Machine Learning algorithms and public decision-making
A conceptual overview

van Krimpen, F.J.; de Bruijn, J.A.; Arnaboldi, M. (Michela)

DOI
10.4324/9781003295945-12

Publication date
2023

Document Version
Final published version

Published in
The Routledge Handbook of Public Sector Accounting

9

MACHINE LEARNING ALGORITHMS AND PUBLIC DECISION-MAKING

A conceptual overview

Floris van Krimpen, Hans de Bruijn and Michela Arnaboldi

Introduction

This chapter is about the use of machine learning (ML) algorithms in public sector decision-making. Much has been published in recent years on the role ML can play in decision-making processes. Practical examples are decision-making processes to identify tax fraud, to support decision-making regarding early release from prison (Berk, 2017), and to prioritize supervision activities (Lorenz, Erp, & Meijer, 2022). Public decision-making has several special characteristics, which can have a significant impact on the effectiveness or legitimacy of algorithm usage.

Why public decision-making matters. First, decisions by governments are often high-impact decisions for citizens and organizations. A licence may or may not be granted or an organization may or may not be faced with an enforcement agency. Second, citizens and individuals often have no exit option (they cannot divert to other parties) because these decisions are based on authority, which only governments have. Third, because of this dominance of governments, public decision-making is bound by particular procedural principles, including transparency, explainability and accountability. Algorithmic decision-making can be at odds with these principles – for example, it is not always sufficiently transparent about which data led to which decision (Burrell, 2016).

The multiple layers of public decision-making. Public decision-making can involve at least two levels: 1) An individual level. Government decisions can relate to individual citizens or organizations. A government may impose a tax charge, provide a benefit or grant a permit. Decision-making on these matters, especially when large numbers of decisions have to be made, is often data-intensive and algorithms play an important role. 2) A collective level. This involves strategic decisions or policy decisions. An enforcement agency has limited capacity and thus has to make strategic choices about which potential lawbreakers to target. Algorithmic decision-making can play an important role in these choices. The same applies to policy-making processes. The development of policies is data-intensive and so algorithms can be an important vehicle for effective decision-making.

The ambiguities of public decision-making. Some public decisions have an unambiguous structure: if one or more conditions are met, a particular decision follows. Decision-making is self-executing. For instance, anyone breaking the speed limit will be fined a certain amount. But many decisions – both individual and collective – are not self-executing. They are decisions that come with ambiguity: the data and information used can be questioned, and normative questions about fairness and equality, for example, often come into play. When a law enforcement agency decides to

 DOI: 10.4324/9781003295945-12

conduct additional scrutiny of certain companies, the question can be asked whether this decision is based on the right data and whether the decision is fair to these companies. When a welfare benefit is denied, the question can also be asked whether the right information was available and whether the denial was proportionate. These ambiguous issues are also known as "wicked problems" (Head, 2019; Hisschemöller & Hoppe, 1995; Rittel & Webber, 1973). Public decisions can thus have the above characteristics – high impact, no exit option, special procedure requirements – and also involve wicked issues. This makes the question of the effectiveness and legitimacy of algorithm use even more urgent.

The importance of public decision-making characteristics. The literature related to ML algorithms in the public sector is growing. Some scholars have started to study algorithms for public decision-making (Hartmann & Wenzelburger, 2021; van der Voort, Klievink, Arnaboldi, & Meijer, 2019) and have made comments about legitimacy issues with the introduction of ML algorithms (König & Wenzelburger, 2021; Veale & Brass, 2019), but the current literature lacks fails to recognize the importance of the characteristics of public decision-making and to discuss how these peculiarities need special attention when ML algorithms are used in decision-making processes. Furthermore, the inherent tension that exists between the wickedness of public decision-making and the use of ML algorithms is underexplored.

This chapter focuses on the question of *what the wicked nature of public decisions means for the use of ML algorithms***.** To answer this question we carried out a narrative literature review of the state of the art of ML algorithms in the public sector. We squared this review with literature that discusses the specificities of public sector decision-making discussed in section 2. This literature includes studies such as those by Cohen, March and Olsen (1972). De Bruijn and Ten Heuvelhof (2018), and Teisman, 2000. This chapter aims to synthesize the knowledge related to our main question and bring additional theoretical concepts to the world of public sector decision-making with ML algorithms.

The remainder of this chapter is structured as follows. In the next section, we elaborate our theoretical framework, pointing out characteristics important for public decision-making. The third section illustrates our methodology. Then, the results of our review are presented. The fifth section discusses the meaning of those results in light of current literature. Finally, the sixth section concludes the chapter.

Machine learning and public decision-making: type 1 and type 2 decisions

ML is a technology that has gained a lot of popularity in the past decade and is being discussed as an instrument that will have a profound effect on the economy and workforce (Brynjolfsson & Mitchell, 2017). ML is defined as "a core branch of AI that aims to give computers the ability to learn without being explicitly programmed" (Samuel, 2000). Essentially, ML is the capability of software or a machine to improve the performance of tasks through exposure to data and experience (Luxton, 2016). A typical ML model first learns the knowledge from the data it is exposed to and then applies that knowledge, for example, to make predictions. While the adoption of ML-based technologies is already becoming widespread in the commercial sector, the public sector is lagging behind in comparison (Desouza, Dawson, & Chenok, 2020). However, attention has been growing and investment in technologies that are based on ML has been one of the most important strategies of the public sector in several countries around the world in recent years (Sousa, Melo, Bermejo, Farias, & Gomes, 2019).

Regardless of the type of application, ML implies the definition of a decision-making process concerning a problem. Even in simple applications, such as automated responses, ML applications "decide" how to respond in the face of specific questions posed by the users. The level of criticality of the decision-making process is, however, higher or lower depending on the type of problem faced.

The literature on public decision-making distinguishes between two types of problems, which we call type I and type II problems here. We hereby briefly explain the distinction between the two problems to set the background for the literature review. An overview of these distinctions related to decision-making is provided in Table 9.1. The premise of the research reported in this chapter is that ML-based decision-making might lend itself mainly to type I decisions. However, decision-making often involves type II problems– and the question is what this means for valuing algorithmic decision-making. There are four fundamental differences between the two when making decisions (De Bruijn & Ten Heuvelhof, 2018).

First, type I decision-making involves structured problems. A structured problem can be unambiguously defined and therefore often has one best solution. Type II decision-making involves unstructured or “wicked” problems. The problem is ambiguous and can sometimes be defined in completely different ways – or even be seen as a non-problem (Head, 2019; Hisschemöller & Hoppe, 1995).

Second, type I decision-making involves a single problem owner. One-actor decision-making means that this one actor can decide how to define a problem and what the right solution is. Type II decision-making is multi-actor decision-making. Multiple actors are involved in this decision-making, who have different values and interests. Moreover, when a problem is wicked, there is a lot of room for these actors to come up with their own interpretation and preferred solution (Adam & Kriesi, 2007; Klijn & Koppenjan, 2000).

Third, in type I decision-making, problems are often stable, meaning that there is a problem definition, which does not change. When there are wicked problems in a multi-actor context, problems are often dynamic – type II decision-making. When power relations between actors shift, the problem definition may shift – and hence the desired solution (De Bruijn & Ten Heuvelhof, 2018).

Fourth, type I decision-making often proceeds linearly, and type II decision-making non-linearly. Decision-making is often a power struggle between actors who adhere to different problems and solutions. Such a power struggle involves iterations, accelerations and delays, and redefinition of problems and solutions. Decision-making is a messy process – the literature talks about “garbage can decision-making”, “windows of opportunity”, and governance by randomness (Cohen et al., 1972; Kingdon, 2011; Teisman, 2000).

The essential difference between type I and type II decision-making (see Table 9.1) is the structured versus the unstructured or wicked nature of problems. When a problem is wicked, a multi-actor context becomes problematic and decision-making will be much more dynamic. We, therefore, elaborate on the concept of wicked problems here to point out the criticality concerning ML applications.

There are several definitions of wickedness in the literature (Alford & Head, 2017; Head, 2019; Rittel & Webber, 1973) that all have their specific perspectives or peculiarities. Most of these definitions have in common that they discuss ambiguity related to facts and norms. Therefore, we use the definition developed by Hisschemöller and Hoppe (1995) because it captures the essence of what wicked problems entail, namely, that they are ambiguous. Table 9.1 has two axes.

1. Consensus or dissensus about norms. The question of whether something is a problem or not is partly determined by normative views. There may be consensus or dissensus about these normative views. Take abortion, for example – there are completely different normative views on abortion,

Table 9.1 Differences between type I and type II decision-making

Type I decision-making	*Type II decision-making*
Involves structured problems	Involves wicked problems
Involves a single problem-owner	Involves multi-actor decision-making
Deals with stable problems	Deals with dynamic problems
Proceeds linearly	Proceeds non-linearly

roughly speaking: the pro-choice and the pro-life view. An example of normative agreement: the value of "equal opportunities" is generally endorsed by all – there is consensus that this is a key value in democratic societies. Consensus or dissensus may also concern the weighing of a set of different values. Usually, in public sector issues, different values have to be weighed (e.g. fairness, transparency, affordability, accessibility and cost-effectiveness). The parties involved may or may not agree on the right trade-off between these values. This issue becomes a crucial point in ML development, where criteria for decision-making and variables of inclusion or exclusion need to be set.

2. The unambiguity or ambiguity of facts. Do the facts allow only one conclusion? Or are they multi-interpretable – and therefore more ambiguous? If someone drives through a red light and is flashed, there is an unambiguous set of facts: it is clear which car drove through which red light, and when. The constellation of facts might also be more ambiguous. What were the consequences of a tax cut? This might be open to debate because there are many intervening variables. Some will say that it promoted inequality, and others will say that it led to economic growth and more opportunities for everyone. In this case, the impact for ML applications can also appear during the operational running of the application, when a new input is arriving or because a new "ambiguous category" of user emerges.

Table 9.2 summirizes the two axes, pointing out four quadrants that are discussed below.

In quadrant I, we find structured problems, which in most cases have one right answer (De Bruijn & Ten Heuvelhof, 2018). If someone drives through a red light (problem), this person will be fined (solution). In quadrant II, the main disagreement is normative – we know what an abortion factually means, but we disagree about the normative question. In quadrant III, the situation is reversed: normatively speaking, there is consensus, but the disagreement is about factual causes and consequences.

The quadrant on the lower right (quadrant IV) contains wicked problems – the essence of which is that they have no unambiguous problem definition and no unambiguous solution (Hisschemöller & Hoppe, 1995). Every problem definition and solution can be questioned. For example, suppose a TSO (an electricity transmission system operator) wants to roll out high-voltage cables. Different stakeholders may have different views on the facts: on the necessity of the cables, given the number of customers; on the effects on health; on the impact or the market value of houses close to the cables. Furthermore, there are different values at stake – economy, ecology, health – that require a trade-off. Different stakeholders will have different opinions on the right trade-off.

Methodology

To investigate and synthesize the knowledge related to the question of what the wicked nature of public decisions means for the use of algorithms and to bring additional theoretical concepts to the world of public decision-making with ML algorithms, we carried out a narrative literature review (Baumeister & Leary, 1997; Webster & Watson, 2002), which was articulated in these phases: initial

Table 9.2 Summary of wicked problems

	Facts unambiguous	*Facts ambiguous*
Consensus about norms	I Running a red light	III Equal opportunities
Dissensus about norms	II Abortion	IV Roll-out of overhead power lines

search and selection of papers; extension of papers portfolio with a snowball approach (Wohlin, 2014); qualitative paper analysis; discussion among authors and synthesis.

For the selection of papers, we used Scopus, focusing on peer-reviewed journals, book chapters and conference papers; written in English; within the social science area. This latter choice is appropriate for the objective of the paper and to have an interdisciplinary range, as this area allows. A systematic search for eligible studies was not carried out, on purpose, given the blurred boundary of the topic under study. However, we comprehensively covered the topic in question.

Search terms included three areas of terms: (i) "machine learning", "algorithms", "big data"; (ii) "public sector", "public administration"; (iii) "decision-making", "decision". These terms were connected with the Boolean elements "AND" and "OR" iteratively to maximise an ample but focused search. The choice to include the third area (iii) was made to make sure the set of literature was focused enough to obtain relevant results.

After the first iteration abstracts were read by us, the authors, we started another search with the snowball method (Wohlin, 2014). The selection terms were integrated; we also integrated papers appearing in the reference list of the first-round documents (after the screening of the title and abstract); we followed authors that emerged as relevant for the topic in the first round. The inclusion within the final set was also based on the criteria used for the inclusion in the starting set. However, in this step in the process, the abstract, introduction and conclusions of the papers were scanned to decide on inclusion or exclusion. Further, exemptions were made for more technically oriented papers cited by already included papers. These technically oriented papers serve to better describe the concept of ML. Also, exemptions were made for literature that describes core concepts from public administration literature helpful to further clarify challenges. Finally, we also searched Google for any other material related to the topic under investigation (through the final set of keywords); they included policy papers and practitioner reports. This was key to further clarifying concepts with empirical examples. The focused selection led to 28 articles.

Figure 9.1 reports the chronological spread. In particular, it indicates that especially from 2016 onwards, attention has been growing for ML algorithms in the public sector. Based on these statistics, it might be argued that after a peak of interest in 2019, attention has declined. However, this could be because, especially in the early phases, there has been attention to the broader topics of algorithms or AI in the public sector instead of ML specifically.

Data analysis was articulated in two steps. Firstly, we carried out a concept-centric analysis, guided by the elements highlighted in the previous section. The contributions were first classified according to their main objectives. The main concepts that appeared within the individual contributions were

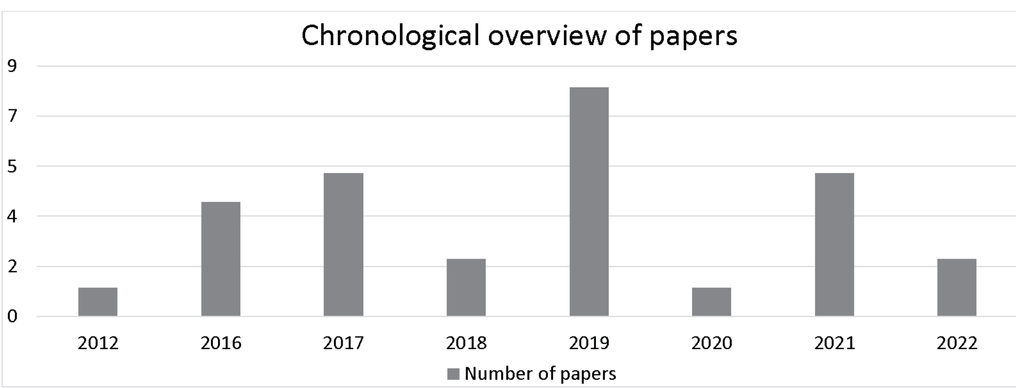

Figure 9.1 Chronological overview of papers.

then identified, within the final boundary of the aim of this research. Next, a cross-sectional analysis of the contributions was done to identify recurring and common concepts.

The second step of the data analysis was carried out with a theory-building aim. We entered more specifically into the topic under investigation guided by the framework provided in section 2 concerning the wicked context of public decision-making. In this step, researchers need several confrontations in pairs and altogether, given that the wicked nature was sometimes present although not underlined with this term.

The final discussion led to an overview that distinguishes between different perspectives related to decision-making and ML algorithms in the wicked context of public decision-making. Firstly, decision-making with algorithms, and secondly, decision-making about algorithms, divided between decision-making about the development of the ML algorithm and the subsequent use of an ML algorithm.

Results

In this section, we discuss the results of the narrative literature review. The issue of decision-making, wicked problems and ML algorithms is a multilayered phenomenon, where a distinction can be made between:

(1) decision-making with algorithms: an algorithm is used to make a decision;
(2) decision-making about algorithms: decision-making about (2a) the development of an algorithm and (2b) the subsequent use of an algorithm.

Decision-making with ML algorithms

Most benefits and challenges are described in more generic literature that discusses at both relatively higher and broader levels, such as Katzenbach and Ulbricht (2019); Wirtz, Weyerer, and Geyer (2019); and Zuiderwijk, Chen, and Salem (2021). However, also more focused technical papers such as Alexopoulos et al. (2019) or Domingos (2012) and papers from legally oriented scholars such as Barocas and Selbst (2016); Coglianese and Lehr (2017, 2018); Kroll et al. (2017); and Liu, Lin, and Chen (2019) are found within this analysis.

ML algorithms result in better decision-making

Here we discuss the potential benefits. Within the reviewed literature, the benefits mostly mentioned are those of efficiency and accuracy; learning processes; objectivity and innovation; and trust.

Efficiency and accuracy. Two of the most mentioned benefits are efficiency and accuracy. Many indicate that AI, in our case ML, can make public decision-making more accurate (Alexopoulos et al., 2019; Eggers, Schatsky, & Viechnicki, 2017) and more efficient (Alexopoulos et al., 2019; Domingos, 2012; Mehr, 2017; Zuiderwijk, Chen, & Salem, 2021). Suppose an enforcement agency, with its limited capacity, has to select the inspectees they pay attention to. This is a continuous and dynamic process. An algorithm can help select the inspectees because it can prioritize cases based on a risk score. In this example, the work becomes more accurate because the ML algorithm helps select those cases that are most interesting. In essence, the algorithm helps strategically choose these cases. Also, the work becomes more efficient because the ML algorithm saves time and resources in the process of determining the cases to inspect. This prioritization based on risk score is also one of the benefits mentioned by Zuiderwijk et al. (2021) as “risk identification and monitoring benefits”.

Learning processes. ML systems contribute to organizational learning processes. ML systems can continuously improve and, in addition, enable team-based and mixed-initiative learning. The continuous improvement relates to the fact that ML systems can self-improve by being fed with new historical data (Alexopoulos et al., 2019). Team-based and mixed-initiative learning entails that ML methods now have the capability of working together with humans. Machines and humans can learn together in a mixed way. For example, a machine can extract information from data sets, while humans suggest hypotheses to be tested based on the extracted data sets (Alexopoulos et al., 2019). Thus, learning occurs in a joint effort between the human and the machine. Often, such benefits are a means to an end of becoming more accurate and efficient. However, the notion of public decision-making that can be continuously improved and in which new learning mechanisms can occur between algorithms and humans should be highlighted.

Objectivity and innovation. The use of ML can also have the benefit of the appearance of objectivity (Lorenz et al., 2022). An example here is the use of "robot judges". Human judges can be biased and different judges can, in equal situations, give different rulings. A judge may be tired – and this may affect a ruling (Crootof et al., 2019). The robot judge can contribute to objectifying judicial decisions. Further, this might also give the appearance of innovativeness (Kuziemski & Misuraca, 2020). The logic goes that a government that relies on innovative, technical tools becomes more objective because of the nature of these tools.

Trust. Objectivity and innovativeness, in turn, might increase satisfaction and trust in the government (Dwivedi et al., 2021). Governments are bound by being objective. Citizens expect governments to treat them equally. The premise is that an objective government treats people equally. Thus, the appearance of objectivity and innovativeness contributes to the legitimacy of the government.

Related to benefits, studies points out that ML has the potential to increase the accuracy and efficiency of decision-making. Further, ML algorithms give the possibility to continuously improve public decision-making by facilitating team-based learning between algorithms and humans. Lastly, by the appearance of innovativeness and objectivity, the legitimacy of government and trust in government can be increased.

ML algorithms result in poorer decision-making

There is also much evidence in the literature that the use of algorithms has significant harmful effects. Dwivedi et al. (2021) mention that challenges range from ethical issues, such as the possibility of discrimination, to matters at the level of technology and technology implementation or data challenges. This section discusses the relevant challenges separately. Multiple scholars also mention legally related issues such as Dwivedi et al. (2021), Wirtz et al. (2019),and Zuiderwijk et al. (2021). Our analysis as well includes this legal perspective, also through elaborating on the work of scholars who discuss particular challenges of ML concerning the rule of law, such as Barocas and Selbst (2016); Coglianese and Lehr (2017, 2018); Kroll et al. (2017); and Liu, Lin, and Chen (2019).

Discrimination. Generally, ethical challenges are mentioned by all the studies we reviewed, and discrimination is often the most critical issue. Discrimination refers to making a distinction between people or groups of people based on the group, class or category to which those people belong. Thus, the challenge of discrimination relates to preventing inequality and unfairness (Thierer, Castillo, & Russell, 2017; Wirtz, Weyerer, & Geyer, 2019). Inherent to the nature of ML, if the data the model is trained on is discriminatory, then the model will become discriminatory. The prejudice in the data becomes "baked into the model"(Barocas & Selbst, 2016). A straightforward example is biased data that has racial prejudices. Consequently, the model might be discriminatory.

Lack of transparency. Transparency is also mentioned from multiple perspectives. The argument is that the opacity of ML systems is a threat to the legitimacy of public decision-making processes (Danaher, 2016). The question becomes whether decisions made or supported by a potentially opaque system or algorithm are accepted as rightful by those affected by the decisions. Furthermore, it raises the issue of whether the current rules of accountability are "fit for purpose". That is, can they cope with new requirements due to ML? Responses from the legal literature provide some insight concerning these questions. Coglianese and Lehr (2018) try to answer the question whether governing utilizing algorithmic systems can be squared with legal principles of governmental transparency. They conclude that the relative inscrutability of ML algorithms does not pose a legal barrier to their responsible use by governmental authorities. Also in an earlier study, Coglianese and Lehr (2017) state that if ML algorithms are properly understood, their use by governmental agencies can fit within legal parameters. However, the opposite response comes from Liu et al., 2019. From an examination of the *State vs. Loomis* case in the US, they conclude that the algorithmization of government functions poses a threat to, among other things, transparency. They show that in this particular case, there was a lack of understanding regarding the effects of the implementation of an ML tool and consequently, the necessity to open up the legal black box was not present for the defendants. Naturally, also for practitioners, transparency is mentioned as a challenge. ML algorithms are often adopted as a black box. The algorithms are opaque and it is unclear how the algorithms work, while they are being used for making socially consequential predictions (Burrell, 2016).

Lack of accountability. Accountability is a key legal principle in a democracy. An argument often made is that transparency is a way of making algorithms accountable. Legal scholars like Kroll et al. (2017) mention that they want to challenge the dominant position in the legal literature. They indicate that solving the challenge of accountability by making algorithms transparent is undesirable and not always possible. Rather, we can use technological tools, more specifically computational methods, to adhere to legal standards of accountability. They mention the methods of software verification, cryptographic commitments, zero-knowledge proofs and fair random choices. These methods can guarantee that decisions are made in an accountable way, without the necessity of a fully transparent model. Concerning accountability, Liu et al. (2019) also have their say, noting that to solve this challenge, we need to treat seriously the black box problem of these algorithms. The black box symbolizes the idea of not being able to look inside the ML algorithm (Burrell, 2016).

Breaching privacy. Is the data that is being used actually in line with privacy regulations? Concerning this point, Wirtz et al. (2019) mention that this is particularly about whether data from individuals is collected and processed with consent from these individuals and in line with existing regulations.

Value conflicts. A last challenge, which is particularly relevant but mentioned only scarcely, is that of value conflicts. In the case of ML algorithms, conflicting challenges and benefits have to be weighed against each other. Choices have to be made that are value-laden (Veale & Brass, 2019). Take, for example, the values of accuracy and data privacy (Arnaboldi, de Bruijn, Steccolini, & Van der Voort, 2022). The use of personal data can contribute to the accuracy of public decision-making, but there is also the value of privacy that needs to be safeguarded. This value conflict requires a trade-off, which is partly contextual: different contexts may result in different trade-offs (De Bruijn, 2021). Other values might be part of the trade-off, for example, safety, sustainability and equity.

Some of the challenges, such as discrimination (or bias), are related to the data that is the basis of the ML algorithms. Data is one of the main building blocks of ML algorithms. A lack of quality and quantity of input data (Dwivedi et al., 2021) can affect the quality of public decision-making. The algorithm is only as smart as the data from which it learns (Wirtz et al., 2019). A way to deal with discrimination might be to have more and better data that is used to train better algorithms. However, with the increasing complexity of ML algorithms, accountability and transparency issues increase. This illustrates that challenges are not necessarily solved by better data. Trade-offs are involved in dealing with such challenges. To complicate matters even more, these trade-offs are also dynamic, meaning that they can change over time.

There is overwhelming evidence that ML algorithms for public decision-making raise a host of issues that need to be addressed. As illustrated, the notion should also be taken into account that many of the challenges are closely connected and should be considered interdependent (Wirtz et al., 2019).

Decision-making about ML algorithms

As stated, the issue of algorithms and decision-making is a multilayered phenomenon. Algorithms can be used to make decisions about wicked problems. Before this, decisions were made on the development and use of algorithms.

The development and use of an algorithm is a process in which multiple stakeholders are involved (Lorenz et al., 2022; van der Voort et al., 2019; Zweig, Wenzelburger, & Krafft, 2018). Zweig et al. (2018) can be described as an eight-step process: 1) algorithm development, 2) algorithm implementation, 3) algorithm selection, 4) data collection, 5) data selection, 6) design and training of the system, 7) the embedding of the system in the societal process, and 8) feedback. We might be tempted to pay particular attention to steps 7 and 8 when the ML algorithm is used for decision-making – but these steps are preceded by a whole process of development of the ML algorithm. To understand the use in the final steps (7 and 8), the entire process is relevant. At every step in that process, different stakeholders are involved. For example, a data scientist has to make a decision about the data to include and the algorithm to use. The data has been labelled by someone with domain knowledge. Then, a model will be trained on the labelled data. The type of model has to be chosen by someone. It might be that a different stakeholder has to determine how the ML tool is embedded into the actual daily practice of decision-makers. With a finished model, a person responsible for decision-making processes in the organization has to decide whether and how the model will be integrated into an actual decision-making process that also involves many actors with different perspectives.

The development of ML algorithms is the result of countless interactions like these and other interactions between actors. Because of the many interactions and many ways in which a final ML algorithm can be developed, decision-making processes related to algorithms are very dynamic. The "right way" to use an algorithm is a wicked problem; actors will disagree about the right way; and actors might gain or lose power, which can make the decision-making process a very dynamic endeavour.

On top of this, algorithms are not stand-alone objects but are relational. They belong to a larger algorithmic system "which involves an intricate, dynamic arrangement of people and code, where multiple insiders cooperate and work on the design and implementation of the algorithmic logic" (Janssen & Kuk, 2016, p. 374). The more complex this algorithmic system, the less predictable algorithm development will be. Algorithms are developed in organizational contexts with different intentions and actual effects (Meijer, Lorenz, & Wessels, 2021). In one organization, there might be a higher level of digital discretion than in the other (Young, Bullock, & Lecy, 2019). Digital discretion refers to the distinction between human and machine agents and the relationship between these two. Numerous other organizational variables might have an impact on the use and impact of ML algorithms – for example, the degree of operational autonomy, digital competence or organizational checks and balances. Thus, algorithms and their effects should be understood as an organizational outcome rather than solely based on the technological features of algorithms (Meijer et al., 2021).

A story of amplification

So far so good. Now let's bring in the element of wickedness. Decision-making using algorithms can lead to better and more poor decision-making. The use of ML algorithms for type II decision-making amplifies many of the shortcomings already prevalent within ML algorithms and public decision-making.

Amplification of bias. Algorithms are often biased, and bias might become "baked into the model" (Baro-cas & Selbst, 2016). Now relate this to the concept of wickedness. Decision-making processes related to wicked problems require the weighing of different values. These processes are multi-actor processes and can be power games par excellence, in which the values of the most powerful actors will dominate. Suppose a powerful actor values a rigorous approach towards fraud related to governmental benefits and wants to use an ML system to investigate potential fraud. This rigorous approach can imply that the value of preventing fraud outweighs other values, for example, proportionality. Because of this rigorous approach, there is little room for doubt and someone may be easily marked as potentially fraudulent. This rigorous approach will be reflected in the data that an ML model is trained on. Thus, because a powerful actor values this rigorous approach, it will be embodied in the algorithm.

Amplification of the lack of transparency. Section 2 shows that decision-making processes are often non-linear – decision-making is an opaque, capricious process and it is not transparent concerning what actors have what impact on the ultimate decisions. This is inherent to the wicked character of problems and the many actors involved in the decision-making process. Now, suppose that in this process one or more actors use algorithms – this does not make the process less opaque. On the contrary, since the development of these algorithms was also an opaque process, the decision-making only becomes less transparent. The lack of transparency is amplified.

Amplification of the lack of accountability. Accountability issues might also be amplified. Accountability issues in regular public decision-making relate to the multi-actor nature of decision-making and the problem of "the many hands". As described before, the introduction of ML algorithms adds another layer of hands, namely, that related to the algorithm. In public decision-making, it is not always clear who should be held accountable for particular decisions. When an ML algorithm is present in the decision-making processes, that makes it even harder to talk about accountability. Thus, the introduction of the ML algorithm leads to an amplification of the accountability challenge.

Amplification of disputes on value trade-offs. There is a need for value trade-offs. The literature shows that algorithmic decision-making comes with debatable value trade-offs, in a context where wicked problems must be solved. Wicked problems are already characterized by dissensus about the underlying values and value trade-offs. So there is already a conflict of values when wicked problems are on the agenda. And again, this conflict of values is amplified by the use of algorithms, because these algorithms represent a value trade-off that is often disputable. So, an algorithm with a disputable value trade-off is used to solve a wicked problem, which requires a value trade-off that is inherently disputable.

This amplification of the drawbacks of algorithmic decision-making also finds a cause in the prior process of development. If algorithms are ultimately used for decision-making on wicked issues, then decision-making on the development and use of algorithms is also wicked.

To illustrate this, we can take two benefits of the use of ML algorithms: they can result in more efficient and more accurate decision-making. However, when it comes to wicked problems, there is a fundamental issue: there is no unambiguous definition of a problem – and therefore efficiency and accuracy cannot be defined unambiguously. Suppose an inspection agency wants to detect fraudulent inspectees. This is a wicked problem: one of the questions is, for what measure of accuracy will developers optimize the algorithm? Will they optimize the algorithm to catch as many fraudulent inspectees as possible? Or to make as few as possible wrong accusations? These two approaches may lead to completely different outcomes. Or let's take the question of who should be inspected and who should not be inspected. Research shows that there is a variety of criteria relevant here (Goosensen, 2021). These criteria require a trade-off, and the weighing of the criteria is not unambiguous – or, put differently, is a wicked activity.

To conclude, a story of amplification emerges when we confront the literature about public decision-making with ML algorithms. There is an amplification of many issues that are already prevalent when we are dealing with multi-actor decision-making on wicked issues. This has mainly to do with the fact that ML algorithms themselves can be considered multi-actor and wicked, while they are also coming into the context of multi-actor decision-making on wicked issues.

Discussion

Our findings suggest that applying ML algorithms for public decision-making creates an amplification of many of the shortcomings prevalent in both. Especially public decision-making often concerns decision-making about type II problems. The characteristics of these problems, squared with the drawbacks of ML algorithms create amplification. The amplified shortcomings regard an amplification of bias, an amplification of the lack of transparency, an amplification of the lack of accountability, and finally an amplification of disputes of value trade-offs. The findings suggest that this has partially to do with the fact that before algorithms are used in the wicked multi-actor context of public decision-making, there is decision-making about the development of ML algorithms, which is also a multi-actor and wicked process.

This is a finding that builds on previous literature that started to acknowledge the importance of the development process of ML algorithms (van der Voort et al., 2019; Zweig et al., 2018). Within this literature, it is emphasized that stakeholders such as data scientists, designers and decision-makers all play a critical role in decision-making processes that are based on data or algorithms (van der Voort et al., 2019). This study concludes similarly. What this study specifically adds is the additional claim that shortcomings of either public decision-making or ML algorithms are amplified when ML algorithms are used for public decision-making. Our study brings forward that the involvement of multiple actors in the development process and the wickedness of the decisions involved, is one of the key issues, rather than technical issues with the algorithms. In that sense, our study also relates to previous literature that discusses technical solutions for challenges such as transparency or accountability (Belle & Papantonis, 2021; Kroll et al., 2017). The message we offer is not that these technical approaches have no value. On the contrary, what we do emphasize is that when algorithms are being applied for so-called type II decision-making, much more is going on and only technical solutions are problematic because they do not go to the root of the problem.

There might be multiple potential avenues for dealing with ML algorithms when applied to type II decision-making. What these suggestions all have in common is the importance of the human factor. We emphasize that awareness of the importance of the human factor is something that should be prevalent in all public decision-making dealing with ML algorithms.

Co-produce decision-making. The first suggestion is connected to the question of what role ML can play in decision-making processes. If ML becomes a substitute for human-based type II decision-making, we might expect the outcome of ML to be heavily criticized, often because of the high personal importance of the problems (Wenzelburger, König, Felfeli, & Achtziger, 2022). Thus, we argue for ML that facilitates human-based decision-making. Two variants might be possible. 1) A variant of competition, where ML-based decision-making does not replace human decision-making but competes as an additional tool. When ML-based decision-making leads to other outcomes than human-based decision-making, it can trigger learning processes – actors may reconsider their original decision and subsequently take a better or better-substantiated decision. 2) A variant of cooperation. For example, ML-based decision-making only concerns a limited number of aspects of the decision to be taken, especially those aspects that are less wicked. Decision-making remains human-based, but human decision-makers are partly supported.

Organize for a variety of perspectives. Secondly, wickedness implies the existence of multiple perspectives. Using only one algorithm in such a context is remarkable. A variety of perspectives requires variety in algorithms. We argue that a variety of algorithms is introduced in public decision-making. For example, when dealing with algorithms, we can have ML algorithms based on different data sets, or use different types of algorithms that support the same decision. Also, variety in the integration into decision-making processes is needed. Instead of having an algorithm supporting decision-making, we can aim for variety in decision-making mechanisms. Thus, an ML algorithm supports decision-making, but also decision-making without algorithms or an algorithm that is verified by a human instead of vice versa. In line with "the algorithmic colleague" from the study of Meijer et al. (2021), we argue that the algorithm should be an instrument of knowledge rather an instrument that should be followed.

Build institutions. Thirdly, a focus is needed on developing the institutions within and outside of organizations that can deal with the complexity of developing, implementing and using ML algorithms. Suppose an inspectorate is shifting from a situation in which solely inspectors are selecting inspectees, towards a situation in which inspectors are supported by ML algorithms in selecting their inspectees. To ensure transparency, accountability or fairness, a big organizational challenge awaits this inspectorate. How do we organize in the new situation? How do we make sure that the inspector uses ML algorithms? There might be great importance in building checks and balances in the chain from algorithm development to algorithm use (Arnaboldi et al., 2022). There is already institutional complexity in the old situation, and even more so in the new situation. By institution building, we can start to create some "rules of the game" for ML algorithms in a public sector context.

Naturally, the present study has its limitations. This study brings together different types of literature. It connects literature on algorithms in the public sector to literature focusing on public decision-making. As a consequence, it brings together many concepts that are not often connected. A limitation of this study is especially this connection of many concepts, as it might create confusion.

Finally, we provide several suggestions for future research. As we have tried to conceptually bring together a variety of concepts, our main suggestion is to bring the study of ML algorithms to the empirical context, as also suggested by others (Veale, Van Kleek, & Binns, 2018). Firstly, future research can focus on empirically studying the use and development of ML algorithms for public decision-making. This chapter brought forward the importance of the human factor. But how are public sector professionals dealing with these types of ML algorithms for public decision-making? Secondly, future research can focus on empirically studying the institutions present in the entire process from design to use, or on studying institutions that are effective in ensuring transparency, accountability and fairness.

Conclusion

In the introduction, we asked what the wicked nature of public decisions means for the use of ML algorithms. Our study suggests that the wicked nature of public decision-making has consequences for how we should look at ML algorithms, how we develop ML algorithms, and how we use ML algorithms in the context of public decision-making, especially when ML algorithms are applied for type II decision-making. The main finding of the study is the existence of the amplification of challenges such as a lack of transparency, a lack of accountability and disputes on value trade-offs. This amplification of challenges finds its cause partially in the development process of ML algorithms, which we can consider a wicked and multi-actor process. In turn, the ML algorithm is then applied in a wicked and multi-actor context as well. The contribution of our study to the literature is mainly related to this notion of amplification. Earlier scholars did not specifically take into account the core principles of public decision-making. Since this study brings together a variety of concepts from different disciplines, there is also a potential pitfall, namely that the number of concepts obfuscates the central message of this chapter. Specifically, the human with and within the ML algorithm matters. However, since our study is based on a review of the literature, our suggestion is to start empirically studying the wicked and multi-actor nature of ML algorithms in their actual contexts (Veale et al., 2018). As stated, the growing use of ML algorithms in the public sector amplifies many common challenges in public decision-making. Fortunately, there are multiple ways forward to deal with these challenges. By employing co-production, organizing for variety, and by institution building, some of these amplifications can be condensed.

References

Adam, S., & Kriesi, H. (2007). The network approach. In P. A. Sabatier (Ed.), *Theories of the Policy Process* (2nd ed., pp. 129–154). Routledge.

Alexopoulos, C., Lachana, Z., Androutsopoulou, A., Diamantopoulou, V., Charalabidis, Y., & Loutsaris, M. A. (2019). How machine learning is changing e-government. In *Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance* (Vol. Part F1481, pp. 354–363). ACM. https://doi.org/10.1145/3326365.3326412

Alford, J., & Head, B. W. (2017). Wicked and less wicked problems: A typology and a contingency framework. *Policy and Society, 36*(3), 397–413. https://doi.org/10.1080/14494035.2017.1361634

Arnaboldi, M., de Bruijn, H., Steccolini, I., & Van der Voort, H. (2022). On humans, algorithms and data. *Qualitative Research in Accounting and Management, 19*(3), 241–254. https://doi.org/10.1108/QRAM-01-2022-0005

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review, 104*(671), 671–732. https://doi.org/10.2139/ssrn.2477899

Baumeister, R. F., & Leary, M. R. (1997). writing narrative literature reviews – Bausmeister & Leary. *Review of General Psychology, 1*(3), 311–320.

Belle, V., & Papantonis, I. (2021). Principles and practice of explainable machine learning. *Frontiers in Big Data, 4*(July), 1–25. https://doi.org/10.3389/fdata.2021.688969

Berk, R. (2017). An impact assessment of machine learning risk forecasts on parole board decisions and recidivism. *Journal of Experimental Criminology, 13*(2), 193–216. https://doi.org/10.1007/s11292-017-9286-2

Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications. *Science, 358*(6370), 1530–1534. https://doi.org/10.1126/science.aap8062

Burrell, J. (2016). How the machine "thinks": Understanding opacity in machine learning algorithms. *Big Data and Society, 3*(1), 1–12. https://doi.org/10.1177/2053951715622512

Coglianese, C., & Lehr, D. (2017). Regulating by robot: Administrative decision making in the Machine-learning era. *Georgetown Law Journal, 105*(5), 1147–1223.

Coglianese, C., & Lehr, D. (2018). Transparency and algorithmic governance. *Administrative Law Review, 71*(1), 18–38.

Cohen, M. D., March, J. G., & Olsen, J. P. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly, 17*(1), 1. https://doi.org/10.2307/2392088

Crootof, R., Bernstein, D., Bloch-Wehba, H., Dayrit, J., Pasquale, F., Re, R., … Surden, H. (2019). Columbia Law Review Forum technological-legal lock-in. *Columbia Law Review, 119*(2016), 233–251.

Danaher, J. (2016). The threat of algocracy: Reality, resistance and accommodation. *Philosophy and Technology, 29*(3), 245–268. https://doi.org/10.1007/s13347-015-0211-1

De Bruijn, H. (2021). *The governance of privacy*. Amsterdam University Press.

De Bruijn, H., & Ten Heuvelhof, E. (2018). *Management in networks* (2nd ed.). Routledge–Taylor & Francis Group. https://doi.org/10.4324/9781315453019

Desouza, K. C., Dawson, G. S., & Chenok, D. (2020). Designing, developing, and deploying artificial intelligence systems: Lessons from and for the public sector. *Business Horizons, 63*(2), 205–213. https://doi.org/10.1016/j.bushor.2019.11.004

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM, 55*(10), 78–87. https://doi.org/10.1145/2347736.2347755

Dwivedi, Yogesh K. et al. Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management, 57*(2021), 101994.

Eggers, W. D., Schatsky, D., & Viechnicki, P. (2017). AI-augmented government: Using cognitive technologies to redesign public sector work. *Deloitte Center for Government Insights*, 1–24. www2.deloitte.com/content/dam/insights/us/articles/3832_AI-augmented-government/DUP_AI-augmented-government.pdf%0Ahttps://www2.deloitte.com/us/en/insights/focus/cognitive-technologies/artificial-intelligence-government.html

Goosensen, H. R. (2021). *Inzicht in de praktijk van het toezicht: Een empirisch onderzoek naar het verloop van operationele inspectieprocessen in de luchtvaart en zeevaart*. [Doctoral dissertation, TU Delft University]. https://doi.org/10.4233/uuid:62e7441c-72c7-4445-bf72-fb8ef047308e

Hartmann, K., & Wenzelburger, G. (2021). Uncertainty, risk and the use of algorithms in policy decisions: A case study on criminal justice in the USA. *Policy Sciences, 54*(2), 269–287. https://doi.org/10.1007/s11077-020-09414-y

Head, B. W. (2019). Forty years of wicked problems literature: Forging closer links to policy studies. *Policy and Society, 38*(2), 180–197. https://doi.org/10.1080/14494035.2018.1488797

Hisschemöller, M., & Hoppe, R. (1995). Coping with intractable controversies: The case for problem structuring in policy design and analysis. *Knowledge and Policy, 8*(4), 40–60. https://doi.org/10.1007/BF02832229

Janssen, M., & Kuk, G. (2016). The challenges and limits of big data algorithms in technocratic governance. *Government Information Quarterly, 33*(3), 371–377. https://doi.org/10.1016/j.giq.2016.08.011

Katzenbach, C., & Ulbricht, L. (2019). Algorithmic governance. *Internet Policy Review, 8*(4), 1–18. https://doi.org/10.14763/2019.4.1424

Kingdon, J. W. (2011). *Agendas, alternatives, and public policies* (2nd ed.). Longman.

Klijn, E. H., & Koppenjan, J. F. M. (2000). Public management and policy networks. *Public Management: An International Journal of Research and Theory, 2*(2), 135–158. https://doi.org/10.1080/14719030000000007

König, P. D., & Wenzelburger, G. (2021). The legitimacy gap of algorithmic decision-making in the public sector: Why it arises and how to address it. *Technology in Society, 67*(July). https://doi.org/10.1016/j.techsoc.2021.101688

Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., & Yu, H. (2017). Accountable algorithms. *University of Pennsylvania Law Review, 165*(3), 633–705.

Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications Policy, 44*(6), 101976. https://doi.org/10.1016/j.telpol.2020.101976

Liu, H. W., Lin, C. F., & Chen, Y. J. (2019). Beyond state v loomis: Artificial intelligence, government algorithmization and accountability. *International Journal of Law and Information Technology, 27*(2), 122–141. https://doi.org/10.1093/ijlit/eaz001

Lorenz, L., Erp, J. Van, & Meijer, A. (2022). Machine-learning algorithms in regulatory practice agencies. *Technology & Regulation,* 1–11. https://doi.org/10.26116/techreg.2022.001

Luxton, D. D. (2016). *An introduction to artificial intelligence in behavioral and mental health care*: *Artificial intelligence in behavioral and mental health care*. Elsevier. https://doi.org/10.1016/B978-0-12-420248-1.00001-5

Mehr, H. (2017). Artificial intelligence for citizen services and government. *Harvard Ash Center Technology & Democracy* (August), 1–16. https://ash.harvard.edu/files/ash/files/artificial_intelligence_for_citizen_services.pdf

Meijer, A., Lorenz, L., & Wessels, M. (2021). Algorithmization of bureaucratic organizations: Using a practice lens to study how context shapes predictive policing systems. *Public Administration Review, 81*(5), 837–846. https://doi.org/10.1111/puar.13391

Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences, 4*(2), 155–169. https://doi.org/10.1007/BF01405730

Samuel, A. L. (2000). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development, 44*(1/2), 206–226. https://doi.org/10.1147/rd.441.0206

Sousa, W. G. de, Melo, E. R. P. de, Bermejo, P. H. D. S., Farias, R. A. S., & Gomes, A. O. (2019). How and where is artificial intelligence in the public sector going? A literature review and research agenda. *Government Information Quarterly, 36*(4), 101392. https://doi.org/10.1016/j.giq.2019.07.004

Teisman, G. R. (2000). Models for research into decision-makingprocesses: On phases, streams and decision-making rounds. *Public Administration, 78*(4), 937–956. https://doi.org/10.1111/1467-9299.00238

Thierer, A. D., Castillo, A., & Russell, R. (2017). Artificial intelligence and public policy. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3021135

van der Voort, H. G., Klievink, A. J., Arnaboldi, M., & Meijer, A. J. (2019). Rationality and politics of algorithms. Will the promise of big data survive the dynamics of public decision making? *Government Information Quarterly, 36*(1), 27–38. https://doi.org/10.1016/j.giq.2018.10.011

Veale, M., & Brass, I. (2019). Administration by algorithm? Public management meets public sector machine learning. In K. Yeung & M. Lodge (Eds.), *Algorithmic regulation* (pp. 1–30). Oxford University Press. https://doi.org/10.31235/osf.io/mwhnb

Veale, M., Van Kleek, M., & Binns, R. (2018). Fairness and accountability design needs for algorithmic support in high-stakes public sector decision-making. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). ACM. https://doi.org/10.1145/3173574.3174014

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly, 26*(2), xiii–xxiii. https://doi.org/10.1.1.104.6570

Wenzelburger, G., König, P. D., Felfeli, J., & Achtziger, A. (2022). Algorithms in the public sector. Why context matters. *Public Administration* (November), 1–21. https://doi.org/10.1111/padm.12901
Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial intelligence and the public sector –applications and challenges. *International Journal of Public Administration, 42*(7), 596–615. https://doi.org/10.1080/01900692.2018.1498103
Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/2601248.2601268
Young, M. M., Bullock, J. B., & Lecy, J. D. (2019). Artificial discretion as a tool of governance: A framework for understanding the impact of artificial intelligence on public administration. *Perspectives on Public Management and Governance*, 301–313. https://doi.org/10.1093/ppmgov/gvz014
Zuiderwijk, A., Chen, Y., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly* (May 2020), 101577. https://doi.org/10.1016/j.giq.2021.101577
Zweig, K. A., Wenzelburger, G., & Krafft, T. D. (2018). On chances and risks of security related algorithmic decision making systems. *European Journal for Security Research, 3*(2), 181–203. https://doi.org/10.1007/s41125-018-0031-2