) or, if one prefers: politics is defined by taking decisions when the most resounding criteria have been considered and one still has to decide the most important part. Algorithmic procedures do reduce complexity, but it does not seem that they can fully suppress the uncertainty under which many political decisions are adopted, thus failing to keep their promise of objectivity. This comes as a relief to those who feared the end of political pluralism.

Democracy is organized uncertainty (Przeworski 1991, 13). Politics in a democratic society means managing the contingency of things, which could always be different. Being conscious of this contingency allows us to see the political circumstances as the product of historical processes and not as a destiny, the results of human, configurable and modifiable actions. For that reason, there is no political life without participation, competition or conflict. Tocqueville spoke of the permanent uneasiness of democracy (1835, 219), while Luhmann addressed it as a continuous irritation (1987, 129). This explains its openness, indeterminacy and discontinuity.

The fact that politics is an especially contingent activity means that it refers to a series of topics that cannot be

resolved through an influx of information and knowledge that would have made the decision unnecessary; politics appears when we know everything that can be known, after having discussed everything possible, demanded the authorized opinion of the experts, having resorted to all the techniques of datification available... and what needs to be done is still not completely clear. Therefore, a decision is taken, the best informed decision possible, but still a decision, in other words, not overwhelming, but refutable and changeable. The procedures of artificial intelligence cannot exonerate us from that decision. Politics is the place where, in spite of all the sophistication of the calculations, we find ourselves obliged to take a decision that is not preceded by overpowering reasons nor carried out by infallible technologies. It is true that all the processes of technification tend to model or automatize so that the “human factor” is less relevant; this objective does not lack justification. Humans have never stopped dreaming about “the perfect technology of justice” (Lessing 1999), but neither have we stopped feeling the burden of our decisions bearing the ultimate responsibility for making society just.

Contemporary societies require a great deployment of knowledge, more accurate tools of analysis, better technologies and more effective management. It is very debatable that algorithms and automatization will be able to take control of the *entire* decision-making process but, even if that were the case, in a democracy, we cannot establish whether decisions are correct without asserting a political logic whose legitimacy ultimately resides in our free political will. A democracy usually produces better decisions than its alternatives, but it does not owe its ultimate legitimacy to the goodness of its decisions but to the popular authorization that underpins those decisions. Although my case for democracy is normative, I share the thesis that ultimately the rationale for democracy is not a function of its outcomes, which is perfectly compatible with the fact that, for epistemic reasons, it often performs better than its alternative models. That the justification for democracy is not instrumental is not incompatible with the consideration of its frequent instrumental superiority.

There is an additional argument in favour of an epistemic argument but in the opposite direction and which cuts off our eventual substitution by objective machines. In some ways, it could be affirmed that democracy does not exist because we know what needs to be done or do what is correct, but neither is it in spite of not knowing. Democracy exists because of not knowing. The inevitability of deciding is the definitive justification for democracy being a form of government in which laypeople, rather than the experts, have the last word. It does not seem there is currently a mechanism, whether analogue or digital, that completely frees us from this necessity to decide.

5 Conclusion: democratic uncertainty

There are times when a piece of evidence brings a political controversy to a close, but what is most frequent is that, even when we have evidence that claims to be scientific, it is confronted with other supposed evidence. Similarly, different epistemic modalities are asserted in political debates. Calling for a neutrality of procedures or objectivity in authorized expert opinions does not make the moment of political decision-making unnecessary, because politics is almost never preceded by overwhelming reasons. The neutrality that algorithmic procedures seem to offer can be very tempting in fastidiously ideologised environments. Understanding this seduction, we should remember that neither the objectivity they promise is so indisputable nor subjectivity has to be arbitrary.

In addition to that, there is a connection between the epistemological environment that is described with the categories of both ambiguity and contingency and our democratic institutions. We can lament not having greater evidence and accuracy, but perhaps, we should consider that democracy exists precisely because our knowledge is so limited, and we are so prone to error. If we think about it this way, uncertainty can be an unexpected factor of democratization. Precisely where our knowledge is incomplete, we have greater need for institutions and procedures that favour reflection, debate, criticism, independent advice, reasoned argumentation and the competition of ideas and visions (Majone 1989). Our democratic institutions are not an exhibition of how much we know but a recognition of our ignorance.

A decision is political when, even following a long process of deliberation and preceded by all the objective analyses within our reach, the ultimate option is still not fully clear. Those who do not understand this will view politics as arbitrary and opportunistic and will be easily seduced by any promise of exactitude formulated by the experts, the machines or algorithms, but they will not have understood what politics does, especially what politics does in a democratic society. A human world needs to be a negotiable world.

Acknowledgements Curmudgeon Corner Curmudgeon Corner is a short opinionated column on trends in technology, arts, science and society, commenting on issues of concern to the research community and wider society. Whilst the drive for super-human intelligence promotes potential benefits to wider society, it also raises deep concerns of existential risk, thereby highlighting the need for an ongoing conversation between technology and society. At the core of Curmudgeon concern is the question: What is it to be human in the age of the AI machine? -Editor.

Funding Open access funding provided by European University Institute - Fiesole within the CRUI-CARE Agreement.

Data availability The empirical data referred to in this paper are available on request from the corresponding author, but are not public due to privacy restrictions.

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

- Bächle TC (2016) *Digitales Wissen. Daten und Überwachung zur Einführung*. Junius, Hamburg
- Becker C, Seubert, S (2020) "Die Selbstgefährdung der Autonomie. Eckpunkte einer Kritischen Theorie der Privatheit im digitalen Zeitalter". In: Kruse JP, Müller-Mall S (eds) *Digitale Transformationen der Öffentlichkeit*, Velbrück Wissenschaft, Weilerswist
- Dennett DC (2006) *Cognitive wheels. The frame problem of AI*. In: Ford KM, Glymour J, Hayes P (eds) *Thinking About Android Epistemology*. MIT Press, pp 147–169
- Herzog L (2021) *Algorithmisches Entscheiden, Ambiguitätstoleranz und die Frage nach dem Sinn*. *Deut Z Philos* 69(2):197–213
- Koster A-K (2021) *Das Ende der Politischen? Demokratische Politik und Künstliche Intelligenz*. *Zeitschrift Für Politikwissenschaft*. <https://doi.org/10.1007/s41358-021-00280-5>
- Lessing L (1999) *Code and other Laws of Cyberspace*. Basic Books, New York
- Luhmann N (1987) *Soziologische Aufklärung 4. VS Verlag für Sozialwissenschaften, Opladen, Beiträge zur funktionalen Differenzierung der Gesellschaft*
- Majone G (1989) *Evidence, Argument, & Persuasion in the Policy Process*. Yale University Press
- Mau S (2017) *Das metrische Wir: Über die Quantifizierung des Sozialen*. Suhrkamp, Berlin
- McCarthy J, Hayes PJ (1969) *Some Philosophical Problems from the Standpoint of Artificial Intelligence*. In: Meltzer B, Michie D (eds) *Machine Intelligence*, vol 4. Edinburgh University Press, pp 463–502
- O'Neil C (2016) *Weapons of Math Destruction. How Big Data Increases Inequality and threatens Democracy*. New York, Crown Publishing Group
- Przeworski A (1991) *Democracy and the market. Political and economic reforms in Eastern Europe and Latin America*. Cambridge University Press
- Tocqueville A (1835) *De la démocratie en Amérique*, II, 3. Michael Lévy. Libraires Éditeurs, Paris

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