

From Rule of Law to Algorithmic Rule by Law

In Chapter 3, I concretised the six principles that constitute the rule of law in the EU legal order in order to develop a normative analytical framework for the purpose of this book's discussion. Drawing on this framework, in this chapter I can now revisit each of these principles and carry out a systematic assessment of how public authorities' reliance on algorithmic regulation can adversely affect them (Section 4.1). I then propose a theory of harm that conceptualises this threat, by juxtaposing the *rule of law* to the *algorithmic rule by law* (Section 4.2). Finally, I summarise these findings and outline the main elements that should be considered when evaluating the aptness of the current legal framework to address this threat (Section 4.3).

4.1 ALGORITHMIC REGULATION AND THE RULE OF LAW

How do the six rule of law principles fare under the increased use of algorithmic systems to inform and adopt administrative acts? In this section, I analyse respectively how such use affects the principle of legality (Section 4.1.1); legal certainty (Section 4.1.2); the prohibition of arbitrariness (Section 4.1.3); equality before the law (Section 4.1.4); judicial review of government action (Section 4.1.5); and the separation of powers (Section 4.1.6). While I assess each of these principles separately, it should be noted that their entwined nature and common purpose renders many observations relevant across the board.

In my evaluation, I draw not only on relevant scholarship, but also on concrete illustrations of how algorithmic regulation is already used by public authorities today. It should be noted that the variety of algorithmic systems deployed in the public sector is enormous, both in terms of technique and purpose, and in terms of application domain. In the sections below, in line with the research aims of this book, I have deliberately selected examples of algorithmic regulation that pose a risk to the rule of law (without claiming the inexistence of illustrations which do not demonstrate such risk). Moreover, I have specifically selected examples of

algorithmic regulation in liberal democracies, to highlight that the risks posed thereby are not limited to autocratic regimes. Many of the examples concern the US and the UK, not only because they are frontrunners in the adoption of algorithmic regulation but also for the simple reason that, over the years, information about their adoption of algorithmic regulation has become publicly available. In fact, until today, in many EU countries information about public authorities' use of algorithmic systems is lacking, and research about their effects has not or barely been conducted. For this analysis, I hence selected my examples based on three criteria: (1) the algorithmic system is deployed by a public authority from a liberal democracy, (2) the examples represent uses of algorithmic regulation in different public sector domains and (3) there is some level of information available about the system's use and effects.

Based on my analysis of relevant scholarship and concrete illustrations, I conclude that, in certain situations, public authorities' reliance on algorithmic regulation can indeed hamper the six rule of law principles. This does not mean that all uses of algorithmic regulation *necessarily* lead to an adverse impact on the rule of law – or, more precisely, such generalisation cannot be concluded from my casuistical evaluation. However, it does imply that algorithmic regulation *can* lead to an adverse impact on the rule of law, and that this needs to be taken into account if the aim is to protect this value and the protective role of the law in liberal democracy.

4.1.1 Legality

As noted above, the primary requirement of the legality principle entails that public authorities and officials comply with the law, and that the measures they adopt for its implementation are congruent therewith, as well as proportionate. At first sight, reliance on algorithmic systems to inform or take administrative acts could contribute to the better fulfilment of this requirement. To the extent legal rules are structured around *if-then* premises, they could in theory lend themselves rather well to a transformation from text to code.¹ Moreover, reliance on algorithmic systems for the adoption of administrative acts could prevent that public officials deviate from the codified requirements and hence that they deviate from the law, since the relevant legal requirements can be codified straight into the architectural design of the system (*law-by-design*).²

¹ See Lawrence Lessig, *Code: And Other Laws of Cyberspace* (Basic Books 1999); Dag Wiese Schartum, 'From Legal Sources to Programming Code: Automatic Individual Decisions in Public Administration and Computers under the Rule of Law' in Woodrow Barfield (ed), *The Cambridge Handbook of the Law of Algorithms* (1st edn, Cambridge University Press 2020). See also Pascal D König 'Dissecting the Algorithmic Leviathan: On the Socio-Political Anatomy of Algorithmic Governance' (2020) 33 *Philosophy & Technology* 467, 470.

² See, e.g., Daria Gritsenko and Matthew Wood, 'Algorithmic Governance: A Modes of Governance Approach' (2022) 16 *Regulation & Governance* 45.

Indeed, the very ontology of code leads to the fact that, once legal rules are codified, their prescriptive nature actually becomes descriptive.³ Legal rules are then no longer normatively guiding the actions or decisions of public officials, but they are applied almost in real time by an algorithmic system. This can be a desirable feature if the aim is to counter rule-deviating behaviour. Moreover, even where an algorithmic system is only used to inform an administrative act rather than to adopt one, public officials that intend to deviate from the algorithmic suggestion may face an additional hurdle to do so, as the deviation from a pre-codified norm typically requires an additional step, for instance in the form of a justification that enables one to override the system. Consequently, reliance on algorithmic systems could deter the deviation from codified norms in both direct and indirect ways, and thus theoretically also contribute to the deterrence of illegal or corrupt behaviour by public officials.⁴ Unfortunately, these very ‘advantages’ can also be considered as problematic, and as potentially hindering the fulfilment of the legality principle.

4.1.1.a Lost in Translation

The aspiration to codify legal rules and concepts in order to automate administrative acts is not as forthright as it seems. Transforming legal text to code requires a translation process, as the law rests on linguistic concepts that embody social constructs, which are open ended in terms of their interpretability.⁵ These concepts are naturally understood by human interpreters as belonging to a broader societal context, and as being multi-interpretable. Moreover, within legal texts, a wide variation exists in the use of language, from very open-ended and abstract principles to more specific and prescriptive formulations (evoking the well-known distinction between rules and principles, and their respective merits and pitfalls).⁶

Indeed, legal rules are inherently indeterminate, a feature that Julia Black notes as arising “*in part from the nature of language, in part from their anticipatory nature,*

³ Laurence Diver, ‘Interpreting the Rule(s) of Code: Performance, Performativity, and Production’ [2021] MIT Computational Law Report 2 <<https://law.mit.edu/pub/interpretingtherulesofcode/release/1>> 6. As regards the law and its effects, it can be noted that Mireille Hildebrandt draws a distinction between the *descriptive* nature of [law as] speech acts (whereby legal terms can be used to *describe* a particular situation) and the *performative* nature of speech acts (whereby legal terms, when certain conditions are met, can give rise to ‘*legal effect*’). Both can, however, be distinguished from the mere *prescriptive* nature of legal terms, serving to provide *normative guidance* on how one ought to act, without immediately attaching legal effects to behaviour that deviates from such guidance. See in this regard Mireille Hildebrandt, *Law for Computer Scientists and Other Folk* (Oxford University Press 2020) 20–21.

⁴ See, e.g., Per Aarvik, ‘Artificial Intelligence – A Promising Anti-Corruption Tool in Development Settings?’ (Chr Michelsen Institute 2019).

⁵ See also HLA Hart, *The Concept of Law* (3rd ed., Oxford University Press 2012) 129.

⁶ See in this regard Julia Black, *Rules and Regulators* (Clarendon 1997). See also Bronwen Morgan and Karen Yeung, *An Introduction to Law and Regulation: Text and Materials* (1st edn, Cambridge University Press 2007) 153 and following.

and in part because they rely on others for their application.”⁷ She therefore underlines the need for a ‘sympathetic interpreter’ of legal rules to ensure they are applied in the way that was intended. According to her, “*problems of inclusiveness and determinacy or certainty can be addressed by interpreting the rule in accordance with its underlying aim. By contrast, the purpose of the rule could be defeated if the rule is interpreted literally, if things suppressed by the generalization remain suppressed.*”⁸ To make this more concrete, consider the example I provided in Section 2.3 regarding the legal rule in the area of Belgian migration law, which grants migrants in Belgium the possibility to apply for a residence permit under the condition that ‘*exceptional circumstances*’ justify the submission of such application.⁹ The legislator purposelessly used a broad legal term rather than making a list of all situations that are considered exceptional, thereby enabling the accommodation of circumstances that might not have been foreseen when the rule was adopted, yet which subsequently present themselves as exceptional and justify the granting of a residence permit. The interpreter of the rule, namely the relevant public authority, can hence interpret the term ‘exceptional circumstance’ in various ways,¹⁰ as long as such interpretation is congruent with the law’s purpose and with other legal norms.

Code, in contrast, is far more rigid. Ultimately it is expressed in zeroes and ones, and it hence requires substantially more precision.¹¹ This means that certain interpretative decisions need to be made *upfront*, since the rich openness of text cannot fully be captured in code. Accordingly, in the course of the codification process, some nuances and potential modes of interpretation will inevitably get lost in translation.¹² The legal concepts that are codified often concern complex social phenomena that cannot be readily expressed in a ‘data-fiable’ and computable

⁷ Black, *Rules and Regulators* (n 6) 10.

⁸ See *ibid* 12. See also the discussion *supra*, in Section 3.3.1. Julia Black puts it as follows,

in forming the generalization, which is the operative basis of the rule, only some features of the particular event or object are focused on and are then projected onto future events, beyond the particulars which served as the paradigm or archetype for the formation of the generalization. The generalizations in rules are thus simplifications of complex events, objects or courses of behaviour. Aspects of those events will thus be left out, or ‘suppressed’ by the generalization. Further, the generalization, being necessarily selective, will also include some properties which will in some circumstances be irrelevant.

(See *ibid* 7.)

⁹ Article 9 bis of the Law of 15 December 1980 on access to the territory, residence, settlement and removal of foreign nationals.

¹⁰ While the term ‘exceptional circumstance’ is fairly broad, as stressed *supra* in Section 2.3.3, also more precise legal concepts are open to multiple interpretations and might see their meaning change over time.

¹¹ See also Bert-Jaap Koops, ‘The (In)Flexibility of Techno-Regulation and the Case of Purpose-Binding’ (2011) 5 *Legisprudence* 171.

¹² Nathalie A Smuha, ‘The Human Condition in an Algorithmized World: A Critique through the Lens of 20th-Century Jewish Thinkers and the Concepts of Rationality, Alterity and History’ (Institute of Philosophy, KU Leuven 2021) 32.

format which, as already discussed, means that reliance on (inevitably imperfect) proxies is typically warranted, and that certain interpretative choices in this regard need to be made.¹³ In the example of the migration law rule, if its application were to be automatised in the context of a hypothetical algorithmic recommendation system, a translation will need to be made from law to code as to what is considered an ‘exceptional circumstance’ so as to qualify for a residence application, and which proxies can help determine this. Inevitably, this translation will *include* circumstances that the interpreter is able and willing to anticipate when the system is developed, and *exclude* circumstances which the interpreter did not consider (or did not wish to consider). These normatively relevant choices will then be embedded into the system, and automatise the rule’s interpretation as decided at that point in time.

The question is then: how exactly does this translation and interpretation process occur in the context of algorithmic regulation? Through which procedure is it decided how textual concepts are essentialised into binary code, and how is it decided which quantifiable proxies are adequate to capture non-readily codifiable phenomena? Who is responsible for these choices? How can it be ensured that they are made, in Black’s words, by a ‘sympathetic interpreter’? And who oversees these choices and makes sure that they comply with the law, and that the algorithmic rules through which they are implemented are congruent and proportionate?

In non-algorithmic context, the CJEU already cautioned against reliance on quantitative criteria to assess complex qualitative phenomena, stating that this threatens to reduce the protection that individuals may need. The case at hand concerned an application for subsidiary protection lodged by an individual, based on the asserted existence of a “*serious and individual threat by reason of indiscriminate violence in situations of armed conflict*”.¹⁴ The relevant public authority rejected the application based on a single quantitative criterion (the ratio between the number of casualties in the relevant area and the total number of individuals composing the population of that area) rather than conducting a comprehensive assessment of the particular circumstances of the individual case.¹⁵ While this case

¹³ See *supra*, Section 2.2.5.

¹⁴ See in this regard Case C-901/19, CF, *DN v Bundesrepublik Deutschland*, 10 June 2021, ECLI: EU:C:2021:472, §15. See also the analysis of M van Ham and KM Zwaan, ‘Kwantificeren is geen kwalificeren: de uitspraak van het Hof van Justitie inzake de vaststelling van willekeurige geweld (art. 15c-situaties)’ (2021) 27 *Nederlands tijdschrift voor Europees Recht* 211.

¹⁵ The Court stated in §35 of the judgment that

The systematic application by the competent authorities of a Member State of a single quantitative criterion, which may be of questionable reliability in view of the specific difficulty of identifying objective and independent sources of information close to areas of armed conflict, such as a minimum number of civilian casualties injured or deceased, in order to refuse the grant of subsidiary protection, is likely to lead national authorities to refuse to grant international protection in breach of the Member States’ obligation to identify persons genuinely in need of that subsidiary protection.

did not involve algorithmic regulation, it demonstrates how the same quantitative and restrictive logic that underpins algorithmic regulation can undermine the law's protection. Moreover, with the use of algorithmic systems, this problem only risks being exacerbated. First, such systems enable decision-making on a much wider scale. And second, the fact that such systems can rely on *multiple* quantitative criteria might give a false impression of higher accuracy and objectivity, even if these criteria might still concern factors that do not necessarily relate to the individual subjected to the system. What happens if this quantification becomes not only ubiquitous and normalised, but also automated? How do we ensure that the difference between 'calculating' and making a judgment or assessment is not forgotten precisely because of this normalisation?

If the system is primarily knowledge-driven, the choices made by the system's developers are in principle rendered explicit into the model, as they need to reflect on the criteria they will use before codifying. For instance, in the Belgian migration law example, choices would need to be made in advance as regards the type of circumstances or events that qualify as sufficiently 'exceptional' to justify a residence permit, and which datapoints can demonstrate the presence of such circumstances. Yet the question remains: who makes this choice? On which basis? How is it justified? How can it be ensured that this choice is made appropriately and in compliance with the law? And who verifies the interpretation of the law?

If the system is primarily data-driven, the codification of relevant criteria and categories is not always delineated in advance, but can be suggested by the system itself based on patterns it identifies in the data it is fed. The contours of the system's apprehended reality hence depend upon the patterns it may or may not pick up, which in turn depends on how the system was designed and how the data was gathered and selected in the first place.¹⁶ Moreover, data-driven categorisations can subsequently be incorporated into knowledge-driven systems, using them as a basis for reasoning and inference-drawing. Accordingly, regardless of the type of system, the choices made by its developers regarding the dataset to be used, the design of the algorithm and the optimisation function will be highly relevant to influence the outcomes of the administrative act.

Importantly, information that may be evident for human beings (given the broader knowledge and 'common sense' that human beings have about society) may not necessarily be 'understood' by the algorithmic system, which can only rely on the concepts that it was trained on or programmed with, and has no understanding of the meaning behind those concepts.¹⁷ This limitation can be problematic when it leads to an algorithmic outcome that is not congruent with the legal rule it is meant to apply, and hence does not align with the principle of legality.

¹⁶ Smuha, 'The Human Condition in an Algorithmized World' (n 12) 32.

¹⁷ See also Gary Marcus and Ernest Davis, *Rebooting AI: Building Artificial Intelligence We Can Trust* (Penguin Random House 2020).

An illustration of this problem can be found with Idaho's algorithmic system that was used to determine benefits budgets for disabled adults, as briefly mentioned in Chapter 2.¹⁸ In the course of the class action litigation that was initiated by those who suddenly saw their budgets being cut, despite their medical needs, several deficiencies of the system came to the surface, including one that defied any logical human reasoning. As explained by Restrepo Amariles:

Likely a product of the multi-collinearity issues, there were several regression coefficients wherein the algebraic sign was the opposite of that expected (that is, an input decreased the budget when one would expect it to increase instead). For example, an indication that a participant has "other neurological impairment(s)" reduced a self-directed budget by \$8,095, and high-level needs for Total Support with Laundry and Assistance in Feeding similarly had negative impacts of \$4,201 and \$5,715, respectively. Decreasing a budget in response to more severe needs seems deeply counter-intuitive and indicates a structural flaw with the prediction tool.¹⁹

While for a human case assessor it would have been evident that people with additional impairments would require *more* rather than *less* monetary aid, the algorithmic system's flaw resulted in the exact opposite result, demonstrating that, inevitably, certain elements can get lost in translation.²⁰ Indeed, "*when a computer learns and consequently builds its own representation of a classification decision, it does so without regard for human comprehension*".²¹

4.1.1.b From Legality to Legalism

I observed in Chapter 3 how the rule of law is, in essence, a middle-ground between the rule *by* law (where rules are applied without any form of discretion and nuance, even if they may be unjust or lead to unjust results) and the total absence of rules. In this context, I also explained that there is a good reason why, to achieve the rule of law and the principle of legality, rules and discretion ought to coexist. However, an

¹⁸ See *supra*, Section 2.2.2. See also Jay Stanley, 'Pitfalls of Artificial Intelligence Decisionmaking Highlighted in Idaho ACLU Case' (*American Civil Liberties Union*, 2 June 2017) <www.aclu.org/blog/privacy-technology/pitfalls-artificial-intelligence-decisionmaking-highlighted-idaho-aclu-case>.

¹⁹ David Restrepo Amariles, 'Algorithmic Decision Systems: Automation and Machine Learning in the Public Administration' in Woodrow Barfield (ed), *The Cambridge Handbook of the Law of Algorithms* (Cambridge University Press 2020) 289.

²⁰ As previously noted, an algorithmic system will not be able to interpret legal concepts as social constructs that only reflect a partial view of the world, and will not understand the *meaning* behind the syntax. See in this regard also David Cole, 'The Chinese Room Argument' in Edward N Zalta (ed), *The Stanford Encyclopedia of Philosophy* (Winter 2020, Metaphysics Research Lab, Stanford University 2020) <<https://plato.stanford.edu/archives/win2020/entries/chinese-room/>>.

²¹ Jenna Burrell, 'How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms' (2016) 3 *Big Data & Society* 1, 10.

overly rigid codified translation of the law risks reducing legality to a form of legalism,²² precisely because of the lack of room for discretion and critical reflection about the underlying purpose that the rule should serve. As noted by Laurence Diver, code tends to be legalistic in light of its inherent ‘ruleishness’.²³ And whereas in non-algorithmic settings the legalistic application of rules can be offset by accommodating nuanced interpretations and modes of application where need be, the rigidity of code does not easily enable such interpretative flexibility.²⁴

Consider the example of Indiana, where a plan was launched in 2006 to out-source and automate the eligibility checks for several welfare programs, including Medicaid and food stamps. As explained by Virginia Eubanks, the tender request specified that the automation process aimed to “*reduce fraud, curtail spending, and move clients off the welfare rolls*”.²⁵ This move followed the finding that two employees of the Family and Social Service Administration (FSSA) of Indianapolis had committed fraud, which lead politicians to claim that the welfare system was fraudulent and “*irretrievably broken*”.²⁶ The hope was not only that automation would reduce the risk of fraud, cut costs and increase process efficiency (especially given the high workload on public officials and increasing backlogs), but it was also claimed that this would free up time for the remaining public officials to work more closely with clients.²⁷

The algorithmic system was however riddled with system failures and technical errors, which led to erroneous denials of benefits, and made the application process far more difficult. Besides technical glitches and integration problems, Eubanks explains how one of the main causes of error was “*the result of inflexible rules that interpreted any deviation from the newly rigid application process, no matter how inconsequential or inadvertent (including missing a phone call from a caseworker) as an active refusal to cooperate*”.²⁸ And once a refusal to cooperate was determined, this led to the straight denial of eligibility. Accordingly, through an overly legalistic interpretation and codification of rules, persons in need were automatically denied benefits, and caseworkers did not have the flexibility to easily remedy this problem where needed. Such use of algorithmic regulation hence led to a disproportionate application of the law, which runs counter to the principle of legality, which – as

²² Definitions of legalism vary, yet Judith Shklar has influentially defined it as “*the ethical attitude that holds moral conduct to be a matter of rule following, and moral relationships to consist of duties and rights determined by rules*.” See Judith N Shklar, *Legalism: Law, Morals, and Political Trials* (Harvard University Press 1986) 1.

²³ Diver (n 3) 3.

²⁴ See also Zenon Bankowski and Burkhard Schafer, ‘Double-Click Justice: Legalism in the Computer Age’ (2007) 1 *Legisprudence* 31.

²⁵ Virginia Eubanks, *Automating Inequality – How High-Tech Tools Profile, Police and Punish the Poor* (Picador 2019) 46.

²⁶ *ibid.*

²⁷ *ibid* 47.

²⁸ *ibid* 51.

discussed above – requires public authorities to make “*a proper balance between any adverse effects which their decision has on the rights or interests of private persons and the purpose they pursue*”.²⁹

One might contend that this can be avoided by simply codifying better or plural interpretations of the law, based on the variability of potentially applicable situations. However, this would require the coders of the algorithmic system to foresee each and every possible situation that might emerge in the future in advance – which is impossible.³⁰ Choices therefore inevitably need to be made, and since algorithmic systems are not only based on pre-programmed algorithms but also lack any human understanding, it is not possible to question or challenge them by explaining that the situation at hand is not accommodated by the codified rules, or that it requires a different legal application. Moreover, even when the algorithmic system only *informs* an administrative act rather than adopting one, public officials may still be discouraged from deviating from the proposed outcome, thereby reinforcing a legalistic approach. This discouragement may stem from the pressure of KPI's that public officials need to meet in light of efficiency goals, such as the speedy handling of case files, from the lack of sufficient information or understanding about the system's operations to challenge its merits, or from more general deference to the system's ‘cognitive superiority’ and ‘air of authority’ arising from the ‘objective’ mathematical rules underlying its functioning.³¹

Notwithstanding the need to find a middle ground between rules and discretion, the ‘ruleishness’ of code risks skewing the balance entirely towards the *rules* side of the spectrum. This also leads to a problematic tension with the principle of proportionality, which is part of the legality principle and acts as a necessary softener of legal rules’ rigidity. Indeed, public authorities have the responsibility to ensure that the measures they take when they apply general laws to individual cases are proportionate, and that they take into account the factors relevant to the case.³² If no such proportionality assessment takes place and the law is mechanically applied, the hard edges of the over- and under-inclusive nature of law are left untouched, opening the door to adverse effects on those subjected to it.

This is precisely what happened in the Dutch child allowance case, which I briefly mentioned in Chapter 3.³³ After it was revealed in 2013 that a criminal

²⁹ Committee of Ministers of the Council of Europe, ‘Recommendation CM/Rec(2007)7 of the Committee of Ministers to Member States on Good Administration’ (2007) 8. See also *supra*, Section 3.3.1.

³⁰ Recall in this regard also the quote by HLA Hart that, after all, “*we are men, not gods*”. See Hart (n 5) 128.

³¹ See *supra*, Section 2.2.6.

³² See also Committee of Ministers of the Council of Europe, ‘Recommendation No. R (80) 2 of the Committee of Ministers Concerning the Exercise of Discretionary Powers by Administrative Authorities’, Adopted by the Committee of Ministers on 11 March 1980 at the 316th meeting of the Ministers’ Deputies, 1980.

³³ See *supra*, Section 3.1.5. See also Parliamentary Inquiry Committee, ‘Unprecedented Injustice’ (House of Representatives of the States General 2020) 35 510, no. 1.

scheme had defrauded the Dutch state of social aid payments for years, the Dutch tax authority doubled down on tackling fraud and started taking a more severe stance,³⁴ including in the area of childcare allowance, which is a means-based type of allowance. Applications for such allowance were only minimally verified, so that nearly any applicant would receive an ‘advance’ on the allowance, after which it would be verified if any revision of the paid allowance was needed and if any recovery was to be paid back to the state. At the time, the relevant law stated that: *“If a revision of an allowance or a revision of an advance results in an amount to be recovered or if a settlement of an advance with an allowance leads to this, the person concerned shall owe the entire amount of the recovery.”*³⁵ As the Venice Commission noted, *“this provision was interpreted by the Tax and Customs Administration as the so-called ‘all or nothing approach’, so that even if a parent had acted in good faith but neither the parent or the childminder could provide proof of hours used or parental contribution etc., the parent had to repay the full amount for the whole year”*.³⁶ Evidently, the public authority’s narrow interpretation of the legal rule led to high demands for repayment.³⁷ Yet this problem was exacerbated by the fact that it relied on an algorithmic system, enabling it to significantly scale up its investigations and targeting practices.³⁸

When the dramatic consequences of this legalistic application of the law came to light, a Parliamentary investigation was conducted. Many families underwent significant financial hardship and, in some cases, children were even taken away from their parents who could no longer afford to take care of them, which resulted in child neglect. Commenting on how things could go so wrong for so long, the committee in charge of the investigation noted that *“the administrative justice system neglected its important function of safeguarding the legal rights of individual citizens”* inter alia by *“perpetuating the ruthless application of the legislation on childcare allowance, over and above what was prescribed by law”*.³⁹ In its opinion on this case, the Venice Commission was particularly critical of the public authority’s refusal to conduct a proportionality test⁴⁰ regarding the measures they were imposing and

³⁴ Richard Barrett and others, ‘The Netherlands: Opinion on the Legal Protection of Citizens’ (Venice Commission 2021) Opinion no. 1031/2021 CDL-AD(2021)031 4.

³⁵ See Article 26 of the General Act on Means-Tested Benefits, which entered into force in 2005. See *ibid.*

³⁶ *ibid.* 4.

³⁷ According to the Venice Commission’s Opinion, based on the Parliamentary Investigation, 15 per cent of the parents were subjected to repayment requests. See *ibid.* 4.

³⁸ Which, as will be discussed *infra* in Section 4.1.4, also appear to have been potentially discriminatory.

³⁹ Parliamentary Inquiry Committee (n 33) 7.

⁴⁰ In addition, it should also be noted that the Venice Commission deplored the fact that the courts who reviewed the administrative decision did not apply a proportionality test – even though it is required by virtue of international law and the rule of law principles – for the mere reason that the legislator had excluded the application of such a test in the relevant legislation. See Barrett and others (n 34) 21.

their effect on the affected families. Instead, they blindly applied the law in a rigid and legalistic manner, leading to a rule *by* law approach rather than respecting the rule of law.⁴¹

While this case shows that public authorities do not need to rely on algorithmic systems to adopt an overly legalistic interpretation of the law, it also demonstrates that the use of such systems can drastically increase the scale of the law's legalistic application, and that it eliminates the possibility for a sound proportionality assessment of individual cases. Furthermore, besides making the scaled application of legalistic rules cheaper, the opaque nature and use of algorithmic systems also makes the identification and problematisation thereof more difficult, thus perpetuating the law's disproportionate application. In sum, by codifying a particular interpretation of text-natured law into code, the law's meaning becomes fixed and unitary, and is thereby essentialised. Without the possibility to correct the adverse effects of such over- and under-inclusivity, it becomes much more difficult to ensure a proportionate application of the law, and to ensure the alignment of algorithmic regulation with the rule of law.

4.1.1.c Loss of Process Transparency

Besides matters of interpretation and correctness, the question also arises of how the translation from law to code can take place in a manner that respects the principle of transparency. As noted in Chapter 3, transparency is a recurring requirement that can be found under the principle of legality (pertaining to the transparency of the law-making process), the principle of legal certainty (pertaining to the transparent character of the law itself) and the principle of non-arbitrariness (pertaining to the transparent nature of the law's application and the justification of its manner of implementation). My focus here is on the first of these. The translation from law to code can be seen as part of the law-making process, given that it constitutes the first step before public authorities can start applying it to legal subjects. The codification process of algorithmic regulation is typically not a public endeavour, but a rather 'technical' matter. Simply put: it belongs to the realm of coders,⁴² by which I intend to denote technical experts (rather than trained public officials) who make choices about the system's design and development, including the datasets that will be used and the labels assigned to them, the algorithmic model that is deployed and its

⁴¹ In its opinion about the childcare allowance case, it pointed to the Venice Commission's Rule of Law Checklist, which states that an 'exercise of power that leads to substantively unfair, unreasonable, irrational or oppressive decisions violates the Rule of Law'. It seems safe to conclude, as several Dutch authorities have done – including the Council of State's Administrative Jurisdiction Division in two judgments in 2019, that applying the 'all or nothing' approach falls under this definition. (See *ibid* 24.)

⁴² See in this regard also the conceptualisation of the 'coding elite' by Jenna Burrell and Marion Fourcade, 'The Society of Algorithms' (2021) 47 *Annual Review of Sociology* 213.

optimisation function, and the translation process from legal text or other linguistic concepts to code.

These choices are often opaque, as the coders do not necessarily make their decisions (and the justification for those decisions) explicit. However, if there is no transparency about this process, it is far more difficult to exercise oversight over the executive's interpretative choices and ensure they respect the rule of law. More generally, transparency is also needed to ensure that the system does not contain errors, bugs, or erroneous translations that may be unintended yet can have an adverse impact.⁴³ Mistakes can happen, and errors also occur without any mediation from algorithmic systems. Yet errors can be reproduced in the systems that humans build, and in that case, the speed and scale of the system's decision-making process can render the error's effects much more problematic. In addition, the larger the amount and variety of data that public authorities gather and process about citizens, the more room there is for mistakes. As stated by Peeters and Widlak: "*There are few barriers for an error to diffuse via data exchange and exclude a citizen from services and overwhelm a citizen with administrative burdens. There are, however, many barriers for a correction to have the same, but opposite, automatic effect.*"⁴⁴

While transparency is no panacea, providing insight about the interpretative choices of the translation, the data and proxies that are being used and the manner in which the system functions, at least makes it easier to identify and correct errors, and to discuss and possibly contest the validity of the assumptions underlying the algorithmic model. Yet public authorities do not always provide transparency about the systems they use, let alone about the translation process preceding their use. Consider, in this regard, the algorithmic system used in Belgium to assist with the identification of social security fraud, known as the OASIS tool.⁴⁵ The system is primarily data-driven and relies on large databases to profile citizens, including the categorisation of potential fraudsters.⁴⁶ As noted by Elise Degrave, the system targets "*mainly the poorest people*", focusing, for instance, on the detection of fraudulent labour providers and bankruptcies, and domicile fraud aimed at securing more social assistance than one is entitled to.⁴⁷ Degrave explains how her many attempts

⁴³ See also Karl de Fine Licht and Jenny de Fine Licht, 'Artificial Intelligence, Transparency, and Public Decision-Making' (2020) 35 AI & Society 917.

⁴⁴ Rik Peeters and Arjan Widlak, 'The Digital Cage: Administrative Exclusion through Information Architecture – The Case of the Dutch Civil Registry's Master Data Management System' (2018) 35 Government Information Quarterly 175, 176.

⁴⁵ Elise Degrave, 'The Use of Secret Algorithms to Combat Social Fraud in Belgium' (2020) 1 European Review of Digital Administration & Law 167.

⁴⁶ For an overview of algorithmic fraud detection in Belgian federal administrations, see Anthony Simonofski, Thomas Tombal and Pauline Willem, 'Policy Report on Big Data Policy of the Belgian Federal Administrations' (BELSPO 2020), <<https://soc.kuleuven.be/io/digi4fed/doc/di-2-policy-report-on-big-data-policy-of-the.pdf>>.

⁴⁷ Degrave (n 45) 170–71. See also Elise Degrave and Amelie Lachapelle, 'Le droit d'accès du contribuable à ses données à caractère personnel et la lutte contre la fraude fiscale' (2014) 5 Revue générale du contentieux fiscal.

to gain more information about the system (through the exercise of her right to access to administrative documents) were unsuccessful, either because the authority considered “*the document could be a ‘source of misunderstanding’ within the meaning of the law on the publicity of the administration*” or because, upon appeal, the Commission for Access to Administrative Documents found that “*the authority addressed was not an administrative one and that the requirements of transparency did not therefore apply*”.⁴⁸ She concludes that “*in short, information about OASIS seems to be hidden*”, as “*we have not been able to access an official document clearly explaining this tool*”.⁴⁹ If a highly educated law professor who is specialised in public information law does not manage to obtain information about an algorithmic system, one can imagine how much more difficult it must be for people who are in a more vulnerable position and even more likely to be adversely impacted by such a system.

In addition to the unwillingness to provide information, the non-transparency of algorithmic regulation can also be related to the private nature of the system’s development process.⁵⁰ The example of the ‘Children’s Safeguarding Profiling System’ used by the Hackney County Council (UK) is telling in this regard.⁵¹ The system was developed in cooperation with Ernst & Young and tech company Xantura to identify children at risk of maltreatment and families who need additional support.⁵² When researchers⁵³ submitted a Freedom of Information Request to the London Borough of Hackney, they received the following response:

Xantura and London Borough of Hackney are working together to develop the system as development partners, but Xantura anticipates operating on a commercial basis. We believe that to reveal detailed workings of the system would be damaging to their commercial interests and, while the project is in pilot phase, of limited public use. We therefore believe that the public interest in seeing any operating

⁴⁸ Degrave (n 45) 171.

⁴⁹ *ibid* 172.

⁵⁰ See also Jenna Burrell, ‘How the Machine “Thinks”: Understanding Opacity in Machine Learning Algorithms’ (2016) 3 *Big Data & Society* 205395171562251, 639; Abe Chauhan, ‘Towards the Systemic Review of Automated Decision-Making Systems’ [2021] *Judicial Review* 1, 4.

⁵¹ Niamh McIntyre and David Pegg, ‘Councils Use 377,000 People’s Data in Efforts to Predict Child Abuse’ *The Guardian* (16 September 2018) <www.theguardian.com/society/2018/sep/16/councils-use-377000-peoples-data-in-efforts-to-predict-child-abuse>.

⁵² Luke Stevenson, ‘Artificial Intelligence: How a Council Seeks to Predict Support Needs for Children and Families’ (*Community Care*, 1 March 2018) <www.communitycare.co.uk/2018/03/01/artificial-intelligence-council-seeks-predict-support-needs-children-families/>.

⁵³ See Lina Dencik, Arne Hintz, Joanna Redden and Harry Warne, ‘Data Scores as Governance: Investigating Uses of Citizen Scoring in Public Services’ (Data Justice Lab, Cardiff University 2018) 59, <<https://datajustice.files.wordpress.com/2018/12/data-scores-as-governance-project-report2.pdf>>. See also Joanna Redden, Lina Dencik and Harry Warne, ‘Datafied Child Welfare Services: Unpacking Politics, Economics and Power’ (2020) 41 *Policy Studies* 507.

manuals is outweighed by Xantura's commercial interests and exempt this part of the request under Section 43 of the Freedom of Information Act.⁵⁴

Accordingly, the commercial interests of a private company, in charge of highly impactful decisions taken by a public authority in the public interest, were prioritised over transparency towards citizens.

Besides the risk of errors or unconscious bias, transparency can also help counter the risk of *intended* mistranslations. Indeed, as Plato's story of the ring of Gyges shows: a lack of transparency can not only cover up mistakes, but it can unfortunately also be exploited to increase power.⁵⁵ The deliberate codification of a rule into an unduly narrow, arbitrary or otherwise illegal interpretation can take place *either* by the private developers of the system, for instance with the aim of translating the rules in their private interests, *or* by the public authority, for instance with the aim of codifying an interpretation that helps it consolidate power or disadvantage political opponents, minorities or others that may challenge its actions. Countering this risk requires oversight during the design and translation phase, rather than only during the period that the system is already being deployed at scale.

At the same time, it must be acknowledged that the risk of abuse is an inevitable reality, and that any oversight and transparency measures will always be subject to limitations. Therefore, given the vastness of the adverse consequences in case things go wrong, there may be administrative acts for which the deployment of algorithmic regulation is undesirable altogether. Pursuant to the principle of legality's requirement to ensure a *participatory* law-making process,⁵⁶ the desirability of algorithmic regulation in light of its potential impact should be the subject of public deliberation.⁵⁷ This also implies that citizens should have a say about the administrative acts that can or should be algorithmised *prior* to the implementation of algorithmic regulation.

4.1.2 Legal Certainty

Legal rules need to be clear and sufficiently precise to make the way in which they will be applied predictable. Furthermore, they must be applied consistently, enabling legal subjects to have legitimate expectations about the rules they will be

⁵⁴ Aleksi Knuutila, 'Documents Relating to the Children's Safeguarding Profiling System – A Freedom of Information Request to Hackney Borough Council' (*WhatDoTheyKnow*, 15 January 2018) <www.whatdotheyknow.com/request/documents_relating_to_the_childr>.

⁵⁵ See Plato, *The Republic* (Benjamin Jowett tr, Dover Publications, Inc 2000) Book 2.

⁵⁶ See *supra*, Section 3.3.1.

⁵⁷ See also Venice Commission, 'Rule of Law Checklist' (Council of Europe 2016) Study no. 711/2013 CDL-AD(2016)007rev 13. Moreover, the need for public deliberation about algorithmisation has also been stressed in Nathalie A Smuha and others, 'How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act' (Social Science Research Network 2021) <<https://papers.ssrn.com/abstract=3899991>> 8.

subjected to and to plan their lives accordingly. At first glance, the application of legal rules by an algorithmic system rather than public officials can contribute to this requirement. As noted in Section 2.3, ensuring the consistent application of a rule is not straightforward when this occurs by mediation of countless public officials who may each have their own way of interpreting a rule, especially if they operate in a decentralised organisational structure.⁵⁸ While the routinisation of decision-making by promulgating detailed guidelines can diminish the risk of diverging interpretations, it cannot prevent this from happening altogether. Yet delegating the application of legal rules to an algorithmic system can in principle ensure greater consistency, since the codification process occurs in a centralised manner, with only one ‘interpreter’ – namely the machine, or rather, the coder – acting as mediator between the rule and *all* legal subjects. In practice, there are, however, several drawbacks that need to be pointed out.

4.1.2.a Fanciful Foreseeability

First, the consistent application of rules is but one of the requirements of the principle of legal certainty and should not be prioritised over the rule of law’s overarching aims. It is possible that a change in circumstances or new societal developments require the adapted application of a rule to maintain its original purpose. This is why the rule of law requires that a balance be struck between stability and flexibility. However, once a particular interpretation of the law is codified and its execution is automatised, this balance risks being skewed, as the codification process stabilises and thereby essentialises the legal rule’s interpretation.⁵⁹ This, in turn, makes it harder for public officials to apply the law to individual cases and adapt the law’s interpretation to changing circumstances when needed.

One could counterargue that the system can be programmed to identify and apply different rules based on different situations, yet this still requires a codification of all possible situations in advance, as well as an *ex ante* decision of the way in which the rule will be applied in those cases. Moreover, there will inevitably be cases that fall outside these codified situations, and that will not be adequately dealt with based on those prior decisions.⁶⁰ Accordingly, those who will not fall into the pre-established categories of persons and rule applications risk encountering a significant disadvantage.⁶¹ Contrary to a public official, the algorithmic system that informs or adopts the

⁵⁸ See also Nadine van Engen, Bram Steijn and Lars Tummens, ‘Do Consistent Government Policies Lead to Greater Meaningfulness and Legitimacy on the Front Line?’ (2019) 97 Public Administration 97; See also Andrew B Whitford, ‘Decentralization and Political Control of the Bureaucracy’ (2002) 14 Journal of Theoretical Politics 167.

⁵⁹ See also Diver (n 3) 9.

⁶⁰ See *supra*, Section 3.3.1.

⁶¹ See in this regard also Peeters and Widlak (n 44); Smuha, ‘The Human Condition in an Algorithmized World’ (n 12) 32.

administrative act will not be able to take a more flexible and case-by-case approach to address this concern. It is therefore important that consistency is not fetishised in the name of legal certainty. Instead, it must be interpreted as one of various elements that can contribute to the rule of law's overarching aim, tailored to the specific context.

Consider, for instance, an algorithmic system deployed by public authorities in Poland to profile the unemployed and, based on their categorisation, decide on their eligibility for specific programmes aimed at helping them back to the job market through 'labour market programs'.⁶² Essentially, the system calculated the 'employment potential' of unemployed individuals. The Polish Ministry of Labor and Social Policy hoped that the system would lead to a more efficient use of limited resources, by allocating more funds for *"those who are particularly distant from the labor market, and less for those who are able to handle finding a job easier"*.⁶³ Besides 'efficiency', one of the reasons cited for the system's introduction was the fact that, previously, local labour offices were already undertaking a form of 'profiling', but they were doing so in an unstructured and inconsistent way. Therefore, *"it could have been the case that the standard or principles of assigning specific active labor market programs to the unemployed varied in different offices."*⁶⁴ To remedy this, the algorithmic system henceforth centrally determined the questions that public officials should ask during their interviews with the unemployed, and subsequently automatically profiled them based on the inputted answers.

However, the Polish NGO Panoptikon who interviewed public officials working with the system and who conducted an extensive investigation, concluded that the proportion of people assigned to particular profiles still varied strongly from one labour office to another, even after the system's introduction.⁶⁵ This was, for instance, because public officials did not always interpret or explain the questions uniformly to the citizens *"especially considering high caseloads and time limits for one interview"*,⁶⁶ nor did the citizens always interpret these questions uniformly, hence leading to differing types of answers and outcomes in similar situations and vice versa. Accordingly, the system failed to lead to 'consistency', demonstrating that the mere use of automated data processing in itself does not necessarily lead to such

⁶² See AlgorithmWatch, 'Automating Society Report 2020' (2020) 186, <<https://automatingsociety.algorithmwatch.org/wp-content/uploads/2020/12/Automating-Society-Report-2020.pdf>>.

⁶³ Jędrzej Niklas, Karolina Sztandar and Katarzyna Szymielewicz, 'Profiling the Unemployed in Poland: Social and Political Implications of Algorithmic Decision Making' (Panoptikon 2015) 7, <www.ohchr.org/sites/default/files/Documents/Issues/Poverty/DigitalTechnology/LSE_appendix2.pdf>.

⁶⁴ *ibid* 8.

⁶⁵ As noted by the NGO: *"It turned out that the percentage share of the unemployed assigned to Profile II in the entire group of the unemployed varies among particular labor offices and ranges from 33% to as much as 96% in some offices. In the case of Profile III, this ratio ranges from 4% to 33%."* See *ibid* 14.

⁶⁶ *ibid* 28.

a result.⁶⁷ At the same time, the researchers of Panoptikon also observed that “the use of algorithmic decision-making can help mask the shortcomings of a given public policy (such as an objective shortage of resources) by limiting options that are available to some categories of citizens and making the management of public resources less transparent”.⁶⁸

It should, moreover, be considered that the concrete effects of a codified rule are not always foreseen or foreseeable in advance. Unintended and unknown bugs in the code might hamper legal certainty, and might lead to unwanted adverse consequences that the coders of the algorithmic system did not predict, let alone the people subjected to the system.⁶⁹ In that sense, foreseeability of the rule’s application through codification may turn out to be a mere illusion. Furthermore, due to the scaled nature of the system’s application, unpredicted adverse effects can affect a vast number of people at the same time. The example of an algorithmic system deployed by the Swedish Public Employment Service is telling in this regard. The system was used to verify whether people who received certain unemployment benefits remained eligible, thereby aiming to ‘increase efficiency’.⁷⁰ However, due to a flaw in the system, about 70,000 unemployed individuals erroneously stopped receiving their benefits,⁷¹ a flaw that had certainly not been anticipated by the coders when they programmed the system, or it would have been addressed *prior* to its use. Therefore, it is not because a line of code is executed consistently, that it also effectively leads to foreseeability and legal certainty. Also in Austria, public authorities’ reliance on a similar algorithmic profiling system to categorise the unemployed based on their job prospects faced significant criticism.⁷²

⁶⁷ Note that, in the meantime, the system has been scrapped. See Jędrzej Niklas, ‘Poland: Government to Scrap Controversial Unemployment Scoring System’ (*AlgorithmWatch*, 16 April 2019) <<https://algorithmwatch.org/en/poland-government-to-scrap-controversial-unemployment-scoring-system/>>.

⁶⁸ Niklas, Sztandar and Szymielewicz (n 63) 35. Also here, a request for public information submitted by an NGO, Panoptikon Foundation, was answered negatively, stating that information about the logic of the profiling is not considered ‘public information’. It ultimately decided to challenge this decision in court until it managed to obtain information about the system’s predetermined questions. See Niklas (n 67).

⁶⁹ See also Diver (n 3) 9.

⁷⁰ See ‘DN Debatt. “Är Sverige Redo Att Låta Maskiner Fatta Besluten?”’ (*Dagens Nyheter* 17 February 2019) <www.dn.se/debatt/ar-sverige-redo-att-lata-maskiner-fatta-besluten/>.

⁷¹ Tom Wills, ‘Sweden: Rogue Algorithm Stops Welfare Payments for up to 70,000 Unemployed’ (*AlgorithmWatch*, 25 February 2019) <<https://algorithmwatch.org/en/rogue-algorithm-in-sweden-stops-welfare-payments/>>.

⁷² See, e.g., Doris Allhutter and others, ‘Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective’ (2020) 3 *Frontiers in Big Data* 1 <www.frontiersin.org/article/10.3389/fdata.2020.00005/full>. The system also faced heavy critique for being allegedly discriminatory, for instance against women and people with a disability, see Nicolas Kayser-Bril, ‘Austria’s Employment Agency Rolls out Discriminatory Algorithm, Sees No Problem’ (*AlgorithmWatch*, 2019) <www.algorithmwatch.org/en/austrias-employment-agency-ams-rolls-out-discriminatory-algorithm/>.

Furthermore, when it comes to enhancing legal certainty, a distinction can be made between knowledge-driven and data-driven systems, as the aspired enhancement of consistency will primarily apply to the former (barring, of course, the abovementioned problems, as well as the fact that even in knowledge-driven systems rules can be incorporated that randomise certain outcomes, thereby diluting predictability). As discussed in Section 2.1, data-driven systems do not rely on an *ex ante* codified model, but instead hinge on a dataset in which patterns are identified and to which weights are assigned, based on which a model is subsequently derived. Accordingly, the rule's application relies on – and is adapted to – the data. The main strength of data-driven systems thus lies in their adaptability to new situations based on new data.⁷³ While this feature can counter the concern of over-stability and inflexibility, reliance on a data-driven system risks pushing the pendulum entirely the other way, towards too much adaptability and too little predictability. Persons subjected to a data-driven system may find it difficult to know in advance to which category the system will correlate them, and hence to which 'personalised' application of the rule they will be subjected to.

In addition, the rule's application will not hinge on a causal relationship between the rule and the person's behaviour or situation but based on the extent to which that person falls into one of the patterns that the system identified. This in turn depends on the behaviour and situations of persons with a similar profile (where such similarity could depend on factors that are entirely irrelevant to the rule itself). Yet this undermines the entire logic behind the law and its application. In principle, one should derive rights and obligations based on one's own action or situation, not based on how much one shows similarities with other people.⁷⁴ To give an example: the entitlement of an ill person to healthcare benefits should hinge on her specific needs rather than on the needs of those people that an algorithm happened to identify as showing certain similarities with her. As cautioned by Restrepo Amariles, the application of the law thereby "*ends up being transformed into a normative correlation of facts*", as citizens are subjected to rules based whether they fall into identified patterns and correlations.⁷⁵

Furthermore, data-driven systems, too, carry certain limitations in terms of flexibility, as they are after all dependent on their programming language and the choices made by coders in terms of the system's design, datasets and model.⁷⁶ They are hence still predetermined to some extent, which opens the door to the same concerns, including the risk of bugs in the code that may render the system

⁷³ See *supra*, Section 2.1.3.

⁷⁴ See also Emre Bayamlioglu and Ronald Leenes, 'The "Rule of Law" Implications of Data-Driven Decision-Making: A Techno-Regulatory Perspective' (2018) 10 Law, Innovation and Technology 295, 296.

⁷⁵ Restrepo Amariles (n 19) 298.

⁷⁶ See also Elena Popa, 'Human Goals Are Constitutive of Agency in Artificial Intelligence (AI)' (2021) 34 Philosophy & Technology 1731.

unpredictable and potentially harmful. The fact that system developers typically lack the practice of meticulously tracing and documenting who made which design choices, and for which reasons, only adds to the problem. It complicates the identification of errors, but it also renders the system more vulnerable to (subsequent) non-documented tweaking. Even the system's designers may no longer remember which choices they made. On the one hand,⁷⁷ the agility of algorithmic systems and the inherent malleability of software and databases can be seen as a strength, since it allows for their continuous adaptation and improvement. Yet, on the other hand, such agility comes with a price, as the openness of the algorithmic system to continuous adaptation – without adequate traceability – undermines the ability of public authorities and those affected alike to have information about the system's (mal)functioning.⁷⁸

This is an undervalued problem which undermines legal certainty, as also demonstrated by the 'means-testing' algorithmic system introduced in the UK to calculate and allocate welfare benefits, under the heading 'Universal Credit'.⁷⁹ Based on 'real-time' information on citizens, drawn from a range of sources and continuously updated, the system calculates each month how many benefits a person is entitled to receive. However, the *"calculation fails to factor in how frequently people are paid, leading it to overestimate their earnings in some months and underestimate them in others. This design flaw has caused irrational fluctuations and reductions in how much benefit people [] receive from month to month"*,⁸⁰ hence leading to anything but 'certainty' for people about their rights. These fluctuations not only forced certain people to rely on food banks and take on debt to make ends meet, but the uncertainty of how many benefits they will receive each month – hinging on a non-transparent processing of datapoints – has also led to mental health problems and heightened anxiety.⁸¹ In sum, lest public authorities pay due attention to these issues when they implement algorithmic systems for administrative acts, the aspired benefit of increased legal certainty and predictability may be merely fanciful.

4.1.2.b Problematic Preservation of the Past

As noted in Chapter 3, while legal certainty requires a certain level of stability, this requirement must not become so rigid that it undoes the potential for adaptability

⁷⁷ See *supra*, Section 2.1.

⁷⁸ Consider in this regard also the discussion *supra* on algorithmic systems as socio-technical infrastructure in Section 2.2, and on the inherent malleability of software in Section 4.1.2.a.

⁷⁹ See in this regard Richard Pope, 'Universal Credit – Digital Welfare' (PT2 2020) <<https://digitalwelfare.report/>>.

⁸⁰ 'Automated Hardship: How the Tech-Driven Overhaul of the UK's Social Security System Worsens Poverty' (Human Rights Watch 2020) <www.hrw.org/report/2020/09/29/automated-hardship/how-tech-driven-overhaul-uks-social-security-system-worsens>.

⁸¹ *ibid.*

when a public official notices that the rule's application in a specific situation would undermine the intended purpose of the law. This can occur, for instance, because once the law is actually applied, it becomes clear that it raises unforeseen unintended consequences, or because the context and circumstances in which the law is applied changed in the meantime.⁸² After all, the world is not static.

In the previous section, I discussed how the use of algorithmic regulation by public authorities might foster legalism. An additional characteristic of legalism as defined by Judith Shklar, is that it can be associated with a conservative ideology.⁸³ The rules we ought to conform to were inevitably established in the past, which in turn risks leading to a commitment to the preservation of this past. This can be problematic if those past rules are no longer apt to help us deal with changing developments in the here and now, including new insights about the adverse impact of previously ill-conceived rules, or the impact of new technological developments on society. Note how, in the context of algorithmic regulation, this problem is prevalent not only with knowledge-driven systems where rules are explicitly codified, but also in data-driven systems where rules are 'found' based on patterns identified in a dataset. As this dataset necessarily contains data from the past, a novel interpretation that does justice to evolving situations is near impossible to achieve without the intervention of human assessment and judgment.⁸⁴ It is this very ability of interpretative adaptability that risks getting lost, as the automation of administrative action prevents public officials to understand and apply the rules in a manner that meets the changing insights or needs of society.

As a counterargument, one can contend that the adaptation of legal interpretation to changing circumstances should not occur by officials who work for public administrations, but rather by legislators who can revise existing laws through the applicable legislative procedure. It must indeed be acknowledged that the separation of powers can come under pressure if government officials unilaterally decide to change the interpretation of the law as put down by the legislator *contra legem*, merely because they consider that certain circumstances changed. Such an action would be contrary to the rule of law. Yet that is not what this argument is about, for an adapted interpretation need not be *contra legem*, and the changed circumstance need not manifest itself at the level of the general rule but can arise at the level of a concrete situation to which public officials must apply the general rule.⁸⁵ It is at the level of the *latter* that discretion should be used – within the confines of the law and

⁸² See for instance Joanne Conaghan, 'Law, Harm and Redress: A Feminist Perspective' (2002) 22 Legal Studies 319.

⁸³ See Shklar (n 22) 10. See also Robin West, 'Reconsidering Legalism' (2003) 88 Minnesota Law Review 119, 120.

⁸⁴ See also *supra*, Section 2.2.3.

⁸⁵ See *supra*, Section 3.3.2.

in a manner appropriate for the particular case – to ensure the law’s purpose remains attainable.⁸⁶

Finally, one might also contend that an adapted interpretation of a legal rule can be secured through litigation before a court rather than by public officials. While courts can indeed play an important role in this regard, the principle of the separation of powers persists – meaning that this interpretation cannot go *contra legem*, unless it violates a hierarchically higher norm (such as EU primary or secondary law). Moreover, courts can only intervene *ex post* when the damage of a problematically applied law has already occurred. This is why public authorities need to act diligently when they adopt measures to implement general laws, and need to ensure their measures are proportionate in the case at hand *before* the law’s application.⁸⁷ While there is certainly a collective responsibility of all branches to adopt, apply and interpret laws in a manner that does not lead to unjust hardship,⁸⁸ this does not dilute the responsibility of public authorities whose actions most directly affect legal subjects.

4.1.2.c Loss of Implementation Transparency

When discussing the impact of algorithmic regulation on the legality principle, I explained how transparency regarding the rule-making process can be compromised. Yet in addition to such procedural transparency, also substantive transparency of the rules’ content and the way they are implemented and applied may be at risk, which is precisely what the principle of legal certainty aims to protect. The law must be sufficiently clear, intelligible and precise, so that legal subjects can predict how it will affect them and how they need to change or adapt their behaviour to ensure it is in line therewith.⁸⁹ This also implies the need for clarity about how the law is *applied* by public authorities. However, if one recalls the earlier discussion about the opacity that often surrounds the use and parameters of algorithmic systems,⁹⁰ it is clear that this requirement can come under pressure.

If the law’s application is mediated through algorithmic regulation, and if the way in which this system operates is not communicated (or, in case of certain data-driven systems, if its operations are unintelligible) how can the principle of legal certainty

⁸⁶ As the Venice Commission stated in its rule of law checklist: “*stability is not an end in itself: law must also be capable of adaptation to changing circumstances. Law can be changed, but with public debate and notice, and without adversely affecting legitimate expectations.*” Venice Commission, ‘Rule of Law Checklist’ (n 57), §60.

⁸⁷ See also *supra*, Section 3.3.3.

⁸⁸ This collective responsibility of all state institutions was also stressed by the Venice Commission in its Opinion on the Dutch child care allowance case. See, e.g., Barrett and others (n 34) 27.

⁸⁹ See *supra*, Section 3.2.2. See also Joseph Raz, ‘The Rule of Law and Its Virtue’, in his *The Authority of Law: Essays on Law and Morality* (Oxford University Press 1979).

⁹⁰ See *supra*, Sections 2.4.4 and 4.1.1.c.

be met? How can one have certainty not only about the law that will be applied, but also about the way in which public officials make use of their discretion when they apply the law? Without such transparency, there can be no oversight, whether by citizens or by other branches, of public authorities' actions that rely on algorithmic regulation, and hence no assurance that these actions are in line with human rights, democracy and the rule of law.

Though the need for transparency about the implementation of laws and policies by public authorities seems evident, the reluctance of providing information about algorithmic systems that are used *for this very purpose* shows this is not a given. This is evidenced by the various cases brought before courts due to the refusal of public authorities to offer information about the algorithmic systems they are using – including those that inform and take administrative acts. I have already discussed the opacity surrounding the Belgian OASIS tool,⁹¹ yet one can also point out the obstacles faced by individuals who sought information about the abovementioned Polish unemployment profiling algorithm.⁹² Accessing information was also an obstacle in the context of the Swedish Trelleborg algorithm dealing with the allocation of welfare benefits,⁹³ and the controversial French Admission Post Bac algorithm, which automatically assigned students to higher education institutions.⁹⁴ Also in the latter case, dismayed students were forced to sue the relevant public authority to enforce their public information right, after obtaining negative replies to their requests for information. Note that there is currently no uniform answer across the EU as to whether the source code of algorithmic systems deployed by public authorities is considered as public information that can be requested pursuant to an access to information request.

Of course, transparency on how public authorities implement legislation is *always* a challenge, given the inherent asymmetry of information between the government and its citizens. However, the use of algorithmic systems as intermediaries between public authorities and citizens can further diminish transparency by adding an additional layer of opacity, one that is not easily pierced.⁹⁵ Furthermore, as demonstrated by the example of Hackney's Child Risk Assessment System, the problem can be aggravated when the commercial or intellectual property rights of the private

⁹¹ Degrave (n 45).

⁹² Niklas, Sztandar and Szymielewicz (n 63).

⁹³ This example is not discussed in this book, but see in this regard Anne Kaun, 'Suing the Algorithm: The Mundanization of Automated Decision-Making in Public Services through Litigation' [2021] *Information, Communication & Society* 1.

⁹⁴ Restrepo Amariles (n 19) 284. See also Madeleine Thompson, 'The French Educational Algorithm of Inefficiency' (Brown Political Review 8 November 2016) <<https://brownpoliticalreview.org/2016/11/french-educational-algorithm/>>.

⁹⁵ See also Ada Lovelace Institute, AI Now Institute, and Open Government Partnership, 'Algorithmic Accountability for the Public Sector' (2021) 7 <www.opengovpartnership.org/documents/algorithmic-accountability-public-sector/>.

company who developed the system are invoked.⁹⁶ This essentially comes down to the supremacy of a private interest (which can already be called dubious given the influence it implies of a private party over public policy), over a public interest. Beyond the concerns this might raise for individuals who are directly affected by the system, this problem is a societal one, as it reduces the possibility for government control more generally.⁹⁷

These elements collectively reveal that the implementation of algorithmic systems to inform or take administrative acts does not unassumingly enhance legal certainty. Rather, it raises several challenges for the attainment of this principle, and risks making it more difficult to achieve the delicate balance between stability and adaptability, thereby potentially exacerbating the tensions that are already part of the rule of law.

4.1.3 Non-arbitrariness

As discussed in Section 3.3, the principle of non-arbitrariness requires public authorities to act in a non-arbitrary fashion, meaning impartially, reasonably, efficiently, fairly and timely. They should be able to justify their actions, and use their discretion in a way that balances the various interests involved, guided by the effects of the measures they adopt and taking into account the factors relevant to the case.⁹⁸ Moreover, public authorities should only use their power to attain the specific purpose for which it was granted to them,⁹⁹ and must put in place mechanisms against the risk of corruption and the potential abuse or misuse of discretion – including the abuse or misuse of (personal) information retained by the authorities.¹⁰⁰ With this recap in mind, how are these requirements affected by reliance on algorithmic systems to inform or take administrative acts?

⁹⁶ See *supra*, Section 4.1.1.c.

⁹⁷ For this reason, France introduced new legislation providing inter alia that – as a default – the use of algorithms in public decision-making should be made transparent to the public, and setting out minimum information that should be provided by public authorities. See LOI no. 2016-1321 du 7 octobre 2016. See also OECD, 'A Data-Driven Public Sector: Enabling the Strategic Use of Data for Productive, Inclusive and Trustworthy Governance', vol 33 (2019) OECD Working Papers on Public Governance 6, 40.

⁹⁸ See Denis James Galligan, 'Discretionary Powers in the Legal Order', in his *Discretionary Powers: A Legal Study of Official Discretion* (Oxford University Press 1990).

⁹⁹ See *supra*, Section 3.3.3.

¹⁰⁰ Recall in this regard also Committee of Ministers of the Council of Europe, 'Recommendation No. R (87) 15 of the Committee of Ministers to Member States Regulating the Use of Personal Data in the Police Sector', Adopted by the Committee of Ministers on 17 September 1987 at the 410th meeting of the Ministers' Deputies (1987); Committee of Ministers of the Council of Europe, 'Recommendation No. R (91) 10 of the Committee of Ministers to Member States on the Communication to Third Parties of Personal Data Held by Public Bodies', Adopted by the Committee of Ministers on 9 September 1991 at the 461st meeting of the Ministers' Deputies (1991).

Prima facie, one might purport yet again that the introduction of algorithmic regulation can contribute to the attainment of this rule of law principle, for reasons already explored in previous sections. Reliance on algorithmic systems diminishes discretion at the level of individual public officials – which could potentially be used arbitrarily or in a way that overly deviates from the law – thereby also diminishing the risk of its arbitrary use. Instead, public officials’ discretionary power could be replaced by ‘evidence-based’, streamlined and centralised automated suggestions and decisions. While these aspirations sound promising in theory, their promise entails more than one catch.

4.1.3.a Optimising Efficiency over Justice

A first catch relates to the aspiration of increased efficiency, which is an important goal of bureaucratic organisation more generally. Yet as hinted at above, the very alignment of bureaucratic and algorithmic logic can also obscure that an increase in ‘efficiency’ might come at the cost of substantive values that public authorities should strive for. An efficient administration is but one of the requirements under the non-arbitrariness principle and should be seen as a means rather than an end in itself – the actual end being: serving citizens in the public interest. By unduly focusing on efficiency, the underlying normative aims of public policies risk being pushed to the sideline, and this can have problematic consequences for the persons involved.

Recall the example of Indiana discussed earlier, where eligibility processes for several welfare programs were automated in an overly narrow manner. In addition to the automated denial of benefits, each time the system perceived a ‘lack of cooperation’ through something as banal as a missed phone call, Eubanks noted that *“performance metrics designed to speed eligibility determinations created perverse incentives for call centre workers to close cases prematurely”*.¹⁰¹ Indeed, *“timelines could be improved by denying applications and then advising applicants to reapply, which required that they wait an additional 30 or 60 days for a new determination”*.¹⁰²

This measuring of success by the number of cases that have been handled (which is quantifiable) rather than by whether people were adequately helped (which is a more difficult metric), also adds to the pressure placed on public officials. It prevents them from deviating from the system and reducing ‘efficiency’, even if this would contribute to other normative values. This problem was also highlighted by the officials who worked with the Polish unemployment system, which can be seen as another example of efficiency gone rogue. As mentioned above, the Polish unemployment agency deployed an algorithmic system to assist in decisions about whether and which unemployment aid and programmes would be offered to

¹⁰¹ Eubanks (n 25) 51.

¹⁰² *ibid.*

citizens, based on their ‘employability’.¹⁰³ Also here, the aim of the system’s implementation was the optimisation of public resources by rendering the process of resource allocation more efficient, and diminishing the risk of arbitrary decisions.

In practice, this ‘efficiency’ aim translated itself to profiling and categorising citizens in an automated way, including a category of people who were seen as ‘lost cases’ (in casu, Profile III). People categorised as such would not be eligible for the labour market programs designed to help them find employment, based on opaque and potentially discriminatory criteria.¹⁰⁴ Furthermore, despite the fact that the algorithmic system was introduced to reduce disparities among local offices by ‘standardising’ the process, empirical evidence suggests that this did not diminish the arbitrary nature of the categorisation. One reason for this was the fact that, during conversations with unemployed citizens, public officials had to deal with answers to questions that were not anticipated by the coders and hence not programmed in the algorithmic system. Based on interviews that Panoptikon conducted with public officials, it appeared that one of those questions concerned “*reasons making it difficult to take up work*”, where answers like “*homelessness*” or “*criminal record*” were lacking and could hence not be processed by the system.¹⁰⁵ To remedy this problem:

The first interviewed counselor suggested that if the unemployed admitted to being homeless, she would either chose ‘too much competition’ or ‘health restrictions’ and ‘lack of job-seeking skills and self-presentation’, depending if the obstacle is only the lack of formal place of residence (‘employer does not want to hire persons without a residence address’ [PUP 3]) or hygiene (a person ‘is dirty and stinks’ [PUP 3]). The second counselor explained that – since homelessness is usually accompanied by other difficulties – she would try to identify other relevant answers to this question, ignoring homelessness as a specific cause: for instance, ‘health restrictions’ or ‘a lack of conviction about the necessity to take up a job’ [PUP 6]. Another suggested solution was to make sure that a person is eventually included in Profile III as a person ‘distant from the labor market’, no matter what the result of the automated scoring will be

[PUP 3].¹⁰⁶

To defend this profiling practice, and the exclusion of Profile III persons from receiving any help, the authorities emphasised it would advance evidence-based decision-making, based on ‘scientific methods’ through the “*combination of an individual ‘examination’ of a person and econometric elements*”.¹⁰⁷ The narrative that algorithmic regulation can supplant potentially arbitrary decisions by public

¹⁰³ See *supra*, Section 4.1.2.a.

¹⁰⁴ Niklas (n 67).

¹⁰⁵ Niklas, Sztandar and Szymielewicz (n 63) 28.

¹⁰⁶ *ibid.*

¹⁰⁷ *ibid* 7.

officials with ‘objective data analysis’ based on ‘science’ is a recurring theme, notwithstanding the fact that both the data and the indicators relied upon by an algorithmic system remain the result of human choices, and can hence likewise be biased.

Accordingly, both in the case of Indiana’s welfare eligibility system and Poland’s unemployment system, the aim of efficiency overshadowed the aim of ensuring that people who need help get help. In both cases, quantifiable economic targets were prioritised over social policies and values. And while the outcome of an automation process does not necessarily need to lead to such problematic prioritisation, these examples do demonstrate that the implementation of automation tools requires extra attention to this risk, especially in a bureaucratic environment which already lends itself to over-emphasising procedural rationality.¹⁰⁸

4.1.3.b Reducing Explainability

To counter the risk of arbitrariness and to ensure compliance with the principle of legality, public authorities must justify the administrative acts they take. In the context of algorithmic regulation, this means that transparency about the underlying choices as regards the system’s parameters, data and model design should be provided, so that individuals can understand the reasons behind the decision.¹⁰⁹ If this is lacking, those subjected to the system’s outcomes are unable to assess the lawful nature of the action and to challenge it where need be. This also counts for the legislative and executive branch, who should be able to ‘check and balance’ the executive’s power.

As explained in Chapter 2, when public authorities deploy algorithmic regulation that is based on knowledge-driven approaches, the system’s operations are typically more intelligible and explainable.¹¹⁰ In principle, this renders it more straightforward for authorities to provide an explanation.¹¹¹ However, even when knowledge-driven approaches are used, a meaningful explanation can still be missing if authorities neglect to provide information about the system’s functioning. Indeed, as noted in Chapter 2, opacity need not be technical in nature, but can also stem from human choices. More importantly, in some situations, the public officials who deploy the system may not necessarily know or understand how the system functions, as they are typically not part of the development process. The example of the

¹⁰⁸ See *supra*, Section 2.3.

¹⁰⁹ See in this regard Melanie Fink and Michèle Finck, ‘Reasoned A(I)Administration: Explanation Requirements in EU Law and the Automation of Public Administration’ (2022) 47 *European Law Review* 376.

¹¹⁰ See *supra*, Section 2.1.4.

¹¹¹ However, see also Mike Ananny and Kate Crawford, ‘Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability’ (2018) 20 *New Media & Society* 973.

Polish unemployment algorithm was telling in this regard. Pursuant to a deliberate choice from the Polish Ministry, the public officials received no insight into the system's operations and the precise parameters that led to the recommended outcomes.¹¹² The interviews with the officials also revealed that citizens who requested an explanation about the system's functioning were treated with suspicion based on the mere fact that they asked for more information.¹¹³

When the decision-making process hinges on a data-driven model that provides recommendations based on deep learning or other 'non-explainable' methods, public authorities' duty to state the reasons for their administrative acts is even more difficult to fulfil. In those situations, even if they want to comply with their obligation, public officials may be unable to provide a meaningful explanation of the system's outcomes. The use of such systems therefore seems even more difficult to reconcile with the core tenet that public authorities should be able to motivate their decisions, and with the more general requirement of transparency that accompanies the practice of automated data processing pursuant to the GDPR. In past case law, the CJEU therefore distinguished algorithmic systems that deploy 'predetermined criteria' to profile citizens from systems that rely on machine learning approaches, given the opacity of the latter and the fact that "*it might be impossible to understand the reason why a given program arrived at a positive match.*"¹¹⁴

4.1.3.c Diminishing Discretion

Besides obscurity about what the system is optimised for, one can also raise questions over who decides about such optimisation, and how this happens. While the introduction of algorithmic regulation diminishes discretion at the level of individual public officials, it does not eliminate discretion entirely. Instead, discretion shifts to those who code the algorithmic system.¹¹⁵ This means that normatively relevant decisions about the aims that public policies should be optimised for, or the quantitative criteria that should be considered in the context of taking an administrative act, are henceforth not taken by specialised public officials who are responsible for the administrative act, but by coders.¹¹⁶ In many cases, these coders are private company employees to which the development of the algorithmic system was

¹¹² See Niklas, Sztandar and Szymielewicz (n 63) 32.

¹¹³ *ibid* 31.

¹¹⁴ See Case C-817/19, *Ligue des droits humains v Conseil des ministres* (PNR Case), ECLI:EU:C:2022:491, §195.

¹¹⁵ See Mark Bovens and Stavros Zouridis, 'From Street-Level to System-Level Bureaucracies: How Information and Communication Technology Is Transforming Administrative Discretion and Constitutional Control' (2002) 62 *Public Administration Review* 174, 181.

¹¹⁶ See Burrell and Fourcade (n 42) 223.

outsourced, since public authorities today often still lack the relevant know-how to do so in-house. In that case, discretion is essentially also outsourced.¹¹⁷

This means that algorithmic regulation radically alters the nature of discretion, as it can no longer function as a potential correction to the downsides of the law's general nature, and as a tool to ensure that, on a case-by-case basis, the administrative acts through which legislation is implemented are appropriate for the specific situation.¹¹⁸ Instead, discretion is centralised, hierarchised and – literally and figuratively – systematised, thereby arguably losing its essence. If discretion is only present in decisions about how an algorithmic system that implements the law is codified, and if it is absent when it comes to the law's actual application to individual cases, then it is no longer capable to play its corrective role. Not when the system accidentally creates adverse effects for individuals and society. But also not when the system has been deliberately codified in a way that is incompatible with the rule of law's principles, including the legality principle which requires compliance with hierarchically higher norms such as human rights and EU law more generally.

Furthermore, embedding rules into a coded infrastructure can remove the possibility of deviating from the codified rule, even in cases where this may be necessary from a legal or moral perspective. This undermines public officials' agency and forces obedience to the rule through the infrastructure's architecture.¹¹⁹ I have already discussed the risk of automation bias, and the relative ease with which individuals tend to rely on the authority of algorithmic systems, particularly given their superior computational skills.¹²⁰ This occurs even more in contexts of time pressure, when people do not have the time to double-check the system's suggestion, or in contexts of scarcity of information, when people lack the data or knowledge to assess the system's reliability. As pointed out by Hildebrandt, *"even in the case of decision-support instead of decision-making, human intervention becomes somewhat illusory, because those who decide often do not understand the 'reasons' for the proposed decision. This induces compliance with the algorithms, as they are often presented as 'outperforming' human expertise."*¹²¹ Even when a system is 'merely' meant to inform an administrative act, it will in practice still render it difficult for

¹¹⁷ See in this regard also Paul R Verkuil, *Outsourcing Sovereignty: Why Privatization of Government Functions Threatens Democracy and What We Can Do about It* (Cambridge University Press 2007).

¹¹⁸ See also *supra*, Section 2.3.3.

¹¹⁹ See in this regard also Arre Zuurmond, *De infocratie: Een Theoretische en empirische heroriëntatie op Weber's idealtipe in het informatietijdperk* (Phaedrus 1994) 1. See also Karen Yeung, 'Can We Employ Design-Based Regulation While Avoiding Brave New World' (2011) 3 *Law, Innovation and Technology* 1.

¹²⁰ See *supra*, Section 2.2.6.

¹²¹ Mireille Hildebrandt, 'Algorithmic Regulation and the Rule of Law' (2018) 376 *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 20170355, 2.

officials to deviate therefrom, as this typically requires an additional step of explanation to someone higher up the hierarchy to justify why a suggestion is not followed.

The constraints to deviate from the system can also come from other factors. Money will have been invested in the system's implementation to gain time and efficiency.¹²² If public officials are still required to spend time on making their own assessment of a case, regardless of the system's recommendation, that investment will be less cost-efficient and might undermine possible KPIs that officials are required to fulfil. Indeed, through the impetus of the New Public Management approach, which introduced indicators and KPIs into public decision-making, officials may be even more incentivised to focus on the number of cases they can close, or the number of decisions they were able to take – regardless of whether these decisions also do justice to the situation at hand from a normative perspective.¹²³ Deviating from the algorithmic suggestion rather than simply ratifying might hence endanger the achievement of those KPIs.

The tendency to follow the algorithm's advice has also been corroborated, for instance, in the context of the KrimPro system used by the German Police in Berlin. KrimPro is a predictive policing system which displays predicted areas on a map showcasing the geographical distribution of the probable 'high-risk' areas in and around Berlin for domestic burglaries and other crimes, based on which decisions are taken regarding the optimisation of resources.¹²⁴ When researchers conducted in-depth interviews about the system's use and utility within the police forces, this revealed a relatively strong pressure to conform to the recommendations of the system. As one interviewee put it: *"I do not risk anything because even if I find it stupid and nothing happens there or even if something happens, it will not be my responsibility. I have not done anything wrong."*¹²⁵ In this regard, Lorenz et al. note that

if the police professionals who are responsible for fighting domestic burglaries reject the prediction and therewith additional units and a crime is committed that might

¹²² See also Maciej Kuziemski and Gianluca Misuraca, 'AI Governance in the Public Sector: Three Tales from the Frontiers of Automated Decision-Making in Democratic Settings' (2020) 44 *Telecommunications Policy* 101976.

¹²³ In this regard, reference can be made to the vast academic literature which addresses the perverted impact of performance management in the public sector. For a general overview, see, e.g., E. Buschor, 'Performance Management in the Public Sector: Past, Current and Future Trends' (2013) 11 *Tékhné* 4. See also Ulrik Hvidman and Simon Calmar Andersen, 'Impact of Performance Management in Public and Private Organizations' (2014) 24 *Journal of Public Administration Research and Theory* 35.

¹²⁴ Lukas Lorenz, Albert Meijer and Tino Schuppan, 'The Algocracy as a New Ideal Type for Government Organizations: Predictive Policing in Berlin as an Empirical Case' (2021) 26 *Information Polity: The International Journal of Government & Democracy in the Information Age* 71, 77.

¹²⁵ Albert Meijer, Lukas Lorenz and Martijn Wessels, 'Algorithmization of Bureaucratic Organizations: Using a Practice Lens to Study How Context Shapes Predictive Policing Systems' (2021) 81 *Public Administration Review* 837, 842.

have been prevented by these units, they put themselves in a bad light. On the other hand, the heads of the inspections do not risk anything when they just comply with the assessment provided by the KrimPro report and deploy additional units even if these extra efforts appear to be ineffective.¹²⁶

A comparative study was conducted of an algorithmic system used by the Dutch police, where the relationship between the police officers and the algorithmic system appeared to be more collaborative rather than hierarchical. Meijer et al. note on this basis that *“two patterns of algorithmisation of government bureaucracy can be identified and that these patterns depend on dominant social norms and interpretations rather than the technological features of algorithmic systems.”*¹²⁷ In other words, reliance on algorithmic regulation need not necessarily lead to a heavy curbing of agency. However, in environments that are bureaucratic in nature and that already leave more limited scope for critical reflection, the use of such systems can reinforce these tendencies, including the push towards obedience to authority.

In this regard, Endicott and Yeung stressed that public agency is an important corollary of the government's legal accountability.¹²⁸ They conceptualise it as follows: *“the community must make itself capable of deciding and acting responsibly, by empowering and requiring officials and institutions to undertake demonstrably reasoned action on its behalf in certain crucial respects”*.¹²⁹ As they convincingly argue, public agency is a prerequisite for a responsible government and hence for the rule of law, since *“no community can be ruled by law unless public agencies are empowered by the law to take reasoned decisions to make and to apply the law”*.¹³⁰ However, when public authorities delegate administrative acts to algorithmic systems, either indirectly by uncritically relying on their recommendations, or directly by adopting an automated decision-process, such agency gets eroded. It not only gets eroded at the level of public officials, but it can also get eroded at the level of the public authority itself when it outsources the system's development to coders who work for a private company or who are in any case untrained to make judgments about administrative acts that can significantly affect individuals.

Importantly, a set-up which leaves little scope for judgment or critical reflection by public officials (and in fact discourages it) also risks detaching the human decision-maker from both the decision and its consequences. Recall in this regard the parallels with Milgram's experiment, which were discussed in Section 2.2.6, and particularly his warning that individuals tend to adopt certain ‘buffers’ to shield themselves from a sense of responsibility when their actions lead to adverse

¹²⁶ Lorenz, Meijer and Schuppan (n 124) 81.

¹²⁷ Meijer, Lorenz and Wessels (n 125) 837.

¹²⁸ Timothy Endicott and Karen Yeung, ‘The Death of Law? Computationally Personalized Norms and the Rule of Law’ [2022] University of Toronto Law Journal 72(4), 378.

¹²⁹ *ibid.*

¹³⁰ *ibid.* 379.

consequences for those subjected thereto, especially if mediated by a machine.¹³¹ This emotional detachment can be enhanced with physical distance between the decision-maker and the subject, as well as the many hands problem¹³² discussed above. In sum, the diminished discretion and agency of public officials can – in the name of efficiency and an alleged reduction of arbitrary decision-making – undermine public officials’ sense of responsibility for the administrative acts they take with algorithmic systems, and hence undermine the rule of law’s overarching telos.

4.1.4 *Equality before the Law*

The principle of equality before the law requires public authorities to treat persons equally and in a non-discriminatory way. Natural and legal persons can only be treated unequally in case there is a justifiable ground or motivation for their differentiation.¹³³ As discussed in Chapter 3, one of the main challenges arising from the principle of equality centres on the question of *when* a differentiation is justifiable, and *when* the very lack of a differentiation may be considered unjustifiable.¹³⁴ In the context of algorithmic regulation, one can recall that the very purpose of algorithmic systems consists of making automated differentiations and categorisations between various types of data (including data about individuals) in order to apply legal rules to particular (categories of) cases in a more efficient, objective and speedy way. The question is therefore: how does reliance on algorithmic regulation contribute to this principle’s attainment or to its challenge?

4.1.4.a Risk of Scaled Bias

Legal rules are abundant with categorisations and, as I discussed above, many of these categories are over- and under-inclusive given the law’s general nature. One could hence contend that algorithmic regulation might help refine the law’s overly rudimentary categorisations by conducting a more ‘personalised’ analysis of citizens’ data and thereby contributing to *substantive* equality – especially when based on data-driven techniques that can identify distinctions that public officials could not easily perceive by themselves. Several scholars have put forward arguments along

¹³¹ See *supra*, Section 2.2.5.

¹³² *ibid.* See also, e.g., Dennis F Thompson, ‘Designing Responsibility: The Problem of Many Hands in Complex Organizations’ in Jeroen van den Hoven, Seumas Miller and Thomas Pogge (eds), *Designing in Ethics* (1st edn, Cambridge University Press 2017). See also Helen Nissenbaum, ‘Accountability in a Computerized Society’ (1996) 2 *Science and Engineering Ethics* 25.

¹³³ See *supra*, Section 3.3.4.

¹³⁴ See also Sandra Wachter, Brent Mittelstadt and Chris Russell, ‘Why Fairness Cannot Be Automated: Bridging the Gap between EU Non-Discrimination Law and AI’ (2021) 41 *Computer Law & Security Review* 105567.

these lines.¹³⁵ Ben-Shahar and Porat have, for instance, argued in favour of a technology-driven personalisation of the law.¹³⁶ The promise of algorithmic regulation in this context, as Endicott and Yeung observed, “*is that it could take social ordering beyond the crude, impersonal techniques of law, with its clumsy dependence on general rules*”.¹³⁷ However, they also point out that this would pose significant challenges for the rule of law.¹³⁸

In the context of the principle of equality, one can question to which extent the (more refined) categorisations or distinctions proposed by an algorithmic system are *justifiable* from a legal perspective, since the validity of categorisations (whether in the law or in the law’s application) hinges upon their justifiability. It is here that the limits of algorithmic regulation come to the fore, for whether a distinction based on a certain criterion is justifiable or not cannot be determined by an algorithm. Algorithmic systems can propose categorisations based on data they receive, and hence based on data about how things *are*, but they will not be able to say anything about how things *should* be. Otherwise, this would constitute a conflation between the normative and the positive, thereby committing the error of the is–ought fallacy.¹³⁹

When speaking of equality, it is also important to recall the discussion in Chapter 2 about algorithmic systems’ risk of biased decision-making. While their machine-like nature and reliance on data may make it seem like they are neutral and objective decision tools, algorithmic systems merely reflect the values and value-laden choices that are embedded in their components and environment. This means that the systems’ coders have a significant influence on the potential (unjust) bias that may be reflected in the system’s outcomes, and on the validity of the distinctions and categorisations that the system will make (whether they pre-programmed these distinctions or whether these distinctions are derived from the

¹³⁵ Casey and Niblett noted, for instance that “*as technologies associated with big data, prediction algorithms, and instantaneous communication reduce the costs of discovering and communicating the relevant personal context for a law to achieve its purpose, the goal of a well-tailored, accurate, and highly contextualized law is becoming more achievable.*” See Anthony Joseph Casey and Anthony Niblett, ‘A Framework for the New Personalization of Law’ (2019) 86 The University of Chicago Law Review 333, 335. See also Horst Eidenmüller and Gerhard Wagner, *Law by Algorithm* (Mohr Siebeck 2021).

¹³⁶ Omri Ben-Shahar and Ariel Porat, *Personalized Law: Different Rules for Different People* (Oxford University Press 2021). For an important critique on their arguments, see, e.g., Endicott and Yeung (n 128).

¹³⁷ Endicott and Yeung (n 128).

¹³⁸ As discussed *supra*, Section 2.2.5.

¹³⁹ The articulation of the is–ought problem is most notably ascribed to David Hume. See David Hume, *A Treatise of Human Nature: Being an Attempt to Introduce the Experimental Method of Reasoning into Moral Subjects and Dialogues Concerning Natural Religion* [1739] (LA Selby-Bigge ed, Clarendon Press 1896). See also Max Black, ‘The Gap Between “Is” and “Should”’ (1964) 73(2) *The Philosophical Review* 165–81; Nicoletta Bersier Ladavac, Christoph Bezemek and Frederick Schauer (eds), *The Normative Force of the Factual: Legal Philosophy between Is and Ought*, vol 130 (Springer International Publishing 2019).

data they gathered and labelled). A vast scholarship exists on how algorithmic regulation impacts the principle of equality and the right to non-discrimination, which I will not be repeating here.¹⁴⁰ Yet suffice it to note that reliance on biased algorithmic systems can lead to unjustifiable discrimination, regardless of how such bias manifests itself.

Consider the example of the algorithmic system used by Allegheny's child welfare agency in Pennsylvania, which was meant to enable the more efficient identification of families where children ran a risk of being neglected or abused, and hence to optimise resources by prioritising further investigations for those flagged families in particular.¹⁴¹ Due to a biased design, the system flagged "*a disproportionate number of black children for a 'mandatory' neglect investigation, when compared with white children*".¹⁴² By deploying this system, Allegheny's child welfare agency not only inflicted harm at the level of the individual families that were affected thereby, but it also undermined the general principle of equality which is an essential societal interest. There can be no rule of law if the law is applied unequally to people based on the mere colour of their skin – and a breakdown of the rule of law is problematic for all members of society rather than just for those who are targeted by a specific system.

An additional problem in the context of (primarily data-driven) algorithmic systems concerns the risk of discrimination through proxies. Even if prohibited discrimination grounds such as ethnicity or gender are purposely *not* taken into consideration when training or deploying an algorithmic system, these grounds can still *implicitly* contribute to a biased outcome by virtue of their strong relationship with seemingly neutral datapoints.¹⁴³ Since the elimination of discrimination through proxies is notoriously difficult (eliminating data that is related to prohibited discrimination grounds can in fact result in a lack of sufficient data to carry out an analysis in the first place), this risk should always be considered when algorithmic regulation is used. In the case of Allegheny's system, researchers who received access to the relevant data and conducted an investigation of the system's deployment found that the algorithmic system "*on its own was more racially disparate than*

¹⁴⁰ See also Toon Calders and Indrė Žliobaitė, 'Why Unbiased Computational Processes Can Lead to Discriminative Decision Procedures' in Bart Custers and others (eds), *Discrimination and Privacy in the Information Society*, vol 3 (Springer Berlin Heidelberg 2013); Parliamentary Assembly of the Council of Europe (PACE), 'Recommendation 2183 (2020) on Preventing Discrimination Caused by the Use of Artificial Intelligence' (PACE 2020) Recommendation.

¹⁴¹ See Dan Hurley, 'Can an Algorithm Tell When Kids Are in Danger?' *The New York Times* (2 January 2018) <www.nytimes.com/2018/01/02/magazine/can-an-algorithm-tell-when-kids-are-in-danger.html>.

¹⁴² Sally Ho and Garance Burke, 'An Algorithm That Screens for Child Neglect Raises Concerns' *Associated Press* (29 April 2022) <<https://apnews.com/article/child-welfare-algorithm-investigation-9497ee937e0053ad4144a86c68241ef1>>.

¹⁴³ Frederik J Zuiderveen Borgesius, 'Discrimination, Artificial Intelligence, and Algorithmic Decision-Making' (Council of Europe – Directorate General of Democracy 2018) 13.

workers, both in terms of screen-in rate and accuracy”.¹⁴⁴ While it is not disputed that human administrators can be biased too, this case illustrates that an algorithmic system can in some situations be even more biased, with as an additional feature the fact that its biased administrative decisions can be implemented instantaneously and at population scale.

This point can also be illustrated by revisiting the algorithmic system used in the Dutch childcare allowance case.¹⁴⁵ I have already noted how the public authority’s narrow interpretation of the existing legal rule marked an excessive number of families as potential fraudsters even without any fraudulent intent, and led to high demands for repayment. In addition, the tax authority also relied on a data-driven algorithmic system that helped flag potential cases of fraud which it had learned from a dataset with examples of past correct and incorrect applications.¹⁴⁶ Problematically, “one of the many indicators used to identify fraud cases was citizenship, and applicants with foreign origin were selected by the system for detailed scrutiny of their applications”.¹⁴⁷ Accordingly, the combination of a system that (1) codified a legal rule in an overly legalistic manner due to the excessively narrow interpretation thereof by the public authority, and that (2) relied on discriminatory criteria, led to a disproportionate targeting of people with a foreign background – many of which went through financial and social dramas.¹⁴⁸

Also here, it can be pointed out that the Dutch tax authority need not have recourse to an algorithmic system to rely on a problematic proxy to identify fraudulent behaviour. Human beings are perfectly capable of using discriminating factors in their decision-making without any assistance from a machine. However, the automation of the process renders potentially *beneficial* deviations from the discriminatorily codified system more difficult, in addition to the vast scale on which a problematic proxy can be applied.

4.1.4.b Exacerbating Societal Inequality

It is precisely this scale that enables algorithmic systems to not only reproduce but also to exacerbate discriminatory tendencies in society. When data-driven systems are deployed to categorise natural and legal persons (whether as (in)eligible for benefits, or as (un)likely fraudsters, as criminals or as having other undesirable

¹⁴⁴ Hao-Fei Cheng and others, ‘How Child Welfare Workers Reduce Racial Disparities in Algorithmic Decisions’, *CHI Conference on Human Factors in Computing Systems* (ACM 2022) 3 <<https://dl.acm.org/doi/10.1145/3491102.3501831>>.

¹⁴⁵ See also Parliamentary Inquiry Committee (n 33).

¹⁴⁶ See Autoriteit Persoonsgegevens, ‘Onderzoeksrapport Belastingdienst/Toeslagen: De verwerking van de nationaliteit van aanvragers van kinderopvangtoeslag’ (2020) 22018–22445, 14 <https://autoriteitpersoonsgegevens.nl/uploads/imported/onderzoek_belastingdienst_kinderopvangtoeslag.pdf>.

¹⁴⁷ Barrett and others (n 34) 4.

¹⁴⁸ See Parliamentary Inquiry Committee (n 33).

features) these systems necessarily rely on datasets that reflect a state of play from the past. This means that societal inequalities that are reflected in the datasets risk being perpetuated in the algorithmic system. Importantly, this risk is not limited to data-driven systems, since also in knowledge-driven systems the algorithmic categories can be based on biased assumptions. In addition, precisely due to historical inequalities, there is less data available about certain population groups (so-called data gaps), which tends to reduce the accuracy of algorithmic models and outcomes as concerns these populations.¹⁴⁹

Algorithmic systems in administrative decision-making are often also deployed in contexts that are inclined to focus disproportionately on individuals or groups that are more vulnerable.¹⁵⁰ This is not surprising. After all, many of the public welfare programmes organised by public authorities are precisely targeted at helping the most vulnerable in society and ensure they can enjoy a range of social and economic rights.¹⁵¹ At the same time, this very vulnerability can also turn them into targets when it comes to the assessment of risks – which can also be seen in examples of algorithmic regulation, such as the system used by Allegheny's child welfare agency. As explained by Eubanks, one of the risk assessment factors on which that system relied concerned the use of social services, like a parent's access to mental healthcare services in a clinic funded by Medicaid.¹⁵² These clinics are obligated to report medical records to the state, which means their patient data can be analysed by state-deployed algorithmic systems – such as the Allegheny system. Importantly, Eubanks points out that private clinics – which are typically more expensive – are *not* obligated to share their records with the state. The algorithmic system would hence *not* pick up a parent's access to mental healthcare services – and the stigma and potential risks associated therewith – if the parent is wealthy enough to afford a private clinic.¹⁵³ Evidently, this also means that the system can disproportionately target people based on their financial situation.

For another example, consider the Harm Assessment Risk Tool (HART) used by law enforcers in Durham. The system was developed by researchers of the University of Cambridge together with Durham Constabulary, to help custody offers take decisions about offenders' eligibility for the so-called Constabulary's Checkpoint

¹⁴⁹ See in this regard also Rónán Kennedy, 'The Rule of Law and Algorithmic Governance' in Woodrow Barfield (ed), *The Cambridge Handbook of the Law of Algorithms* (1st edn, Cambridge University Press 2020) 216.

¹⁵⁰ See also Sofia Ranchordas and Luisa Scarcella, 'Automated Government for Vulnerable Citizens: Intermediating Rights' (2021) 30 William & Mary Bill of Rights Journal 373.

¹⁵¹ This very vulnerability, however, often also leads to higher administrative burdens. See in this regard Julian Christensen and others, 'Human Capital and Administrative Burden: The Role of Cognitive Resources in Citizen-State Interactions' (2020) 80 Public Administration Review 127.

¹⁵² See Eubanks (n 25) 147.

¹⁵³ See also J Khadijah Abdurahman, 'Birthing Predictions of Premature Death' [2022] Logic Magazine <<https://logicmag.io/home/birthing-predictions-of-premature-death/>>.

programme.¹⁵⁴ This programme essentially seeks to deal with an offence outside court prosecution, with the aim of reducing future offences by dealing with the underlying reasons of why a person may be committing a crime (such as drug or alcohol abuse, homelessness or mental health).¹⁵⁵ The system categorises offenders as low, medium or higher risk of re-offending, whereby those presenting a ‘medium’ level of risk can be eligible for the programme. However, concerns arose that the algorithm was “*discriminating people from poorer areas*”,¹⁵⁶ for one of the factors taken into consideration by the model concerned a person’s postal code.

A review of HART notes that

the primary postcode predictor is limited to the first four characters of the postcode, and usually encompasses a rather large geographic area. Yet even with this limitation, one could argue that this variable risks a kind of feedback loop that may perpetuate or amplify existing patterns of offending. If the police respond to forecasts by targeting their efforts on the highest-risk postcode areas, then more people from these areas will come to police attention and be arrested than those living in lower-risk, untargeted neighbourhoods. These arrests then become outcomes that are used to generate later iterations of the same model, leading to an ever-deepening cycle of increased police attention.¹⁵⁷

The mere fact that one resides in a given postcode only affected the system’s outcome indirectly, in combination with other predictive criteria. Nevertheless, the system was altered to address the criticism that it risked categorising a disproportionate number of people from poorer neighbourhoods as high-risk and hence as ineligible for the rehabilitation programme.

To tackle these issues, a lot of research currently focuses on making algorithmic systems ‘fairer’ and ‘eliminating bias’.¹⁵⁸ There are, however, no straightforward solutions to this problem, especially given the fact that ‘fairness’ can be conceptualised and defined in numerous ways.¹⁵⁹ Moreover, the fact that algorithmic bias is virtually always an emanation of underlying structural problems and societal inequalities indicates that mere technical fixes will be unable to sustainably address

¹⁵⁴ Marion Oswald and others, ‘Algorithmic Risk Assessment Policing Models: Lessons from the Durham HART Model and “Experimental” Proportionality’ (2018) 27 Information & Communications Technology Law 223, 225.

¹⁵⁵ Matt Burgess, ‘UK Police Are Using AI to Inform Custodial Decisions – But It Could Be Discriminating against the Poor’ (*Wired UK*, 1 March 2018) <www.wired.co.uk/article/police-ai-uk-durham-hart-checkpoint-algorithm-edit>.

¹⁵⁶ *ibid.*

¹⁵⁷ Oswald and others (n 154) 228.

¹⁵⁸ However, see also Ivo Emanuilov and Katerina Yordanova, ‘Do You Believe in FAIR-y-Tales? An Overview of Microsoft’s New Toolkit for Assessing and Improving Fairness of Algorithms’ (*CITIP blog*, 14 July 2020) <www.law.kuleuven.be/citip/blog/do-you-believe-in-fair-y-ales-an-overview-of-microsofts-new-toolkit-for-assessing-and-improving-fairness-of-algorithms/>; Virginia Dignum, ‘The Myth of Complete AI-Fairness’ in Allan Tucker and others (eds), *Artificial Intelligence in Medicine* (Springer International Publishing 2021).

¹⁵⁹ See Wachter, Mittelstadt and Russell (n 134).

this problem. In that sense, the HART example is almost ironic: while the aim of the programme is to tackle the structural problems underlying crime, being part of a neighbourhood with structural problems risked reducing one's eligibility for this very programme. This showcases once again the limits of quantification as a substitute for qualification.

However, even when deploying a more basic algorithmic system that does not rely on elaborate data-analytics, public authorities can carry out *scaled* discriminatory decision-making. The example of a Dutch algorithmic system to detect welfare fraud – which was basically implemented through a sophisticated excel sheet – is a sad case in point. For years on end about 158 communities who looked for an efficient way to identify and investigate welfare fraud profiled individuals based on parameters that were plainly discriminatory. Although in 2020 the Ministry of Social Affairs spurred communities to stop using this system, noting that it breached the GDPR, researchers¹⁶⁰ revealed that – up until 2022 – a number of communities were still relying on it.¹⁶¹ Indicators of potential fraud included factors like employment as a taxi driver or a hairdresser, residing in a low-income area, or having a low level of education.¹⁶² As the researchers note, these variables have no statistical evidence but are merely a collection of past prejudices.¹⁶³ Furthermore, by analysing the source code, they also stumbled upon hidden fields with the option to profile individuals based on the indicators 'native' or 'foreigner', hence suggesting that, in the past, these systems might also have been used to discriminate people based on nationality. This makes it painfully clear that no complex machine learning systems are needed to adversely affect a large group of citizens with algorithmic regulation.

4.1.4.c Loss of Comparability

A final point to raise under the principle of equality, is the fact that the use of algorithmic systems to inform or take administrative acts also renders it more difficult to know that the equality principle is being infringed in the first place. One can recall the opacity that often surrounds the use of algorithmic regulation – including the fact that such technology is deployed, the parameters it considers, the

¹⁶⁰ 'Junk Science Underpins Fraud Scores' (*Lighthouse Reports*, 25 June 2022) <www.lighthousereports.nl/investigation/junk-science-underpins-fraud-scores/>.

¹⁶¹ See also 'Verboden fraudescores bleven in gebruik bij gemeenten' NRC (25 June 2022) <www.nrc.nl/nieuws/2022/06/25/profileren-verboden-fraudescores-bleven-in-gebruik-bij-gemeenten-a4134660>.

¹⁶² An interactive overview of the 'fraudescorekaart' was made available by the researchers who uncovered the algorithm's continued use, and can be found here: <https://fraudescorekaart.lighthousereports.nl/#/>.

¹⁶³ Lighthouse Reports, 'People All over the Netherlands Have Been Flagged for Welfare Fraud Based on a Crude Spreadsheet & Some Unscientific Prejudices. How Do We Know? We Took the Algorithm Apart & Rebuilt It. Score Yourself & See' (2022), <<https://twitter.com/LHreports/status/1540598392300179456>>.

data it relies on, and the way in which textual legal concepts and rules have been translated to code. The lack of information about these elements also renders it challenging to assess – both by the people subjected to the system and by those who deploy it – whether the principle of equality is being respected.¹⁶⁴ Which type of data was fed into the algorithmic system, and upon which parameters does it rely to inform or take administrative decisions? If those parameters involve distinction grounds that directly or indirectly relate to a prohibited discrimination ground, or that are otherwise problematic, how can their use be challenged, and (how) can public authorities justify these distinctions? The asymmetry of information between those subjected to the system and those who design and develop it requires that the answer to these questions is made explicit, ideally proactively, in order to at least somewhat restore this imbalance.¹⁶⁵

Potential clashes with the principle of equality are moreover not limited to the way in which an algorithmic system is developed, but can also arise from the way in which the system is used. The example of the system deployed by the UK's Office of Qualifications and Examinations Regulation (Ofqual) during the Covid-19 pandemic can illustrate this problem. After exams were cancelled in 2020 in light of the pandemic, high-school teachers were asked to predict what the A-level results of their high-school students would likely have been. Anticipating that those results would be overly optimistic, given that "*evidence suggests that estimated grades will tend towards over-estimation*",¹⁶⁶ Ofqual also deployed an algorithmic system which relied on basic statistical modelling to consider student grades from previous years as well as grades at national level for the same subjects, in order to 'objectivise' the process.¹⁶⁷ As a result, almost 40 per cent of the A-level grades predicted by teachers in England were downgraded.¹⁶⁸ Needless to say, the system's impact on students was significant, especially considering that their grade also determined whether they met the admission requirements set by higher education institutions.

Besides various types of criticism about the system's accuracy and design (on which Ofqual only provided transparency after the grades were assigned), one concern that raised tensions with the equality principle was the fact that the system was only used for schools with more than fifteen children taking an A-level subject.

¹⁶⁴ Emre Bayamlioğlu, 'Contesting Automated Decisions: A View of Transparency Implications' (2018) 4 European Data Protection Law Review 433.

¹⁶⁵ See also Katherine Fink, 'Opening the Government's Black Boxes: Freedom of Information and Algorithmic Accountability' (2018) 21 Information, Communication & Society 1453.

¹⁶⁶ Ofqual, 'Awarding GCSE, AS, A Level, Advanced Extension Awards and Extended Project Qualifications in Summer 2020: Interim Report', Ofqual/20/6656/1, (13 August 2020), 16, <https://assets.publishing.service.gov.uk/media/5f3571778fa8f5173f593d61/6656-1_Awarding_GCSE_AS_A_level_advanced_extension_awards_and_extended_project_qualifications_in_summer_2020_-_interim_report.pdf>.

¹⁶⁷ For an extensive report setting out Ofqual's methodology, see *ibid*.

¹⁶⁸ Will Bedingfield, 'Everything That Went Wrong with the Botched A-Levels Algorithm' [2020] *Wired UK* <www.wired.co.uk/article/alevel-exam-algorithm>.

Where a school had five or fewer children taking an A-level subject, the grades were primarily based on the teacher's predictions, and where a school had between five and fifteen children taking an A-level subject, the grades were based on a combination of the teacher's prediction and the system's prediction.¹⁶⁹ Given the alleged over-estimation of teachers' scores, scores were thus higher for smaller classes – thereby disadvantaging state schools which typically have larger groups of students.¹⁷⁰ At the same time, reliance on the national average grades also led to the penalisation of students at excellent schools, who saw their results downgraded by virtue of an overall lower average.¹⁷¹

Given the public backlash against the system's use, the government ultimately decided to ignore the algorithmic predictions and rely only on the teacher's estimates. This was not a perfect solution either, as Ofqual pointed to the fact that:

studies of potential bias in teacher assessment suggest that differences between teacher assessment and exam assessment results can sometimes be linked to student characteristics, including gender, age within year group, ethnicity, special educational needs, and having English as an additional language. However, such effects are not always seen, and when they are, they tend to be small and inconsistent across subjects.¹⁷²

While human bias hence remained a risk, this was ultimately still preferred over the vastness of the risks presented by the use of the algorithmic system. Latent to this algorithmic antagonism was not only the wider scale of its (problematic) impact, but also its overreliance on statistical modelling, and the fact that it assessed students not only based on their own capabilities, but also on various parameters that did not directly relate to them – such as the size of their classroom and the level of their peers in other schools. In addition, the fact that Ofqual did not previously share detailed information about the system's methodology and parameters, and failed to conduct a prior public consultation to seek feedback, was strongly criticised, as this could have prevented or at least mitigated some of the concerns.

I already discussed the lack of transparency several times.¹⁷³ Yet with respect to the principle of equality, an additional risk can be pointed out, namely the potential loss of comparability between various individuals or groups of individuals. Challenging an administrative act in light of a presumed breach of the principle of equality requires that the subject of the act has information about the fact that other persons were treated differently, despite similar circumstances – or the same, despite

¹⁶⁹ See also the detailed analysis provided by Jeni Tennison, 'How Does Ofqual's Grading Algorithm Work?' (2020) <<https://rpubs.com/JeniT/ofqual-algorithm>>.

¹⁷⁰ Bedingfield (n 168).

¹⁷¹ *ibid.*

¹⁷² Ofqual (n 166) 16.

¹⁷³ The lack of transparency on the rule-making process was discussed in Section 4.1.1.c, and the lack of transparency about the law's application was discussed in Section 4.1.2.c.

different circumstances.¹⁷⁴ However, legal subjects are not always aware of the fact that a particular differentiation is taking place. While laws and regulations of general application are in principle rendered public – in view of the transparency required by the principle of legal certainty and legality – the individual acts taken by public authorities when they apply these laws to specific cases are not always public, which means that it is not always easy for persons to ascertain that they may be treated differently than others who are in a similar situation.

I therefore noted that the principle's requirement of an effective remedy against the discriminatory application of legislation also requires transparency by public authorities on how they interpret a general rule and how they intend to apply it. In a non-algorithmic context, such transparency can take the shape of the publication of administrative guidelines on the methodology used by public officials to apply legal rules to different categories of persons.¹⁷⁵ However, if these guidelines are replaced by more 'refined' or 'personalised' distinctions or categorisations undertaken by virtue of an algorithmic assessment, this means that subjects will not be able to easily compare the methodology applied in their case with the methodology that was applied to other persons – and thus the extent to which a potential difference in treatment was justified. As a consequence of this loss of comparability, the unequal treatment of persons through the differentiated application of general legislation might remain hidden under the algorithmic surface. Evidently, this also affects the possibility of challenging a breach of the principle of equality through the judicial review of administrative action, which I discuss in what follows.

4.1.5 *Judicial Review*

The judicial review of administrative acts aims at ensuring effective judicial protection against executive action that does not comply with the rule of law, including, as part of the legality principle, actions that infringe human rights.¹⁷⁶ In the previous sections, I started my discussion of how algorithmic regulation can impact the principles of the rule of law with an optimistic note, noting that, theoretically at least, it might in fact advance the fulfilment of the principle. How algorithmic systems could enhance the judicial review of administrative acts is not immediately evident, other than the abovementioned aim of preventing officials to deviate from the codified rule or from the algorithmically proposed outcome (which is not the same as preventing them from taking arbitrary or unlawful decisions).

Recall that the principle of judicial review is part of the broader principle of effective judicial protection and access to justice, and that it serves as an overarching point of oversight to ensure that any infringements that occur, despite all the rule of

¹⁷⁴ See *supra*, Section 3.3.4.

¹⁷⁵ See in this regard also Bovens and Zouridis (n 115).

¹⁷⁶ See *supra*, Section 3.3.5.

law requirements imposed on public authorities, can be remedied by means of a review by an independent judge.¹⁷⁷ The judge is hence the last soldier guarding respect for the various other principles that were already discussed, namely legality, legal certainty, non-arbitrariness of executive power and equality before the law.¹⁷⁸ The implementation of the principle of judicial review can, however, face a number of challenges.¹⁷⁹ Are these challenges aggravated by the deployment of algorithmic systems to inform or adopt administrative acts? For the reasons set out below, I believe the answer to that question is positive.

4.1.5.a Informational Limits for Review

To carry out the judicial review of an administrative act, the reviewing judge must have access to the necessary information to assess that act, including the legal basis on which it was grounded, the way in which it was adopted, and the reasons or justification for its adoption.¹⁸⁰ I have already explained that the deployment of algorithmic regulation can hinder the transparency of government action. Such transparency is not only important for those affected by the public authority's system (to decide whether or not they should challenge its use), but also for judges (to assess the action's conformity with the principles of the rule of law). Admittedly, depending on the type of administrative act and the level of discretion that the public authority has, judicial review may be limited in scope to avoid that a judge substitutes a public decision-maker.¹⁸¹ However, even when judicial review is limited to a legality check, the judge still needs to be able to ensure that the way in which the public authority implemented and applied a general legal rule

¹⁷⁷ See in this regard also the Judgment of the ECtHR in *Case of Klass and Others v Germany*, Application no. 5029/71, 6 September 1978, where it stated in §55 that

One of the fundamental principles of a democratic society is the rule of law, which is expressly referred to in the Preamble to the Convention (see the Golder judgment of 21 February 1975, Series A no. 18, pp. 16–17, para. 34). The rule of law implies, inter alia, that an interference by the executive authorities with an individual's rights should be subject to an effective control which should normally be assured by the judiciary, at least in the last resort, judicial control offering the best guarantees of independence, impartiality and a proper procedure.

It also underlined the importance thereof in the more recent case of *Breyer v Germany*, Application no. 50001/12, 30 January 2020, which likewise dealt with state surveillance, in §102.

¹⁷⁸ See in this regard Guobin Zhu, *Deference to the Administration in Judicial Review: Comparative Perspectives* (Springer International Publishing 2019).

¹⁷⁹ See *supra*, Section 3.3.5.b.

¹⁸⁰ The provision of such information can be seen as part of the cooperation that the different branches of the state must engage in to ensure checks and balances. See also Harry Woolf, Jeffrey Jowell and Andrew P Le Sueur, *De Smith, Woolf and Jowell's Principles of Judicial Review* (Sweet & Maxwell 1999).

¹⁸¹ Hasan Dindjer, 'What Makes an Administrative Decision Unreasonable?' (2021) 84 *The Modern Law Review* 265.

occurred in accordance with the law, with respect for human rights, and in a non-arbitrary and proportionate manner.

If administrative action is taken or informed by an algorithmic system, the above-mentioned loss of transparency (both as regards the process of the rule-making and the rule's application to the concrete situation) is thus problematic for the proper exercise of judicial review. A judge that does not know the parameters that led to a decision, and the reasons grounding the decision, will not be able to assess whether these parameters contain an unlawful ground of discrimination, or whether those reasons are arbitrary. In this regard, merely providing the judge with the source code of the algorithmic system may not be of much help.¹⁸² The overall majority of judges will not be able to interpret this code, and even if they could, this would still not offer them insight into how a concrete decision about the individual who challenges the act was adopted. The judge must hence be able to review the underlying logic of the system, including the parameters on which it is based.

Note that, in some cases, ensuring transparency about the system's operations might also imply the need to share information about how the general rules lying at the basis of the specific administrative act were translated from text to code, and whether this translation complies with the various rule of law principles. After all, this translation process plays an important role in the allocation (or potential narrowing down) of rights for natural and legal persons. These can also include specific rights that individuals derive from EU law, and for which they have a fundamental right to effective protection.¹⁸³ Public authorities must hence duly document this translation process and the various choices that were made in that context, since without such documentation, the principle of effective judicial review may be impeded.¹⁸⁴ Unfortunately, many public authorities currently do not proactively keep track of the normative choices that underlie the system's design and development (especially if the system's design was outsourced). The absence of such documentation practice is problematic not only from a (judicial) transparency perspective, but also from a security perspective. If there is no trace of *which* developer took *which* translation decision, the system is more vulnerable to (subsequent) traceless tweaking.

The problem lies, however, not only in the *ability* of judges to review the system, but also in their *willingness* to do so, which requires sufficient knowledge of the risks associated therewith and the openness to carry out a critical examination rather than

¹⁸² See in this regard also Ananny and Crawford (n 111).

¹⁸³ Recall in this regard the importance of ensuring sufficient transparency to enable individuals to invoke their right to an effective judicial remedy, as also stressed by the CJEU in Case C-817/19, *Ligue des droits humains v Conseil des ministres* (PNR Case) (n 114) §195.

¹⁸⁴ This is important in the context of all algorithmic systems, but may be even more problematic for systems that are meant to be 'dynamic' and 'agile'. Consider in this regard also the problematic example I mentioned *supra*, under Section 4.1.2, regarding the 'continuously updated' algorithmic system used for the UK's Universal Credit programme.

a mere deference to public authorities.¹⁸⁵ Consider in this regard the case brought by Edward Bridges, a civil liberties campaigner, against the South Wales Police Force (SWP) in the UK in order to challenge the police's use of automated facial recognition technology in public (through a pilot called 'AFR Locate').¹⁸⁶ Besides claiming that the use of this technology breached the right to privacy and data protection law, Mr Bridges also argued that it affected the right to equality and constituted an infringement of the 'Public Sector's Equality Duty',¹⁸⁷ since the authority "*failed to have regard to the possibility that use of the AFR software would produce a disproportionately higher rate of false positive matches for those who are women or from minority ethnic groups, such that use of AFR Locate would indirectly discriminate against those groups*".¹⁸⁸

The High Court's assessment of this claim is telling of the difficulties one can encounter when seeking judicial review of public authorities' use of algorithmic regulation, and of the burden of proof one may need to meet. It stated that: "*In our view, and on the facts of this case there is an air of unreality about the Claimant's contention. There is no suggestion that as at April 2017 when the AFR Locate trial commenced, SWP either recognised or ought to have recognised that the software it had licenced might operate in a way that was indirectly discriminatory*".¹⁸⁹ It continued by stating that

even now there is no firm evidence that the software does produce results that suggest indirect discrimination. Rather, the Claimant's case rests on what is said by Dr Anil Jain, an expert witness. In his first statement dated 30th September 2018, Dr Jain commented to the effect that the accuracy of AFR systems generally could depend on the dataset used to 'train' the system. He did not, however, make any specific comment about the dataset used by SWP or about the accuracy of the NeoFace Watch software that SWP has licensed. Dr Jain went no further than to say that if SWP did not know the contents of the dataset used to train its system 'it

¹⁸⁵ More generally, reference can also be made to the specific skills and virtues that judges should have in the context of adjudication, as well as their (ethical and legal) duties. See also, e.g., Iris van Domselaar, 'The Perceptive Judge' (2018) 9 *Jurisprudence* 71.

¹⁸⁶ See *R (Bridges) v CCSWP and SSHD*, 4 September 2019, High Court, Case no.: CO/4085/2018, [2019] EWHC 2341 (Admin). As the judgment states at §28, SWP deployed AFR Locate on about fifty occasions between May 2017 and April 2019 at a variety of large public events.

¹⁸⁷ This argument was based on section 149(1) of the UK's Equality Act 2010, which sets out the so-called 'Public Sector Equality Duty' and requires that

a public authority must, in the exercise of its functions, have due regard to the need to (a) eliminate discrimination, harassment, victimisation and any other conduct that is prohibited by or under this Act; (b) advance equality of opportunity between persons who share a relevant protected characteristic and persons who do not share it; (c) foster good relations between persons who share a relevant protected characteristic and persons who do not share it.

¹⁸⁸ *R (Bridges) v CCSWP and SSHD* (n 186) §21.

¹⁸⁹ *ibid* §153.

would be difficult for SWP to confirm whether the technology is in fact biased'. The opposite is, of course, also true.¹⁹⁰

Further in the judgment, the High Court also included the statement by Dr Jain: "I cannot comment on whether AFR Locate has a discriminatory impact as I do not have access to the data sets on which the system is trained and therefore cannot analyse the biases in those data sets. For the same reason, the defendant is not in a position to evaluate the discriminatory impact of AFR Locate."¹⁹¹

It is hard to find a more blatant example of how informational limitations can have an impact on the judicial review of public authorities' use of algorithmic systems. Furthermore, the High Court's stance aggravated the problem. It considered Mr Bridges' claim that the SWP insufficiently considered the risk that the system might suffer from bias and lead to indirect discrimination to have "*an air of unreality*" since he could not provide evidence of such discrimination, all the while acknowledging that he did not get access to the relevant datasets and that any evidence in this regard was hence unfeasible. Given this reasoning, it is hence no surprise that the High Court rejected Mr Bridges' arguments. The Court of Appeal, however, took the opposite stance. On the particular issue of the duty of equality, it stated: "*With respect to the Divisional Court, we do not consider that there is 'an air of unreality' about the Appellant's contention that there has been a breach of the PSED [Public Sector Equality Duty]. On the contrary, it seems to us to raise a serious issue of public concern, which ought to be considered properly by SWP.*"¹⁹² Moreover, it underlined the informational limitations associated with the use of the algorithmic system by noting that "*Dr Jain cannot comment on this particular software but that is because, for reasons of commercial confidentiality, the manufacturer is not prepared to divulge the details so that it could be tested*".¹⁹³ Furthermore, the Court of Appeal found that the current legal framework was *not* sufficient to constitute a legal basis for the use of such technology and suffered from fundamental deficiencies.¹⁹⁴ Especially when compared to the findings of the High Court, the Court of Appeal's stance is a substantial improvement. It has, nevertheless, been argued that the Court of Appeal's judgment still failed to grasp the full significance of the technology's capabilities, and that its evaluation of the legal arguments was hence inadequate.¹⁹⁵ In sum, the courts' engagement with the capabilities and limitations of algorithmic regulation in the context of judicial review (and its impact on individual and societal interests) cannot

¹⁹⁰ *ibid.*

¹⁹¹ *ibid.*, §155.

¹⁹² *R (Bridges) v CCSWP and SSHD*, 11 August 2020, Court of Appeal (Civil Division), Case no.: C1/2019/2670, [2020] EWCA Civ 1058, §173.

¹⁹³ *ibid.*, §199.

¹⁹⁴ *ibid.*, §91.

¹⁹⁵ See the discussion of this case in Karen Yeung, 'Constitutional Principles in a Networked Digital Society' [2022] SSRN Electronic Journal, 10 <www.ssrn.com/abstract=4049141>.

be taken for granted, and hinges on the judges' understanding of, and willingness to pay attention to, the specific risks arising therefrom.

Finally, recall that judicial review must also be available when public authorities delegate their tasks to a private entity.¹⁹⁶ Accordingly, whenever the development (or deployment) of an algorithmic system has been outsourced to a private company, public authorities should not be able to escape the need to provide information about the algorithmic system by reference to the private company's intellectual property rights. Instead, they must ensure that, where such rights exist, these do not hinder the judge's review of how the challenged administrative act came into being.

4.1.5.b Difficult Access to a Remedy

Before a legal dispute concerning an administrative act is heard in court, the individual subjected to the act must first have taken the step to challenge it and thus to request its judicial review. This means that she must already have knowledge of the fact that the act was potentially arbitrary or unlawful, or have sound arguments to make this case. If she lacks information on how the algorithmic system works, it will be difficult for her to make such arguments. It is here that the asymmetry of information between the individual subjected to the system and the system's developer or deployer can also lead to a stronger asymmetry of power, which in the context of judicial proceedings might also hamper access to justice and the equality of arms principle.¹⁹⁷ An individual may not always know that an algorithmic system lies at the basis of an administrative act affecting her. And even if she knows this fact,

¹⁹⁶ See Venice Commission, 'Rule of Law Checklist' (n 57), §45.

¹⁹⁷ See also Council of Europe Ad Hoc Committee on Artificial Intelligence – CAHA, 'Feasibility Study' (Council of Europe 2020) CAHA(2020)23, §23. In this regard, reference can also be made to the AERIUS rulings of the Dutch Council of State, concerning a software used by the relevant authority to calculate the deposition of nitrogen in the context of environmental impact assessments for agricultural projects (case references: Raad Van State, 17 May 2017, ECLI:NL:RVS:2017:1259; Raad Van State, 18 July 2018, ECLI:NL:RVS:2018:2454; and Hoge Raad, 17 August 2018; ECLI:HR:2018). That calculation formed the basis for a partially automated decision process to determine whether a project "*is likely to cause a significant deterioration or disturbance of a nitrogen-sensitive habitat located in a Natura 2000 site*", and hence whether a permit ought to be provided (see also Joined Cases C-293/17 and C-294/17, *Coöperatie Mobilisation for the Environment*, 7 November 2018, ECLI:EU:C:2018:882), in light of a preliminary reference procedure that arose in the case before the Council of State. When the outcomes of the software – which operated in an opaque manner – were challenged, the Dutch Council of State had the opportunity to stress the importance of the equality of arms principle whenever public authorities rely on algorithmic systems. Particularly, both at the lower and higher instance it was stated that whenever the decision of an administrative authority is the result of an automated process – even if only in part – and the affected party wishes to verify the correctness of the choices made during that automated process (including the verification of the data and assumptions used for this analysis) to potentially challenge it, the administrative authority must ensure that its choices, assumptions and data are *transparent* and *verifiable*.

she may still be in the dark as regards the system's functioning – and particularly how its suggestions or decisions come about. Yet as long as such information is not accessible, it will be more difficult for individuals to seek the judicial review of an administrative act that adversely impacts them and to obtain a remedy. For instance, in case a data-driven system is deployed, individuals would ideally need to have information about the potential patterns and categorisations that the system picked up, and how the system correlated their data with that of other persons. How else can they assess and – if need be – argue before a court that such correlations may be spurious or discriminatory?

I have already discussed several instances where public authorities refused to provide insight into the algorithmic systems they are using when informing or taking administrative acts. That such refusal can take extreme proportions is also illustrated by the STIR system (short for 'System Teleinformatyczny Izby Rozliczeniowej') used in Poland automatically to detect suspicious bank activities. The system, described by the Polish government as a 'warehouse of data', was adopted in 2017 and used by the Polish National Revenue Administration which was established that same year, when the government declared tax fraud to be a priority.¹⁹⁸ As explained by AlgorithmWatch, "*STIR can be accessed by analysts working in a special unit of KAS. Every day, reports from STIR land on their desks, which include information on transactions that were automatically labelled as suspicious as well as 'entities classified as high-risk groups using the financial sector for tax evasion'*".¹⁹⁹ Based on this information, it can then be decided to freeze the bank accounts of companies suspected of tax fraud – an administrative act that various companies already sought to challenge in court. These companies were, however, unable to challenge the outcomes of the algorithmic system as such, given a lack of information about its operations.

In 2017, the Polish government adopted a law that serves as the legal basis for this system, mandating the National Revenue Administration to establish relevant risk indicators. It, however, only did so in general terms.²⁰⁰ While the opacity about the algorithm and its risk indicators was raised early on by civil society organisations, the law goes as far as qualifying the provision of information about the algorithm as an offence. In fact, "*the law introducing STIR states that disclosing or using algorithms or risk indicators, without being entitled to do so, is an offense. A person who discloses algorithms can be imprisoned for up to five years (if the act was unintentional, then a fine can be given)*".²⁰¹ On the one hand, one can understand, in the context of fighting tax fraud, the reluctance of the government to provide overly detailed information about how the algorithmic system operates, lest fraudsters can use this

¹⁹⁸ AlgorithmWatch (n 62) 179.

¹⁹⁹ *ibid* 180.

²⁰⁰ *ibid* 181.

²⁰¹ *ibid* 185.

information to avoid being caught. On the other hand, however, the freezing of a company's bank account can have far-reaching consequences, and natural and legal persons should have sufficient information to be able to challenge the system in case doubts arise as to the legality of its outcomes. When sharing information about an algorithm's risk indicators is penalised with a prison sentence, one can question whether the balance between the government's legitimate aim to combat tax fraud and ensuring the possibility for judicial review is appropriately struck.

Evidently, the government should not be able to justify its decision to freeze a bank account merely based on the fact that an algorithmic system flagged one or more transactions as suspicious, without also explaining where the suspicion stems from. And the reviewing judge will still be able to ask the government to substantiate its decision – whether it was based on an algorithmic system or not. Yet such substantiation becomes more difficult if the system's flagging, for instance, hinges on the detection of 'unusual patterns' on a bank account, without an explanation as regards the causality between the pattern and the potential fraud. To which extent can a judge defer to the public authority's discretion to use algorithmic tools to identify a risk of fraud, even if the tool is opaque? What is the scope of the review that the judge must carry out? Does it suffice that the public authority – in this case, the Polish National Revenue Administration – provides a general list of potential risk indicators that may have been triggered? Or should the judge also be able to review the choices underlying the algorithmic decision-making tool? These are open questions that each judge might answer differently, yet they matter for the provision of an effective remedy for those subjected to the system and wishing to challenge not just its outcome as put forward by a public administration, but also its underlying assumptions, choices and inner workings.

An additional point to raise as regards the complication for natural and legal persons to obtain a remedy against an administrative act is that, in some jurisdictions, access to judicial review can be conditional upon the submission of a prior formal complaint with the public authority that took the challenged act. For instance, under Belgian law, an individual seeking the judicial review of an administrative act first needs to submit a formal complaint with that administration whenever a complaint mechanism is foreseen.²⁰² Note that the existence of a complaint procedure depends on the particular public authority, and that its modalities may differ from authority to authority.²⁰³ Yet the mediation of algorithmic regulation might render the filing of a formal complaint more difficult. As noted elsewhere,²⁰⁴ when an individual is adversely impacted by algorithmic regulation (for instance because of a miscategorisation, or because of the absence of a category that fits her case) the possibilities for re-interpretation, contestation and adaptation are close to

²⁰² See in this regard Steven Van Garsee (ed), *Handboek Bestuursrecht* (Politeia 2016).

²⁰³ *ibid.*

²⁰⁴ See also Smuha, 'The Human Condition in an Algorithmized World' (n 12) 36–37.

zero, as the system itself cannot be ‘reasoned’ with, and the relevant design choices have typically been made not by the public officials using the system, but by its coders.²⁰⁵ Merely adding a public official as a ‘human in the loop’ will therefore not be of much help if that official subsequently uncritically refers back to the system.

4.1.5.c Lack of Systemic Review

When an individual does manage to successfully challenge a problematic administrative act in court and submit it to judicial review, the damage will already be done, in some cases irreversibly so. Recall in this regard the example of the UK Post Office scandal I mentioned in Chapter 2, where people wrongly accused of theft due to reliance on a flawed algorithmic system spent years in prison or even died before a court was able to set the record straight.²⁰⁶ Or recall the example of the Dutch child care allowance scandal, where families were driven into poverty and went through emotional dramas before the problem was addressed.²⁰⁷ This is why, in the context of algorithmic regulation, *ex post* judicial review is a necessary but insufficient safeguard.²⁰⁸ Nevertheless, and setting the irreversibility of certain types of damage aside, judicial review does enable a judge to remedy the situation *ex post* by indemnifying those adversely affected and condemning the state if the rule of law’s principles were breached.

When the administrative act that is being challenged was informed or taken through an algorithmic system, there is, however, an additional element to consider: judicial review in principle only applies to the administrative act pertaining to the individual that brought the case to court, and not to other administrative acts taken by the same, potentially flawed, system. This is, however, problematic, as the system’s flaws may be systemic in nature, and risk causing systemic harm rather than mere individual harm.²⁰⁹ As long as the root of the problem (namely the faulty design, development or deployment of the algorithmic system) is not addressed, adverse effects will remain, whether to other individual interests or to societal interests more generally. As observed by Abe Chauhan, “*deciding on individual cases distances courts from the root of the systemic error in decisions made by the*

²⁰⁵ See also Burrell and Fourcade (n 42).

²⁰⁶ See *supra*, Section 2.2.2.

²⁰⁷ See *supra*, Sections 4.1.1 and 4.1.4.

²⁰⁸ See also Joe Tomlinson, Kate Sheridan and Adam Harkens, ‘Judicial Review Evidence in the Era of the Digital State’ (2020) 4 Public Law 740.

²⁰⁹ Karen Yeung, ‘Responsibility and AI: A Study of the Implications of Advanced Digital Technologies (Including AI Systems) for the Concept of Responsibility within a Human Rights Framework’ (Council of Europe, 2019) 63, <<https://rm.coe.int/a-study-of-the-implications-of-advanced-digital-technologies-including/168096bdab>>; Nathalie A Smuha, ‘Beyond the Individual: Governing AI’s Societal Harm’ (2021) 10 Internet Policy Review 3; Bart van der Sloot and Sascha van Schendel, ‘Procedural Law for the Data-Driven Society’ (2021) Information & Communications Technology Law 1.

relevant department or authority as to the design and implementation of such systems. Each of these issues is exacerbated by the evidential difficulties created by opacity and the effect of automation bias.”²¹⁰ Due to the jurisdictional limitations of the judicial review process, judges may hence not always be in the position to remedy the broader adverse impact of the system’s problematic use. While some courts have started to accept the review of upstream decisions made in relation to algorithmic regulation,²¹¹ such a remedy is not uniformly available, and the lack thereof renders the halting of public authorities’ (intentional or unintentional) systemic infringement of the rule of law’s principles more difficult. Once again, this emphasises the need for ex ante safeguards in the context of algorithmic regulation, but also for a reconsideration of mechanisms that would allow for more structural judicial remedies, so as to ensure that ex post review does not hinge on the shoulders of individual citizens. Systemic problems, after all, require systemic solutions.²¹²

Finally, it must also be stressed that the judiciary depends on the executive branch to uphold its judgments. This led Montesquieu to state that “*of the three powers above mentioned, the judiciary is next to nothing*”.²¹³ It is hence better to prevent rather than cure when it comes to keeping the executive power in check, especially if one recalls that in some EU Member States, authoritarian tendencies have already resulted in an erosion of the judiciary’s independence.²¹⁴

4.1.6 Separation of Powers

The last of the six rule of law principles concerns the separation of powers. In essence, this principle is aimed at avoiding a concentration of power by ensuring adequate checks and balances amongst the different branches.²¹⁵ In constitutional liberal democracies, each power should only exercise the functions that it is legally ascribed, with due regard for the protection of the rights and liberties of citizens, who should also be able to hold their government accountable.²¹⁶ In addition, the separation of powers can also be said to imply a separation of public power from private power, ensuring that public power is used in the public interest, rather than

²¹⁰ Chauhan (n 50) 5.

²¹¹ *ibid.*

²¹² See also Jennifer Cobbe, Michelle Seng Ah Lee and Jatinder Singh, ‘Reviewable Automated Decision-Making: A Framework for Accountable Algorithmic Systems’, *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (Association for Computing Machinery 2021).

²¹³ See Baron de Montesquieu, *The Spirit of the Laws* (1748) (Thomas Nugent tr, Hafner Publishing Company 1949). This extract was later also quoted by Alexander Hamilton in Federalist Paper 78 (1788).

²¹⁴ See *supra*, Section 3.2.5.

²¹⁵ See *supra*, Section 3.3.6.

²¹⁶ Andreas Schedler, ‘Conceptualizing Accountability’ in Andreas Schedler, Larry Diamond and Marc F Plattner (eds), *The Self-Restraining State: Power and Accountability in New Democracies* (Lynne Rienner Publishers 1999).

in the interest of private parties. If we now examine how this principle fares in light of the public authorities' increased reliance on algorithmic regulation, one can intuitively imagine that this may affect the power dynamics between the different branches of power and between the state and its citizens. Considering all that has already been discussed under the previous principles, let me focus on three points in particular.

4.1.6.a Strengthening the Executive

To begin with, one can note that the use of algorithmic regulation strengthens the power of the executive branch in several ways. First, it provides the executive with the possibility to adopt administrative acts at a faster speed and scale, thereby enabling it to exercise its decision-making power on many more natural and legal persons at the same time.²¹⁷ Second, it enables the executive to decide how text-based general laws are translated into code, which implies both an interpretation process and a reduction process, since the rich openness of text is necessarily reduced to a specific reading of the text, which is then turned into machine-readable code.²¹⁸ Third, it also ensures that executive policies can be executed in a centralised and systemic way, whereby discretion that can be used to deviate from the codified policy is eliminated. In addition, once the infrastructure for algorithmic regulation has been implemented, the executive will have the ability to singlehandedly rewrite the code, which is malleable and can be adapted instantaneously. Underlying the deployment of algorithmic systems is an entire technical infrastructure comprising hardware, software and databases, which will likewise be controlled by the executive branch of power, and which will cement the automation of decision-making for years to come.

Consider, in this regard, the example of the algorithmic system used by the US Immigration and Customs Enforcement (ICE) to help assess whether an illegal immigrant should be detained. Under the Obama administration, such detainment typically only occurred if immigrants were caught crossing the border illegally or when they had a serious criminal record. Otherwise, they were in principle released on bond.²¹⁹ However, when Trump became president, an executive order was issued to put an end to this practice, and to instead insist on the detention of immigrants regardless of any criminal record. As a consequence, the ICE proceeded to modify the algorithmic tool and removed the possibility for the system to recommend 'release'.²²⁰ From one day to the next, all public officials, by virtue of the change

²¹⁷ Restrepo Amariles (n 19).

²¹⁸ See *supra*, Section 4.1.1.

²¹⁹ Mica Rosenberg and Reade Levinson, 'Trump's Catch-and-Detain Policy Snares Many Who Call the U.S. Home' *Reuters* (20 June 2018) <www.reuters.com/investigates/special-report/usa-immigration-court/>.

²²⁰ *ibid.*

in the algorithmic system, were only able to get a negative recommendation out of the system. Certainly, Trump's executive order in itself already required them to act accordingly, yet the layer of automation further diminished their agency to act differently in cases where a release on bond might nevertheless have been warranted in view of particular circumstances. Accordingly, it must be kept in mind that the decision to automate (part of) certain administrative acts in one legislature will inevitably also create affordances for the next legislature, as the infrastructure that enables it remains.

Furthermore, the lack of transparency that often surrounds algorithmic regulation reduces the ability of the other branches of power to exercise checks and ensure that power is balanced. Checks can also be complicated by the lack of technical expertise within the other branches, who lack skilled personnel to understand and scrutinise these systems. Since the beginning of the *trias politicas* doctrine and even more so in recent years,²²¹ scholars have argued that an asymmetry of power between the branches exists, in favour of the executive.²²² The efficiency and scalability of executive decision-making through the use of algorithmic regulation can further exacerbate this power asymmetry. Moreover, the fact that in virtually all states, the adoption of algorithmic systems is primarily an affair of the executive branch of power also means that its power increase is not counterbalanced by the potential use of such systems by the other branches (assuming that such counterbalancing is at all feasible).²²³

It can hence be concluded that the executive's use of algorithmic regulation reinforces existing power structures rather than rebalancing power, and that it seems to skew the balance further in favour of the executive. This holds true regardless of whether the executive branch purposely implements algorithmic systems to consolidate power, or whether it implements it with the intention to better serve the public interest.

4.1.6.b Privatising Legal Infrastructure

To ensure that public authorities uphold the principles of the rule of law, it is also important that private entities do not exercise undue power over public matters. This can be referred to as the separation of public power from private power. Many public authorities, however, still lack technical know-how to design and develop algorithmic systems themselves.²²⁴ Accordingly, more often than not, the

²²¹ See *supra*, Section 3.3.6.

²²² See also Susan Rose-Ackerman, *Democracy and Executive Power: Policymaking Accountability in the US, the UK, Germany, and France* (Yale University Press 2021).

²²³ See Reijer Passchier, *Artificiële intelligentie en de rechtsstaat* (Boom 2021).

²²⁴ See Antonio Cordella and Leslie Willcocks, 'Outsourcing, Bureaucracy and Public Value: Reappraising the Notion of the "Contract State"' (2010) 27 *Government Information Quarterly* 82.

development of algorithmic systems used by public authorities is outsourced to private entities.²²⁵ Given the stakes I described above (the translation exercise from text to code, the various interpretation choices, and the normative consequences), one might ask whether such outsourcing, in practice, risks providing private actors with an undue ability to shape public policy.²²⁶ This pertains not only to the choices of interpretation, but also the choices of optimisation, model selection, data gathering, labelling and cleaning, and so on. To which extent does the outsourcing of these normatively relevant choices imply a privatisation of legal interpretation and application? And how can the translation process be verified, controlled and legitimised by the public officials who are actually in charge of the task, if their insufficient familiarity with algorithmic systems is what drove the outsourcing in the first place?

One can take this line of questioning a step further, and also inquire into the choices relating to the underlying infrastructure of the algorithmic system²²⁷ which is, almost by definition, likewise controlled by private entities. As noted above,²²⁸ infrastructural questions may sound boring, yet they matter a great deal. Together with the data and model of the system, the choice of infrastructure on which it is built likewise bears normative consequences, and can have an impact on individuals, groups and society at large.²²⁹ While in-house knowledge to develop algorithmic systems is rare but existing, there are almost no public authorities that are also the full controller of the underlying infrastructure on which these systems operate. This implies a vulnerability, since once the system is in place and relied upon by public authorities, they will not only become dependent on the system's adequate functioning, but also on the adequate functioning of the infrastructure that enables it, and that can be altered by private entities who are not subjected to democratic oversight. In sum, one must heed the risks associated not only with the outsourcing of legal interpretation to private entities, but also with dependencies on 'legal infrastructure' more generally.

4.1.6.c Citizen Surveillance

Finally, public authorities' reliance on algorithmic systems also has an impact on the role that citizens and civil society can play to ensure that the separation of powers is upheld. Civil society contributes to the functioning of checks and balances by seeking information about how its representatives act, and by holding its representatives to account if they do not respect the rule of law, for instance during democratic

²²⁵ See also *supra*, Section 3.3.6.

²²⁶ See also Verkuil (n 117).

²²⁷ See Diver (n 3) 20–21.

²²⁸ See *supra*, Section 2.2.1.

²²⁹ See also Julie E Cohen, 'Affording Fundamental Rights: A Provocation Inspired by Mireille Hildebrandt' (2017) 4 Critical Analysis of Law 76.

elections, or in court when seeking the judicial review of specific government actions.²³⁰ However, when public authorities deploy algorithmic regulation, public scrutiny by civil society, media and the public at large can become more difficult, for the same reason that scrutiny by the other branches of power can become more difficult.²³¹ Many citizens also lack the technical skills to understand how algorithmic systems function. Public authorities hence need to ensure that, if information about the system and its functioning is provided, such explanation is understandable for non-experts while remaining sufficiently meaningful.

In addition, increased reliance on algorithmic systems is also accompanied by increased data-gathering on citizens, to enable the system to profile them and take administrative acts relating to them.²³² Beyond the risk that such information is reflected incorrectly or in a biased manner, and beyond the increased risk of data-leaks and other vulnerabilities, it is also possible that such information, along with its decision-making infrastructure, is at some point deliberately used against them. Consider the concerns that arose when Poland proclaimed it would henceforth keep a centralised registry of healthcare data of citizens, including data about whether a woman is pregnant, with the stated aim of enabling a faster and more personalised delivery of health services based on such information.²³³ While the Polish government at the time sought to emphasise the beneficial goal behind such data collection,²³⁴ civil society organisations were concerned that, in a country where abortion is near banned, such information could also be used to monitor women's compliance with abortion laws, and potentially lead to the establishment of automated red flags when women are no longer pregnant prior to their due date.

Regulations can be altered, and laws can change. Under a new government or a reversal of precedent case law,²³⁵ actions that were once deemed a legal exercise of a fundamental right can become criminalised and vice versa. Yet through it all, data that was previously collected from citizens remains, as does the infrastructure that enables automated decision-making based on such data. The phenomenon of function creep that might accompany such infrastructure is well illustrated by an application of algorithmic regulation that is widely used in Belgium today, namely automated number-plate recognition (ANPR) cameras, which are essentially mass

²³⁰ See *supra*, Section 3.3.6.

²³¹ See also Fink (n 165); Fink and Finck (n 109).

²³² See, e.g., Peeters and Widlak (n 44); Kennedy (n 149); Heather Broomfield and Lisa Reutter, 'In Search of the Citizen in the Datafication of Public Administration' (2022) 9 *Big Data & Society* 1.

²³³ See "Tool of Repression": Anger over Poland's New "Pregnancy Register" *euronews* (6 June 2022) <www.euronews.com/2022/06/06/poland-s-government-criticised-over-pregnancy-register-amid-strict-abortion-laws>.

²³⁴ Karolina Kocemba, 'Pregnancy Registry in Poland' (*Verfassungsblog*, 22 June 2022) <<https://verfassungsblog.de/pregnancy-registry/>>.

²³⁵ Regarding the same subject matter, one can point to the US Supreme Court's overruling of *Roe v Wade* after fifty years, on 24 June 2022. See *Dobbs v Jackson Women's Health Organization*, 597 U.S. 2022.

surveillance tools.²³⁶ These cameras have been deployed on the Belgian roads since many years to read number plates of passing cars and crossing them with a database containing the number plates of wanted vehicles. The cameras were initially installed after the terror attacks that took place in 2016, with the mere purpose of catching terrorists and other criminals. The infrastructure, for which substantial public investments were made, not only raises significant privacy concerns, but has thus far also not been effective, primarily due to a large percentage of false-positive alerts (in some cases up to 80 per cent)²³⁷ and a lack of personnel to actually go after a car once it has been flagged.²³⁸ This did not stop the government from incrementally extending the offences for which the cameras could be deployed, from the identification of stolen vehicles, and vehicles of which the owner did not pay a traffic fine, to most recently the identification of vehicles of which the owner did not pay off debts with the Ministry of Finance, including income tax, corporation tax, VAT or overdue alimony.²³⁹ Another Belgian example concerns the installation of security cameras in the Jewish neighbourhood in Antwerp during the terrorist threat in 2015 and 2016 to protect the Jewish community. A few years later, during the Covid-19 pandemic in 2020, those same cameras were used to snoop on the community's compliance with the lockdown that was imposed, and especially with the ban on (religious) gatherings.²⁴⁰

Taking these examples one step further, one can thus imagine that the same infrastructure which facilitates scaled algorithmic decision-making in the so-called public interest can, under a worst case scenario, also be used by subsequent governments to oppress those very citizens, and specifically to target minorities, marginalised communities or political opponents. What happened in Afghanistan in the aftermath of the Taliban's return to power in August 2021 is telling in this regard. As reported by Human Rights Watch, before the Taliban's return

foreign governments such as the United States, and international institutions, including United Nations agencies and the World Bank, funded and in some cases built or helped to build vast systems to hold the biometric and other personal data of

²³⁶ Ian Warren and others, 'When the Profile Becomes the Population: Examining Privacy Governance and Road Traffic Surveillance in Canada and Australia' (2013) 25 *Current Issues in Criminal Justice* 565.

²³⁷ See in this regard the investigation conducted by Comité P, Belgium's Official Police Monitoring Committee: Comité permanent de controle des services de police, 'Le fonctionnement de la police intégrée en matière de traitement des hits ANPR de plaques d'immatriculation volées' (2022) 15.

²³⁸ Belga, 'Duur cameraschild langs snelwegen werkt amper: Tot 80 procent vals alarm' *De Morgen* (2 February 2022) <www.demorgen.be/gs-b84b968d>.

²³⁹ See Belga, 'Belastingzondaars kunnen voortaan van de weg geplukt worden' *De Morgen* (19 August 2022) <www.demorgen.be/gs-b56328c5>. See particularly 'Wet houdende diverse fiscale bepalingen' of 5 July 2022, which introduces these extended grounds, accessible at www.ejustice.just.fgov.be/eli/wet/2022/07/05/2022032714/justel.

²⁴⁰ See Matthias Verbergt, 'Camera's in joodse wijk controleren nu synagogegeangers' *De Standaard* (13 March 2021) <www.standaard.be/cnt/dmf20210312_98151173>.

various groups of Afghans for official purposes. In some cases, these systems were built for the former Afghan government. In others, they were designed for foreign governments and militaries.²⁴¹

It is believed that several of these systems are now used by the Taliban with the aim of targeting journalists and political opponents.²⁴² While this example does not concern the use of algorithmic regulation in a liberal democracy, the algorithmic systems enabling it were placed there by the public authorities of liberal democracies who believed they were acting in the public interest.

Despite the stronger legal protection mechanisms and higher political stability in the EU, it would be short-sighted to assume that infrastructures built in European countries would be immune from the same fate if, over the longer term, authoritarian tendencies further increase, especially in Member States where the rule of law is already under threat. These examples hence show how important it is to consider a *long-term* perspective when rolling out algorithmic regulation infrastructures with large databases, as the normative pillars underpinning liberal democracies are inherently fragile. The deployment of algorithmic systems should hence go hand in hand with an assessment of the longer-term risks for individual and societal interests, and with mechanisms to rebalance the increased asymmetry of power that the unchecked use of such systems imply.²⁴³

4.2 ALGORITHMIC RULE BY LAW

In the previous section, I conducted a systematic analysis of how public authorities' reliance on algorithmic regulation can adversely impact each of the six rule of law principles, drawing on concrete illustrations. Let me reiterate that, while I do not claim that these adverse effects *always* manifest themselves, my analysis shows they *can*, and that this risk should therefore be pre-empted and addressed. Many of the

²⁴¹ 'New Evidence that Biometric Data Systems Imperil Afghans' (*Human Rights Watch*, 30 March 2022) <www.hrw.org/news/2022/03/30/new-evidence-biometric-data-systems-imperil-afghans>. The article lists six systems that were built by private companies for or with foreign governments or international institutions: (1) Afghan National Biometric System, used to issue Afghan national identity cards, known as e-Tazkira; (2) US Defense Department Automated Biometric Identification System (ABIS), used to identify people whom the US believed might pose a security risk as well as those working for the US government; (3) Afghan Automated Biometric Identification System (AABIS), used to identify criminals and Afghan army and police members; (4) Ministry of Interior and Defense Afghan Personnel and Pay Systems (APPS) for the army and police, into which the AABIS was integrated in early 2021; (5) payroll system of the National Directorate of Security, the former state intelligence agency; and (6) payroll system of the Afghan Supreme Court.

²⁴² *ibid.* See also the statement by Aziz Rafiee, the executive director of the Afghan Civil Society Forum, who commented that "*the international community might have thought it was helping us, but instead it played with our fate and ended up creating systems more dangerous than they were helpful*".

²⁴³ I will further discuss these mechanisms *infra*, in Sections 4.3 and 6.2.2.

identified concerns are recurring across the six principles and are interlinked, since they stem from the combination of the risks inherent to algorithmic regulation on the one hand, and the role of the rule of law to tame public power on the other hand. In this section, I will therefore consolidate and summarise my findings, by proposing a theory of harm that conceptualises the adverse impact of algorithmic regulation on the rule of law. Conceptualising this harm can not only foster a better understanding of what is at stake, but it can also facilitate the evaluation of the legal framework's ability to counter it.

As announced in the Introduction, I propose to denote this theory of harm as *algorithmic rule by law*, to stress its deviation from the rule of law's ideal. Under the rule of law, public power is tamed by law, yet public authorities acknowledge the internal and external tensions that are inherent thereto, as well as the need to safeguard other EU values like respect for human rights and democracy. In contrast, under algorithmic rule by law, the law's power is channelled into a centralised algorithmic infrastructure that can be shaped and changed opaquely by a handful of people, and is prone to be wielded in a way that undermines the rule of law's very purpose – whether deliberately or not.

My analysis has brought to the surface at least five overarching problematic elements that characterise the threat of algorithmic rule by law. First, the illustrations indicated a prioritisation of algorithm-induced efficiency and procedural rationality over normative values like human rights and administrative justice (“*primacy of techno-rationality*”) (Section 5.2.1). Second, rather than trained public officials, the outcome of administrative actions is determined by the handful of people who design and develop the algorithmic systems, who thereby gain significant influence over public decision-making (“*supremacy of coders*”) (Section 5.2.2). Third, the analysis demonstrated how reliance on algorithmic regulation can reduce law's inherent openness and ambiguity to an overly formalised and narrow shape, leading to a legalistic approach instead, without the possibility to correct its hard edges where needed (“*automation of legalism*”) (Section 5.2.3). Moreover, the opacity accompanying the systems' design and implementation processes tends to diminish the possibility to exert oversight over the executive's operations, and to ensure that constitutional checks and balances are maintained (“*deficit of accountability*”) (Section 5.2.4). Finally, since algorithmic regulation rests on an underlying technical infrastructure, this introduces an important vulnerability in the legal system, which – as well as being instantly malleable – can also be deployed in a way that systemically undermines EU values (“*systemic vulnerability*”) (Section 5.2.5).

While each of these elements is worrying in and of itself, they are interrelated and reinforce each other. Collectively, they can therefore be seen as symptoms of the broader problem that lies at the heart of this book: the risk that algorithmic regulation, under the guise of implementing law, actually serves inadvertently or deliberately to undermine the law's protective power and foster a rule *by law* approach instead, hence meriting the term *algorithmic rule by law*. I deliberately opt for a

conceptualisation that focuses on the perversion of the *law*, rather than on the use of algorithms – therefore foregoing the use of terms like ‘rule by algorithms’ or ‘algorithmic rule’. The core problem revealed by the analysis above stems from the way in which those responsible for the design, development and deployment of algorithmic systems may – under the veneer of legality – undercut its value and open the door to illiberal and authoritarian practices. In what follows, I conceptualise this threat by setting out its five problematic features, and outline how they erode the law’s protective role.

4.2.1 Primacy of Techno-rationality

In Chapter 3, I described how the rule of law provides both procedural and substantive protection for citizens, by ensuring that public authorities duly consider their rights and interests and by empowering public officials to make appropriate trade-offs in between rules and discretion, thereby safeguarding individual justice. However, when algorithmic regulation is used, we can observe that the law’s implementation is portrayed as a techno-scientific endeavour rather than a normative one.²⁴⁴ Law is seen as an expression of rationality, and its application becomes a matter of mathematics rather than judgment and evaluation. Open-ended legal concepts such as ‘exceptional’ or ‘reasonable’, and even normative values like equality, are reflected into mathematical calculations and programmed into algorithmic systems.²⁴⁵ Yet, as noted above, these concepts are not always uniformly understood, and their interpretation and codification embodies certain normative choices.²⁴⁶ Notwithstanding this fact, pursuant to the logic of algorithms, the law’s application is handled by a problem-solving approach, driven by efficiency rather than justice. By identifying optimal models and codifying optimal computations for the law’s application, the ‘solution’ can be automated at scale, rendering individual judgment and assessment, and the time and resources that such assessments might require, redundant. In sum, the adoption of administrative acts, and of all the preparatory decisions to support this act, is reduced to a techno-scientific enterprise.

As a consequence, the law that is being algorithmically applied by public authorities is decoupled from the broader normative framework that it is part of, which in turn risks decoupling it from the overarching normative ends it should serve. Procedural rationality is hence favoured over substantive rationality, and normativity is being replaced by techno-rationality. This risk does not arise solely in algorithmic

²⁴⁴ See Diver (n 3); Marijn Janssen and George Kuk, ‘The Challenges and Limits of Big Data Algorithms in Technocratic Governance’ (2016) 33 *Government Information Quarterly* 371; Niklas Andreas Andersen, ‘The Technocratic Rationality of Governance – The Case of the Danish Employment Services’ [2020] *Critical Policy Studies* 1.

²⁴⁵ See Schartum (n 1) 315.

²⁴⁶ See *supra*, Sections 3.3.1 and 4.1.1.

context, but in bureaucratic organisation more generally.²⁴⁷ However, undeniably, algorithmic regulation can significantly exacerbate it.

Reliance on algorithmic regulation gives law, and the legal text that is being translated from law to code, an “*unwarranted aura of objectivity*”.²⁴⁸ While the notion of objectivity fits very well with the bureaucratic ideals of impersonality, rationality and efficiency, it is misplaced in the context of the law’s application, and can be at odds with the ideal of individual justice. As the above illustrations made clear, the law’s application is never truly ‘objective’, as open-ended legal concepts allow for a variety of legal interpretations.²⁴⁹ Yet by essentialising a given interpretation and acting as if it is an objective one, public authorities not only reduce the role of the law but also sweep their underlying normative choices under the rug, all the while maintaining the aura of legality.

Accordingly, the positive and normative become conflated. Algorithmic regulation might present the application of a legal rule as something that is a positive interpretation of legislation: the law as it is. However, in the legal context, a purely positive interpretation only rarely exists. There is always some element of normativity in the way one interprets legal rules. And such interpretation, explicitly or implicitly, always requires a trade-off between different values and interests. Even Weber already emphasised that no rational scientific procedure exists to tell us which trade-offs to make between competing values.²⁵⁰ Such trade-offs are always an inherently normative choice, and no layer of algorithmic modelling can change that, although it can obscure it.

Furthermore, a techno-scientific approach to law is inherently reductionist, since the richness of language, and the reality it represents, can never be wholly captured by mathematical models (whether knowledge- or data-driven). Yet public authorities’ techno-optimism, coupled with the pressure of achieving efficiency gains and making budget cuts, might make them overlook this fact, even if it has important consequences for the individuals subjected to the system. Individuals who fall outside the model, for instance because they do not fit into any of the programmed categories of a knowledge-driven system, or because their situation was not picked up as a distinct pattern by a data-driven system, may be treated as an outlier, both statistically and legally. Crucially, in the context of public authorities, “*being*

²⁴⁷ I have already discussed how scholars like Arendt and Bauman cautioned against it. See *supra*, Section 2.3.2.

²⁴⁸ Hildebrandt, ‘Algorithmic Regulation and the Rule of Law’ (n 121) 3.

²⁴⁹ Interpretations of legal rules are inherently contestable, and can hence be seen as a subjective exercise. As also noted by Julia Black, “*there is no inherent, fixed meaning to rules or to language; the meaning, and hence the application, of a rule is not an objective fact but is contingent on the interpretive community reading the rule*”, in *Rules and Regulators* (n 6).

²⁵⁰ See *supra*, Section 2.3.2. See also Michael W Spicer, ‘Public Administration in a Disenchanted World: Reflections on Max Weber’s Value Pluralism and His Views on Politics and Bureaucracy’ (2015) 47 *Administration & Society* 24, 30.

excluded from the system also means being excluded from public services”²⁵¹ with highly problematic consequences, hence requiring the anticipation of this risk and the availability of alternative access routes to such services. Equally problematically, besides exclusionary classifications, individuals may also have been classified erroneously or based on discriminatory grounds. In addition, especially with data-driven systems, instead of being based on a *causal* relationship between the facts and the law, administrative acts can be taken based on how certain facts about an individual *correlate* with other facts that are not linked to the law at all. As noted by Langford, “an individual’s rights may be determined on the basis of predictions derived from the behavior of a general population group”,²⁵² thereby undermining the notion of individual justice.²⁵³

More generally, legal subjects are literally and figuratively dehumanised by being reduced to faceless datapoints subjected to mathematical rules.²⁵⁴ This creates further distance between the public officials responsible for the adoption of an administrative act and the person subjected thereto, which in turn can diminish the sense of responsibility and empathy²⁵⁵ that can help counter the excesses of procedural rationality. As discussed above, this distance (which is present in any bureaucratic form of organisation, but is significantly extended when relying on algorithmic regulation) undermines the ‘internal morality’ of public authorities. It may even deliberately be exploited to apply the law in an overly rigid manner, with the adverse consequences being felt especially by those already in a vulnerable situation.²⁵⁶ Simultaneously, the fact that algorithmic regulation renders the contestability of administrative acts more difficult also makes it challenging to correct potential wrongs in the system.²⁵⁷

Finally, the logic of efficiency that underlies algorithmic regulation is likely to favour the optimisation of the operation of public authorities rather than the optimisation of the rights of individuals. As noted by Scharf, “in mass administrative systems, choices of interpretations may easily be affected by expected effects on government budgets – for instance, by pushing interpretation of concepts to extremes

²⁵¹ Peeters and Widlak (n 44) 176.

²⁵² Langford (n 7) 142.

²⁵³ As defined by Reuben Binns, individual justice “refers to the notion that each case needs to be assessed on its own merits, without comparison to, or generalization from, previous cases”. See Reuben Binns, ‘Human Judgment in Algorithmic Loops: Individual Justice and Automated Decision-Making’ (2022) 16 Regulation & Governance 197, 198.

²⁵⁴ See also Li Huang, Zhi Lu and Priyali Rajagopal, ‘Numbers, Not Lives: AI Dehumanization Undermines COVID-19 Preventive Intentions’ (2022) 7 Journal of the Association for Consumer Research 63; Broomfield and Reutter (n 232).

²⁵⁵ Sofia Ranchordas, ‘Empathy in the Digital Administrative State’ (2022) 71 Duke Law Journal 1341.

²⁵⁶ Ranchordas and Scarcella (n 150).

²⁵⁷ Karen Yeung, ‘Why Worry about Decision-Making by Machine?’ in Karen Yeung and Martin Lodge (eds), *Algorithmic Regulation* (Oxford University Press 2019).

to make possible reuse of data”.²⁵⁸ Recall in this regard also the pressure on public officials to meet KPIs, at the cost of ensuring individualised justice for persons subjected to administrative acts. The logic of efficiency and the logic of respect for individual rights and human dignity are therefore not necessarily aligned. Unfortunately, in the many illustrations discussed above, we must agree with Galligan that “*in the very nature of bureaucratic administration*”, and a fortiori in the nature of algorithmic regulation, “*the logic of efficiency is more powerful than that of rights*”.²⁵⁹

In sum, by relying on algorithmic regulation, the application of the law is reduced to a quantitative endeavour rather than a qualitative one. It is turned into a mathematical instrument and stripped away from its substantive notions, all in the name of efficiency, objectivity and consistency. Yet this can undermine the law’s protective role, creating a semblance of legality without leading to justice. When the implementation of algorithmic regulation is framed as a mere positive interpretation and application of the law rather than a normatively relevant translation exercise, in the best case, public authorities risk remaining blind for these adverse consequences and, in the worst case, they deliberately use this blindness to push through problematic interpretations. Accordingly, when opting for reliance on algorithmic regulation, it is crucial that its normative role be acknowledged, and that appropriate mechanisms exist to curb the primacy of techno-rationality over justice.²⁶⁰

4.2.2 Supremacy of Coders

Under the EU conceptualisation of the rule of law, the principle of legality foresees that the law is adopted through a pluralistic democratic process based on political deliberation and civic participation. Subsequently, it should be applied by public authorities in a manner congruent thereto and in line with the authoritative interpretation of the law by independent courts. Especially in highly regulated societies, this typically implies that a wide range of competences are delegated to public authorities,²⁶¹ including discretionary powers to decide the optimal course of action to attain broadly formulated policy goals.²⁶² Yet the use of these delegated powers must occur in line with the rule of law’s principles and, at least in theory, public officials are trained to do so. They are in principle hired based on their skills

²⁵⁸ Schartum (n 1) 308.

²⁵⁹ Denis J Galligan, ‘Public Administration and the Tendency to Authoritarianism’ in András Sajó (ed), *Out of and into Authoritarian Law* (Brill–Nijhoff 2002) 193.

²⁶⁰ I will touch upon those mechanisms *infra*, in Sections 4.3 and 6.2.

²⁶¹ See *supra*, Section 2.3.3.

²⁶² Denis James Galligan, ‘Senses of Discretion’ in his *Discretionary Powers: A Legal Study of Official Discretion* (Oxford University Press 1990); Tony Evans, ‘Professionals and Discretion in Street-Level Bureaucracy’ in Peter Hupe, Michael Hill and Aurélien Buffat (eds), *Understanding Street-Level Bureaucracy* (Bristol University Press 2015).

and expert knowledge, and their ability to implement legislation and apply it to concrete cases based on their reasoned judgment and experience.

While there is no need to idealise the output produced by *all* public officials,²⁶³ the fact that they are trained to carry out their tasks, given the significant impact of their actions on individual and societal interests, is important to stress.²⁶⁴ This is particularly relevant given the power that public officials, as part of an organisation vested with public authority, can wield. Public officials are therefore typically also bound by administrative rules and specific deontological procedures relating to their professional and moral behaviour, to safeguard that their duties are carried out in the public interest.²⁶⁵ Echoing the influential work of Jerry Mashaw, these rules and procedures are aimed at enabling ‘bureaucratic justice’, which includes not only bureaucratic rationality, but also the professional treatment of administrative cases and the exercise of moral judgment – thereby institutionalising normative values within administrations.²⁶⁶ Recall in this regard also the discussion about the ‘internal morality’ within public authorities, which aims to ensure that procedural rationality does not overtake substantive rationality to the detriment of the rights and interests of individuals and society.²⁶⁷ Public officials also act under the political responsibility of members of the executive, who exercise political oversight over their actions and enable democratic accountability, in addition to having their actions subjected to legal review by courts.

However, when public authorities rely on algorithmic regulation, they essentially re-delegate their decision-making power to what I have referred to above as ‘coders’: people who may have the technical skills to design and develop algorithmic systems, but who are not necessarily trained in public decision-making, nor in the responsibilities and delicate trade-offs this implies. These coders suddenly become the intermediary actors between public authorities and citizens.²⁶⁸ While this shift of

²⁶³ As discussed *supra*, in Section 1.2.3, my focus on the risks of algorithmic regulation does not imply a blind faith in human beings, or the misguided belief that decision-making by human beings is always free from prejudice or error.

²⁶⁴ See also Evans (n 262).

²⁶⁵ See *supra*, Section 2.3.

²⁶⁶ See Jerry L Mashaw, *Bureaucratic Justice: Managing Social Security Disability Claims* (Yale University Press 1983). Mashaw also underlined the importance of internal governance and control mechanisms, and he considered judicial review – as an external mechanism of control – to be irrelevant or impertinent in most of the cases, given the remoteness of judges from bureaucratic realities. See in this regard Robert A Kagan, ‘Varieties of Bureaucratic Justice: Building on Mashaw’s Typology’ in Nicholas R Parrillo (ed), *Administrative Law from the Inside Out* (1st edn, Cambridge University Press 2017) 248. Indeed, he found that “*the task of improving the quality of administrative justice is one that must be carried forward primarily by administrators*”. See also Morgan and Yeung (n 6) 245.

²⁶⁷ See Galligan, ‘Public Administration and the Tendency to Authoritarianism’ (n 259).

²⁶⁸ See also David Freeman Engstrom and others, ‘Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies’ (Administrative Conference of the United States 2020) <www.acus.gov/sites/default/files/documents/Government%20by%20Algorithm.pdf>.

power away from public officials has been denoted by some as *algocracy* or rule of algorithms,²⁶⁹ I am wary of such conceptualisation, since it wrongly suggests that power is wielded by algorithmic systems.²⁷⁰ In truth, power (implying here all the normative and political choices that relate to the implementation and application of the law and the adoption of administrative acts) lies in the hands of those who develop and design these algorithmic systems, or the coders. Speaking of the supremacy of coders would hence be more accurate, since the affordances of the technology are entirely shaped by the decisions underlying the algorithms' design, and hence by their coders.

As previously stressed, the transformation of text-based law to code implies a myriad of morally and politically relevant choices. Algorithmic systems "*only follow unambiguous rules, and there is no room for doubt or discretion*", even if "*it will almost always be possible to claim that other results are correct and legally valid, and thus there may be grounds to disagree that the interpretations embedded in the code should be held as correct*".²⁷¹ In the context of algorithmic regulation, discretion about the law's interpretation is centralised and moves upstream, away from public officials, to the handful of coders who translate, interpret and operationalise legal rules through the algorithms they design.²⁷²

Moreover, this translation process typically occurs in a frictionless manner, as the normative choices underlying it remain invisible, not only for the citizens subjected to the system, but also to the public officials that rely thereon for the purpose of taking administrative acts.²⁷³ While this invisibility might give a semblance of impersonality and objectivity, it reduces the possibility for contestability.²⁷⁴ Recall in this regard the claim by Peeters and Widlak that the 'digital cage' of public administration can hence not only extend to citizens, but also to public officials who see their discretionary actions constrained by the technology's architecture,²⁷⁵ which is determined by coders. As noted by Yeung,²⁷⁶ this absence of friction also

²⁶⁹ See *supra*, Section 2.3.4.

²⁷⁰ Similarly, regulation through algorithms has also been conceptualised through notions like *technological management*, the *rule of algorithms* or *algocracy*. See respectively Roger Brownsword, 'Technological Management and the Rule of Law' (2016) 8 Law, Innovation and Technology 100; Michael Meyer-Resende and Marlene Straub, 'The Rule of Law versus the Rule of the Algorithm' [2022] (Verfassungsblog, 28 March 2022) <<https://verfassungsblog.de/rule-of-the-algorithm/>>; John Danaher, 'The Threat of Algocracy: Reality, Resistance and Accommodation' (2016) 29 Philosophy & Technology 245; Lorenz, Meijer and Schuppan (n 124).

²⁷¹ Schartum (n 1) 306.

²⁷² See Bovens and Zouridis (n 115) 181. See also Burrell and Fourcade (n 42) 217.

²⁷³ Recall in this regard, for instance, the example of the Polish unemployment algorithm, where the public officials using the system stated that they did not know the underlying logic of the system. See *supra*, Sections 4.1.2 and 4.1.3.

²⁷⁴ Hildebrandt, 'Algorithmic Regulation and the Rule of Law' (n 121) 3.

²⁷⁵ See Peeters and Widlak (n 44) 181.

²⁷⁶ Yeung, 'Can We Employ Design-Based Regulation While Avoiding Brave New World?' (n 119) 16.

undermines public officials to exercise their agency and use their judgment, pursuant to their duty of acting in the public interest and in line with their deontological codes, for the seamlessness of the technology's design, in the name of user friendliness, might obliterate the possibility to do so.

At the same time, the outsourcing of discretion to coders occurs under a coat of legality, since, formally speaking, public officials are the ones who remain accountable for the decisions they make, even if they are no longer able to exercise much judgment in this regard. It is for this reason that the power wielded by coders can be seen as part of the larger threat posed by algorithmic rule by law. The fact that the translation of legislation to algorithms is considered as a mere techno-scientific enterprise rather than a normative task minimises the moral and political consequences attached to this process. In practice, however, the delegation of the law's implementation to coders also implies the delegation of public authority and moral responsibility.²⁷⁷ Yet such delegation occurs without guarantees of adequate training, legal expertise, subjection to deontological codes, or even awareness of such responsibility and, as we shall see further below, without adequate accountability mechanisms for the power that comes with it.²⁷⁸

In sum, reliance on algorithmic regulation by public authorities entails a shift in power, whereby the interpretation of legal rules is delegated to coders rather than public officials hired based on their domain expertise and constrained to safeguard the public interest. It therefore also opens the door for these coders – or rather, for those who pay the salary of these coders – to opt for design choices that are problematic from a democracy and human rights-perspective, under the guise that it concerns a purely technical matter. Under a best case scenario, those problematic choices concern errors that can be rectified and hopefully do not cause irreversible damage (though we have seen that the scaled nature of the systems can make the adverse consequences extensive). Under a worst case scenario, those problematic choices deliberately use the veneer of the law, albeit in algorithmic shape, to implement illiberal and authoritarian practices at scale. Consequently, if public authorities wish to rely on algorithmic regulation, it is crucial to ensure that the upstream decisions of coders, regardless of whether they work for a private company to which the design of the system is outsourced or whether they work in-house, be subjected to review and oversight, both internally (to maintain public agency) and externally (to preserve public accountability).

4.2.3 Automation of Legalism

A third way in which algorithmic regulation can undermine the rule of law, concerns the way in which it disrupts the balance between rules and discretion.

²⁷⁷ See also *supra*, Section 2.2.6. Moreover, see Yeung, 'Can We Employ Design-Based Regulation While Avoiding Brave New World?' (n 119); Yeung, 'Responsibility and AI' (n 209).

²⁷⁸ See *infra*, Section 4.2.4.

This balance is indispensable for the law to carry out its protective function, as an overly rigid application of rules without discretion to ensure their contextualisation can lead to unjust outcomes.²⁷⁹ The exercise of discretion, or the autonomous application of reasonable judgment,²⁸⁰ should be aligned with the rule of law's principles and exercised based on an examination and assessment of the facts at hand.²⁸¹ Importantly, in undertaking this assessment, public officials rely not only on their specific expertise, but also on their *implicit* knowledge of society and human beings more generally.²⁸² Furthermore, this discretion can also serve as a correcting factor (and 'little goodness', as borrowed from Emmanuel Levinas²⁸³) in a situation where the provision of public services has been institutionalised and systematised, and might fail to deliver individual justice.²⁸⁴ As also stressed by Binns, the need for individual justice or "*the notion that each case needs to be assessed on its own merits, without comparison to, or generalization from, previous cases*", requires a certain level of discretion to enable individualised assessments.²⁸⁵ This is particularly relevant for the adoption of (individual) administrative acts, where public officials are required to apply general rules to individual cases and also when they rely on algorithmic systems to do so.²⁸⁶

However, the above analysis demonstrated that reliance on algorithmic regulation, which requires unambiguous and precise rules, can foster an overly strict interpretation of the law, which shifts the pendulum entirely away from 'discretion' all the way towards 'rules' instead of promoting their marriage. This does not result in the algorithmic system's conformity with *legality*, but in an automated form of *legalism*, with several problematic consequences. As defined above, legalism is characterised by a strict adherence to the law, based on the law's letter rather than its spirit.²⁸⁷ While a rigid application of the law, regardless of its substantive ends or its concrete effects, can also be opted for without algorithmic systems, reliance on these systems *induces* a legalistic approach, in light of the reductive translation

²⁷⁹ See *supra*, Section 2.3.3.

²⁸⁰ See Galligan, 'Senses of Discretion' (n 262) 8.

²⁸¹ See in this regard also Lars Tummers and Victor Bekkers, 'Policy Implementation, Street-Level Bureaucracy, and the Importance of Discretion' (2014) 16 Public Management Review 527.

²⁸² See Diver (n 3) 8.

²⁸³ See Emmanuel Levinas, *Is It Righteous to Be?: Interviews with Emmanuel Levinas* (Jill Robbins ed, Stanford University Press 2001) 116.

²⁸⁴ See *supra*, Section 2.3.3.

²⁸⁵ Binns (n 253) 198.

²⁸⁶ Binns points out both epistemic and normative limitations of algorithmic decision-making systems when it comes to ensuring individual justice, without denying that the need for individual justice may at times conflict with other values, such as, for instance, consistency. See *ibid* 201 and following. Note that this potential for conflict is simply a reflection of the tensions that are inherent in a society organised on the basis of the rule of law, as discussed *supra*, in Section 3.4.

²⁸⁷ See Shklar (n 22). See also Bankowski and Schafer (n 24). Note that HLA Hart also recalls in this regard the notion of 'formalism', see Hart (n 5) 129.

exercise it requires from open-ended legal concepts to codifiable rules. This condenses the law's pluralistic conceptualisation into a monistic straightjacket, which will be codified and essentialised. The only 'discretion' exercised in this context are the choices made by the coders when they take upstream decisions about the system's design and the law's translation. In doing so, they need to anticipate all situations to which the law may be applied, and the effects that their translation will have downstream, bearing in mind that all information that the system relies on must be rendered *explicit*. Yet, as noted above, reliance on algorithmic regulation often *disguises* the fact that interpretative choices are made, since all of these choices occur upstream and prior to the system's use.

Furthermore, it can also disguise the potential 'creative compliance' of the law by those developing the system, which McBarnet and Whelan conceptualise as a manipulation of the law *"to turn it – no matter what the intentions of legislators or enforcers – to the service of their own interests and to avoid unwanted control"*.²⁸⁸ Indeed, *"creative compliance thrives on a narrow legalistic approach to rules and legal control, on a formalistic conception of the law"*,²⁸⁹ which is precisely the risk identified with algorithmic regulation. While creative compliance can certainly also occur without reliance on algorithmic systems, their opaque and automated nature can both camouflage and facilitate this practice, on a very wide scale.

The above illustrations also demonstrated that discretion at the 'street-level' is significantly reduced, and one can hence no longer speak of discretion as the exercise of autonomous judgment based on a particular situation. Instead, judgment is replaced by mathematical rules and functions. Some might argue that discretion can be codified into the system, for instance by anticipating and programming different variations of a legal rule based on different criteria. Yet this can hardly be referred to as the act of 'judging', but should rather be seen as a *"replacement of discretion with a series of fixed cumulative criteria; that is, criteria that could be solved by collecting relevant machine-readable data"*.²⁹⁰ As explained by Scharf, modelling open-ended concepts – for instance, 'suitable employment', in the context of the evaluation of unemployment benefits – *"would require access to an unrealistically large number of types of data"*.²⁹¹

More generally, since reality and the human condition are not characterised only by a limited number of features, it would be impossible to anticipate all possible

²⁸⁸ Doreen McBarnet and Christopher Whelan, 'The Elusive Spirit of the Law: Formalism and the Struggle for Legal Control' (1991) 54 *The Modern Law Review* 848.

²⁸⁹ *ibid* 848. See in this regard also Karen Yeung's discussion of creative compliance, *"whereby technical compliance with rules may be achieved yet the underlying spirit and purpose of those rules might be simultaneously undermined"*, in *Securing Compliance* (Hart Publishing 2004) 11. Yeung therefore draws a distinction between 'rule compliance', on the one hand, and 'substantive compliance', on the other hand, whereby the latter pays attention to the collective goals *behind* the rule rather than merely to the literal formulation of the rule itself.

²⁹⁰ Scharf (n 1) 316.

²⁹¹ *ibid*.

situations in which the law may be applied in advance. As remarked by Hart, if “*everything would be known*” and if “*for everything, since it would be known, something could be done and specified in advance*”, this would “*be a world fit for ‘mechanical’ jurisprudence*”.²⁹² However, “*plainly this world is not our world; human legislations can have no such knowledge of all the possible combinations of circumstances which the future may bring*”.²⁹³ Accordingly, both from a practical and technical perspective, it may be difficult or unfeasible to automate the replacement of ‘discretion’, especially in a way that avoids the risk of “*creating an ‘echo chamber’ where old points of views become decisive even in new cases with new contexts*”.²⁹⁴ A legalistic approach thus overlooks or ignores the infinite variability of social contexts and interpretations, thereby hiding, but not undoing, the clash between the indeterminacy of rules and their algorithmic application.

Worryingly, algorithmic regulation not only reduces discretion at the street level, but it also creates a technological architecture that renders deviations from the law (or from its codified interpretation) technically unfeasible.²⁹⁵ Accordingly, this eliminates the possibility for public officials to remedy potential adverse consequences of the law’s rigid application, and restricts them to apply the rules in accordance with how they were programmed.²⁹⁶ While algorithmic regulation can hence stand in the way of the ‘softening’ of the law’s hard edges, it also prevents public officials to make corrections in case of an unjust situation. Put simply, the ‘little goodness’ that can correct the rough edges of an institutionalised legal system no longer has a place, and public officials can also no longer take up their role in ‘speaking truth to power’.²⁹⁷ It is moreover important to stress that reliance on algorithmic regulation instead does not prevent public officials to take arbitrary or unlawful decisions. It merely prevents them from taking decisions that deviate from whichever rules have been codified in the algorithmic system, without guaranteeing that those rules and outcomes in and of themselves are not arbitrary or unlawful.

Yet the problem goes further still. The prolonged attrition of public officials’ autonomy – and the concomitant absence of the possibility to exercise discretion – might numb their ability to make a critical evaluation of the law’s application in concrete cases, until their motivation and skill for reasoned judgment becomes superfluous.²⁹⁸ Without the space to practise human agency, the question of

²⁹² Hart (n 5) 128.

²⁹³ *ibid.*

²⁹⁴ Schartum (n 1) 316.

²⁹⁵ See in this regard Yeung, ‘Can We Employ Design-Based Regulation While Avoiding Brave New World?’ (n 119); Brownsword (n 270); Peeters and Widlak (n 44).

²⁹⁶ Recall that adherence to the outcomes of an algorithmic system does not necessarily equal adherence to the law as such, even if that system is meant to have codified legal provisions.

²⁹⁷ See in this regard also Thomas Elston and Gwyn Bevan, ‘New Development: Scarcity, Policy Gambles, and “One-Shot Bias” – Training Civil Servants to Speak Truth to Power’ (2020) 40 Public Money & Management 615.

²⁹⁸ See *supra*, Section 2.3.2.

whether a certain legal interpretation or application leads to an unjust situation might not even pose itself.²⁹⁹ As argued above, this, in turn, might lead to a problematic discharge of moral involvement and responsibility, for which agency is a precondition.³⁰⁰ One might argue that this problem can be solved by ensuring that algorithmic regulation is relied upon only informatively rather than decisively. However, the above illustrations have shown that, even in those cases, in practice the space for agency is marginal, due to a high case load, the pressure to meet KPIs, the limited understanding of the system's operations, and more generally the impossibility of verifying the validity of recommendations pertaining to thousands or even millions of citizens. Accordingly, algorithmic regulation and the legalistic approach it induces might lead to the mindless execution of rules,³⁰¹ thereby reinforcing the hierarchical obedience to authority that already permeates bureaucratic organisation, and ultimately also to a *banalisation*³⁰² of its potentially adverse consequences.

It would be a mistake to ignore or make invisible the normative tensions that are inherent to the law. Yet it would be as problematic to make them invisible by disguising them as a techno-rational optimisation exercise, or to eliminate them altogether by opting for the codification of *one* interpretation over and above others, without an avenue for the reasoned judgment and potential contestability of such interpretation when inappropriate for the particular situation.³⁰³ In the best case, the automation of legalism is an unintentional by-product of algorithmic regulation,

²⁹⁹ See Yeung, 'Can We Employ Design-Based Regulation While Avoiding Brave New World?' (n 119). See also the emphasis on public agency in Endicott and Yeung (n 128).

³⁰⁰ See also *supra*, Section 2.2.5.

³⁰¹ Diver (n 3) 13.

³⁰² This refers to Hannah Arendt's conceptualisation of the 'banality of evil', arguing that evil need not stem from a deliberately destructive intention, but that it can also arise from the execution of tasks without critical judgment or reflection about their morality, which can lead to a dangerous discharge of responsibility and thereby enable evil actions, particularly in a bureaucratic setting. See Hannah Arendt, *Eichmann in Jerusalem: A Report on the Banality of Evil* (Viking Press 1963). Recall in this regard also the experiment carried out by Milgram, discussed *supra*, in Section 2.2.6.

³⁰³ While my analysis focuses on automation in the context of administrative actions, it should be noted that similar problems also arise in other contexts. Consider, for instance, the mass automation of content filtering on social media channels, where broad legal concepts such as 'hate speech' and 'illegal content' are codified into algorithmic systems in order to detect and take down problematic messages. It goes beyond the purpose of this book to analyse the use of such algorithmic systems in detail and compare their challenges with those used in the public sector, yet I nevertheless wish to point out that certain parallels can be drawn. These include, for instance, the need to cater for scaled and speedy decision-making on the one hand (the legal rules which provide protection against harmful content need to be upheld, despite the vast scale and speed at which data can be shared online), and the need to ensure individual justice on the other hand (the rights of individuals who share content and exercise their freedom of expression also need to be upheld, and any 'filtering' needs to occur in a transparent and contestable manner). These parallels might give rise to certain lessons or best practices that can be drawn from one area to the other – a subject that merits further research. For a discussion on some of the challenges of automated content filtering, see, e.g., Emma J Llansó, 'No Amount of "AI" in Content Moderation Will Solve Filtering's Prior-Restraint

and one that public authorities seek to remedy by safeguarding the discretion and autonomy of public officials. Yet in the worst case, the elimination of discretion, and the subsequent erosion of responsibility, can be used to prevent internal criticism, and to prevent the deviation from a problematic (or problematically codified) rule, despite its adverse impact. This approach might reinforce and automatise, at scale, illiberal and authoritarian interpretations of the law, while stifling the opportunity for critical reflection and remedial action. Therefore, if public authorities wish to rely on algorithmic regulation, they need to ensure that the realisation of the rule of law, which hinges on the sustainment of the tensions inherent thereto, rather than their dissolution, maintains the law's openness.

4.2.4 Deficit of Accountability

Ensuring public accountability is a central function of the rule of law. The legal system ensures that public authorities can be held to account whenever their actions infringe the principles of the rule of law, from the principle of equality to the prohibition on the arbitrary use of executive power, and secures the possibility of judicial review to challenge and remedy government actions whenever such infringement occurs.³⁰⁴ It hence requires that public authorities comply with the law, and that they be held to account when they do not. However, as my analysis has shown, this function of the law can become more difficult to uphold in the context of algorithmic regulation.

First of all, the decisive aspect of public authorities' decisional action has shifted from 'street-level' to 'system-level',³⁰⁵ as the choices that determine administrative acts have in fact been outsourced to coders. Yet this upstream move of discretion is not followed by an upstream move of accountability, given that these choices remain largely invisible. Indeed, as noted by Peeters and Widlak, the information architecture that enables algorithmic regulation "*is a less 'visible' form of rationalisation*".³⁰⁶ This invisibility or opacity is not necessarily limited to the normative choices underlying the system's design and the law's translation, but often also encompass the system's mode of operation (especially in the case of complex data-driven systems) and at times even its existence.³⁰⁷

Such reduced transparency, along with the 'rational' framing, renders it much more difficult to contest certain normative decisions relating to algorithmic systems,

Problem' (2020) 7 Big Data & Society 1; Niva Elkin-Koren, 'Contesting Algorithms: Restoring the Public Interest in Content Filtering by Artificial Intelligence' (2020) 7 Big Data & Society 1.

³⁰⁴ See also *supra*, Sections 3.3.5 and 3.3.6.

³⁰⁵ Recall in this regard the discussion *supra*, in Section 2.3.4, and the conceptualisation proposed by Bovens and Zouridis regarding the wide-spread uptake of algorithmic systems by public authorities as 'system-level bureaucracy'. See Bovens and Zouridis (n 115).

³⁰⁶ Peeters and Widlak (n 44) 177–78.

³⁰⁷ See also *supra*, Sections 2.2.4, 4.1.1.c and 4.1.2.c.

and hence diminishes the opportunity to hold public authorities to account for their outcomes, especially when public officials *themselves* might not know how the systems function. Given the scale at which algorithmic regulation can operate, oversight over the systems' functioning to ensure no errors are made at the level of individual decisions is in any case challenging.³⁰⁸ As noted by Schartum, "*if millions of individual decisions are made by the system, in the blink of an eye, it will generally not be feasible to manually check each output from the system, because it would take an army of case officers and extraordinary budgets to exercise meaningful controls*".³⁰⁹ Moreover, by the time a problematic decision is taken, whether directly or based on an algorithmic recommendation, it may already be too late to counter potential adverse effects that may ensue therefrom. This renders the need for upstream control and oversight even more pressing,³¹⁰ not only to avoid potential errors, but also to ensure that the coders did not take too many liberties in the translation process from law to code, whether at their own initiative, the initiative of their private employer or the initiative of the executive that hired them (and may seek to entrench its power).

Yet this very need also raises a fundamental question: how can such upstream oversight be organised, and who can fulfil this role? Coders typically do not have domain expertise the way traditional public officials do, nor are they trained in or bound by the public sector's deontological codes if they are part of a private company to which the system's development is outsourced. In the case of the latter, they can still be said to act *on behalf of* public authorities, and public authorities could hence – through contractual means – hold them to account when they do not deliver what was agreed.³¹¹ Yet for that to happen, the public authority first needs to know something is *off*, which is not easy if the relevant choices pertaining to the system's design are implicit and invisible. Accordingly, *internal* review and oversight is not always straightforward, even though the public authority that relies on

³⁰⁸ See also Stefan Buijsman and Herman Veluwenkamp, 'Spotting When Algorithms Are Wrong' (2023) 33 *Minds and Machines* 541.

³⁰⁹ Schartum (n 1) 307.

³¹⁰ I will come back to this need for institutionalised oversight *infra*, in Section 4.3 and in subsequent chapters.

³¹¹ Consider in this regard the example of the Concentrix System used in the UK, designed to add capacity to the Revenue and Customs (HMRC) services to prevent or detect error and fraud in personal tax credits awards. Like many of the other examples, this system was adopted to save costs and reduce personnel, which could in theory be deployed elsewhere. However, the system was fraught with errors and decreased the intended 'customer service standard', and backlogs were significant. The HMRC actually had to reallocate public officials to start carrying out the tasks manually again. Furthermore, as an investigation by the National Audit Office noted, "*Concentrix stopped or amended tax credit awards in around 12% of cases investigated, of which 32% of these decisions were overturned following a mandatory reconsideration*". In November 2016, the HMRC decided to end the contract. See Report by the Comptroller and Auditor General, National Audit Office, 'Investigation into HMRC's Contract with Concentrix' (HM Revenue & Customs 2017) <www.nao.org.uk/wp-content/uploads/2017/01/Investigation-into-HMRCs-contract-with-Concentrix.pdf>.

algorithmic regulation is in theory publicly accountable for its functioning, regardless of whether it was developed in-house or procured. The difficulty of carrying out internal oversight, along with the fact that public officials have reduced agency over their decisions, is a worrisome combination of factors.

Moreover, while internal review is challenging, *external* review is even less straightforward. In theory, the legislative branch of power should be able to exercise democratic oversight of the executive's action to ensure it is aligned with democratically adopted legislation, and the judicial branch of power should be able to exercise judicial oversight over those actions in court.³¹² Yet both types of oversight are difficult to achieve if the centre of gravity of the executive's action lays in the invisible normative design choices made by a set of coders through the system's architecture. Additionally, it should be borne in mind that, even if certain aspects of the system's design *are* visible, its operation is still not necessarily intelligible for non-technical experts (including most members of the legislative and judiciary branch of power). I also noted above how oversight and contestability are complicated for the natural and legal persons affected by the system and, more importantly, how the lack of the availability of *systemic* rather than mere individual review undermines the ability to challenge the societal harm that may be engendered through problematic algorithmic systems.³¹³

Let me clarify that the risk of diminished accountability goes beyond the mere bypassing of the democratic process; it can also entail a deliberate misuse of 'democracy' (narrowly conceived as the will of the majority) to erode constitutional protections of minorities. Algorithmic regulation could potentially enable a tyrannical majority, under the guise of democracy and legality, to codify the interpretation of certain rights in a manner that erodes the law's protective role in constitutional liberal democracies. Recall in this regard that some EU Member States have indeed been relying on oppressive yet 'legally' adopted laws to erode the rights of minorities, and that the translation of these laws to code would hence enable them to apply such laws at scale, while simultaneously reducing visibility over their application, even if infringing EU law and human rights law.

To conclude, the fact that oversight by the legislator, the judiciary, the public and even at times the executive itself is made difficult risks diluting the constitutional checks and balances that tame the executive's power, and creates a problematic accountability deficit. Left unaddressed, over time, this deficit might further enlarge the asymmetry of power and information between the executive and the other branches, and thereby exacerbate the problem.³¹⁴ Unless algorithmic systems, and the normative processes underlying their design and deployment, are rendered

³¹² See also Richard Bellamy (ed), *The Rule of Law and the Separation of Powers* (Ashgate/Dartmouth 2005).

³¹³ See *supra*, Section 4.1.5.

³¹⁴ See also Passchier (n 223); Malcolm Langford, 'Taming the Digital Leviathan: Automated Decision-Making and International Human Rights' (2020) 114 AJIL Unbound 141.

intelligible and controllable for non-coders, the contestability of the administrative acts they inform and adopt is undercut, as is the public accountability for their effects.³¹⁵ Once again, the protective role of the law, serving as a means to keep the executive's power in check and to protect human rights, risks being undermined. The difficulty to carry out internal and political oversight over these – normatively relevant – upstream design choices, despite their techno-rational coating, is problematic not only if the aim is to avoid erroneous translations and applications of the law, but also if the aim is to counter potentially abusive or arbitrary implementations, especially over the longer term. It must hence be ensured that algorithmic regulation cannot become a tool to bypass the democratic process and shortcut the principle of participation in law and policymaking, by securing accountability not only for the system's individual outcomes but also for the upstream decisions that shape these outcomes.

4.2.5 Systemic Vulnerability

There is one further characteristic of algorithmic rule by law that needs to be examined, which pertains to the underlying digital infrastructure that enables algorithmic regulation. Once such infrastructure has been put in place to implement and apply the law, by informing or adopting administrative acts, it should be kept in mind that it can also be altered, openly or behind the scenes. Unlike legal texts and policy implementation guidelines, which are often published in an official journal or a government website, software is inherently malleable, and can be changed by coders (or by hackers) with a few mouse clicks. When one considers the consequences in the long term, including the realisation that governments and policies change over time, one must also face the risk that this infrastructure may be used to implement policies that are normatively dubious, or plainly disregard EU or human rights law. Recall the example I mentioned above, about the algorithmic system deployed by the US Immigration and Customs Enforcement to help evaluate whether illegal immigrants should be detained or released on bail, and how from one day to the next, the system's functioning was altered following a change in policy by the Trump office.³¹⁶ The example of Belgium's reliance on ANPR cameras and the function creep accompanying their use is likewise a case in point.³¹⁷

Let me complement those examples with a hypothetical illustration that builds on an existing use of algorithmic regulation, namely automated risk assessments to

³¹⁵ As argued by Hildebrandt “*Administrative decisions taken by code-driven regulation must thus always be contestable on the double basis of: ‘the decision is based on legal conditions that do not apply because the system got the facts wrong’, and ‘the decision is based on a wrong interpretation of the relevant legal norms’*” in ‘Algorithmic Regulation and the Rule of Law’ (n 121) 3.

³¹⁶ Rosenberg and Levinson (n 219).

³¹⁷ See *supra*, Section 4.1.6. See also Belga (n 239).

detect and predict potential child abuse or neglect. As previously explained, such systems are used to identify families where (typically child welfare) authorities will prioritise their investigations, and may ultimately lead to the potential displacement of children away from their parents. Both in Europe and in the US, algorithmic regulation is already used for this purpose.³¹⁸ Now let me consider a development that in first instance seems unrelated: the rising state-sanctioned discrimination against LGBTQ+ persons in countries that are supposedly adhering to liberal democratic values, including Hungary and Romania.³¹⁹ All of these countries are, at least on paper, committed to human rights, democracy and the rule of law, yet by adopting legislation that curtails the rights and visibility of LGBTQ+ persons (for instance based on the view that it may cause 'damage' to children) they show that no algorithms need to be relied on to act in contradiction with those values.³²⁰ Is it too far a stretch to hypothesise that information relating to such orientation (e.g. a registered same-sex partnership or marriage, or related proxies) could in these countries, at some point, be considered a risk-relevant parameter that should be added to the aforementioned algorithmic system, based on the reasoning that this information may contribute to a 'better' assessment of risks for children?

In such an example, one can imagine how much further-reaching discriminating policies could be if they are supported by an infrastructure of algorithmic regulation which allows the automated and systemic application of a policy, rather than being a mere piece of text that still needs to be implemented by (potentially critical) public officials. Add to this the other elements discussed above, namely that the choice to add this discriminatory risk-factor may be disguised as a merely techno-rational one, the supremacy of coders who can make these choices with little to no visibility and oversight, the deficit of accountability in terms of constitutional checks and balances, and the automation of obedience which side-lines critical reflection and technically prevents any correction by public officials who oppose such discriminating policies. Immediately, it becomes clear that what is at stake is not merely the risk of isolated instances of discrimination at the level of individuals, but the risk of a systemic breach of the rule of law, enabled by an algorithmic legal infrastructure that allows for instantaneous mass decision-making that can directly affect the population at large.

In EU legal doctrine, the concept of a 'systemic' deficiency of the rule of law has been developed to denote a situation in which a Member State infringes the law in a

³¹⁸ Reference can, for instance, be made to the Gladsaxe model used in Denmark to detect risk indicators for children in vulnerable families. This system has been put on hold after critical coverage in the press. See in this regard Kaun (n 93); AlgorithmWatch (n 62).

³¹⁹ See *supra*, Section 3.2.5 for a discussion thereof.

³²⁰ See also Zoltán Kovács, 'Portrayal and Promotion – Hungary's LGBTQI+ Law Explained' *Euractiv* (24 June 2021) <www.euractiv.com/section/non-discrimination/news/portrayal-and-promotion-hungarys-latest-anti-lgbt-law-explained/>.

structural manner and at scale, rather than ‘merely’ episodically.³²¹ The systemic nature of the breach reflects the high threshold that needs to be reached before the procedure of Article 7 TEU (aimed at protecting EU values by suspending a Member States’ rights deriving from the Treaties, including voting rights in the Council) can be triggered.³²² The consequences of this mechanism are severe: it has at times been referred to as the ‘nuclear’ option.³²³ Occasional violations of the law therefore do not qualify for such standard, but only violations that are persistent or structural. As noted by Toggenburg and Grimheden, “*a solid debate on systemic deficiencies cannot stare at single legalistic elements in isolation but has to look at the ‘combined effects of many developments’. Against a specific political background various legal developments can lead to a situation where ‘the whole is greater than the sum of its parts’.*”³²⁴

The threshold of a ‘systemic’ deficiency has also been used in the context of the two-step test developed by the CJEU to assess whether a European Arrest Warrant should be executed.³²⁵ National judicial authorities can use this test to determine, based on the presence of systemic deficiencies relating to the independence of the judiciary in a given Member State, whether there are substantial grounds for believing that the person in respect of whom a European arrest warrant has been issued by that Member State, “*if surrendered, runs a real risk of breach of his or her*

³²¹ See, for instance, Armin von Bogdandy and Michael Ioannidis, ‘Systemic Deficiency in the Rule of Law: What It Is, What Has Been Done, What Can Be Done’ (2014) 51 Common Market Law Review 59.

³²² It should be noted that Article 7 TEU in fact speaks of a ‘serious and persistent’ breach, yet the CJEU and other EU institutions – along with legal scholars – typically have also used the term ‘systemic’. See also Joined Cases C-354/20 PPU and C-412/20 PPU, 17 December 2020, EU: C:2020:1033, §69; Case C-216/18 PPU, LM, 25 July 2018, EU:C:2018:586, §§47–75; Joined Cases C-562/21 PPU and C-563/21 PPU, 22 February 2022, ECLI:EU:C:2022:100, §50. See also the references to ‘systemic’ threats and ‘systemic’ risks to the rule of law by the European Commission, ‘Communication from the Commission to the European Parliament, the European Council and the Council: Further Strengthening the Rule of Law within the Union – State of Play and Possible Next Steps’ (2019) COM/2019/463 final.

³²³ See José Manuel Durão Barroso, ‘State of the Union 2012 – Address Plenary Session of the European Parliament’ (European Commission 2012), <https://ec.europa.eu/commission/press-corner/detail/en/SPEECH_12_596>. See also Sonja Priebe, ‘Watering down the “Nuclear Option”? The Council and the Article 7 Dilemma’ [2022] Journal of European Integration 1.

³²⁴ Gabriel N Toggenburg and Jonas Grimheden, ‘Managing the Rule of Law in a Heterogeneous Context: A Fundamental Rights Perspective on Ways Forward’ in Werner Schroeder (ed), *Strengthening the Rule of Law in Europe: From a Common Concept to Mechanisms of Implementation* (Hart Publishing 2016) 225.

³²⁵ See also Stefano Montaldo, ‘On a Collision Course! Mutual Recognition, Mutual Trust and the Protection of Fundamental Rights in the Recent Case-Law of the Court of Justice’ (2017) 2016 1 European Papers – A Journal on Law and Integration 965; Leandro Mancano, ‘You’ll Never Work Alone: A Systemic Assessment of the European Arrest Warrant and Judicial Independence’ (2021) 58 Common Market Law Review 683.

*fundamental right to a fair trial before an independent and impartial tribunal previously established by law”.*³²⁶

The need for an elevated threshold does not stem from the idea that sporadic infringements of the law are not problematic, but hinges on the fact that, in a well-functioning liberal democracy, with a legal system based on the rule of law, such infringements can in principle be overcome. That is, after all, the role of the law: providing both *ex ante* and *ex post* protection against violations of the law, and ensuring that governments who violate their legal obligations can be held to account. However, in a context where violations have become *systemic*, citizens can no longer count on the fact that the law will be able to fulfil this role, leading to a loss of trust in the legal system and in public institutions more generally.³²⁷

If we consider the adverse impact of algorithmic regulation on the rule of law based on the analysis carried out above, we can observe that this is precisely what is at stake here: the threat of a systemic deficiency in the rule of law, both literally and legally speaking. Literally, because the law's inability to properly play its protective role is exacerbated by the use of an algorithmic *system*, embedded in a networked infrastructure that enables its opaque automation and systematisation in a way that can undermine the rule of law's spirit. Legally, because the sheer scale at which this practice can take place, precisely due to the automation that enables mass decision-making, and the fact that it touches upon the very foundations of a Member State's legal system, can meet the threshold of a *systemic* breach of EU law.

Recall in this regard also the conceptualisation by Huq and Ginsburg of the broader phenomenon of ‘constitutional retrogression’:³²⁸ a development that happens piecemeal by introducing gradual changes in the legal system that undermine liberal democratic values. They note that, whereas “*each of these changes may be innocuous or even defensible in isolation*”, “*it is only by their cumulative, interactive effect that retrogression occurs.*”³²⁹ Similarly, the gradual introduction of algorithmic regulation in an increasing range of domains in which public authorities make impactful decisions on individuals, without adequate safeguards, can result in such cumulative adverse impact on the rule of law, and on EU values in general.

³²⁶ See, e.g., Case 480/21, *W O and J L v Minister for Justice and Equality*, Order of the Court of 12 July 2022, ECLI identifier: ECLI:EU:C:2022:592. Indeed, pursuant to Article 1(3) of the Framework Decision of the European Arrest Warrant, Member States' obligations to respect fundamental rights are not modified by anything contained in that Decision. See 2002/584/JHA: Council Framework Decision of 13 June 2002 on the European arrest warrant and the surrender procedures between Member States – Statements made by certain Member States on the adoption of the Framework Decision 2002 (OJ L).

³²⁷ von Bogdandy and Ioannidis (n 321) 73.

³²⁸ Aziz Huq and Tom Ginsburg, ‘How to Lose a Constitutional Democracy’ (2018) 65 *UCLA Law Review* 78, 83.

³²⁹ *ibid* 97.

There is another, deeper, issue at stake here, which can be clarified by revisiting the concept of the ‘little goodness’ proposed by Emmanuel Levinas.³³⁰ Recall that this ‘little goodness’ is juxtaposed to the systematised Goodness – with capital G – which relies on the legal and political system to enforce a set of ideas of which the current office-holder is convinced that it is the ‘Good’. As history has shown, however, all systematisation of ideologies of the ‘Good’, no matter how benevolent, risk becoming a tool to do wrong precisely in the name of the good.³³¹ In the context of algorithmic regulation, such systematisation can take place literally, by codifying ideals into algorithmic systems that regulate the entire population. Yet this tends to essentialise *one* view of the good over others, and may have no place in a democratic and pluralistic society, especially if one considers the long-term consequences thereof.³³²

Accordingly, when we stop looking at the adverse effects that *one* problematic algorithmic system might have on the rights of one individual, or of one collective of individuals, and we start looking at the cumulative impact of algorithmic regulation across society, we are forced to confront the risk of a systemic threat to the rule of law, and to the normative foundations of liberal democracies more generally.³³³ While these foundations are always fragile and do not require the use of algorithmic systems to be undermined, the intrinsic malleability of algorithmic regulation, which allows for scaled and instant decision-making at the level of the entire population, introduces a systemic vulnerability in the legal system. Consequently, the risk that this vulnerability is used in a way that inadvertently or deliberately undermines the rule of law in a ‘serious and persistent’ manner needs to be considered, ideally *before* the large-scale implementation of algorithmic systems in the public sector.³³⁴

4.3 CONCLUDING REMARKS

In this chapter, I have carried out a systematic analysis of how algorithmic regulation, when used by public authorities in the context of administrative acts, can

³³⁰ See *supra*, Section 2.3.3.

³³¹ The book by Grossman, based on which Levinas’ juxtaposition of the systematised Goodness versus the little goodness is based, particularly focuses on Nazism and Stalinism. See Vasily Grossman, *Life and Fate* (1980) (Robert Chandler tr, Vintage Classic 2017). See also Luc Anckaert, ‘Goodness without Witnesses: Vasily Grossman and Emmanuel Levinas’ in Michael Fagenblat and Arthur Cools (eds), *Levinas and Literature* (De Gruyter 2020).

³³² Consider also in this regard Isaiah Berlin’s opposition against so-called ‘monism’, or “*the old perennial belief in the possibility of realising ultimate harmony*”, whereby those in power are driven by a single ideal of perfection, both at individual and societal level, without any space for a pluralistic view of ‘the good’. See Isaiah Berlin, *The Crooked Timber of Humanity: Chapters in the History of Ideas* (Henry Hardy ed, Fontana 1991) 8 and more generally Isaiah Berlin, *Liberty: Incorporating Four Essays on Liberty* (Oxford University Press 2002). See also Cécile Hatier, ‘Isaiah Berlin and the Totalitarian Mind’ (2004) 9 *The European Legacy* 767.

³³³ This point has already strongly been made in Yeung, ‘Responsibility and AI’ (n 209).

³³⁴ Simuha and others (n 57) 39.

adversely impact each of the six rule of law principles, by drawing on illustrations from existing applications. When conceptualising the rule of law and its principles in Chapter 3, I already noted that meeting their requirements entails inherent challenges, regardless of any use of algorithms. Yet the above analysis has demonstrated that reliance on algorithmic regulation can significantly exacerbate these challenges, and make compliance therewith even more difficult. As a consequence, the rule of law risks turning into *algorithmic rule by law*. The veneer of legality remains: algorithmic regulation, after all, aims to merely implement and apply the law in an optimised and more efficient way. However, the protective role of the law is hollowed out, and opens up weaknesses that can be exploited to undermine the rule of law's very purpose.

I outlined five problematic characteristics of algorithmic rule by law, thereby consolidating the common findings of the principle-by-principle impact analysis I conducted. Summarised, the law's application is being reduced to a techno-rational exercise (*primacy of techno-rationality*); its interpretation and translation to code is centralised and delegated to a handful of people with technical expertise who have the impossible task to anticipate all potential downstream situations, and whose upstream choices shaping the technology's affordance are largely invisible (*supremacy of coders*); discretion at the street-level is eliminated, leaving public officials without agency to counter the problem of law's over-generality and technically constrained to uncritically defer to the algorithmic outcomes (*automation of legalism*); public accountability mechanisms for the legality of the law's interpretation and application are eroded (*deficit of accountability*); and the infrastructure enabling algorithmic regulation introduces a significant vulnerability in the legal system, whereby a particular vision of the Good – despite potential adverse effects – can be systematised and risk leading to a systemic deficiency of the rule of law (*systemic vulnerability*).

Altogether, these characteristics can also exacerbate authoritarian elements in society, from the centralisation of power to the loss of transparency and accountability and the erosion of a sense of critical thinking and moral responsibility. Moreover, as the illustrations have shown, algorithmic regulation may also undermine human rights and foster illiberal practices, by limiting and infringing individual rights in the name of efficiency. The risk I see is not so much a sudden elimination of the law's protective function, but rather an erosion – and one that goes potentially unnoticed, given the veil of legality that surrounds algorithmic regulation, not least because of the European Commission's promotion of its uptake,³³⁵ through the incremental reliance and dependence thereon for decisions that affect individual and societal interests. Accordingly, it can serve as a tool to attain the constitutional retrogression which Huq and Ginsburg conceptualised.³³⁶

³³⁵ See *supra*, Section 2.3.5.

³³⁶ Somewhat pessimistically, Huq and Ginsburg note that “*the constitutional safeguards against retrogression are weak*”, and argue that “*the near-term prospects of constitutional liberal democracy hence depend less on our institutions than on the qualities of political leadership, popular resistance, and the quiddities of partisan coalitional politics.*” See Huq and Ginsburg (n 328).

I therefore argue that the irresponsible implementation of algorithmic regulation might foster the threat of algorithmic rule by law – whereby irresponsible means to disregard the risks I outlined, or with the deliberate aim of exploiting those risks. Let me stress that human beings are not devoid of error or ill intention, and that I hence do not argue that the regression of democracy and the erosion of the rule of law is *caused by* reliance on algorithmic regulation. Nor do I claim that the use of algorithmic regulation *necessarily* leads to a materialisation of the threat of algorithmic rule by law. My claim is merely that it can be exacerbated thereby, in light of the features inherent to algorithmic systems. Accordingly, if public authorities wish to rely on algorithmic regulation, the threat of algorithmic rule by law needs to be addressed. The fact that this technology provides the executive branch with more power and introduces stronger risks to the rule of law requires appropriate counterbalancing mechanisms.

The question is hence: does the current legal framework have sufficiently strong mechanisms in place to enable such counterbalancing?³³⁷ Certainly, the law has its limits and it would be a mistake to consider legal rules as a panacea to all the identified problems. Yet it is worth asking which safeguards the EU legal order currently provides against the conceptualised threat. It goes beyond the purpose of this book to formulate detailed legal solutions. However, based on my analysis, a number of general conclusions can be made regarding the protection that the legal system should ideally provide.

First, given the vast scale of the harm that can arise from the problematic use of algorithmic regulation and the potential irreversibility of the damage, mere reliance on ex post remedies is insufficient. This does not mean that ex post remedies have no role to play in countering the identified threat. To the contrary, it is equally important to reflect on how they can be strengthened to ensure that not only individual but also systemic review of the executive's action through algorithmic regulation can be carried out. However, ex ante protection mechanisms are also needed, for instance in the form of certain requirements that should be fulfilled before algorithmic regulation can be used. In light of the importance of the decisions made during the design and implementation phase of algorithmic systems, oversight also needs to occur at the upstream level, rather than only at the level of the outcomes proposed or adopted by the system. The translation of legal rules and policies from law to code is not a techno-scientific matter, but an exercise that entails normative and political choices.³³⁸ While the drive to rationality and efficiency might lead public authorities to ignore this fact, the principle of legality can only be secured by ensuring transparency, oversight and contestability over these important upstream choices (*prior and continuous oversight and accountability*).

³³⁷ As my analysis in Chapter 5 will demonstrate, the answer to this question is negative. However, in Section 6.2, I will provide a number of recommendations to secure stronger safeguards against this threat.

³³⁸ See *supra*, Section 4.1.1. See also Hildebrandt, 'Algorithmic Regulation and the Rule of Law' (n 121).

Second, one can argue that the mere choice of introducing algorithmic regulation, especially in certain sensitive domains, is already an administrative act that should be subjected to democratic oversight and judicial review in its own right. Given the potential consequences linked thereto, it is fair to claim that there should be no algorithmisation without representation (to paraphrase American revolutionists).³³⁹ More generally, given the impact of algorithmic regulation on citizens, and given the fact that the principle of participation is increasingly recognised as essential also in public administration, citizens should be able to participate in – and give feedback on – important choices regarding the algorithmisation of the public sector (*public participation in algorithmisation*).³⁴⁰

Third, since the threat of algorithmic rule by law comes from reliance on algorithmic regulation by Member States, it is important that these safeguards, both *ex ante* and *ex post*, do not solely rely on public authorities of those very Member States. Ideally, safeguards can be invoked through both private and public enforcement mechanisms, ensuring that also citizens can play a role to hold the government's use of algorithmic regulation to account. Moreover, given the importance for the EU as a whole that Member States adhere to the rule of law, and given that not only national but also EU law can be inadvertently or deliberately infringed, one should also consider the role that EU institutions might play in mitigating and addressing the risk that Member States infringe the rule of EU law through reliance on algorithmic regulation (*private and public enforcement, at national and EU level*).

Fourth, the protective role of the law needs to be safeguarded by ensuring adequate individual and societal remedies against the scaled risks introduced by algorithmic regulation. Besides ensuring remedies for individuals who can be adversely impacted thereby, the fact that the rule of law's erosion leads to *societal* harm means that citizens and public interest groups should be able to counter this harm also without necessarily demonstrating individual harm. More generally, stronger checks and balances are also needed to ensure that the legislative and judicial branches of power, along with civil society and the public at large, can hold the executive accountable for its actions (*stronger checks and balances*).

Finally, attention should also be given to the role of public officials, and the importance of safeguarding their agency when applying legislation and taking administrative acts. The balance between rules and discretion, rigidity and fluidity,

³³⁹ The phrase 'no taxation without representation' was used by American colonists during British rule in the late eighteenth century.

³⁴⁰ Elsewhere, I drew a parallel with the need to ensure public participation and feedback in the context of environmental impact assessments, comparing the harm of public actions to the environment (for instance by allowing or establishing a polluting activity) to the harm that can ensue from AI's impact on societal interests. Just like environmental legislation provides public participation in decisions that affect the environment, so should such participation be foreseen in decisions that concern the implementation of algorithmic regulation when it can affect societal interests such as the rule of law. See in this regard Smuha, 'Beyond the Individual: Governing AI's Societal Harm' (n 209).

predictability and adaptability needs to be maintained. This means that, rather than operating seamlessly and restrictively, algorithmic regulation should allow for a certain level of friction that enables public officials to exercise critical judgment and maintain both actually and mentally a sense of responsibility for the outcome of public action. Furthermore, rather than viewing algorithmic regulation as a means to systematise ‘the Good’, opportunities for contestation and agonistic interpretations³⁴¹ need to be ensured, including the role of the ‘little goodness’ to soften the law’s hard edges where need be (*contestation and internal critical reflection*).

³⁴¹ See in this regard Mireille Hildebrandt, ‘Privacy as Protection of the Incomputable Self: From Agnostic to Agonistic Machine Learning’ (2019) 20 *Theoretical Inquiries in Law* 83.