

Machine Learning and the Re-Enchantment of the Administrative State

Eden Sarid*  and Omri Ben-Zvi†

Machine learning algorithms present substantial promise for more effective decision-making by administrative agencies. However, some of these algorithms are inscrutable, namely, they produce predictions that humans cannot understand or explain. This trait is in tension with the emphasis on reason-giving in administrative law. The article explores this tension, advancing two interrelated arguments. First, providing adequate reasons is a significant facet of respecting individuals' agency. Incorporating inscrutable algorithmic predictions into administrative decision-making compromises this normative ideal. Second, as a long-term concern, the use of inscrutable algorithms by administrative agencies may generate systemic effects by gradually reducing the realm of the humanly explainable in public life, a phenomenon Max Weber termed 're-enchantment'. As a result, the use of inscrutable machine learning algorithms might trigger a special kind of re-enchantment, making us comprehend less rather than more of shared human experience, and consequently altering the way we understand the administrative state and experience public life.

INTRODUCTION

In recent years, the use of machine learning algorithms by administrative agencies has sparked considerable debate in public law scholarship.¹ On the one hand, these algorithms are generally considered to be able to outperform standard human cognition in managing certain tasks, such as analysing large amounts of data and generating meaningful correlations, distinctions, and predictions, suggesting real promise for the administrative state.² On the other hand, machine learning algorithms introduce new problems and concerns for public law and raise novel issues of regulatory design. For example, some of these

* Essex Law School.

† Hebrew University of Jerusalem, Faculty of Law. We thank Haim Abraham, David Dyzenhaus, Aviv Gaon, Lital Helman, Aziz Huq, Ariel Katz, Mary Mitchell, Daragh Murray, Simon Stern, Oren Tamir, and the MLR reviewers for their helpful comments. Earlier drafts of this article were presented at the Federmann Cyber Security Research Center Cyberlaw Workshop, ICON-S annual conference, CUHK Machine Lawyering Conference, and University of Toronto Innovation Law Workshop. We thank the participants in these workshops for their valuable input.

1 See Aziz Huq, 'A Right to a Human Decision' (2020) 106 *Virginia Law Review* 611, 618; Cary Coglianese, 'Administrative Law in the Automated State' (2021) 150 *Daedalus* 104; Rebecca Williams, 'Rethinking Administrative Law for Algorithmic Decision Making' (2022) 42 *OJLS* 468; Jennifer Cobbe, 'Administrative Law and The Machines of Government: Judicial Review of Automated Public-Sector Decision-Making' (2019) 39 *Legal Studies* 636.

2 See Coglianese, *ibid*; Williams, *ibid*. See also Katherine Strandburg, 'Rulemaking and Inscrutable Automated Decision Tools' (2019) 119 *Columbia Law Review* 1851, 1857.

algorithms are not transparent and they can be ‘gamed’ by external parties or influenced by various biases.³

A key concern on which this article focuses is the issue of *inscrutability*. Some machine learning algorithms employ advanced techniques that generate predictions and recommendations that are inscrutable in the sense that humans cannot understand the reasoning behind them.⁴ Inscrutability is the result of several factors. First, to generate any specific output, these algorithms analyse enormous amounts of information (‘big data’), far beyond the scope of what any single person can take in, let alone seriously contemplate in a lifetime.⁵ Second, to reach their conclusions, these algorithms employ a set of complicated interconnected techniques and methodologies. Though the algorithms’ designers set up the basic model for the algorithms’ operation, in many cases it is still hard if not impossible to evaluate or explain the algorithms’ results, even for experienced systems designers and engineers, not to mention non-technical reviewers.⁶ Third, these algorithms often arrive at conclusions which are counter-intuitive and therefore hard to accept.⁷ When combined, these characteristics make predictions generated by these types of algorithms incomprehensible to the human mind. Simply put, we cannot understand how these algorithms arrived at their conclusions – we do not know which specific data points the algorithm analysed, how they were used, and the weight given to different components. For example, an algorithm ‘might predict a person’s preferred style of shoe based upon the type of fruit the person typically purchases for breakfast’.⁸ These predictions may be accurate, but we cannot fully understand why. For the sake of this article, we will call these types of algorithms ‘inscrutable algorithms’. In a nutshell, inscrutable algorithms are algorithms that make predictions that one cannot verify independently. Below, we elaborate on these inscrutable algorithms as well as other types of machine learning algorithms that do not raise these concerns.⁹

For administrative law, the main problem with inscrutable algorithms is that agencies do not have access to the data and the logical processes by virtue of which these algorithms’ predictions are supposed to be accurate.¹⁰ This issue – also called the ‘black box problem’ – currently generates much scholarly debate, including which solutions work best to mitigate inscrutability, and how to strike a balance between our desire for transparency, accountability, and intelligibility, and the need for effectiveness.¹¹

3 See David Freeman Engstrom and Daniel E. Ho, ‘Artificially Intelligent Government: A Review and Agenda’ in Roland Vogl (ed), *Research Handbook on Big Data Law* (Cheltenham: Edward Elgar Publishing, 2021).

4 See Ashley Deeks, ‘The Judicial Demand for Explainable Artificial Intelligence’ (2019) 119 *Columbia Law Review* 1829.

5 See David Lehr and Paul Ohm, ‘Playing with the Data: What Legal Scholars Should Learn about Machine Learning’ (2017) 51 *UC Davis Law Review* 653.

6 Cobbe, n 1 above, 639.

7 See Andrew Selbst and Solon Barocas, ‘The Intuitive Appeal of Explainable Machines’ (2018) 87 *Fordham Law Review* 1085, 1090.

8 Bernard Bell, ‘Replacing Bureaucrats with Automated Sorcerers?’ (2021) 150 *Daedalus* 89, 90.

9 See text below under the heading ‘Machine Learning and the ‘Black Box’ Problem’.

10 See Selbst and Barocas, n 7 above, 1094.

11 See Yavar Bathaee, ‘The Artificial Intelligence Black Box and the Failure of Intent and Causation’ (2018) 31 *Harvard Journal of Law & Technology* 889; Frank Pasquale, *The Black Box Society: The*

In this article we focus on two major normative concerns regarding inscrutability. The first is that inscrutable algorithms do not provide sufficient reasoning for us to properly understand their predictions; the second is that the growing use of inscrutable algorithms in administrative decision-making generates systemic concerns. The lack of reason-giving and the systemic concerns are in fact interconnected in an important and as yet unexplored manner. Inscrutable algorithms do not simply alter the decision-making process in individual cases. Rather, by incorporating inscrutable components to decisions that shape public life, they may have wide systematic effects on our understanding of the administrative state itself, because more aspects of public life would gradually become regulated by processes that are incomprehensible to human cognition.¹²

More specifically, the article advances two major normative arguments. First, inscrutable algorithms' inability to provide adequate explanations for their predictions in specific cases may amount to a failure of the normative duty to provide reasons for decisions in individual cases. This argument focuses on the short-term, individual level of decision-making. On this level of analysis, the article engages with arguments that defend the view that inscrutability does not pose a problem vis-à-vis a general requirement to give reasons. We contend that these arguments are unpersuasive and that, therefore, a certain normative cost is paid whenever inscrutable algorithms are used for decision-making. This cost might not be high in all cases, and it can be outweighed by other, more substantial reasons, chief among which is usually the efficiency of such algorithms in making predictions.¹³

The second argument focuses on the long-term perspective and considers the system-wide consequences of using inscrutable algorithms from the point of view of reason-giving. Here we argue that the requirement for reasoned state action translates into a broader, systemic concern. Something larger than just the sum of many individual decisions is at stake when we consider using these algorithms: widespread use of inscrutable algorithms could eventually alter the manner in which we understand the administrative state itself.¹⁴

The possible change to which we allude is the gradual receding of the realm of the humanly explainable in public life. At least since the Enlightenment, and perhaps earlier, Western societies have been on a long journey to expand the range of phenomena (both natural and human-created) which we can

Secret Algorithms That Control Money and Information (Cambridge, MA: Harvard University Press, 2015); Katherine Strandburg, 'Adjudicating with Inscrutable Decision Tools' in Marcello Pelillo and Teresa Scantamburlo (eds), *Machines We Trust* (Cambridge, MA: MIT Press, 2021).

12 See Roger Brownsword, 'Law, Liberty, and Technology' in Roger Brownsword, Eloise Scotford and Karen Yeung (eds), *The Oxford Handbook of Law, Regulation and Technology* (Oxford: OUP, 2017) 8.

13 Thus, our argument is not intended as a general case against the use of machine learning algorithms.

14 See Brownsword, Scotford, and Yeung, n 12 above, 4. See also, Roger Brownsword, 'Artificial Intelligence and Legal Singularity: The Thin End of the Wedge, the Thick End of the Wedge, and the Rule of Law' in Simon Deakin and Christopher Markou (eds), *Is Law Computable?: Critical Perspectives on Law and Artificial Intelligence* (Oxford: Hart Publishing, 2020) 136 and 153–155; Roger Brownsword, *Rethinking Law, Regulation, and Technology* (Cheltenham: Edward Elgar Publishing, 2022) 82.

understand. This process of the *disenchantment* of the world, as Weber called it, has been an overarching theme of Western thought for centuries.¹⁵ Disenchantment is the historical process by which the experienced world becomes more understandable and knowable and consequently less mysterious for us.¹⁶

As Weber recognised, bureaucratic governance – what today we call the administrative state – is a clear manifestation of the logic of disenchantment in the public sphere.¹⁷ Administrative bodies are commonly thought of as making decisions based on expertise and reason rather than other factors, such as pure political will, magic, or tradition.¹⁸ By so doing, they shape the public sphere in a manner consistent with accessible reasons for action,¹⁹ and thus make state coercion understandable for the subjects of the state.

When the administrative state is put in this theoretical context, it is easy to see why modern administrative law places an emphasis on reason-giving,²⁰ even if the law does not necessarily require that a reason be given for every agency decision.²¹ As Jerry Mashaw notes, '[t]he path of American administrative law has been the path of the progressive submission of power to reason. The promise of the administrative state ... is the institutional embodiment of the enlightenment project to substitute reason for the dark forces of culture, tradition, and myth.'²²

Substantial adoption of inscrutable algorithms' predictions and suggestions could potentially instantiate a sort of 're-enchantment', ie, the phenomenon of understanding less rather than more of shared human experience. We term this possibility *the re-enchantment of the artificial*. While in Weberian terminology, re-enchantment usually means the renewed introduction of non-rational thinking into the language of natural science and the bureaucracy, we argue that the phenomenon we describe here is different. On the one hand, in the re-enchantment of the artificial, decisions are made on the basis of algorithms that promise more, not less, rationality.²³ After all, a major advantage of machine learning algorithms is their potential accuracy in making predictions. But on the other hand, notwithstanding this promise of greater efficiency, decisions that are guided by inscrutable algorithms' predictions could register as unexplained *to us*, because we have no access to the reasons (the facts and logical processes)

15 Peter Lassman, Herminio Martins and Velody Irving (eds), *Max Weber's 'Science As A Vocation'* (Indianapolis, IN: Hackett Publishing Company, 1989) 13–14.

16 Richard Jenkins, 'Disenchantment, Enchantment and Re-Enchantment: Max Weber at the Millennium' (2000) 1 *Max Weber Studies* 11, 12.

17 Jerry Mashaw, 'Small Things Like Reasons Are Put in a Jar: Reason and Legitimacy in the Administrative State' (2001) 70 *Fordham Law Review* 17, 18.

18 See Yishai Blank, 'The Reenchantment of Law' (2011) 96 *Cornell Law Review* 633, 637.

19 See Thomas Nagel, *Equality and Partiality* (Oxford: OUP, 1995) 141–142; John Rawls, *The Law of Peoples* (Cambridge, MA: Harvard University Press 2001) 54–48.

20 See Frederick Schauer, 'Giving Reasons' (1995) 47 *Stanford Law Review* 633; Jerry Mashaw, 'Reasoned Administration: The European Union, the United States, and the Project of Democratic Governance' (2007) 76 *George Washington Law Review* 99; Timothy Endicott, 'Legal Interpretation' in Andrei Marmor (ed), *The Routledge Companion to Philosophy of Law* (New York, NY: Routledge, 2011) 109, 110.

21 See text below under the heading 'Giving Reasons'.

22 Mashaw, n 17 above, 26.

23 Because machine learning algorithms are better than humans at making distinctions and detecting correlations within datasets.

which make them plausible and accurate in the first place.²⁴ The mere promise that inscrutable algorithms' predictions are in fact accurate could mean little if more and more people experience the administrative state as generating decisions that influence them on the basis of reasons they do not and cannot understand. Our argument relies on the idea that disenchantment should be understood as an existential category – it is not a claim about the world as such but about the world as experienced by humans.²⁵

The article proceeds as follows. The next section briefly introduces key concepts necessary to investigate machine learning inscrutability. We then unpack the argument concerning the potential costs affiliated with the use of inscrutable algorithms in individual decisions made by administrative bodies. Against this background, we argue that machine learning inscrutability introduces a potential systematic concern regarding the re-enchantment of the administrative state. Finally, we contend that the potential re-enchantment of the administrative state presents an important issue for scholars and policymakers at the current crossroads for administrative law and artificial intelligence.

Before elaborating on the article's arguments, we note two qualifications regarding the scope of our argument. First, when we say that administrative agencies 'use' machine learning algorithms, we mean that these algorithms produce predictions that assist administrative agencies in their decision-making, but, for the most part, algorithms do not (as yet) make decisions themselves.

Second, not all machine learning algorithms raise the concerns explored in the article. Our focus is on a particular subset of machine learning algorithms, which we call 'inscrutable algorithms'. Inscrutable algorithms, for the sake of this article, are algorithms that make predictions that one cannot verify independently at the time of receiving the output. These algorithms rely on big data, involve opaque and recursive processing, and may produce counter-intuitive conclusions; human cognition cannot verify the algorithms' predictions independently and we are forced to evaluate the predictions based on the algorithms' output alone. It is important to note that there are algorithms that might employ inscrutable black box models, but their predictions can be independently verified (for example facial recognition algorithms determining a person's identity from a photo can be verified by a human agent); models that can provide insights into how the algorithm arrived at its prediction (for example provide details on the data and calculations it used to predict that a certain factory was likely polluting); or models whose design we can explain (for example explain the expected impact and potential biases of algorithms that suggest certain individuals for tax audit).²⁶ While the workings of those algorithms might remain inscrutable, our ability to validate, interpret, or explain them renders them comprehensible. These types of algorithms do not pose a problem from the

24 See Nick Bostrom, 'The Superintelligent Will: Motivation and Instrumental Rationality in Advanced Artificial Agents' (2012) 22 *Minds and Machines* 71, 72.

25 Duncan Kennedy, 'The Disenchantment of Logically Formal Legal Rationality, or Max Weber's Sociology in the Genealogy of the Contemporary Mode of Western Legal Thought' (2004) 55 *Hastings Law Journal* 1031, 1057.

26 See further discussion below under the heading 'Machine Learning and the 'Black Box' Problem'.

point of view of re-enchantment because they do not trigger the notion of being subject to decisions which we cannot understand.

Examples of the use of inscrutable algorithms by administrative agencies include using such algorithms to determine the probability of an individual committing subsequent offences upon being granted bail,²⁷ to decide which undocumented immigrants to target for deportation,²⁸ the likelihood of a borrower repaying their loan,²⁹ or determining the allocation of inspectors based on predictions of which factories are more prone to deficiencies. We now turn to discuss further examples and the challenges that inscrutable algorithms present.

MACHINE LEARNING AND THE 'BLACK BOX' PROBLEM

The use of machine learning is particularly attractive to designers of delegated, distributed decision systems because it holds the promise of improving consistency, reducing biases, and lowering costs.³⁰ These attributes are driving the continued interest in and use of various types of machine learning algorithms in diverse fields, including employment,³¹ policing and sentencing,³² credit,³³ patent applications,³⁴ tax,³⁵ and even college admissions.³⁶

Administrative law has also seen an upsurge of interest in the use of machine learning algorithms in recent years. Administrative agencies at local and national levels have been adopting machine learning with growing enthusiasm.³⁷ One of the challenges in this regard is that it is often difficult to know the extent to which a particular agency relies on an inscrutable algorithm or whether it uses a simpler form of algorithmic technique with higher degrees of explainability. Different systems may be employed in different contexts, and the extent of explainability will differ accordingly. The general trend that can be observed in administrative agencies globally is a heightened use of algorithms which themselves tend to be more and more complicated and therefore more inscrutable.

27 See Keith Kirkpatrick, 'It's Not the Algorithm, It's the Data' (2017) 60 *Communications of the ACM* 21.

28 See Joint Council for the Welfare of Immigrants, 'We Won! Home Office to Stop Using Racist Visa Algorithm' (JCWI, 2020) at <https://www.jcwi.org.uk/news/we-won-home-office-to-stop-using-racist-visa-algorithm> [https://perma.cc/2AMB-TYAK].

29 See Mayank Anand, Arun Velu and Pawan Whig, 'Prediction of Loan Behaviour with Machine Learning Models for Secure Banking' (2022) 3 *Journal of Computer Science and Engineering (JCSE)* 1.

30 Strandburg, n 2 above, 1857.

31 See Pauline Kim, 'Data-Driven Discrimination at Work' (2017) 58 *William & Mary Law Review* 857.

32 See *State v Loomis* 881 N.W.2d 749, 753 (Wis. 2016).

33 See Matthew Bruckner, 'The Promise and Perils of Algorithmic Lenders' Use of Big Data' (2018) 93 *Chicago-Kent Law Review* 3.

34 See Arti Rai, 'Machine Learning at the Patent Office: Lessons for Patents and Administrative Law' (2019) 104 *Iowa Law Review* 2617.

35 Benjamin Alarie, Anthony Niblett and Albert Yoon, 'Using Machine Learning to Predict Outcomes in Tax Law' in Vogl (ed), n 3 above.

36 Strandburg, n 11 above.

37 Thomas Vogl and others, 'Smart Technology and the Emergence of Algorithmic Bureaucracy: Artificial Intelligence in UK Local Authorities' (2020) 80 *Public Administration Review* 946, 949.

Therefore, the problems we address in this article, while already warranting attention, will only become more relevant as time passes. The use of machine learning algorithms in administrative law is a growing trend including in Australia,³⁸ Canada,³⁹ Estonia, Denmark, and South Korea, to name but a few,⁴⁰ and experts predict that this trend will continue.⁴¹

Examples of the use of inscrutable algorithms in administrative decision-making include the UK Home Office using a 'visa streaming' tool to grade entry visa applications, assigning risk ratings that significantly impact application outcomes.⁴² Another example is the UK's Department for Work and Pensions using an inscrutable algorithm to determine which claims for universal credit to investigate.⁴³ Local authorities in the UK have also been applying inscrutable algorithms to support decisions on transportation, houses in multiple occupation, and children's social care.⁴⁴ The Dutch tax authorities have reportedly used an inscrutable algorithm to identify potential cases of tax fraud.⁴⁵ In the US, Immigration and Customs Enforcement uses an inscrutable risk assessment algorithm to determine which undocumented immigrants to target for deportation.⁴⁶ Additionally, several police forces in a number of US jurisdictions have been using an inscrutable algorithm called COMPAS to predict crime recidivism.⁴⁷

While the use of automated decision-making tools seems to be on the rise, some scholars have suggested that decision-makers in fact do not need to use inscrutable algorithms since alternative models can produce similar results. These alternative models, it is argued, can alleviate inscrutability concerns while achieving many of the advantages associated with automated decision-making.⁴⁸ Interpretable AI and explainable AI are often discussed as two such models.

Interpretable AI generally refers to the ability of a model to provide insights into how it arrived at its output or prediction. This can include identifying which features of the input data were most influential, using a linear model where variables are weighted and added together, or identifying how different

38 Rusul Abduljabbar and others, 'Applications of Artificial Intelligence in Transport: An Overview' (2019) 11 *Sustainability* 189.

39 Jennifer Raso, 'Unity in The Eye of The Beholder? Reasons for Decision in Theory and Practice in The Ontario Works Program' (2019) 70 *UTLJ* 1.

40 Coglianesi, n 1 above, 106-107.

41 Deeks, n 4 above, 1839.

42 See Joint Council for the Welfare of Immigrants, n 28 above.

43 See 'Work and Pensions Committee (Parliamentlive.tv, 24 November 2021), 10:48 at <https://parliamentlive.tv/event/index/d4766433-5e00-4060-8e24-a5e4030da3d3?in=10:47:54> [<https://perma.cc/JE24-78VY>].

44 Vogl and others, n 37 above, 951.

45 Melissa Heikkila, 'Dutch Scandal Serves As A Warning For Europe Over Risks Of Using Algorithms' (Politico, 29 March 2022) at <https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/> [<https://perma.cc/ZH5T-XFW9>].

46 Estefania McCarroll, 'Weapons of Mass Deportation: Big Data and Automated Decision-Making Systems in Immigration Law Notes' (2020) 34 *Georgetown Immigration Law Journal* 705, 710-725 and 728.

47 See Kirkpatrick, n 27 above; Cynthia Rudin and Joanna Radin, 'Why Are We Using Black Box Models in AI When We Don't Need To? A Lesson From An Explainable AI Competition' (2019) 1 *Harvard Data Science Review* 1.

48 Rudin and Radin, *ibid*.

inputs were combined to produce the output.⁴⁹ Other interpretable models, as Cynthia Rudin and Joanna Radin explain, are comprised of simpler models put together, or placing constraints on the model to add a new level of insight.⁵⁰ In other words, such models show us how they work in a way that we can understand. There are examples of the potential of interpretable AI to provide reliable predictions for administrative decision-making. For instance, as mentioned, several police forces in the US have been using the inscrutable algorithm COMPAS to predict crime recidivism,⁵¹ yet researchers have demonstrated that interpretable models reach results that are just as accurate.⁵² Another example is the US Environmental Protection Agency using interpretable AI to assist with environmental monitoring, including predicting noncompliance risks, and identifying facilities operating without required environmental permits.⁵³ Indeed, some scholars view interpretable AI as a potential solution to the black box issue linked to inscrutable algorithms.⁵⁴ Rudin and Radin, for example, argue that there are ‘interpretable models, which provide a technically equivalent, but possibly more ethical alternative to black box models’.⁵⁵

Explainable AI (or XAI for short) is a broader concept that encompasses not only the interpretability of the predictions generated by the algorithms but also the design and development of AI systems that are transparent, accountable, and fair. XAI comprises a set of processes and methods that enable human users to understand and trust the algorithm and its predictions. XAI can describe the AI model, its expected effects, and possible biases; it can draw out the strengths and weaknesses of the process and provide a sense of how the system will behave in the future.⁵⁶ XAI might involve creating models that run counterfactuals to see how results may vary, or creating models that provide visualisations of how the prediction was arrived at.⁵⁷ XAI also has demonstrated potential for administrative agencies. For example, researchers have noted XAI’s potential to support German tax authorities in targeting organisations for auditing purposes, arguing that even if the tax authorities operate a black box algorithm, using XAI can mitigate concerns and comply with transparency and reason-giving requirements.⁵⁸ In a similar vein, the US Department of Defense is currently using

49 See Pantelis Linardatos, Vasilis Papastefanopoulos and Sotiris Kotsiantis, ‘Explainable AI: A Review of Machine Learning Interpretability Methods’ (2021) 23 *Entropy* 18.

50 Rudin and Radin, n 47 above.

51 Kirkpatrick, n 27 above; Rudin and Radin, *ibid*.

52 Rudin and Radin, *ibid*.

53 Robert Denney, ‘Opportunities for Artificial Intelligence in Environmental Compliance’ (2022) 52 *Environmental Law* 99, 109–110.

54 See Rudin and Radin, n 47 above; Cynthia Rudin, ‘Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead’ (2019) 1 *Nature Machine Intelligence* 206.

55 Rudin and Radin, n 47 above, 2.

56 See Arun Rai, ‘Explainable AI: From Black Box to Glass Box’ (2020) 48 *Journal of the Academy of Marketing Science* 137; Greg Adamson, ‘Explainable Artificial Intelligence (XAI): A Reason to Believe?’ (2020) 37 *Law in Context: A Socio-Legal Journal* 23.

57 Hans de Bruijn, Martijn Warnier and Marijn Janssen, ‘The Perils and Pitfalls of Explainable AI: Strategies For Explaining Algorithmic Decision-Making’ (2022) 39 *Government Information Quarterly* 101666.

58 Nijat Mehdiyev and others, ‘Explainable Artificial Intelligence (XAI) Supporting Public Administration Processes – On the Potential of XAI in Tax Audit Processes’ (Wirtschaftsin-

XAI to identify the most cost-effective vendors and expedite its procurement process.⁵⁹ Some scholars argue that XAI can allow administrative agencies to provide the efficiency of machine learning algorithms while upholding important public law principles, including transparency and reasoned decisions.⁶⁰ Still, some scholars argue against the adoption of XAI, particularly in high-stakes scenarios, due to ongoing limitations. These include concerns that XAI explanations may not faithfully reflect the original model's computations or provide sufficient insight into the black box's actions.⁶¹

Algorithms which can be understood in a relatively straightforward manner, such as interpretable AI or XAI models, do not raise the philosophical concerns explored in this article. However, the question of whether interpretable AI or XAI can indeed resolve inscrutability concerns remains highly debated, to say the least. This is because it is currently technologically doubtful whether interpretable AI or XAI algorithms can produce the timely and accurate results that inscrutable algorithms produce. Some scholars contend that despite explainability, in reality 'many advanced AI systems remain black boxes', and that the enthusiasm some display for XAI is 'unreasoned',⁶² while others go so far as dismissing XAI as 'false hope'.⁶³ Other scholars contend that there is a 'clear trade-off between the performance of a machine learning model and its ability to produce explainable and interpretable predictions'⁶⁴ and, so long that this trade-off remains, governments will be reluctant to prioritise interpretable or XAI models, since the primary advantage of using machine learning algorithms is their efficiency and accuracy. And indeed, administrative agencies increasingly use inscrutable algorithms, or do not disclose which algorithms they have used, rendering the algorithms effectively inscrutable. For example, in a recent study, Colin van Noordt and Gianluca Misuraca document 250 cases of use of algorithms in decision-making in public sector organisations across the EU.⁶⁵ Of these 250 cases, a considerable proportion were inscrutable algorithms, and an additional substantial part were ambiguous algorithms (ie, it was unclear which type of algorithm the agency used).⁶⁶ Similarly, in the US, Ryan Calo created a database of use of AI algorithms by the Federal Government, documenting over 402 such cases, a substantial part of which are inscrutable algorithms or cases

formatik 2021 Proceedings, 2021) at <https://aisel.aisnet.org/wi2021/SSmartCity/Track08/5> [<https://perma.cc/Q7BL-ASUW>].

59 Marcin Frackiewicz, 'The Benefits of Explainable AI for Government Agencies' (TS2 SPACE, 2023) at <https://ts2.space/en/the-benefits-of-explainable-ai-for-government-agencies/> [<https://perma.cc/8AD6-U34L>].

60 See Deeks, n 4 above.

61 Rudin, n 54 above, 208–210.

62 Adamson, n 56 above, 23.

63 Marzyeh Ghassemi, Luke Oakden-Rayner and Andrew L. Beam, 'The False Hope of Current Approaches to Explainable Artificial Intelligence in Health Care' (2021) 3 *The Lancet Digital Health* 745 (focusing on the example of XAI in healthcare).

64 See Linardatos, Papastefanopoulos and Kotsiantis, n 49 above, 18. Rudin, n 54 above, however, opposes this assertion.

65 See Colin van Noordt and Gianluca Misuraca, 'Artificial Intelligence for the Public Sector: Results of Landscaping the Use of AI in Government across the European Union' (2022) 39 *Government Information Quarterly* 101714.

66 *ibid.*

in which the type of model used remains unknown.⁶⁷ The use of inscrutable algorithms, at least for the time being, is increasingly a part of administrative decision-making, and therefore the philosophical questions that we raise here remain as relevant as ever.

Notwithstanding the potential benefits of inscrutable algorithms, they also prompt deep concerns.⁶⁸ One such concern includes problems related to *systematic effects*, capturing the possibility of undesirable spillover effects that ‘unfold dynamically and cumulatively’,⁶⁹ such as unfair distribution of social power. Inscrutable algorithms map input data to reach output predictions, and if these data are tainted or biased, the algorithm might reproduce and even exacerbate biases and social inequalities.⁷⁰ Another type of concern relates to *comprehensible reasoning* – inscrutable algorithms do not provide reasons for their predictions. We are asked to trust the algorithm’s predictions without being able to understand, let alone appraise, its reasons.⁷¹

Reason-giving is important for administrative law for several normative reasons. A core value of many legal systems is that decision-makers justify their decisions using socially and legally acceptable reasons.⁷² Having to give reasons promotes accountability, allows external scrutiny, and reflects a degree of respect given to the object of the decision and their dignity and agency.⁷³ Thus, some scholars believe that until machine learning inscrutability problems are properly addressed, administrative bodies should be wary of adopting machine learning.⁷⁴ Other scholars, by contrast, believe that such concerns are mainly misplaced or can be remedied, an argument to which we will now turn.

INSCRUTABILITY AND INDIVIDUAL DECISION-MAKING

The existing literature presents two primary arguments in support of the idea that inscrutability can be compatible with providing reasons: the first argument posits that while we cannot properly explain the predictions produced by inscrutable algorithms, programmers can explain the choices that were made at the ‘ground level’ that gave the algorithm its shape and form. As long as judges

67 Ryan Calo, ‘Fed Gov AI Use Case Inventories’ (Google Drive, 2023) at https://docs.google.com/spreadsheets/u/0/d/1FH-fzqwOsifhG-rp-MB7me6W9_XZiIbRFkwfQRMObfRs/htmlview#gid=0 [https://perma.cc/3TKV-9N26].

68 See Danielle Keats Citron, ‘Technological Due Process’ (2008) 85 *Washington University Law Review* 1249.

69 Huq, n 1 above, 651.

70 Deeks, n 4 above, 1833. See also Amanda Levendowski, ‘How Copyright Law Can Fix Artificial Intelligence’s Implicit Bias Problem’ (2018) 93 *Washington Law Review* 579; Michael Veale and Irina Brass, ‘Administration by Algorithm? Public Management Meets Public Sector Machine Learning’ in Karen Yeung and Martin Lodge (eds), *Algorithmic Regulation* (Oxford: OUP, 2019) 142.

71 Schauer, n 20 above, 635–636.

72 H. Surden, ‘Ethics of AI in Law: Basic Questions’ in Markus Dubber, Frank Pasquale and Sunit Das (eds), *The Oxford Handbook of Ethics of AI* (Oxford: OUP, 2020) 732.

73 See Andrew Le Sueur, ‘Robot Government: Automated Decision-Making and Its Implications for Parliament’ in Alexander Horne and Andrew Le Sueur (eds), *Parliament: Legislation and Accountability* (Oxford: Hart Publishing, 2016).

74 See Surden, n 72 above.

and lawmakers require designers to explain the algorithmic design (rather than the algorithm's predictions), this should satisfy the normative requirement for reasoned decision-making.⁷⁵ We term this argument the 'design transparency view'.⁷⁶

The second argument contends that inscrutability itself is a common trait of both human and machine decision-making. The human mind, the argument goes, is a 'black box' in the same sense that an inscrutable algorithm is: human decision-making happens at deep levels of our brains, concealed from conscious comprehension.⁷⁷ If we are content to live with humans making decisions that cannot be fully explained to their fellow citizens, machine inscrutability should not pose a problem either.⁷⁸ We term this argument 'universal inscrutability'.⁷⁹

Both the design transparency view and the universal inscrutability argument, in our view, fail to establish that inscrutable algorithms satisfy the requirement to provide reasons. This means that using inscrutable algorithms comes at normative costs. These costs should not be downplayed solely because the benefits of using such algorithms might outweigh them. We now turn to discussing the significance of reason-giving in administrative law to clarify what is at risk when using inscrutable algorithms. We then address the design transparency and the universal inscrutability arguments.

Giving reasons

In administrative law, it is generally considered essential that decision-makers offer reasoning for their decisions, whether concerning overall policies or specific applications of those policies.⁸⁰ This idea has been given various normative justifications. First, agencies have consequential grounds for supplying reasons, for example, to allow efficient external scrutiny, to detect biases, and to obtain public cooperation.⁸¹ Second, as David Dyzenhaus, Murray Hunt, and Michael Taggart explain, 'the principle of legality requires a duty on administrative decision-makers to give reasons for their decisions, and judges to defer to the extent that they find that the justification meets the applicable standard.'⁸² Third, reason-giving in administrative law is important for intrinsic, deontological reasons – the duty to give reasons forces administrators to treat subjects who are affected by their decisions as ends in themselves. As Jerry Mashaw observes, 'to be subject to administrative authority that is unreasoned is to be treated as

75 Deeks, n 4 above.

76 Note that this view generally holds that being able to explain how the algorithm is designed (including its flaws) satisfies the reasoning requirement.

77 See Bathaee, n 11 above, 891–892; Bell, n 8 above, 90.

78 See Huq, n 1 above, 643.

79 We do not discuss viewpoints suggesting inscrutability can be mitigated by technological solutions; see above under the heading 'Machine Learning and the 'Black Box' Problem'.

80 See Surden, n 72 above, 732.

81 See Mashaw, n 20 above, 103.

82 David Dyzenhaus, Murray Hunt and Michael Taggart, 'The Principle of Legality in Administrative Law: Internationalisation as Constitutionalisation' (2001) 1 *Oxford University Commonwealth Law Journal* 5, 6.

a mere object of the law or political power, not a subject with independent rational capacities.⁸³

Although reason-giving is an important normative ideal, it does not necessarily translate into an all-encompassing doctrinal requirement. In England and Wales, for example, while there is no general duty in administrative law to provide reasons for a decision, such a duty is often imposed by law or statute in cases of judicial or quasi-judicial decisions and administrative decisions that impact individuals.⁸⁴ Furthermore, the common law, as Lord Justice Elias remarks in *Oakley v South Cambridgeshire*, ‘is moving to the position [that] whilst there is no universal obligation to give reasons in all circumstances, in general they should be given unless there is a proper justification for not doing so’.⁸⁵ This position was further elaborated by the Supreme Court in *Dover District Council v CPRE Kent*, where Lord Carnwath canvassed the various European and domestic sources for an obligation to give reasons, stating that while there is no general duty to give reasons at common law, reasons will be required when they are necessary to allow courts to scrutinise administrative decisions.⁸⁶

In the US, it is generally accepted that courts are required to consider an agency’s rationale for its action, to determine ‘whether the [administrative] decision was based on a consideration of the relevant factors and whether there has been a clear error of judgment’.⁸⁷ Thus, a court may find that an administrative decision is ‘arbitrary or capricious’ (per the Administrative Procedure Act)⁸⁸ if the agency relied on factors which Congress has not intended it to consider, or if it entirely failed to take into consideration an important aspect of the problem.⁸⁹

An important qualification regarding the requirement of reason-giving is that sometimes an agency might satisfy the normative and legal requirement to give reasons if it supplies *second-order* reasons for its actions.⁹⁰ A first-order reason is a reason which describes in a straightforward manner why a certain decision is justified, ie, the fact that a person committed repeated offences in the past justifies that we treat them as dangerous in bail hearings. A second-order reason, on the other hand, does not provide a direct justification for a decision, but rather describes whether one should act on a first-order reason or refrain from acting on it.⁹¹ For example, when agencies face genuine uncertainty with regard to the consequences of their actions, they might decide their course of action on the basis of second-order reasons, such as the precautionary principle, or the need to make a decision quickly. In such situations, the agency might meet the reasoning requirement without explaining why its decision is justified

83 Mashaw, n 20 above, 104.

84 Cobbe, n 1 above, 648.

85 *Oakley v South Cambridgeshire* DC [2017] EWCA Civ 71 at [30]. See also Deeks, n 4 above, 1849.

86 *Dover District Council v CPRE Kent* [2017] UKSC 79.

87 *Citizens to Preserve Overton Park v Volpe* 401 U.S. 402, 416 (1971) (*Citizens to Preserve Overton Park*). See for example Jacob Gersen and Adrian Vermeule, ‘Thin Rationality Review’ (2016) 114 *Michigan Law Review* 1355, 1358.

88 Administrative Procedure Act, 5 USC, s 706(1).

89 *Motor Vehicle Manufacturers Association of the United States, Inc v State Farm Mutual Automobile Insurance Co* 463 U.S. 29, 43 (1983) (*Motor Vehicle Manufacturers*).

90 See Gersen and Vermeule, n 87 above.

91 Joseph Raz, *Practical Reason and Norms* (Oxford: OUP, 1975) 36.

on first-order grounds. That said, we must keep in mind that for the most part reason-giving (first- or second-order) remains a defining feature of administrative law. Inscrutable algorithms challenge these normative underpinnings, requiring us, as Roger Brownsword emphasises, to articulate benchmarks for legality and the compatibility of using new technological capabilities with those benchmarks (what he dubs ‘new coherentism’).⁹²

Having considered the importance of reason-giving in administrative law, we now turn to discuss arguments that attempt to defend the view that inscrutable algorithms can cohere with reason-giving – the design transparency view and the universal inscrutability argument.

The design transparency view and administrative law

The design transparency view holds that while it is hard to understand the algorithm’s *output* (ie, its prediction), it is not impossible to understand its *design*, and this ought to be enough to satisfy the reasoning requirement.⁹³ After all, algorithms are designed by humans, and the activity of programming involves making various choices – some of which are normative – that make the algorithms the way they are.⁹⁴ Thus, ‘it is not the case that machine decisions are bereft of justifying grounds. It is rather that those reasons are supplied at a point in time far removed from state action impinging upon the individual.’⁹⁵ On this view, an algorithm would meet the reasoning requirement if enough information regarding its general design were supplied.

We should note that we understand design-level information as *content-neutral*, ie, information that does not relate to the actual decisions being carried out by the administrative agency. For example, in a regular administrative setting, design-level explanations can specify the statutes that authorise action, procedures decision-makers must follow, requirements regarding oversight, etc. In machine learning settings, design-level information will provide details about the dataset used to train the algorithm, how the algorithmic recursive system was set up, what specific algorithmic technique was being used (for example deep learning, neural networks), etc. This type of data can be contrasted with general information relating more directly to the decisions being made, for example setting out the normative criteria that will be used by administrators. This conception of design-level information is helpful for the purposes of discussing the use of machine learning algorithms because the details regarding the way the algorithm actually made its prediction in a specific case generally remain hidden from us.

We grant that relevant information regarding an algorithm’s design can be supplied, at least in principle. We also agree that this information may help curb some problems associated with machine learning in general, such as

92 Brownsword ‘Artificial Intelligence and Legal Singularity’ n 14 above, 136, and 153–154.

93 See Joshua Kroll, ‘The Fallacy of Inscrutability’ (2018) 376 *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 1, 2.

94 *ibid.*

95 Huq, n 1 above, 674.

accuracy concerns and cognitive biases. Moreover, we do not completely rule out the possibility that such information is helpful in meeting the requirement for reason-giving. Ideally, this question should be examined on a case-by-case basis. However, we argue that in most areas of administrative law, even after this information is supplied, a normative cost is incurred by using inscrutable algorithms. This is because at least some of the relevant normative justifications for reason-giving require that reasons be supplied regarding the *specific decision* under examination. Therefore, while design choices may be relevant and even important, they leave something to be desired.

To understand why a normative cost remains even if information about design choices is supplied, recall that reason-giving is important at least partially for deontological reasons. Supplying reasons is a crucial part of what it means to treat individuals as subjects with independent rational capacities. From this perspective, it is doubtful that in most cases explanations regarding the design of an algorithm can satisfy the requirement of reason-giving. Explaining in broad detail how the algorithm works will not treat the subject of the decision as a rational agent capable of understanding for herself why the system worked the way it did *in this specific case* which affected her. To return to the examples discussed above, deciding whether to release a person on bail, to grant a government loan, or to send an inspector to a specific factory all require making determinations and predictions regarding specific individuals and companies. These subjects may legitimately demand an explanation regarding the rationale employed by the relevant government agency.

To further understand this point, it might be helpful to remove the unique attributes of machine learning for a moment. Imagine, for example, a human administrator who decides to deny a person's request for certain benefits. The administrator decides this in the usual manner, ie, with no algorithms involved. When asked for her reasons, the administrator refuses to explain to the subject of her decision why she decided to deny the request. Instead, she only supplies information regarding when, by whom, and according to what laws she, the administrator, was chosen to serve in her administrative capacity. In addition, the administrator provides information that shows that she is careful to avoid relevant biases. In this scenario, the administrator supplied design-level information. Imagine as well that this information is comprehensive and informative as far as it goes. Has the administrator treated the subject of her decision as a rational being worthy of respect? We doubt it. The information supplied succeeds in justifying the legal *architecture* immediately surrounding the decision, at least partially, but the decision itself was left obscure. Humans as rational agents demand explanations regarding instances in which state force has been wielded against them, and not about design choices in the abstract.

This example is illustrative for our discussion because the design transparency view asks us to settle for the same kind of information in the case of algorithms. Design-level information about an algorithm could theoretically be enough to answer queries about outputs in specific cases *if* its design influenced its predictions in a clear and straightforward manner (as a non-AI algorithm might operate). But inscrutable algorithms do not work in this way. Inscrutable algorithms are recursive or employ other advanced techniques while learning

for themselves in a manner not open for inspection. We cannot simply draw a clear, explainable line between their design and their output. Design-level information leaves the moment of prediction obscured. Yet, shedding light on this moment in the decision-making process is important for treating individuals as rational subjects. This is because treating someone as a rational subject means explaining to her why state force was applied to her *individually*, and not in the abstract.

Second, consider the idea that reason-giving promotes the rule of law and the principle of legality by requiring state actors to show that their decisions conform to the values, rules, and principles of the legal system.⁹⁶ How does design-level information fare on this account of reason-giving? We think that the answer might vary, depending on the specifics of the legal framework at hand. Imagine an administrative decision-maker tasked with assessing a person's threat level to determine their eligibility for bail. Assume also that the law does not mention machine learning algorithms. The principle of legality demands of this administrator an explanation as to how she fulfilled her legal duty in specific cases that came before her. Using a machine learning algorithm to aid her in accomplishing this task will hinder the decision-maker as she will struggle to explain the relationship between her legal duty – the duty to properly ascertain an individual's level of threat – and her decision in a given case. This is because the decision she is required to explain by the principle of legality is a *specific* decision regarding a *specific* person, and not the decision to establish an assisting algorithm that runs in the background and supplies predictions. This last action (the establishment of the algorithm) can indeed be adequately explained by design-level information and can also be justified as a means to the end of fulfilling the actual legal duty. But legality further demands in this case an explanation as to how this facilitating tool helped the administrator meet her legal duty in a specific case, and this further requirement cannot be met without output-level information.

Compare this case with a scenario in which the law itself specifically allows (or requires) the administrator to use machine learning tools to determine levels of danger. In this case, the administrator can easily explain how in each individual case the decision to be assisted by an algorithm to determine a legal outcome relates to her legal duties. This analysis, while short and inconclusive, hints that it might be better, from the point of view of the normative ideal of the rule of law, if administrative use of machine learning algorithms be based on express legal authorisation. Yet, to the best of our knowledge, currently most authorising laws do not expressly mention machine learning algorithms. Furthermore, our analysis also suggests that the rule of law justification for reason-giving might currently find many instances of the use of such algorithms problematic. An exception to the observation that current laws generally do not expressly regulate the use of algorithms is Article 22 of the EU's General Data Protection Regulation (GDPR), which applies also in the UK, and which generally guarantees 'the right not to be subject to a decision based solely on automated

⁹⁶ Dyzenhaus, Hunt and Taggart, n 82 above.

processing'.⁹⁷ Indeed, Article 22 GDPR is more consistent with the reason-giving principle in administrative law, it expressly (even if implicitly) regulates the use of algorithms in decision-making thus balancing (or attempting to balance) competing interests, and advances the normative ideal of the rule of law.⁹⁸ However, it is doubtful whether this norm solves the problem of inscrutability.⁹⁹ We address this question in more detail below.

Other, consequentialist justifications for reason-giving might be neutral in respect of the choice between design-level and output-level explanations. For example, if one worries about public cooperation, it indeed might be enough to supply only robust design-level information (though this is an empirical point which should be investigated further). In any respect, we believe there are sufficient grounds to conclude that at least some justifications for reason-giving are not fully satisfied by using design-level information only.

A possible objection to our argument is that design-level information supplies exactly the kind of second-order reasons to rely on machine learning predictions that is normatively and legally required. This type of information usually shows that the algorithm is accurate, does not encode biases etc, and therefore the agency has a second-order reason to rely on its predictions, even if the agency does not know the first-order reasons that justify the prediction. Vermeule and Gersen argue that sometimes 'agencies ... have excellent reasons to depart from idealized first-order conceptions of administrative rationality'.¹⁰⁰ Is this not a good example of such a case?

In response, consider these two points. First, Vermeule and Gersen show that the main reason to depart from first- to second-order reasoning in administrative law is uncertainty, ie, cases in which the facts underdetermine agency choices, and therefore 'there is no general way to arbitrate among [policy options], no further requirement of rationality that would knock out all but one approach'.¹⁰¹ In these scenarios, agencies can permissibly choose between options on the basis of second-order reasons, without disclosing first-order reasons, because the first-order reasons 'run out'. But with inscrutable algorithms the opposite is true: there *are* first-order reasons that determine that one outcome is decidedly more probable than the other, and those first-order reasons – supplied by the algorithm – supposedly justify a specific prediction. An agency that relies on algorithmic predictions would never claim that it is operating arbitrarily or that there is mirror-image reversibility between its options, as is the case with uncertainty.¹⁰² So even if agencies can permissibly rely on second-order reasons in situations of uncertainty, that in and of itself does not mean that they can do so in this case.

Other justifications for opting for second-order reasons that Vermeule and Gersen discuss (mean-variance trade-offs, speed, asymmetric cost errors, etc) do

97 Regulation (EU) 2016/679 General Data Protection Regulation, Art 22.

98 See also Julia Black and Andrew Douglas Murray, 'Regulating AI and Machine Learning: Setting the Regulatory Agenda' (2019) 10 *European Journal of Law and Technology* 1.

99 See Orla Lynskey and others, 'Machine Learning with Personal Data: Is Data Protection Law Smart Enough to Meet the Challenge?' (2017) 7 *International Data Privacy Law* 1.

100 Gersen and Vermeule, n 87 above, 1357.

101 *ibid*, 1385.

102 *ibid*, 1386.

not apply here either. For example, Vermeule and Gersen contend that agencies sometimes possess ‘tacit knowledge’ that is hard or too costly to convey to third parties.¹⁰³ But the predictions that are generated by inscrutable algorithms cannot be classified as tacit knowledge. In fact, they seem to be quite the opposite: Vermeule and Gersen follow Hayek in explaining that tacit knowledge is generated by ‘circumstances of time and place’ that ‘by [their] nature cannot enter into statistics and therefore cannot be conveyed to any central authority in statistical form’.¹⁰⁴ But machine predictions are based on statistics, understood broadly – the predictions are the result of general computational steps taken by a program with no first-personal ongoing experience or expertise of the kind that would generate tacit knowledge.¹⁰⁵ So, it seems that Vermeule and Gersen’s framework does not easily justify why machine learning predictions may be relied upon using only second-order reasons (ie, design-level information).

Second, even if we accept that there are good second-order reasons to rely on machine learning predictions, this conclusion coheres well with our general argument. Recall that we argue that (1) relying on inscrutable predictions is not without normative cost; and (2) this cost may be worth paying for reasons that have to do with the predictions’ accuracy, the ability to generate them quickly, etc. So, even if second-order reasons eventually point in the direction of mass use of predictions generated by inscrutable algorithms, this does not mean that there is *no* normative problem with using algorithmic predictions – the deontological and other considerations that we discuss above still apply. In fact, note that Vermeule and Gersen’s view can help explain why the normative cost in the case of inscrutable algorithmic decision-making is actually *greater* than in others: in cases of uncertainty, for example, not supplying first-order reasons for subjects of the decision is not a problem from a deontological point of view because there are no relevant first-order reasons to supply. But in the case of inscrutable algorithms, the administrative agency is in fact acting directly because of first-order reasons (supplied by the machine prediction) while simultaneously not explaining them to the subject of the decision. This creates the normative cost outlined above.

Universal inscrutability

The second argument suggesting that inscrutable algorithms do not pose a specific concern in administrative decision-making follows a diametrically opposed path to the design transparency view: instead of arguing that inscrutable algorithms are not ‘black boxes’, the universal inscrutability argument holds that these algorithms are indeed ‘black boxes’ – but so is every other decision protocol, including standard human decision-making. If human cognition is as much a mystery to us as machine learning algorithms, the argument goes, we have no *special* reason to worry about the latter.

103 *ibid*, 1396.

104 *ibid* (quoting F. A. Hayek, ‘The Use of Knowledge in Society’ (1945) 35 *American Economic Review* 519, 524).

105 *ibid*, 1399.

We accept, for the sake of the argument, that we do not have adequate access to the biochemical processes that explain human decision-making. To that we might add that it is not clear, philosophically, if a biochemical explanation can ever ground a proper account of human cognition.¹⁰⁶ But even granting this, we argue that the universal inscrutability argument fails. Our argument, in short, is that equating human thought to machine processing is not persuasive when it comes to making predictions. When humans make predictions, biochemical activities, to which we do not have full access, are certainly taking place in our brains. However, this is beside the point, because biochemical processes only *causally explain* our reasoning but do not *justify* them, and we usually place little weight on such causal explanations. The moral justifications for reason-giving (and doctrinal law too) care little about the causal forces behind a decision. Instead, the focus is on *normative justifications*: we do not ask what causally brought about a decision but what normative reasons support or justify it. As Michael Veale and Irina Brass observe, ‘there is a difference between *explanation* in law and *justification*, and tracing back why a decision occurred ... may not serve to justify it’.¹⁰⁷

Roughly speaking, reasons are facts that explain why an action is good, preferable, or valuable to a certain degree.¹⁰⁸ When making predictions, we often introduce the reasons that justify our conclusions. For example, we say ‘you should bring a bottle of water to this walk because you tend to get thirsty during long walks’. In philosophical terms, the law’s objects of analysis, at least in this regard, are *motivating reasons*. A motivating reason is a reason that the agent believes justifies or favours her action and in light of which she acts. Motivating reasons are commonly distinguished from ‘normative reasons’, which are defined as reasons that *actually* support or justify an action (that is, justify an action from an objective point of view).¹⁰⁹ For example, a person’s motivating reason for not eating a piece of cake might be her mistaken belief that it contains milk, together with the fact that she is lactose intolerant. In this example, we can identify a motivating reason for the action (the putative fact about the cake containing milk which generated a prediction that she would be harmed by eating it), but not a normative reason. Note that both normative and motivating reasons are different from causal explanations: a causal account for the prediction regarding the damage from the cake would take a very different form and importantly would not contain any reasons (ie factors that count in favour of the decision). Rather, it would contain descriptions of certain parts of the brain, certain groups of neurons, etc.

Let us now return to the examples discussed earlier, of the probability of reoffending or repaying a loan. To make predictions regarding these issues in the standard manner, we routinely use past experience and data, employ common statistical methods, etc. These techniques ideally offer *justifications* for our

106 See David Chalmers, ‘Facing Up to the Problem of Consciousness’ (1995) 2 *Journal of Consciousness Studies* 200.

107 Veale and Brass, n 70 above, 131.

108 See Joseph Raz, *Engaging Reason: On the Theory of Value and Action* (Oxford: OUP, 2002).

109 See Maria Alvarez, ‘Reasons for Action: Justification, Motivation, Explanation’ (The Stanford Encyclopedia of Philosophy, 2017) at <https://plato.stanford.edu/archives/win2017/entries/reasons-just-vs-expl/> [<https://perma.cc/G9LH-7KKG>].

predictions: putative facts which make our belief in these predictions justifiable, preferable, etc. We never offer biochemical causal explanations because they are beside the point since they cannot justify our beliefs in certain predictions: facts about the way in which our brains function cannot *justify* our beliefs about crime or loans, but only causally explain them.¹¹⁰ So, we are left with the demand for actual motivating reasons for making our predictions. In the words of Michael Pardo and Dennis Patterson, ‘the difficult questions regarding the law’s role vis-à-vis moral decision-making cannot be resolved by appealing to physical states of the brain’.¹¹¹

A quick glance at case law illustrates administrative law’s focus on motivating reasons and not on causal explanations.¹¹² Take, for example, the seminal case of *Flannery v Halifax Estate Agencies*,¹¹³ where Henry LJ states: ‘[T]he duty to give reasons . . . is a function of due process, and therefore of justice. Its rationale has two principal aspects. The first is that fairness surely requires that the parties especially the losing party should be left in no doubt why they have won or lost The second is that a requirement to give reasons concentrates the mind; if it is fulfilled, the resulting decision is much more likely to be soundly based on the evidence than if it is not.’ Similarly, in the case of *Department of Commerce v New York*¹¹⁴ (*Department of Commerce*), in which the US Supreme Court reviewed a decision by the Secretary of Commerce to reinstate a question regarding citizenship in the 2020 census, one of the key questions the court asked was: What are the reasons that the Secretary believes justifies this decision (ie, what are the decision’s motivating reasons)? Chief Justice Roberts explained that ‘[t]he reasoned explanation requirement of administrative law, after all, is meant to ensure that agencies offer genuine justifications for important decisions, reasons that can be scrutinized by courts and the interested public’.¹¹⁵ The court refused to inspect the normative reasons for the decisions, ie, the reasons that objectively justify it, and instead, following established precedent on this matter,¹¹⁶ focused on the motivating reasons alone,¹¹⁷ thus adhering to the notion that the court’s duty is to inspect the specific decision based on the considerations that the agency actually thought convincing at the time. Note that the court is interested in the *official* motivating reasons behind the decision, ie, administrative bodies must provide an official account of their motivating reasons. But an official motivating reason is not a different type of reason; rather, it is a way of presenting motivating reasons in a manner that is consistent with

110 See Oliver Goodenough and Micaela Tucker, ‘Law and Cognitive Neuroscience’ (2010) 6 *Annual Review of Law and Social Science* 61.

111 Michael Pardo and Dennis Patterson, ‘Philosophical Foundations of Law and Neuroscience’ (2010) 2010 *University of Illinois Law Review* 1211, 1214.

112 For example, see *Motor Vehicle Manufacturers* n 89 above, 57; *Citizens to Preserve Overton Park* n 87 above 420.

113 [2000] 1 WLR 377 (CA) 381, 381.

114 588 U.S. ____ (2019); 139 S. Ct. 2551 (2019).

115 *ibid.*, 27.

116 See *Vermont Yankee Nuclear Power Corp v Natural Resources Defense Council, Inc* 435 U.S. 519, 549 (1978).

117 n 114 above, 25–26.

the requirements of administrative law (for example via affidavits, etc).¹¹⁸ In any case, what the court did not even contemplate doing was to review the Secretary's decision using explanatory, causal mechanisms such as biochemical processes. This is because explanatory modalities simply cannot 'be scrutinized by courts and the interested public' in the manner that Chief Justice Roberts demands – they fail to address the *contents* of the decision they explain and to count in favour of the decision. To take the example from *Department of Commerce*,¹¹⁹ they can only present facts about a certain brain, and not about illegal immigrants and citizens, but only facts about the question at hand which were actually considered – about citizens, illegal aliens, and census data – can count as justifying and motivating reasons. Therefore, such biochemical information cannot do the work of normative justification.

We see, then, that universal inscrutability arguments might incorrectly assume that biochemical causal explanations do (or could do) relevant work in justifying administrative decisions, when, in fact, they do not. This is because biochemical causal explanations can only explain decision-making as it relates to the physical state of a specific brain. Administrative law focuses on motivating reasons, ie, factors that count in favour of the prediction at hand and which were actually considered by the relevant agent. But as far as motivating reasons go, humans are not 'black boxes' – it is very common to ask what reasons motivate a given person's prediction, to accept their answers as revealing their justifying reasons, and to scrutinise that decision on the basis of the answers given. Inscrutable algorithms that rely on advanced methodologies (such as deep learning or neural networks), by contrast, are mostly black boxes for the reasons outlined earlier: the processes they employ are opaque to us, they are chosen by the algorithm after we have already programmed it, and their predictions rely on subtle big data correlations which we cannot ourselves contemplate. Therefore, the difference between human cognition and predictive machine learning remains and the argument for universal inscrutability fails.

Admittedly, humans may also give false reasons for their actions (knowingly or unknowingly), for example if they base their decisions on biases, prejudice, etc. In this limited sense humans are a type of 'black box' themselves. But the fact that we sometimes make mistakes in discerning other people's reasons (or do not understand them) does not make them inscrutable, or beyond our understanding. After all, we tend to believe that people know the reasons that motivate them, and we do not treat explanations regarding human reasoning as generally untrustworthy. Rather, we assume that usually a human agent can reveal her reasoning in a justified manner, and this is the basis of ordinary human interaction both in and outside of government. In other words, error-free judgement is not a requirement for a process or output to be understandable (in this sense, the objection falsely equates the ability to make an error in a specific case with the general characteristic of inscrutability).

118 Furthermore, this example shows that complex bodies involving multiple decision chains can have motivating reasons. cf Haim Abraham, 'Tort Liability, Combatant Activities, and the Question of Over-Deterrence' (2022) 47 *Law & Social Inquiry*, 885.

119 n 114 above.

Moreover, law in general, and administrative law specifically, contains mechanisms that scrutinise decisions for implicit mistakes, including sometimes mistakes about reasoning that the agent makes unwittingly. For example, in American administrative law, a court may inquire more deeply into the reasoning of agencies and set aside the official administrative record if ‘a strong showing of bad faith or improper behavior’ can be demonstrated.¹²⁰ Agencies thus make decisions while knowing that they must be transparent about their reasoning in order to avoid such scrutiny. Such a standard cannot be applied to algorithmic predictions as they never show ‘bad faith or improper behavior’. This and other legal instruments make mistakes about reasoning less common and further complicate the idea that human reasoning is inscrutable in the same manner as machine learning predictions.

One should therefore not equate the possibility of a *mistake* in a specific account of reasoning – a scenario we deal with every day, and which does not make reasoning itself mysterious – with inscrutability. This is true in administrative settings as well, as we note above. Lawyers are highly trained in thinking critically about the reasons given by office holders. They can be very critical of answers given by administrative bodies precisely because office holders think in principle like them. This is not the case with inscrutable algorithms: we usually cannot even begin to critically assess their predictions. These algorithms do not reveal their motivating reasons at all: they generate predictions but fail to explain their reasoning in the sense described above. And so, the difference between human cognition and machine learning algorithms with which we started remains; it cannot be bridged by the universal inscrutability argument.

This section demonstrates that when we choose to use inscrutable algorithms’ predictions we incur a normative cost, because subjects of the law will fail to fully understand the decisions that apply to them. It is reasonable to believe that this cost is worth paying in many cases, given the usefulness of machine learning algorithms. But it is a cost, nonetheless. We now move to discuss a further potential cost that is more connected to the long-term, systemic effects of using these technologies.

THE RE-ENCHANTMENT OF THE ARTIFICIAL

If a normative cost is indeed incurred when using inscrutable algorithms to facilitate state action because of the requirement to give reasons, then when such an algorithm’s prediction is used by an administrative body, something is lost. Namely, a particular individual is unable to satisfyingly understand at least part of the reasons that ground the use of force by the state in her case. This residue of incomprehensibility that is generated by inscrutable algorithms may accumulate. If the use of inscrutable algorithms continues to grow, once enough inscrutable algorithms are used across a wide enough range of policy domains, the residue that each subject experiences for herself in a particular case

120 *Overton Park, Inc v Volpe* 401 U.S. 420 (1971). This exception was used by the court in *Department of Commerce* n 114 above, to determine the agency’s actual reasoning.

could potentially translate into a shared interpersonal phenomenon that would affect society as a whole. Some of us would be denied social benefits without wholly understanding why; others would be deemed dangerous on the basis of predictions they do not understand; still others would have their immigration requests denied on the basis of an algorithmic logic that is not transparent to them. Additionally, more and more people would experience this second-hand because of relatives or friends affected by inscrutable algorithmic predictions. After a certain ubiquity threshold is reached, the *possibility* of it happening to each individual will hover in the background and inform public life. Meeting the ‘black box’ will not be confined to a series of private encounters between a subject and the state, but rather it could become a constitutive part of our shared public existence. It is hard to identify exactly when this ubiquity threshold will be met, as it includes societal elements which are hard to quantify or predict and it contains both quantitative elements (the number of people directly and indirectly affected) and qualitative elements (the extent to which individuals and groups were affected). However, it is not difficult to imagine the impact that reaching this threshold will have on public life and the way subjects of the law experience it. The residue of incomprehensibility might affect our public life not only as a community committed to the primacy of moral reason, but also as a community whose administrative law is committed to respecting individuals’ agency.¹²¹

We stress that this possibility is only a potential danger and not a certainty, and that the price is possibly (or even likely) worth paying for more efficient and accurate public services. However, when contemplating the issue of inscrutable algorithms, we must acknowledge that the public sphere might experience a substantial transformation. In addition, this problem does not apply to verifiable algorithms (like facial recognition software) as such algorithms do not make predictions but rather make determinations that humans can validate.

A good way to approach the issue of the cumulative affect of using inscrutable algorithms to support decisions in individual cases is through the Weberian idea of enchantment. Employing this terminology, we argue that using machine learning algorithms could trigger a new kind of re-enchantment.

As we note earlier, an enchanted world is a world which is at least in part incalculable; some elements of the world are hidden from us, unknowable, and therefore mysterious.¹²² This something ‘more’ which partially governs the natural world or human behaviour was given many names in the past, such as magic or divine will. The important thing is that the process of disenchantment, which Weber claims has been going on for centuries,¹²³ entails a reorientation of human experience – disenchanted people experience the world as un-mysterious *in principle*, even if they themselves have not bothered to study physics, astronomy, or politics.¹²⁴ As Weber himself articulates, ‘[i]t

121 Brownsword, ‘Rethinking the Rule of Law’ n 14, above 83–84.

122 See David Trubek, ‘Max Weber on Law and the Rise of Capitalism’ [1972] *Wisconsin Law Review* 720; Blank, n 18 above, 643.

123 Weber, n 15 above.

124 See Patrick Sherry, ‘Disenchantment, Re-Enchantment, and Enchantment’ (2009) 25 *Modern Theology* 369.

is the knowledge or conviction that if only we wished to understand them [the conditions under which we live] we could do so at any time. It means that in principle, then, we are not ruled by mysterious, unpredictable forces, but that, on the contrary, we can in principle control everything by means of calculation. That, in turn, means the disenchantment of the world.¹²⁵

Armed with this understanding of disenchantment, let us focus on the potential long-term, societal effects of using inscrutable algorithms to facilitate state action. What would it be like to live in potential constant contact with state decisions that exert power on citizens based, at least partially, on inscrutable mechanisms?¹²⁶ Abstracting from any specific administrative decision, how would public life be experienced by the subjects of the law, specifically administrative law?

In our view, the use of inscrutable algorithms could re-introduce into public life an element which would be permanently beyond the reach of human cognition and, therefore, mysterious and incomprehensible. This facet of public life would act as an ‘other’ to us, something which we do not understand, even though it was made by us. We term this phenomenon *the re-enchantment of the artificial*. Surely we would be assured that the processes through which agencies decide how to use state power are rational – the machines (allegedly) do not decide arbitrarily. But this assurance might not be enough; it might not prevent us from *experiencing* social life as governed by forces we do not understand. Weber explains that disenchantment is connected to the notion that if one only wanted to, one could find an explanation at any time.¹²⁷ But this idea is missing in the case of inscrutable algorithms – even given all the time in the world, we could not, in principle, understand why the predictions turned out as they did.

Scholars have already documented many instances of re-enchantment in modern life, both generally and within the law.¹²⁸ Therefore the idea of re-enchantment is not new in itself. However, we argue, the re-enchantment of the artificial is unique because it has a different internal logic than other forms of re-enchantment. While most types of re-enchantment bring some non-rational element (for example magical, mystical) into everyday life, the re-enchantment of the artificial makes facets of life unintelligible for humans *by bringing the rational modality itself to its logical endpoint*. In other words, instead of the usual dichotomy we find in Weberian thought between rational thinking (leading to disenchantment) vs non-rational, ‘magical’ thinking (which encourages re-enchantment), inscrutable algorithms demonstrate that rationality can, by itself, become an agent of re-enchantment, if pushed to the extreme.

Let us take a closer look at the unusual logic of this phenomenon. On the one hand, one might view inscrutable algorithms as the pinnacle of the type of

125 Weber, n 15 above, 12–13.

126 This is a phenomenological question. For a discussion of the importance of the phenomenological perspective, see Omri Ben-Zvi and Eden Sarid, ‘Legal Scholarship as Spectacular Failure’ (2018) 30 *Yale Journal of Law & the Humanities* 1. On Weber’s phenomenological conception of enchantment, see Kennedy, n 25 above, 1057. It is quite clear that Weber sees this category as relating to the human experience of meaning. See Weber, *ibid*, 8, 13, 15, 28, 30.

127 Weber, *ibid*, 12–13.

128 Jane Bennett, *The Enchantment of Modern Life: Attachments, Crossings, and Ethics* (Newark, NJ: Princeton University Press, 2001); Blank, n 18 above.

rational thought that is associated with disenchantment. This is because these algorithms can calculate and predict social phenomena more accurately than ever before. If disenchantment is about controlling important facts of social life 'by means of calculation', inscrutable algorithms could be seen not as an obstacle but as a natural corollary to the idea of disenchantment by bureaucratic governance. This is why inscrutable algorithms push rationality to its logical endpoint: if administrative rationality is at least partly preoccupied with responding to the relevant factors facing agencies and acting on the basis of reasons,¹²⁹ then inscrutable algorithms are from this point of view *hyper*-rational (assuming they work properly, without biases, etc): they take into account millions (if not billions) of factors which humans cannot work through in a timely manner, and they recognise and respond to subtle correlations to generate predictions that are usually much more precise than those generated by humans. The machine learning state is thus the result of the type of thinking that favours rational calculability.

However, this description of machine learning as furthering the cause of calculability is persuasive only on the condition that one ignores the fact that *we are not the ones doing the calculating*. The actual process of rationalisation leading to a machine prediction is hidden from us, in principle, because of inscrutability. This means that humans in the machine learning state would encounter many predictions which are promised to be highly rational but which they do not truly understand.

In this phenomenon of inscrutable algorithms, then, we have two elements: first, the *first-person experience* of meeting an 'other' which we do not understand but which is connected to state power in important ways, because it is used to facilitate administrative decisions; and second, a *promise* that this 'other' is in fact more rational than us and therefore we should be content to leave the predicting tasks to it. Will the mere fact of the promise (perhaps together with the ability to point to past successes) be enough to mitigate the unsettling first-person experience? We cannot be sure, and therefore we only talk of a danger that *might* occur. But we have some reason to believe that at the very least, the first-person experience of encountering and interacting with modes of rationality so opaque to us will leave a mark and in the long run could impact the way we experience social life. This effect is what we signify with the idea of re-enchantment.

Our argument coheres with and continues existing literature that has already started to document the way algorithms affect our shared ways of thinking about social institutions. In this sense, algorithms possess not only computational-functional power but also 'social power'.¹³⁰ As David Beer explains, 'the algorithm can be part of the deployment of power, not just in terms of its function but also in terms of how it is understood as a phenomenon'.¹³¹ Many changes may occur as a result of the 'dissolution of AI and digital technologies into the

129 See Gersen and Vermeule, n 87 above.

130 See David Beer, 'The Social Power of Algorithms' (2017) 20 *Information, Communication and Society* 1.

131 *ibid*, 11.

bloodstream of society',¹³² and the experience of the world as less sympathetic to human cognition and harder to understand is one of them.¹³³

We can already see the idea of re-enchantment due to algorithmic decision-making beginning to have some influence in society, albeit sporadically. For example, Alexander Campolo and Kate Crawford argue that many discussions about machine learning present machine learning techniques as magical and mysterious – a discourse they term 'enchanted determinism'.¹³⁴ Other studies show that artificial intelligence in private life is already thought of in religious and mysterious terms by the engineers themselves,¹³⁵ as well as the general public (ie, the rise in use of the phrase 'blessed by the algorithm' to describe everyday encounters with machine learning algorithms).¹³⁶ These examples suggest that some degree of re-enchantment is already occurring due to everyday encounters with predictive algorithms in private life (for example Siri, Alexa etc).¹³⁷ We argue that the bigger the decision at hand, the more weight should be given to the fact that the decision stems from inscrutable mechanisms. Therefore, it is not surprising that most encounters with AI today do not generate a re-enchantment.

As additional support for our view, consider another type of encounter which can trigger re-enchantment. In her book on the idea of enchantment, Jane Bennett notes that 'cross-species encounters' (as can be found in movies, for example) generate an experience of re-enchantment because they evoke a sense of the mysterious and unexplained in us.¹³⁸ Machine decisions could be viewed as such encounters, only with a different type of object. In our view, the fact that algorithms are man-made should not a priori be considered to reduce the level of enchantment which could be generated. After all, as Nick Bostrom maintains, '[a]n artificial intelligence can be far less human-like in its motivations than a space alien'.¹³⁹

If inscrutable algorithmic predictions determine more and more facets of our social life and are incorporated more strongly in the bureaucracy, this re-enchantment could be intensified. Thus, the idea of being 'blessed' (or 'cursed') by the algorithm – that is, being fortunate or unfortunate, seemingly (to us) at random, according to some obscure algorithm's decision or prediction – could become a popular way in which we mediate to ourselves what goes on in our lives. A society of people who understand themselves as being routinely subject to the 'blessing' or the 'curse' of the algorithm-bureaucracy is therefore experiencing a more enchanted experience of social life.

132 Anthony Elliott, *The Culture of AI: Everyday Life and the Digital Revolution* (Abingdon: Routledge, 2019) 107.

133 See Sofia Ranchordas, 'Empathy in the Digital Administrative State' (2022) 71 *Duke Law Journal* 1341.

134 See Alexander Campolo and Kate Crawford, 'Enchanted Determinism: Power without Responsibility in Artificial Intelligence' (2020) 6 *Engaging Science, Technology, and Society* 1.

135 Roberto Musa Giuliano, 'Echoes of Myth and Magic in the Language of Artificial Intelligence' (2020) 35 *AI & Society* 1009.

136 Beth Singler, "'Blessed by the Algorithm": Theistic Conceptions of Artificial Intelligence in Online Discourse' (2020) 35 *AI & Society* 945.

137 Pasquale, n 11 above.

138 Bennett, n 128 above.

139 Bostrom, n 24 above, 72.

The use of algorithmic predictions to guide state action could bring with it a higher degree of enchantment than is currently experienced, because algorithms that are used by the state are not only *interacting* with us in mundane ways, as machine learning processes currently do; rather, they will be partially responsible for *governing us*, as they will provide partial justification for the use of force by the state in many administrative scenarios. The experience of being *subject to force* because of factors one does not understand is substantially different from simply interacting with a programme that offers suggestions to use in everyday life. Therefore, there is a normative difference between, on the one hand, Netflix or Siri suggesting which movie to watch or what route to take and, on the other, an algorithm that helps in deciding whether we will receive social benefit payments or be granted bail.

Furthermore, a large body of research now suggests that people's acceptance of authorities' decisions, as well as their acceptance of legal rules, is linked with their assessment of the fairness of the procedures and processes through which the decision has been made and the rules applied.¹⁴⁰ While inscrutable algorithms may be more accurate and less biased than human decision-makers, and while experts may even be able to understand them (to varying degrees), as long as they are not *perceived* as fair by the affected individuals and by society at large, the process of re-enchantment will only deepen.¹⁴¹

Still, is such a re-enchantment a problem for us generally? We use the term 'enchantment' descriptively rather than normatively, but note that the ability to understand fully what goes on specifically in public life is considered valuable and morally important: 'as rational beings we cannot but want our lives to have made rational sense, to add up to a story not only of whats but of *whys*.'¹⁴² Centering public life around explanations which we understand is also important to society as a whole: 'one central aspect of the common good lies in what we might call the moral intelligibility of our lives. A community is worse off to the extent that its members are unable to make moral sense of the lives that they and their fellow citizens lead.'¹⁴³ A more re-encharmed society is therefore normatively lacking to some extent, at least in this respect.¹⁴⁴

We emphasise, again, that this is a concern rather than a deterministic outcome, and that it may be wholly rational to prefer paying the price of inscrutability in exchange for its advantages. But the choice must be a conscious one. Our discussion presents another facet of this choice and shows that the reasoning requirement in administrative law is crucial to public life and should not be summarily discarded.

As we note above, some forms of regulation today attempt to deal with the potential problems raised by algorithmic decision-making. For example,

140 Tom R. Tyler and E. Allan Lind, 'Procedural Justice' in Joseph Sanders and V. Lee Hamilton (eds), *Handbook of Justice Research in Law* (New York, NY: Springer, 2001) 65.

141 See Lily Morse and others, 'Do the Ends Justify the Means? Variation in the Distributive and Procedural Fairness of Machine Learning Algorithms' (2021) 181 *Journal of Business Ethics* 1083.

142 John Gardner, 'The Mark of Responsibility' (2003) 23 *OJLS* 157, 158.

143 David Luban, Alan Strudler and David Wasserman, 'Moral Responsibility in the Age of Bureaucracy' (1992) 90 *Michigan Law Review* 2348, 2354.

144 See Brian Sheppard, 'Warming Up to Inscrutability: How Technology Could Challenge Our Concept of Law' (2018) 68 *UTLJ* 36.

Article 22 of the GDPR states that: ‘The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.’¹⁴⁵ While this response presents a coherent attempt to confront other problems which machine learning algorithms raise (mainly the ‘person in the loop’ normative requirement),¹⁴⁶ we believe it is lacking as far as the problems of inscrutability are concerned. For giving natural persons the right to a human decision only reproduces the dilemmas of inscrutability on the level of the human decision-maker. The decision-maker must decide whether to accept or reject an algorithmic prediction and then to justify this decision. But how can she justify relying on predictions that are based on an algorithm which she does not understand? Thus, whenever an administrative body places much weight on an algorithmic prediction in making a decision (and, after all, making predictions which will be utilised is the point of having these algorithms in the first place), the problems of inscrutability return in full, even if a human was in the loop. How can reliance on an inscrutable algorithm be explained to the subjects of the law in a manner which respects them fully? It is true that it was a human who made the decision to rely on the algorithm, but on what grounds? After all, the decision-maker herself does not understand the basis of the prediction. And so, we return to the same type of dynamic we elaborated above.

CONCLUSION

Society has long been on a journey of disenchantment – the process of understanding and explaining for ourselves the natural world and all areas of human experience. Administrative law provided a clear manifestation of this process in public life, as administrative bodies are thought to make reasoned decisions based on expertise and reason rather than other factors such as higher powers, magic, or pure political will. In this article, we have argued that if the use of inscrutable machine learning algorithms in administrative decision-making continues to grow, it might trigger a shift away from this path and towards making decisions based on unexplainable mechanisms – predictions made by machine learning algorithms which we cannot comprehend or explain. We call this process *the re-enchantment of the artificial*. Much more needs to be said about this topic before one can determine if the price of a possible re-enchantment is not worth the gains of using machine learning by administrative bodies. Yet, we must be aware that the use of inscrutable algorithms by the administrative state not only alters the way in which decisions are made; it also presents a possible fundamental shift in our understanding of administrative law and the public sphere.

145 See n 97 above.

146 See Huq, n 1 above, 622–624.