2

Algorithmic Regulation

As its name indicates, algorithmic regulation relies on the automation of regulatory processes through algorithms. Examining the impact of algorithmic regulation on the rule of law hence first requires an understanding of how algorithms work. In this chapter, I start by focusing on the technical aspects of algorithmic systems (Section 2.1), and complement this discussion with an overview of their societal impact, emphasising their societal embeddedness and the consequences thereof (Section 2.2). Finally, I examine how and why public authorities rely on algorithmic systems to inform and take administrative acts (Section 2.3), before drawing some conclusions for subsequent chapters (Section 2.4).

2.1 TECHNICAL ASPECTS

To get a better grasp of the technical processes that underlie algorithmic regulation, I will start by explaining what algorithms are (Section 2.1.1). Subsequently, I examine two commonly distinguished categories of algorithmic approaches, namely those based on reasoning, also referred to as knowledge-driven systems (Section 2.1.2) and those based on learning, also referred to as data-driven systems (Section 2.1.3). After comparing their respective strengths and weaknesses, I emphasise that the distinction between both should not be seen as strict, since many systems are composed of a combination of techniques, and the trend in the field is evolving towards hybrid systems (Section 2.1.4). I conclude with a discussion on the relationship between algorithmic systems and artificial intelligence, and a technical conceptualisation of algorithmic regulation for the purpose of this book (Section 2.1.5).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

2.1.1 *Algorithms*

Algorithms are most commonly described as finite sequences of defined instructions, aimed at solving a problem.[1] These instructions can be expressed in computer code, which enables the algorithm's execution by a computer program. In this way, algorithms enable human beings to automate processes that they would otherwise need to undertake themselves, from a simple repetitive task to a complex calculation, or the analysis of heaps of information. An algorithm is hence shaped by whichever is the human-defined objective behind it, or the problem one aims to solve.[2] Over time, researchers have sought to address problems of varying complexity through algorithmic systems – understood as systems that are comprised of multiple algorithms – from automatically translating text from one language to another, or distinguishing between different human voices, to calculating the individual tax rates of the entire population. Algorithmic systems are increasingly used to take over tasks which would be considered as requiring 'intelligence' if carried out by a human being.[3] Accordingly, the term artificial intelligence – which I will discuss further in Section 2.1.5 – can be seen as an umbrella term for different algorithmic techniques aimed at carrying out 'intelligent' functions.[4]

In its simplest form, an algorithm relies on certain *input* that is provided, after which the algorithm – or the sequence of step-by-step *instructions* expressed in code – is executed, which ultimately results in an *output*.[5] Figure 2.1 on the next page provides a visual abstraction of this process.

[1] The concept of 'algorithm' dates back to antiquity and was primarily associated with the computation of mathematical functions. The first algorithm is considered to have originated with Euclid around 300 BC and served to compute the greatest common divisor of two integers. Etymologically, the word 'algorithm' stems from the ninth-century Persian Muḥammad ibn Mūsā al-Khwārizmī (Latinised as 'Algorithmi'), a polymath who – amongst many other achievements – popularised algebra and presented the first systematic solution of linear and quadratic equations. It is, however, only in the nineteenth century that algorithms also started to be considered as a function that could be executed by a computer. The first such algorithm was written by writer and mathematician Lady Ada Lovelace sometime between 1842 and 1843, long before the first modern computer was built. See WW Rouse Ball, *A Short Account of the History of Mathematics* (4th edn, Dover Publications 1908) 129; Stuart Jonathan Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (3rd edn, Pearson 2016) 8. With the advent of modern computers in the twentieth century, algorithmic functions started being used as an automation technique. For a further discussion of the definition of algorithms, see, e.g., Thomas H Cormen and others, *Introduction to Algorithms* (3rd edn, MIT Press 2009).

[2] See in this regard also Michael Veale, 'Governing Machine Learning That Matters' (PhD thesis, University College London, 2019), 29.

[3] See also Nils J Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (Cambridge University Press 2009).

[4] See also Russell and Norvig (n 1).

[5] See, e.g., Cormen and others (n 1). See also Woodrow Barfield and Jessica Barfield, 'An Introduction to Law and Algorithms' in Woodrow Barfield (ed), *The Cambridge Handbook of the Law of Algorithms* (Cambridge University Press 2020) 4.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

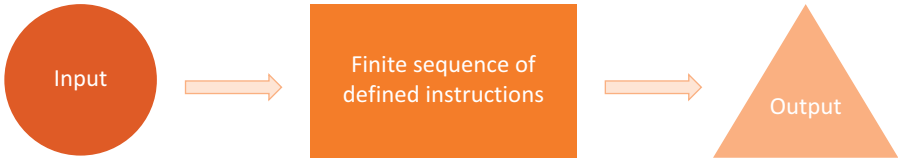


FIGURE 2.1 Abstraction of an algorithm.

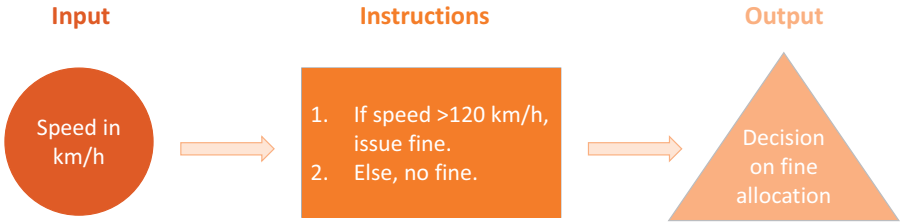


FIGURE 2.2 Abstraction of an algorithmic system to automate the decision to issue a fine.

An algorithm can be compared to a cooking process: it requires ingredients (input), step-by-step directions of how to prepare and cook these ingredients (instructions), resulting in a – hopefully delicious – dish (output). The shape and content of the input, instructions and output will depend on the algorithm's context and purpose, just as the ingredients, cooking instructions and resulting dish will depend on what one intends to cook.

To provide an example, I may wish to automate the issuing of a fine to highway drivers who drive too fast. They should, however, only be fined when they exceed the maximum speed limit, which on Belgian highways is 120 km/h. To this end, I could develop an algorithmic system that uses as input the measured km/h that the car drives. If I do not want to manually insert those km/h into the system, I could equip the system with a sensor that enables it to automatically measure the speed and thereby facilitate the provision of its input. In terms of instructions, I will need to write out in code that, once the speed limit is exceeded, a fine should be issued. Finally, the resulting output would be a concrete decision to issue a fine (or not) to the driver in question. Of course, I could also decide to include additional input and instructions to the system so that I not only automate the decision of issuing a fine but also the subsequent step of looking up the driver's address, and notifying her that a fine should be paid within a certain timeframe – yet in its simplest form, this algorithmic process comes down to Figure 2.2.

The outcome of an algorithmic system can thus be used to make a decision without the need for further human intervention, such as in the above example. This can be referred to as *algorithmic decision-making*, since it concerns a decision-making process

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

that is automated through an algorithm.[6] Although I may be physically present when a driver crosses the speed limit, I will not necessarily be prompted to first verify the system's input, nor will I need to confirm that a fine should indeed be issued. Instead, I have outsourced that decision to the algorithmic system. The outcome of such system, however, need not necessarily be a final decision. Algorithms can also be used for a range of other functions or sub-tasks, from executing a simple calculation, to computing a prediction, recommendation or categorisation,[7] which can then inform a subsequent human or algorithmic decision.

The above algorithmic system could, for instance, also be tweaked to produce as output a recommendation of whether or not to issue a fine, which a human being then needs to approve, rather than an actual fining decision. This is sometimes denoted as an *algorithmic recommendation system*. Moreover, it is also possible that an algorithmic system is used to merely support a sub-part of a decision, based on which a human being can then decide to take a particular action. For instance, an algorithmic system can be used to automatically determine the speed at which a car was driven, without also being programmed to issue a recommendation to fine the driver when that speed reaches the limit. Nevertheless, I may use that piece of information to determine whether the speed limit was exceeded and whether a fine should be issued. This can be referred to as a more general *algorithmic decision-support system*, since the system is used to support a decision, without making an explicit recommendation or taking a final decision.[8]

Collectively, the algorithmic decision-making system, recommendation system and support system described above enable *algorithmic regulation*, since they are deployed by a public authority to inform or take an administrative act (namely, issuing a fine). In this book, I will hence deploy the term *algorithmic regulation* to cover all three types of systems. However, when discussing concrete examples, and whenever the distinction is relevant, I will indicate whether it concerns an algorithmic system that *adopts* (takes) or that *informs* (recommends or supports) an administrative act.

An obvious advantage of using an algorithmic system instead of manually measuring the driving speed and issuing corresponding fines is the fact that the process can be applied at much larger scale and speed, so that many more drivers can be controlled and fined where needed and unwanted behaviour can be reduced. Moreover, I no longer always need to be physically present at the place of the event,

[6] See, for instance, Theo Araujo and others, 'In AI We Trust? Perceptions about Automated Decision-Making by Artificial Intelligence' (2020) 35 AI & Society 611, 611–12.

[7] This variety of functions is, for instance, reflected in the definition of 'AI' of the European Commission, the European Commission's High-Level Expert Group on AI, the AI Act and the OECD, to which I will return in subsequent sections.

[8] See, for instance, Anne Kaun, 'Suing the Algorithm: The Mundanization of Automated Decision-Making in Public Services through Litigation' [2021] Information, Communication & Society 1, 5.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

but can instead be at my office and work on another matter in the meantime. However, since the delegation of tasks to algorithmic systems also entails, to a certain degree, the delegation of responsibility for tasks originally undertaken by human beings, reliance thereon can give rise to several ethical and legal questions, depending on the task at stake. As long as I am using an algorithmic system to automatically activate my heater when the temperature falls below 20°C, this algorithmisation does not lead to much consternation. Yet the algorithmisation of tasks that have more significant consequences for individuals, collectives and society evidently needs to be handled with greater care. In the example above, for instance, all three of the systems – although differentiable in the part of the task that they automate – can entail consequences for drivers who may be subjected to a fine.

Various algorithmic approaches can be used to automate tasks, and these can be classified in several ways. A commonly made distinction concerns algorithmic systems focused on *reasoning* – sometimes referred to as *knowledge-driven* systems – and those focused on *learning* – sometimes referred to as *data-driven* systems.[9] Yet it must be stressed that the boundaries of these categories are not uniformly set, and that algorithmic systems can also draw on different approaches simultaneously or on approaches that do not neatly fall under one or the other category. With this caveat in mind, in what follows I will set out the main features of knowledge-driven and data-driven systems, respectively. Rather than giving a comprehensive account of the wide set of techniques falling under each of these approaches, the discussion below will be succinct, aimed only at providing the reader with a basic understanding of how these systems might operate the context of algorithmic regulation.

2.1.2 *Knowledge-Driven Systems*

As their name implies, knowledge-driven systems start from a certain knowledge model of reality that is represented through code, based on which the algorithm can subsequently reason. Two elements are core in this regard. The first concerns a knowledge base, which can contain facts, observations and other types of data that represent a coded model of reality. Such knowledge must be 'translated' so that it can be processed by an algorithmic system – also known as *knowledge*

[9] Luc De Raedt, 'De artificiële-intelligentierevolutie en de impact ervan' in Pieter d'Hoine and Bart Pattyn (eds), *Wetenschap in een veranderende wereld: Lessen voor de eenentwintigste eeuw* (1st edn, Universitaire Pers Leuven 2020) 286. Note that others have proposed a classification of algorithmic systems based on large paradigms within the research domain of AI, including the symbolic school, the Bayesian school, the analogisers, the neural or connectionist school and the evolutionary school. See Pedro Domingos, *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World* (1st edn, Penguin 2017).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

representation.[10] The second element concerns coded instructions about how the system can analyse, combine and draw inferences based on that knowledge to solve a problem – known as *knowledge manipulation* or simply *reasoning*.[11] These instructions are typically codified in the system through symbolic rules that set out the relationships between the symbols and establish how the system can deduce information, engage in planning and scheduling activities, search through knowledge bases, and optimise solutions.[12] How the developer of the system translates knowledge and rules into code is hence key, as is the extent to which the codified knowledge is correct or based on appropriate assumptions and beliefs. Given the important role that pre-established rules play in these systems, they are sometimes also referred to as 'rule-based systems'.

Rule-based systems were the dominant paradigm up until the 1980s and are therefore also known as 'traditional' or 'good, old-fashioned' AI.[13] They have, for instance, successfully been programmed to win a game of chess from the world chess champion,[14] as the rules of chess-playing are well delineated and can hence be represented in code with relative ease. Such systems can also be used by public services to assist in decision-making on the allocation of social welfare benefits. To this end, information about the income, family situation and legal status of individuals first needs to be represented in code – along with other information that determines whether or not an individual is eligible to receive benefits. In addition, the relevant legislation that stipulates the criteria and conditions under which individuals can receive benefits also needs to be translated to symbols and rules. On that basis, the system can subsequently infer a citizen's benefits eligibility, calculate the corresponding amount and provide as output a recommendation or

[10] See also Ronald J Brachman and Hector J Levesque, *Knowledge Representation and Reasoning* (Elsevier 2004). Knowledge representation can occur through symbolic or non-symbolic methods, though the former is more prevalent.

[11] See also M Michalewicz, ST Wierzchoń and MA Kłopotek, 'Knowledge Acquisition, Representation & Manipulation in Decision Support Systems' (arXiv, 23 May 2017) <http://arxiv.org/abs/1705.08440>. Note that the use of the term 'manipulation' in this technical context is thus not the same as the use of the term in a social or legal context, which, according to the *Cambridge Online Dictionary*, rather connotes "*controlling someone or something to your own advantage, often unfairly or dishonestly*".

[12] See High-Level Expert Group on AI, 'A Definition of AI: Main Capabilities and Scientific Disciplines', Brussels, April 2019, <www.digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>, 3.

[13] Nan Duan, Duyu Tang and Ming Zhou, 'Machine Reasoning: Technology, Dilemma and Future' in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Tutorial Abstracts* (Association for Computational Linguistics 2020) 1 <www.aclweb.org/anthology/2020.emnlp-tutorials.1>.

[14] See also Brian P Bloomfield and Theo Vurdubakis, 'IBM's Chess Players: On AI and Its Supplements' (2008) 24 The Information Society 69; Feng-Hsiung Hsu, *Behind Deep Blue: Building the Computer that Defeated the World Chess Champion* (Princeton University Press 2022).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

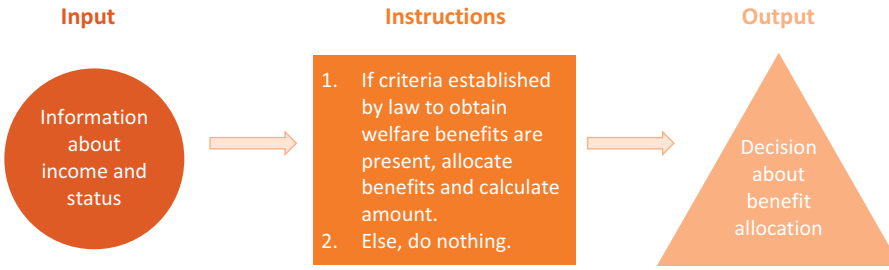


FIGURE 2.3 Abstraction of a knowledge-driven system to automate welfare benefits allocation.

decision on whether benefits should be allocated and how high the amount should be – as represented in Figure 2.3.[15]

Given their reliance on finite and discrete symbols and logical rules, reasoning processes are typically interpretable. The reasoning steps of the algorithm – which can have varying forms of complexity – can be retraced, which makes the algorithmic process relatively transparent and explainable (at least for technical experts, who should then be able to communicate about how this process works to those subjected thereto). This traceability also makes it possible to verify whether some important information has potentially been overlooked, or whether the legal rules about benefits eligibility, which were translated into symbolic rules, have been interpreted correctly. At the same time, reasoning techniques in principle presuppose that the steps or rules that need to be taken to solve a problem are – at least to a certain extent – already known.[16] Similarly, they presuppose the existence of human expertise in the domain to which the problem relates.[17] In situations where the rules that apply are pre-established by law, relatively precise and relatively easily codifiable in symbols, this need not pose major issues – even though, as we shall see in Chapter 4, the translation from legal rules to code poses fundamental challenges.[18] However, in situations where the rules that govern the situation are not (yet) known, vague or uncertain, a knowledge-driven approach that is based on simplistic rules may not be effective.

[15] As will be discussed in Section 4.1, algorithmic systems are already used by several governments for this purpose. See also Paul Henman, 'Digital Technologies and Artificial Intelligence in Social Welfare Research: A Computer Science Perspective' in Michael Adler (ed), *Research Agenda for Social Welfare Law, Policy, Practice and Impact* (Edward Elgar Publishing 2022).

[16] See, e.g., Birte Glimm and Yevgeny Kazakov, 'Classical Algorithms for Reasoning and Explanation in Description Logics' in Markus Krötzsch and Daria Stepanova (eds), *Reasoning Web. Explainable Artificial Intelligence*, vol 11810 (Springer International Publishing 2019).

[17] See also James P Ignizio, *Introduction to Expert Systems: The Development and Implementation of Rule-Based Expert Systems* (McGraw-Hill 1991).

[18] See in this regard also Mireille Hildebrandt, 'Law as Information in the Era of Data-Driven Agency' (2016) 79 The Modern Law Review 1.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

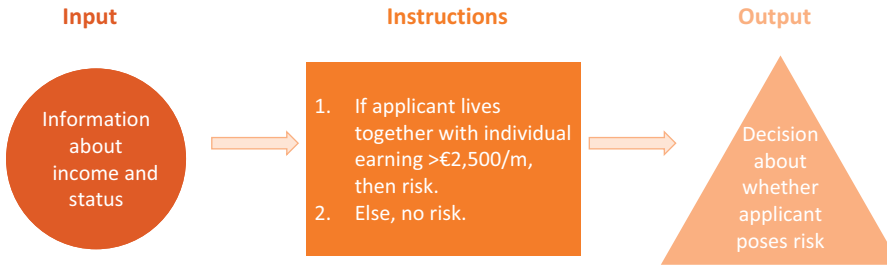


FIGURE 2.4 Abstraction of a knowledge-driven system to identify the risk of social welfare fraud.

Consider, in this regard, an example in the same sphere: an algorithmic process to assess the risk that individuals commit social welfare or tax fraud. While the majority of taxpayers and welfare applicants act in good faith, in every country, there will be individuals who commit fraudulent acts to pay less than they owe, or to receive more than they should.[19] Since public resources are limited, and not every individual can or should be investigated for fraud, one may consider developing an algorithmic system that can help identify or predict the risk of fraud, in order to prioritise those cases for investigation. Yet, while I can probably reuse a lot of the information on citizens that I had gathered for the previous example, this task may be more difficult to delegate to a rule-based knowledge-driven system. What are the rules that I would need to codify to identify a potential risk of fraud? Fraudulent types of behaviour do not always occur in easily describable steps.[20] I could make an attempt by codifying the rule that anyone who applies for welfare benefits yet lives together with someone earning more than 2,500 EUR/month poses a risk of fraud, as shown in Figure 2.4. However, with this type of rule, the output – namely, the decision on whether I need to further investigate that applicant – risks being both overly inclusive and exclusive.

[19] The detection of welfare fraud is a major application of algorithmic regulation. Its deployment aims to offset the costs incurred for the development of an algorithmic fraud detection system by the money that can be saved through the optimisation of fraud detection. See also Marvin van Bekkum and Frederik Zuiderveen Borgesius, 'Digital Welfare Fraud Detection and the Dutch SyRI Judgment' [2021] European Journal of Social Security 13882627211031256; Rikke Frank Jørgensen, 'Data and Rights in the Digital Welfare State: The Case of Denmark' [2021] Information, Communication & Society 1.

[20] This has not stopped public authorities from doing so anyway, as showcased by the example of the 'fraudescore-algorithm' used in the Netherlands up until 2020 (and in some municipalities even thereafter still). The system relied on predetermined criteria to calculate the risk of fraud, which included, for instance, the level of education of citizens, the fact that they worked as a hairdresser or taxi driver, and the neighbourhood they lived in. See also 'Junk Science Underpins Fraud Scores' (*Lighthouse Reports*, 25 June 2022) <www.lighthousereports.nl/investigation/junk-science-underpins-fraud-scores/>; '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>. I will come back to this example in Section 4.1.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

Taking a rule-based approach may pinpoint me somewhat in the right direction, but will still require me to take into account that this rule may be too general to yield effective results. Of course, the more knowledge and experience I gain, the more I can use this knowledge to formulate more tailored rules, and perhaps also to add some exceptions to the rule. However, fraudulent techniques might change over time, so I would constantly need to update those rules.

To remedy the problem of uncertainty, recourse can be made to probabilistic and statistical models that can help make predictions based on the knowledge that I do have, and thereby cover certain knowledge gaps and reduce uncertainty.[21] These models are increasingly being researched and tailored for use in the context of knowledge-driven approaches. By relying on these models, the rigidity of more simplistic 'rule-based' systems that merely apply a static rule to a knowledge base can be diminished. Probabilistic and statistical models can hence be seen as lying somewhere in between knowledge-driven and data-driven systems.

Importantly, the knowledge on which the system reasons needs to be transformed into machine-readable code.[22] Depending on the size of the knowledge base, the potentially informal nature of the knowledge, and the relationships that need to be represented between the various symbols that represent the knowledge, this can be quite laborious. In the context of algorithmic regulation, whenever the legal rules that need to be translated are vague or open to multiple understandings – which is not infrequent, given the inherent openness and interpretability of language – certain interpretative decisions will need to be taken in advance, to the exclusion of others. Furthermore, if knowledge about certain aspects of the problem is lacking, it cannot be represented. Unless this is remedied through approaches that provide more flexibility, this will mean that – when a specific situation arises that deviates from the codified norm and for which no explicit alternative rule has been foreseen – the system is unlikely to provide a solution. Notwithstanding this restraint, knowledge-driven systems are used profusely in virtually all societal domains, and are able to assist in tasks of high complexity.

2.1.3 *Data-Driven Systems*

Researchers also explored other routes to perform 'intelligent' tasks, focusing on *data-driven* approaches instead. Rather than codifying a certain model of knowledge

[21] See, e.g., Judea Pearl, 'Graphical Models for Probabilistic and Causal Reasoning' in A Tucker and others (eds), *Computing Handbook* (2014).

[22] This task has also been referred to as 'handcrafted knowledge', which reflects the manual labour that is typically associated therewith. See John Launchbury, 'A DARPA Perspective on Artificial Intelligence' (DARPA, February 2017) <www.darpa.mil/about-us/darpa-perspective-on-ai>.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

in advance, data-driven systems "*let the data speak*"[23] and infer the model from the data instead. Accordingly, these models typically allow for more flexibility and adaptability since, by *learning* from the data, the system will be able to adjust to new situations. Though already explored earlier, data-driven systems primarily gained traction after the 1980s, when the capabilities of knowledge-driven systems seemed to be hitting their limits.[24] Especially from the twenty-first century onwards, data-driven approaches started booking new successes and led to several major breakthroughs in the field – particularly those relying on machine-learning techniques.[25] It is for this reason that some AI definitions *only* refer to learning-based systems.[26]

The recent successes of those systems are largely linked to the enhanced availability of data and in particular 'big data', which can be defined as "*large volumes of extensively varied data that are generated, captured, and processed at high velocity*".[27] A growing number of tools enable the collection and analysis of data in digital – and hence easily perusable – format, by means of sensor-embedded objects, phones and computers, and the Internet of Things (IoT) more generally.[28] The fact that the processing power of computers significantly improved – thus enabling more and faster computations, and the fact that the storage of large volumes of data became more affordable, likewise contributed to the success of data-driven systems.

There are various types of data-driven techniques. Generally speaking, such techniques aim to infer the rules that should be followed from (a large number of) examples contained in a dataset, which can be labelled or not. The algorithm is thus tasked with inferring the rules (or function) behind the data and to create a

[23] Already in 1973, the French mathematician Jean-Paul Benzécri introduced the idea of "*letting the data speak for themselves*", stressing that "*Le modèle doit suivre les données, non l'inverse*" in J-P Benzécri, *L'analyse des données. 2: L'analyse des correspondances* (Dunod 1973). See also François Husson, Julie Josse, and Gilbert Saporta, 'Jan de Leeuw and the French School of Data Analysis' (2016) 73(6) Journal of Statistical Software 16.

[24] Even earlier, in the late 1950s, seminal papers by – amongst others – Ray Solomonoff laid down the groundwork for inductive inference systems, which helped pave the way for modern machine learning approaches. See also 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).

[25] De Raedt (n 9) 286.

[26] Some authors, for instance, explicitly exclude reasoning-based systems from the term 'algorithmic systems', which they reserve for data-driven systems only (as opposed to 'information systems' which can be reasoning-based), such as in 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.

[27] Wendy Arianne Günther and others, 'Debating Big Data: A Literature Review on Realizing Value from Big Data' (2017) 26 The Journal of Strategic Information Systems 191. See also Roger Clarke, 'Big Data, Big Risks' (2016) 26 Information Systems Journal 77.

[28] Alem Čolaković and Mesud Hadžialić, 'Internet of Things (IoT): A Review of Enabling Technologies, Challenges, and Open Research Issues' (2018) 144 Computer Networks 17.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

model that should identify the relationship between the various datapoints provided to the system (for instance, inputs and outputs). By analysing a large set of examples (*training phase*), and by continuously tweaking the model based on those examples to better fit the dataset's curve, the function can be determined with ever more accuracy. Once the model has been developed based on training data, whenever the system is confronted with new input, the model can be used to predict that new input's corresponding output (*use phase*). In addition, there will typically also be a *testing phase* in between, to see how well the model performs before deploying it at scale.

The fact that data-driven systems are able to 'learn' a model from the data they are provided, does not take away the fact that these systems likewise rely on human coding. After all, the algorithms still hinge on instructions pertaining to the development of the model, and the use of the model to categorise or cluster new data. Yet the representation of reality itself will in principle not be codified in advance, nor will the instructions set out how such representation can be manipulated to achieve a certain result. This absence of the codification of knowledge and rules that set out the relationship between different knowledge elements also means that data-driven systems are typically not able to pick up *causal* relations between input and output (though this is a growing topic of research).[29] Instead, the inferred relationship between input and output is one of *correlation*. As to inferring the model, different techniques exist. A distinction is often made between supervised, unsupervised and reinforcement learning techniques – each of which is briefly discussed below.[30]

Under *supervised learning*, an algorithm is instructed to analyse training data examples that consist of input–output pairs which are explicitly labelled or scored, and to infer a model therefrom. The labelling typically occurs by a human being prior to the dataset's use. Based on the model that is inferred from the original examples, the system will have learned how to classify new examples. The performance of the system will typically increase over time, as the analysis of more data can allow it to infer more granular rules and improve the model.

For an example of such a method, I suggest revisiting the algorithmic system used to assess the risk of social welfare fraud. Instead of using a rule-based algorithm and programming rules that are likely to be under- and over-inclusive, I can instead try to teach the system to distinguish between fraudulent and non-fraudulent applications, as represented in Figure 2.5 on the next page. To develop the system, I could train it on a dataset containing information about past welfare benefits applicants, as well as an indication (or label) of whether they committed fraud in the past or not. Crucially, I would need a sufficiently large dataset, with enough examples of both

[29] See also Judea Pearl, 'The Seven Tools of Causal Inference, with Reflections on Machine Learning' (2019) 62 Communications of the ACM 54; Sema K Sgaier, Vincent Huang and Grace Charles, 'The Case for Causal AI' (2020) 18 Stanford Social Innovation Review 50.

[30] See also Igor Kononenko and Matjaž Kukar, *Machine Learning and Data Mining* (Woodhead Publishing 2007).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

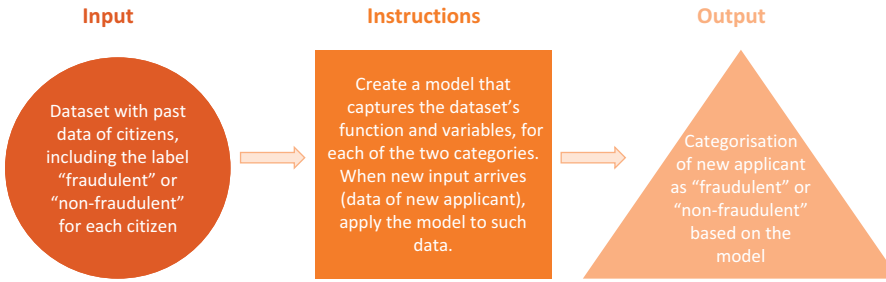


FIGURE 2.5 Abstraction of a supervised data-driven system to predict the propensity of fraud.

types of categories.[31] On that basis, I would instruct the algorithm to create a model that seeks to deduce the relevant variables for each of the two categories. When a new applicant subsequently submits an application to receive benefits, I can then use that model to predict which of the two categories – fraudulent or not – best captures the pattern of the applicant's profile. The final output would consist of the categorisation of that applicant in the corresponding category. If the applicant would be categorised in the 'fraudulent' category, I may treat this application with extra suspicion, and investigate it further before accepting it.

Alternatively, it may well be that I would like the system to discover fraudulent features that I did not think of myself. In that case, I can turn to *unsupervised learning* techniques. Under unsupervised learning, the data based on which the algorithm is trained does not contain pre-assigned labels. Instead, the algorithm analyses the unlabelled dataset by looking for potential patterns therein, and by developing a model based on those patterns. This model can, for instance, result in the clustering of data into categories with similar features, or the detection of certain anomalies within a given dataset. These techniques are hence useful to explore certain rules or clusters that were not previously identified within a dataset.[32]

When looking for welfare fraud, rather than explicitly labelling examples, I could let the algorithm create a model based on past data from citizens, and let it identify clusters of datapoints or anomalies in the dataset that may reveal potential correlations with fraud – as represented in Figure 2.6 on the next page. For instance, the data analysis might reveal that a large majority of past fraudulent applicants refused to provide public authorities with information about their financial situation when

[31] Note that, in practice, sufficient (representative) examples are not always available, which undermines the ability to generalise and can hence stand in the way of the system's proper functioning.

[32] *Deep learning* algorithms are an example thereof. They rely on *artificial neural networks*, modelled loosely after the way in which the neurons in a human brain work. The term 'deep' is a reference to the fact that these networks consist of multiple layers. Such algorithms are prevalently used for the detection of certain features in an unlabelled dataset (though they can also be used with labelled datasets).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005



FIGURE 2.6 Abstraction of an unsupervised data-driven system to predict the propensity of fraud.

asked. A model can then be created based on the identified correlations, which I could subsequently use to assess the fraud propensity of a new applicant by verifying whether one of the identified patterns is also present there. Also in this example, the final output – based on the model's prediction of fraud propensity or not – can be a recommendation to treat a certain application with suspicion. A civil servant may choose to follow this recommendation or not. Alternatively, the recommendation can also be formalised into a decision to automatically take further investigative steps, without leaving any margin of appreciation in this respect.

A beneficial feature of unsupervised learning methods is not only the reduced workload – since the training set need not be labelled in advance – but also the fact that patterns can be identified which human beings did not previously consider or perceive. At the same time, this technique requires a larger dataset to produce successful results, and it can also lead to the identification of patterns that are irrelevant, erroneous or useless, precisely because there are no prior labels that indicate what is of value to solve the problem.[33] In the example above, it may well be that the data analysis reveals that a majority of past fraudulent applicants owned a cat, and a majority of past non-fraudulent applicants owned a dog. It would, however, be a far stretch to automatically treat all new applications from cat-owners as suspicious due to the mere fact that such a correlation was picked up. As I discuss in the next section, beyond the risk of spurious correlations,[34] the lack of labels also means that the data clustering can occur in a manner that deviates from non-discrimination law, by, for instance, taking into account prohibited grounds of discrimination, such as gender or nationality.[35]

[33] Algorithmic systems can also be run on *semi-supervised* learning methods aiming to combine the assets of both techniques. As the name implies, under semi-supervised learning, part of the training data is labelled, while the other part is left unlabelled.

[34] Tyler Vigen, *Spurious Correlations* (Hachette Books 2015).

[35] See in this regard also European Union Agency for Fundamental Rights (ed), *Handbook on European Non-Discrimination Law* (Publications Office of the European Union 2018).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

Finally, I should also mention a third learning technique, called *reinforcement learning*.[36] This approach aims at enabling an algorithmic system (denoted as 'the agent') to learn which decision it should take through a reward mechanism, given uncertainty about the system's environment and given an objective – represented by a value function – that needs to be maximised. The system is instructed to observe the *state* of an environment and to identify the optimal *action* to alter the environment's state in such a way that the receipt of a *reward* signal is maximised. Since the system's aim is to maximise accumulative reward signals rather than an immediate reward, it learns to identify the course of action yielding the best results over the longer term.[37]

Unlike supervised learning, there are no labelled pairs of input–output examples to train the system. Instead, the model considers state–action pairs, and the extent to which they yield a certain reward. This reward feedback – typically in the form of a numerical score – allows the model to be tweaked so as to increase its success in achieving the value. Each time, the system needs to balance the exploitation of knowledge it gained from previous modelling attempts, with the exploration of new action methods that were not yet tried but could yield a (higher) reward.[38] By repeating the process of analysing the result of an action on the environment's state and the corresponding reward, the model can thus come closer to maximise the value that represents the objective over the long term.[39]

An example of a reinforcement learning system is a chatbot deployed by a public authority to provide citizens with answers to some frequently asked questions, as represented in Figure 2.7 on the next page. It can draw on an initial dataset of past questions and answers, but also on a wider collection of information concerning the public authority's functioning that may be of use for citizens. The chatbot can be programmed to seek feedback from the citizen at the end of the conversation to evaluate whether the provided answers satisfied the citizen's inquiries. This feedback can be used to improve the chatbot's answer whenever a similar question is asked by a citizen in the future.

The upside of this technique is the fact that one does not need to undertake the laborious task of labelling input–output pairs, one does not need to have complete information about the environment and one does not need to give continuous feedback about the model's performance, since this is only done for the totality of the outcome (in the example above, the quality of the chatbot's final answer). An important downside, however, is the risk that the algorithm may compute a

[36] See for instance Richard S Sutton and Andrew G Barto, *Reinforcement Learning: An Introduction* (2nd edn, The MIT Press 2018).

[37] See also LP Kaelbling, ML Littman and AW Moore, 'Reinforcement Learning: A Survey' (1996) 4 Journal of Artificial Intelligence Research 237.

[38] Sutton and Barto (n 36) 3.

[39] Csaba Szepesvári, *Algorithms for Reinforcement Learning* (Morgan & Claypool Publishers 2010).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

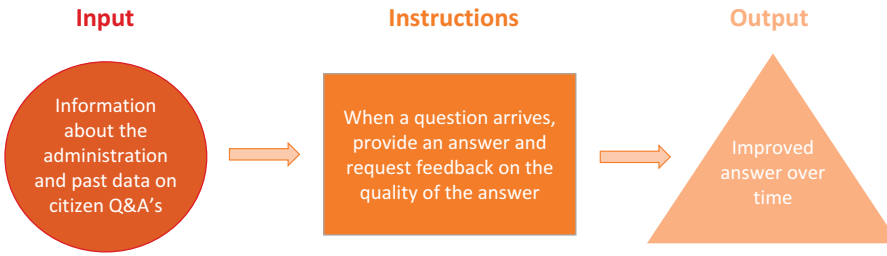


FIGURE 2.7 Abstraction of a data-driven system to automate and improve the answering of citizen questions.

model that merely pursues 'reward hacking', namely, the maximisation of the reward signal, while oblivious to the fact that intermediate actions to achieve that reward may be undesirable or cause undesirable consequences.[40] The way in which the objective – or the value function – is defined, is hence essential.

Finally, let me also make a brief note about *large language models* or LLMs, which are learning-based models that are designed to both 'comprehend' and generate text. These models are trained on a very large dataset of text examples, which in some cases span virtually all text that can be found on the internet.

The model is trained to infer and predict which sequence of words constitutes the most sensible response to a prompt provided by the user, based on the prompt's context. Just as with other learning-based models, LLMs consist of large strings of numbers (weights) which are 'interpreted' and executed by code. When the model is trained, it slightly adjusts these weights based on what it has 'learned', allowing it to improve its predictions of the next suitable word(s) in light of the overall context. It is precisely the capability of large language models to take such context into account that explains the success of their predictions, and hence of their performativity. These models have also been referred to as foundational models, as they are typically trained on a broad range of data rendering them capable of performing various general tasks, which makes them a suitable 'foundation' to build new applications in a more cost-effective manner, without the need to develop a model from scratch. Indeed, the models can be finetuned to serve more particular purposes, for instance by feeding them with domain-specific datasets to adapt them to certain contexts (e.g. in the area of law or medicine) or to adapt them to specific language types (e.g. therapeutical conversations).

Since LLMs are typically further trained on the inputs they receive from their users, they heavily rely on reinforcement learning by human feedback. By virtue of their ability to predict relevant word sequences, LLMs can be used to power chatbots, such as ChatGPT (by OpenAI), Bing Chat (by Microsoft), Bard (by

[40] See also Yinlong Yuan and others, 'A Novel Multi-Step Reinforcement Learning Method for Solving Reward Hacking' (2019) 49 Applied Intelligence 2874.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

Google) and Llama (by Meta). Beyond text, similar approaches also drive applications that enable the generation of images, videos, and audio. Since the interfaces of these generative models are specifically designed to enable their use by non-experts, the wide availability of the systems – often free or at low cost – arguably lead to a major breakthrough in the public's awareness of the technology's transformative impact on society. Jumping on this hype, it should come as no surprise that public authorities too have been exploring how they can use LLM-powered chatbots to facilitate their interactions with citizens to more efficiently deliver their services. In October 2023, the UK government, for instance, announced that it would launch an AI chatbot to help Britons pay their taxes and access their pensions, based on a model that would be powered by OpenAI.[41]

2.1.4 *Comparing Approaches*

Bearing in mind the above description of knowledge- and data-driven algorithms, we can now make a few comparisons. Firstly, when knowledge-driven systems merely rely on the codification of symbols and rules, they can be relatively inflexible, since they do not deal very well with new and uncertain situations that deviate from the coded model. This is, however, different for knowledge-driven systems that also draw on probabilistic and statistical methods, as they can extrapolate from the known information to make predictions that can contribute to a solution, despite the potential uncertainty or knowledge gaps. Data-driven systems are typically more flexible, as they enable the inference of models that can evolve over time in light of new data that is fed into the system. Yet the models established through data-driven systems or through statistical methods can subsequently be used for reasoning, thus underscoring the fact that systems need not only focus on one approach or another.[42]

The fact that data-driven systems do not require a coded description of what the data represent by translating knowledge into symbols and symbolic rules has been an important contributor to the successes they booked, especially in domains where the representation of codified knowledge is more difficult. Indeed, over the last decade, learning-based approaches have, for instance, been responsible for significant advances in computer vision (or methods to acquire and analyse images and videos), computer hearing (or methods to acquire and analyse sound data like music or

[41] See James Titcomb and Matthew Field, 'Sunak to Launch AI Chatbot for Britons to Pay Taxes and Access Pensions' The Telegraph (28 October 2023) <www.telegraph.co.uk/business/2023/10/28/rishi-sunak-launch-ai-chatbot-pay-taxes-access-pensions/>. This announcement was seen by some as an answer to the huge pressure on the call centers of the UK's Revenue and Customs departments, as citizens were faced with very long waiting times.

[42] It has also been argued that there exists "*a conceptual continuity between algebraically rich inference systems, such as logical or probabilistic inference, and simple manipulations, such as the mere concatenation of trainable learning systems*"; see Léon Bottou, 'From Machine Learning to Machine Reasoning' (2014) 94 Machine Learning 133.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

speech), and computer touch (or methods to acquire and analyse tactile information). Furthermore, especially when using unsupervised learning techniques, learning algorithms can help identify patterns and categorisations that humans did not previously perceive or consider, which can lead to useful new insights.[43]

However, the use of data-driven systems also has drawbacks. First, the inherently probabilistic nature of their outcomes means that there will always be a margin of error. In some situations, such algorithms can achieve a very high level of accuracy in their predictions, yet it may be that the dataset on which the system was trained did not present an accurate representation of real-life situations, thus reducing the system's accuracy and utility outside the lab.[44] The output may also be false or entirely estranged from reality, which is especially problematic in the context of generative models that produce text or images which come across as highly realistic. It has been noted that data-driven algorithms can be "*statistically impressive*", yet they remain "*individually unreliable*".[45] That said, knowledge-driven algorithms are not free from error either: if the knowledge that the system should reason upon is not correctly represented by the human programmer, the outcome will be equally flawed.

Second, certain data-driven algorithms – especially those relying on deep learning methods[46] – suffer from a lack of transparency and interpretability, which has been referred to as the so-called 'black box' problem. Such systems "*internalise data in ways that are not easily audited or understood by humans*", for instance due to the "*complexity of the algorithm's structure*", or because they rely "*on geometric relationships that humans cannot visualise*".[47] This means that human developers and deployers of such algorithms cannot always explain how a certain outcome was achieved, and why it was achieved.[48] This can be contrasted with the interpretability of knowledge-driven systems which, in light of their typically predetermined symbolic rules, provide better insight into how and why a certain output is provided.

Third, while data-driven algorithms can identify patterns and correlations between data points that humans might not have recognised, they are in principle unable to establish *causal* relationships between those points, which may lead

[43] Yavar Bathaee, 'The Artificial Intelligence Black Box and the Failure of Intent and Causation' (2018) 31 Harvard Journal of Law & Technology 890, 891.

[44] See, e.g., David Thogmartin, 'Ensuring Reliable AI in Real World Situations' (Deloitte 2022), <www2.deloitte.com/content/dam/Deloitte/de/Documents/Innovation/DELO-8505%20Trustworthy%20AI%20Robust%20and%20Reliable_KS6.pdf>.

[45] Launchbury (n 22).

[46] See also Yoshua Bengio, Aaron Courville and Ian Goodfellow, *Deep Learning* (The MIT Press 2016).

[47] Bathaee (n 43) 901.

[48] Cynthia Rudin and Joanna Radin, 'Why Are We Using Black Box Models in AI When We Do Not Need To? A Lesson from an Explainable AI Competition' (2019) 1 Harvard Data Science Review 2 <https://hdsr.mitpress.mit.edu/pub/f9kuryi8/release/6>.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

to spurious correlations, a risk inherent to virtually all statistical models.[49] Furthermore, the lack of symbolically represented knowledge also renders data-driven systems unable to contextualise the information they process. This means that certain logical facts that are evident for human beings ('common sense'), may not be picked up by the system. In addition, the system can also 'learn' in ways that the developers did not predict or intend, which may lead to unforeseen adverse consequences.[50]

Finally, data-driven algorithms have a higher vulnerability to adversarial attacks. By inserting deceptive input into the system – which can go unnoticed by the human eye – one can deliberately confuse the model created by the system and thereby compromise its output.[51] While certain defence techniques exist against those attacks, such as data manipulation constraints, the risk nevertheless remains that the system's vulnerabilities are exploited, and that the results computed by the system become unreliable.[52] The high complexity of the models and the decreased transparency mentioned above, can also render the detection of those attacks more difficult.

In sum, knowledge-driven and data-driven techniques each have a set of capabilities that enable the automation of (intelligent) tasks. When seeking to solve a problem through an algorithmic system, the choice of technique will hence be determined by the type of problem, the existing knowledge and expertise about the problem, the availability of (representative) data, the characteristics one wishes to prioritise when seeking a solution to execute the task and – most importantly – the budget. One does not necessarily need to choose between one technique or another, as various methods exist to remedy the constraints of each, by combining reasoning techniques with learning-based approaches. Data-driven models can be used to help predict how the sought-after data could look like, after which the outcome of that prediction can serve as a basis for further reasoning and analysis through symbolic rules.[53] This demonstrates that the distinction between knowledge- and data-driven techniques is not always clear-cut. In fact, the most recent research trend focuses on hybrid systems, whereby knowledge- and data-driven techniques are combined.[54]

[49] See also Vigen (n 34). With complex data-driven systems, the so-called 'curse of dimensionality' – first coined by Richard Bellman – can make this risk more prominent. See Richard Bellman, *Dynamic Programming* (Princeton University Press 1957).

[50] See, e.g., Gina Neff and Peter Nagy, 'Talking to Bots: Symbiotic Agency and the Case of Tay' [2016] International Journal of Communication 4915.

[51] See, e.g., Marco Barreno and others, 'Can Machine Learning Be Secure?', in *Proceedings of the 2006 ACM Symposium on Information, Computer and Communications Security* (Association for Computing Machinery 2006), <www.dl.acm.org/doi/10.1145/1128817.1128824>.

[52] See, e.g., Luis Muñoz-González and Emil C. Lupu, 'The Security of Machine Learning Systems', in Leslie F. Sikos (ed), *AI in Cybersecurity*, Intelligent Systems Reference Library (Springer International Publishing, 2019), 47 and following.

[53] Duan, Tang and Zhou (n 13) 2.

[54] De Raedt (n 9) 300. See also Gary Marcus and Ernest Davis, *Rebooting AI: Building Artificial Intelligence We Can Trust* (Penguin Random House 2020).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

2.1.5 AI Systems

Before concluding this technical chapter, there is one more term I wish to expand on in the context of algorithmic regulation, namely artificial intelligence (AI) – the importance of which has been described as "*more profound than electricity or fire*".[55] World leaders have stated that "*whoever becomes the leader in this sphere will become the ruler of the world*",[56] and the European Parliament's Special Committee on Artificial Intelligence in a Digital Age even claimed that it can be "*thought of as the fifth element after air, earth, water and fire*".[57] Aside from the hype that these hyperbolic statements represent, what does this technology consist of, and how does it relate to the algorithmic systems I described in the previous sections?

As noted above, AI can be understood as an umbrella term for a range of 'intelligent' technological applications. The pursuit of intelligent machines is almost as ancient as human history,[58] yet the first time the term AI was used dates back to the now-famous 1956 Dartmouth Workshop, considered to mark the birth of 'AI' as a field of research.[59] While all AI systems are composed of algorithms, not all algorithmic systems necessarily fall under the AI umbrella. As the term indicates, to be considered 'artificially intelligent', a system would need to perform a task that is typically associated with *intelligence*, rather than merely constituting the simple *automation* of a task. What is considered an 'intelligent' task may vary significantly based on the person to whom the question is asked, and the discipline they are trained in. A philosopher might define the term quite differently from a schoolteacher or a lawyer. More importantly, AI experts, too, tend to define this technology in different ways. In their seminal AI handbook, Russell and Norvig, for instance, identify definitions that focus on the ability of a machine to think or the ability act

[55] Uttered by Sundar Pichai, CEO of Google, in 2018. See Catherine Clifford, 'Google CEO: A.I. Is More Important than Fire or Electricity' *CNBC* (1 February 2018) <www.cnbc.com/2018/02/01/google-ceo-sundar-pichai-ai-is-more-important-than-fire-electricity.html>.

[56] Uttered by Vladimir Putin, President of the Russian Federation, in 2017. See James Vincent, 'Putin Says the Nation that Leads in AI "Will Be the Ruler of the World" – The Verge' *The Verge* (4 September 2017) <www.theverge.com/2017/9/4/16251226/russia-ai-putin-rule-the-world>.

[57] Stated by the European Parliament's Special Committee on Artificial Intelligence in a Digital Age (AIDA), 'Draft Report on Artificial Intelligence in a Digital Age' (European Parliament 2021) (2020/2266(INI)) 9.

[58] See also Pamela McCorduck, *Machines Who Think* (2nd edn, A K Peters, Ltd 2004).

[59] The workshop proposal described the project as aiming:

> to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.

See John McCarthy and others, 'A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence', 31 August 1955, *AI Magazine*, 27(4), 2006, 12.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

humanly, versus definitions focusing on their ability to think or act *rationally*.[60] The latter definition is more prominently used among AI researchers nowadays, as it can more easily be formalised by virtue of an AI system's optimisation function, through which the system can be tasked to select those actions that are expected to maximise its utility.[61]

In essence, one could consider algorithmic systems on a spectrum. One extreme contains algorithms with only a very minimal set of instructions, too basic or simplistic to be considered 'intelligent'. The other extreme contains algorithms that are highly complex and sophisticated, and would be considered 'intelligent' by virtually all researchers. Yet, when it comes to algorithmic systems situated in the middle part of the spectrum, researchers do not always agree on whether the system would be considered 'intelligent' or not, and the perception thereof also tends to change over time.[62] There is no fixed rule as regards *how many* or *which type of* instructions an algorithmic system must contain in order to be called an 'AI' system.

Since the term AI is currently widely used, and since the European Union's regulatory initiative is titled 'the AI Act', let me discuss some AI definitions put forward in the context of EU policy, and examine how these compare with the description above. In this regard, it must be borne in mind that the conceptualisation of 'AI' has not only an academic or scientific relevance, but also a political one. The scope of the regulatory requirements that legislators may impose on 'AI' systems will after all depend on how they define 'AI'. Furthermore, the definition of AI also has a societal relevance, as it contributes to framing the broader narrative around the impact of AI in society.[63] Consider, for instance, the way in which the European Parliament's Delvaux report frames AI, evoking the context of "*Mary Shelley's* Frankenstein's Monster" and "*the story of Prague's Golem*",[64] and cautioning that "*humankind stands on the threshold of an era when ever more sophisticated robots, bots, androids and other manifestations of artificial intelligence ('AI') seem to be poised to unleash a new industrial revolution*".[65] The Report further states that "*ultimately there is a possibility that in the long-term, AI could surpass human intellectual capacity*",[66] and makes the much-commented suggestion to create "*a specific legal status for robots in the long run, so that at least the most sophisticated*

[60] Russell and Norvig (n 1) 2.

[61] De Raedt (n 9) 286.

[62] This phenomenon relates to the so-called *AI effect*, whereby a type of application that was originally considered as 'intelligent' becomes so commonplace that it is no longer considered sufficiently intelligent to still warrant the designation 'AI'. See also McCorduck (n 58).

[63] See in this regard Bilel Benbouzid, Yannick Meneceur and Nathalie Alisa Smuha, 'Quatre nuances de régulation de l'intelligence artificielle: Une cartographie des conflits de définition' (2022) 232–33 Réseaux 29.

[64] Mady Delvaux, 'Report with Recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL))' (European Parliament 2017) A8–0005/2017 §A.

[65] ibid §B.

[66] ibid §Q.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

autonomous robots could be established as having the status of electronic persons responsible for making good any damage they may cause"[67] – a suggestion that was rejected by the European Commission. However, the Commission did follow up on the report's proposal to establish a definition of AI and to draft ethics guidelines to guide use of the technology.

To this end, it set up a High-Level Expert Group on AI in June 2018 composed of fifty-two experts from various domains, tasked with the drafting of Ethics Guidelines for AI as well as Policy and Investment Recommendations.[68] As part of its tasks, the Expert Group set up a working group to establish a definition of AI – an exercise that took just as long as agreeing on AI ethics principles – which resulted in the following formulation:

> Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans[69] that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions.[70]

The following elements can already be distilled from this definition. First, as will be stressed in the next section, the Expert Group emphasised that systems are designed by human beings. This means that they do not 'overcome' us passively but that they are an active creation of persons who are therefore also responsible for their consequences.[71] Second, the definition takes into account that there is no 'single' AI system.[72] It indicates that AI systems can 'use symbolic rules', as in the case of knowledge-driven systems, thus not limiting the term AI only to data-driven systems. Accordingly, AI is considered an overarching concept covering a range of

[67] ibid §59(f).

[68] European Commission, 'High-Level Expert Group on Artificial Intelligence' (*Shaping Europe's Digital Future – European Commission*) <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>. See also Nathalie A Smuha, 'The EU Approach to Ethics Guidelines for Trustworthy Artificial Intelligence' (2019) 20 Computer Law Review International 97.

[69] The footnote in the definition states that: "*Humans design AI systems directly, but they may also use AI techniques to optimize their design*." It was added by the Expert Group to reflect the fact that AI systems can sometimes also be programmed to develop new algorithms.

[70] High-Level Expert Group on AI, 'A Definition of AI: Main Capabilities and Scientific Disciplines' (n 12).

[71] See also Nathalie A Smuha, 'Laten we intelligenter zijn wanneer we het over Artificiële Intelligentie hebben' (*Knack Data News*, 11 March 2020) <https://datanews.knack.be/ict/nieuws/laten-we-intelligenter-zijn-wanneer-we-het-over-artificiele-intelligentie-hebben/article-opinion-1574905.html>.

[72] Nathalie A Smuha, 'From a "Race to AI" to a "Race to AI Regulation": Regulatory Competition for Artificial Intelligence' (2021) 13 Law, Innovation and Technology 57.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

technologies that have certain properties in common, namely their ability to reason on or learn from the data provided to them, and to "*act in the physical or digital dimension*" based on such data. As regards data-driven systems, the definition indicates that AI systems can 'adapt their behaviour' over time[73] based on what they 'learn',[74] which is in turn dependent on how they have been programmed to do so by a human being. Lastly, the definition also mentions hardware, thereby indicating that knowledge- and data-driven algorithms can also be incorporated into hardware to design machines that can carry out physical tasks, which is the aim of the field of robotics.

While the European Commission endorsed the Expert Group's *Ethics Guidelines*, it did not retain the above definition in its subsequent proposal for a new AI regulation, published two years later in April 2021. Instead, it suggested a somewhat different definition of AI, encompassing "*software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with*".[75] Like the High-Level Expert Group, the Commission's definition stressed that the objectives of AI are 'human-defined', also indicating a difference in style with the Delvaux report.[76] At the same time, the Commission's original proposal suggested to exhaustively list in annex I the specific techniques and approaches that should fall under the AI Act's scope, so as to enhance legal certainty. These concerned:

(a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;
(b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;
(c) Statistical approaches, Bayesian estimation, search and optimisation methods.

[73] They can do this in an autonomous manner, yet only once they are programmed to do so by a human being.

[74] Besides formal definitions provided by governmental organisations, consider also the definition (s) provided by Russell and Norvig in their influential Handbook on AI (n 1).

[75] European Commission, Proposal for a Regulation of the European Parliament and the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts 2021 [COM(2021) 206 final].

[76] This point can be criticised, since AI systems can also be programmed to set objectives autonomously (in light of certain restraints and/or elements of information provided to them). Some hence argued that these systems might fall outside the scope of the Commission's AI definition. However, one can counter-argue that, even for those systems, there is still a programming phase during which a human being sets out the system's objectives at a more abstract level, to be further concretised by the system later on, and hence that they do fall under the definition's scope.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

Accordingly, the Commission originally opted for a broad general definition of AI, while simultaneously meticulously specifying the underlying techniques that would actually be caught. This implies a focus on the technical approach behind the systems and their specific research discipline classification rather than their properties or effects. The idea behind listing these techniques under an annex was to ensure flexibility in the future, in case new approaches come along that do not neatly fit under the existing ones – even if list-based definitional approaches inevitably risk being both over- and under-inclusive. Stakeholders have therefore proactively been seeking to influence the regulation's definition of AI prior to its adoption.[77]

A number of actors have argued that the Commission's original definition was too broad.[78] Some argued that only data-driven approaches – and more specifically machine-learning applications – should be covered by the proposed AI Act,[79] while others proposed that at least the approaches listed under point (c) should be excluded.[80] In addition, it has been argued that not only AI systems but also other technologies that can be used for similar tasks should be included, to ensure a 'level playing field' between different technologies.[81] Evidently, a narrower definition also

[77] During the trilogue negotiations about the AI Act, both the Council of the European Union and the European Parliament have been discussing whether the approach of listing AI techniques in annex I of the AI Act should be maintained or whether this list should be removed in favour of a (broader) definition in the main body of the regulation. See, e.g., Luca Bertuzzi, 'Artificial Intelligence Definition, Governance on MEPs' Menu' (*Euractiv*, 8 November 2022) <www.euractiv.com/section/digital/news/artificial-intelligence-definition-governance-on-meps-menu/>. Ultimately, they decided to opt for the latter option. I will also return to this point *infra* when discussing the AI Act in more detail, in Section 5.4.

[78] Consider, for instance, the feedback by Digital Europe (a leading trade association representing digitally transforming industries in Europe) stating that: "*The definition of 'artificial intelligence' set in the AI Act is too wide. The proposed definition encompasses many software technology applications, even when they pose no major concerns around data, opaqueness, safety and reliability. It notably includes within AI techniques 'logic-based and statistical approaches, Bayesian estimation, search and optimisation methods'*". 'DIGITALEUROPE's Initial Findings on the Proposed AI Act' (DIGITALEUROPE 2021) <www.digitaleurope.org/resources/digitaleuropes-initial-findings-on-the-proposed-ai-act/>.

[79] See, e.g., the position of the Slovenian presidency as reported by AlgorithmWatch, 'European Council and Commission in Agreement to Narrow the Scope of the AI Act' (*AlgorithmWatch*, 23 November 2021) <https://algorithmwatch.org/en/eu-narrow-scope-of-ai-act/>.

[80] Huawei, 'Huawei Response on the European Commission's Proposal for a Regulation of the European Parliament and of the Council Laying Down the Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts' (*European Commission – Have your say*, August 2021) <https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Artificial-intelligence-ethical-and-legal-requirements/F2665442_en>.

[81] 'Feedback to the European Commission's Regulation Proposal on the Artificial Intelligence Act' (DigitalPoland 2021) <https://digitalpoland.org/en/blog/2021/08/feedback-to-the-european-commission-s-regulation-proposal-on-the-artificial-intelligence-act>. As I also noted elsewhere, "*the adoption of a strict rule that imposes burdensome obligations on AI deployers to minimise certain risks, would not cover a manifestation of the same risk by other types of technology, and*

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

narrows down the scope of systems that are subjected to regulatory requirements and scrutiny. It is therefore no surprise that a regulatory battleground emerged around this definition, which continuously shifted in shape during the course of the negotiations on the AI Act between the European Parliament and Council.[82]

The two latter institutions decided to eliminate the Commission's annex with AI approaches and techniques, and to solely include a (slightly altered) definition in the regulation's main text. In the final version of the AI Act, which entered into force on 1 August 2024, Article 3(1) reads: "'*AI system' means a machine-based system designed to operate with varying levels of autonomy, that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments*". This definition has the advantage of being closer to the one proposed by the OECD[83] and is therefore seemingly more reflective of the slowly emerging global definitional consensus. It also lists the 'actions' that AI systems can be programmed to undertake based on the input they are provided, such as making recommendations and predictions, or taking decisions, and it explicitly includes the concept of 'autonomy'.

I will come back to the AI Act's definition of AI in Section 5.4 where I analyse the regulation in more detail. At this stage, suffice it to say that this definition may come across as rather broad, but that the actual regulatory scope of the AI Act is significantly narrowed down by the regulation's Recital 12, which distinguishes it from 'simpler traditional software systems or programming approaches' and from 'systems that are based on the rules defined solely by natural persons to automatically execute operations', neither of which should fall under the AI Act's scope. Furthermore, its scope is also narrowed down by the specific AI-applications listed under its substantive provisions, since not all systems that fall under the AI definition are also subjected to new regulatory requirements.

Importantly, this book is not interested in the definitional conundrums of AI, but rather in the adverse effects that the use of algorithmic regulation – by virtue of the scaled and automated decision-making it enables – may have on the rule of law and its principles, regardless of the underlying technique used for this purpose. Data-driven approaches are often considered to be more problematic than knowledge-driven approaches, given their reliance on potentially biased or erroneous datasets,

might merely push AI deployers towards the use of other tools to achieve the same problematic end", in Smuha, 'From a "Race to AI" to a "Race to AI Regulation"' (n 72) 64.

[82] For a discussion of this battleground, and its various definitional camps, see, e.g., Benbouzid, Meneceur and Smuha (n 63).

[83] In its policy documents, the OECD defines AI systems as follows: "*An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy.*" See OECD, 'Recommendation of the Council on Artificial Intelligence' (2019) OECD/LEGAL/0449 <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

their dynamic and evolving nature, their greater opacity, and their relative novelty – which also means less time has passed to develop appropriate mitigation measures for their risks. I could hence have opted to limit my investigation to the way in which public authorities rely on data-driven systems only, and leave aside the use of knowledge-driven systems. However, as noted above – the distinction between both is not firm, since systems can rely on both approaches, and on intermediary probabilistic approaches. Furthermore, as already hinted at, and as I will discuss further below, knowledge-driven approaches, too, can pose significant risks, even those that are only based on very crude rules. While they typically benefit from a higher degree of intelligibility and a greater level of control and predictability, their use can still cause substantial societal harm. Moreover, public authorities tend to rely both on knowledge- and data-driven approaches to automate administrative acts.

Consequently, in this book I will focus on algorithmic systems in general. Where the distinction between different approaches is relevant (for instance in light of a specific risk or drawback arising therefrom) I will mention whether the system is primarily knowledge- or data-driven. Rather than trying to set boundaries around which techniques are sufficiently 'intelligent' to fall under the AI umbrella, I will limit my use of the term AI altogether by reserving it only for my discussion of the AI Act in Section 5.4, given the Act's heavy reliance on this term. In the remainder of the book, I will instead examine the deployment of algorithmic regulation by public authorities to inform or take administrative acts, regardless of the underlying approach or technique they rely on. Collectively, I will refer to the use of such systems as *algorithmic regulation*. With these technical aspects in mind, let me now consider the societal aspects of those systems.

2.2 SOCIETAL ASPECTS

In the section above, I described the building blocks of algorithmic systems, and focused on their technical aspects in isolation. In reality, algorithmic systems do not operate in a vacuum and cannot be reduced to their technical characteristics. As noted in the Introduction, they can regulate their environments, govern social relationships and steer human behaviour – and are often explicitly designed to do so.[84] This is because they are always already embedded in a broader infrastructure, which is composed not just of the system's software, but also of the wider network of individuals, organisations, cultures, languages, laws and customs in which they are designed, developed and used. In other words, algorithmic systems are 'socio-

[84] In this regard, it can also be pointed out that algorithmic systems can serve as a tool to 'nudge' individuals. Karen Yeung speaks of 'hypernudge', since data-driven systems particularly enable the channelling of "*user choices in directions preferred by the choice architect through processes that are subtle, unobtrusive, yet extraordinarily powerful*", "*due to their networked, continuously updated, dynamic and pervasive nature*". See Karen Yeung, '"Hypernudge": Big Data as a Mode of Regulation by Design' (2017) 20 Information, Communication & Society 118.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

technical' systems, as they have an influence on, and are influenced by, their societal environment.[85] When public authorities implement algorithmic regulation, they do so with a certain purpose, which involves the concrete effects it can have beyond the algorithmic system. For instance, they count on the system to enhance the efficiency of their organisational processes, to improve the accuracy of their decisions, or to help predict or analyse societal phenomena based on which they can optimise their actions.

In what follows, I therefore examine this mutual influencing process further by providing a characterisation of algorithmic systems that emphasises their socio-technical nature (Section 2.2.1). Following this characterisation, I discuss various consequences associated with reliance on algorithmic systems, including the risk of error (Section 2.2.2), the risk of bias (Section 2.2.3), the impact on human agency (Section 2.2.4), the dependency of algorithmic systems on data and proxies (Section 2.2.5) and the potential opacity surrounding their use (Section 2.2.6). Keeping these elements in mind will enable a better understanding, in subsequent chapters, of the impact of algorithmic regulation on societal interests, including on the rule of law.

2.2.1 *Algorithmic Systems as Socio-technical Infrastructure*

The intuition that the design, development and use of a technology is not a mere technical matter, but also a normative and political matter, is far from new. Already in 1964, Lewis Mumford juxtaposed *democratic technics* with *authoritarian technics*. He argued that certain technologies, and particularly systems used for automation, lend themselves more easily to authoritarian tendencies in society given their scale, their centralised and controlling nature, and their focus on efficiency and speed instead of human individuality.[86] Over the past decades, scholarship on the relationship between technology and society has been burgeoning, producing theoretical frameworks that draw on a variety of disciplines.[87]

[85] High-Level Expert Group on AI, 'Ethics Guidelines for Trustworthy AI' (European Commission 2019) <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>; Gordon Baxter and Ian Sommerville, 'Socio-Technical Systems: From Design Methods to Systems Engineering' (2011) 23 Interacting with Computers 4; Andreas Theodorou and Virginia Dignum, 'Towards Ethical and Socio-Legal Governance in AI' (2020) 2 Nature Machine Intelligence 10; Shakir Mohamed, Marie-Therese Png and William Isaac, 'Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence' [2020] 33 Philosophy & Technology, 659; Pekka Ala-Pietilä and Nathalie A Smuha, 'A Framework for Global Cooperation on Artificial Intelligence and Its Governance' in Bertrand Braunschweig and Malik Ghallab (eds), *Reflections on Artificial Intelligence for Humanity* (Springer International Publishing 2021); Gry Hasselbalch, *Data Ethics of Power* (Edward Elgar Publishing 2021).

[86] Lewis Mumford, 'Authoritarian and Democratic Technics' (1964) 5 Technology and Culture 1.

[87] See also Hanseth and Monteiro, who rely on actor-network theory to examine the societal implications of standardisation processes for information infrastructures, and mention other relevant theoretical frameworks, from structuration theory and phenomenology to hermeneutics and Habermas' theory of communicative action: Ole Hanseth and Eric Monteiro,

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

A chief insight of this scholarship, albeit still controversial for some,[88] concerns the fact that technology is not neutral[89] but that it embodies normative and political choices on the part of its designers, developers and users, which shape the technology's affordances.[90] These choices can not only be judged by the extent to which the technology is effective in reaching its goals, but also by how it can alter or entrench power structures and impact human relationships. By virtue of being part of a society, technical systems are inherently interwoven therein and influenced thereby. Yet, once created, they exert an influence on society in turn. This influence can have various modalities. It can be deliberate or accidental, known or obscure, significant or minor, positive or negative (and is often mixed), and it can manifest itself in the short or in the longer term. To expose this influence, both the social forces behind the technology and the technology's particular characteristics should be assessed,[91] as each of these elements can contribute thereto.

A well-known example of a deliberate way in which artifacts have been designed to exert a particular influence on society, concerns Robert Moses' low-hanging bridges on Long Island.[92] Their height prevented public buses – and the people of colour or lower-income passengers they typically carried – from accessing public beaches that were primarily intended for the recreation of white middle- and upper-class citizens, who could pass the low-hanging bridges with their cars.[93] Another example, which is not deliberate yet nevertheless highly worrisome, concerns the way in which cars are designed by using crash-test dummies that reflect the average height and size of men, not taking into account the anatomy of women. This renders it 47 per cent more likely for a woman that is involved in a car crash to be

'Inscribing Behaviour in Information Infrastructure Standards' (1997) 7 Accounting, Management and Information Technologies 183, 185.

[88] Some authors still adhere to the notion that technology – by virtue of its inanimate nature – is inherently neutral, and that one can only make normative or political judgments about the society in which the technology is used or about the person that uses it, rather than about the technology itself. See, e.g., Joseph C Pitt, '"Guns Don't Kill, People Kill"; Values in and/or Around Technologies' in Peter Kroes and Peter-Paul Verbeek (eds), *The Moral Status of Technical Artefacts*, vol 17 (Springer Netherlands 2014). See also Boaz Miller, 'Is Technology Value-Neutral?' (2021) 46 Science, Technology, & Human Values 53.

[89] Melvin Kranzberg, 'Technology and History: "Kranzberg's Laws"' (1995) 15 Bulletin of Science, Technology & Society 5.

[90] In this context, the affordance of a technology can be described as the quality that defines its possible uses or that makes clear how it can or should be used. See in this regard also Mireille Hildebrandt, *Smart Technologies and the End(s) of Law* (Edward Elgar Publishing 2015); Julie E Cohen, 'Affording Fundamental Rights: A Provocation Inspired by Mireille Hildebrandt' (2017) 4 Critical Analysis of Law 76.

[91] Langdon Winner, 'Do Artifacts Have Politics?' (1980) 109 Daedalus 121, 123.

[92] However, see also the critical account of Bernward Joerges, 'Do Politics Have Artefacts?' (1999) 29 Social Studies of Science 411.

[93] Winner (n 91) 124. See also Lawrence Lessig, 'The Law of the Horse: What Cyberlaw Might Teach' (1999) 113 Harvard Law Review 501, 543.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

seriously injured, and 71 per cent more likely for her to be moderately injured, than for a man.[94]

As I will elaborate in the next sections, this holds true not only for physical artefacts, but also for digital technologies – including algorithmic systems. An example of how the combined use of physical and digital technology can deliberately co-contribute to a societal impact, is found in Flanders. Since February 2021, Flemish municipalities have the competence to penalise excessive speed themselves by means of a fine, so as to increase the enforcement of speed limits. However, for smaller towns, the costs to set up automatised speed traps to measure cars' driving speed are prohibitively high. Enter private companies. Through a contractual arrangement with several towns, a number of companies agreed to bear the installation costs of the automatised speed traps in exchange for a lumpsum of €24 for every fine the town issued. Of course, such business model implies that the more fines are issued, the more profit is made.[95] The companies therefore contractually stipulated with the towns that physical infrastructures meant to discourage drivers from driving fast, such as speed bumps and chicanes, as well as priority arrangements for cyclists, are to be removed and not to be reintroduced for at least six years.[96] After media reports and public outrage, the Home Office Minister annulled the contracts, citing the adverse impact on 'the public interest' and the importance of the town's autonomy as a justification for the annulment.[97]

This example illustrates in a disconcertingly clear way how physical and digital infrastructure can be used to optimise one goal (e.g. private profits) over another (e.g. road safety in the public interest). In other words: the design, development, use, interplay, as well as the governance and ownership of infrastructure is not merely of technical but also of societal relevance. Moreover, it shows that the societal embeddedness of technology also translates into physical power by those who design, develop and deploy the system, over those who are subjected to the system. In sum, technological normativity shapes possibilities for action,[98] both for those who actively make use of the system, and for those who are subjected to the system.

I described above how algorithms essentially consist of a given input, instructions and output. Moreover, they depend on the physical infrastructure that enables their

[94] Caroline Criado Perez, 'The Deadly Truth about a World Built for Men – From Stab Vests to Car Crashes' *The Guardian* (23 February 2019) <www.theguardian.com/lifeandstyle/2019/feb/23/truth-world-built-for-men-car-crashes>. See also Caroline Criado Perez, *Invisible Women: Exposing Data Bias in a World Designed for Men* (Vintage Publishing 2020).

[95] Yannick Verberckmoes, 'Mogen verkeersboetes een verdienmodel zijn? Volgens experts is een grens overschreden' *De Morgen* (12 December 2021) <www.demorgen.be/gs-b72e332e>.

[96] Matthias Verbergt, 'Privébedrijf achter trajectcontroles doet verkeersdrempels verdwijnen' (*De Standaard*) (11 December 2021) <www.standaard.be/cnt/dmf20211210_97898448>.

[97] 'Bart Somers vernietigt systeem trajectcontrole Bonheiden' *De Standaard* (14 January 2022) <www.standaard.be/cnt/dmf20220114_94712942>.

[98] 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>.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

functioning and, in particular as regards data-driven systems, the data they are fed. All of these individual elements, including the designers and owners behind it, therefore influence the system's impact on those subjected to its use. Constitutive of these elements are human choices, such as the choice of what to optimise the system for, what data to use and how to categorise and label it, what elements to dismiss, how to translate knowledge to code, how to use and present the system's output, and how to communicate the system's capabilities. These choices also pertain to which programming language to use – bearing in mind that the affordance of language we code in also shapes how we think, for instance in terms of coding solutions – and at the meta-level, how to design and develop programming languages and the Application Programming Interfaces (APIs) through which they can be used.[99]

The implicit and explicit choices regarding all of these elements contribute to the system's influence on society, rendering the human beings that make these choices – and that are positively or adversely affected thereby – part of the system's broader socio-technical infrastructure.[100] These choices are often invisible as they are not always rendered explicit, and tend to correspond to what Susan Star calls 'the study of boring things'.[101] She gives the example of the physical infrastructure of our sewer system which may undoubtedly sound like a 'boring thing' to study and analyse, yet nevertheless constitutes an important element of modern cities' utilities and reveals a lot of information about how we live and how society is organised. We barely notice this infrastructure's existence, unless or until it breaks down and no longer functions as it should.[102] Star suggests that perhaps, "*if we stopped thinking of computers as information highways and began to think of them more modestly as symbolic sewers*", the realm of the functional and value-laden choices behind them might become more visible.[103]

What Star notes about IT systems in general also applies to algorithmic systems. In fact, the term *system* already implies this infrastructural nature, typically referring

[99] ibid 19.

[100] Furthermore, the three elements mentioned above – input, instructions and output – each of which are essentially abstractable to patterns of zeroes and ones, also hinge on a physical infrastructure that enables these patterns to be processed in the first place, including the human labour that goes into creating them: from tangible components such as processors and batteries that make up a computer's hardware, to data storage centres that consume a significant amount of energy, and connections to private and public networks to exchange information (for instance through optic fibre cables running across oceans), all of which are likewise subjected to human design, development and use choices. For a detailed overview, see Kate Crawford and Vladan Joler, 'Anatomy of an AI System: The Amazon Echo as an Anatomical Map of Human Labor, Data and Planetary Resources' (*Anatomy of an AI System*, 2018) <www.anatomyof.ai>.

[101] Susan Leigh Star, 'The Ethnography of Infrastructure' (1999) 43 American Behavioral Scientist 377, 377.

[102] Hasselbalch (n 85) 18.

[103] Star (n 101) 379.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

to a set of things that work interdependently as parts of a broader mechanism or an interconnected network that forms a united whole.[104] This notion hence emphasises that algorithmic systems are not isolated entities, but embedded in an overarching *system*, and part of a broader socio-technical infrastructure.[105]

To conclude, in order to understand the societal impact of algorithmic systems beyond the technical elements described above, it is useful to render visible at least part of the human choices (and hence the power structures) that underlie their design, development and use. Considered holistically, algorithmic systems are more than just the sum of their individual components, as the interplay of these components engenders societal effects.[106] When analysing how algorithmic regulation is used in the public sector, I will therefore keep this broader dimension in mind. In the remainder of this section, I discuss some of the consequences of the societal embeddedness of algorithmic systems.

2.2.2 *Risk of Human Error*

When people apologise for a blunder by stating that they are 'only human', this seems to imply that non-human machines could be programmed without this 'human' tendency of making mistakes. Indeed, one of the ideas that drive the uptake of algorithmic systems in society is the fact that human errors and mistakes could be avoided or minimised, as a machine could carry out tasks without, for instance, becoming tired, drunk, ill or simply overlooking something. However, since the developers of algorithmic systems are 'only human' too, this is wishful thinking. Mistakes can occur throughout various instances of an algorithmic system's lifecycle, for instance when gathering or selecting input, labelling or categorising data, translating knowledge and rules to code, designing the system's optimisation function, interpreting the system's output, or applying it in practice. Some mistakes will be easily noticeable and potentially also easily correctible. Others might be more difficult to spot and, especially if the system is used on a large scale, may cause significant damage before the mistake is rectified.

Consider the example of an algorithmic system to help allocate medical benefits, deployed by the Idaho Department of Health and Welfare. The system was set up to

[104] See the definition of 'system' in the Merriam-Webster online dictionary, <www.merriam-webster.com/dictionary/system>.

[105] There is a related notion that is worth pointing out at this stage as well, namely the *systemic* effects – positive or negative – that algorithmic systems can give rise to. Precisely because algorithmic systems are part of a broader infrastructure that allows them to be used and relied upon in a systemic way, their impact is typically not limited to an isolated node of society, but can affect all of the elements that the system is interconnected with – and vice-versa – be it tangible physical items, individuals, organisations or intangible practices and norms.

[106] 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) 215.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

administer the state's Development Disabilities Waiver programme, which establishes a personalised annual budget for individuals in need to contribute to their living costs. The personalised budget was calculated based on a number of datapoints, including a review of medical records, information gathered during a visit at the individual's place, and an evaluation of the individual's 'Scale of Independent Behavior'.[107] All of this information first had to be manually inputted into the system, which entails a first layer of risk for human error. The information can accidentally be inputted incorrectly, or it can be gathered in an erroneous manner (for instance in case the evaluator misinterprets someone's ability to independently perform a task and hence attaches a wrong score to this criterion). Next, based on the inputted information, the system was programmed to make a budget decision that was supported by a data-driven model which aimed to predict the individual's budget needs. The model in question was, however, deficient. In some instances, it *decreased* the budget of people whose medical needs had *increased*, indicating a structural flaw with serious consequences.[108] Ultimately, the combination of various mistakes in the system led to budget reductions for a large number of individuals in need, without explanation, which forced those who were affected to start a class action to seek redress.[109]

Mistakes can also occur when data pulled from different databases loses its original contextual integrity and thereby no longer accurately reflects reality,[110] or when an organisation relies on data that reveals a spurious correlation[111] yet erroneously interprets this as having a causal relationship. In the previous section, I gave the fictitious example of the fraud propensity of dog- versus cat-owners. Consider, however, the non-fictitious correlation that has been identified between US per capita cheese consumption and the number of people who die by becoming entangled in their bedsheets (correlation of 94.71 per cent between 2000 and 2009).[112] One should hope that no public authority would believe that introducing cheese consumption restrictions will also lead to fewer deaths. More generally, mistakes can also arise when a reasoning process relies on the mistranslation of knowledge to code, or when the model or optimisation function of an algorithmic

[107] See Restrepo Amariles (n 24) 288. See also Ryan Calo and Danielle Keats Citron, 'The Automated Administrative State: A Crisis of Legitimacy' (2021) 70 Emory Law Journal 797, 801.

[108] See *K.W. v Armstrong*, 298 FRD 479 (D. Idaho 2014). For a thorough discussion of this case, see Restrepo Amariles (n 24) 287 and following.

[109] ibid 289–90.

[110] ibid 292.

[111] Spurious correlations can be defined as "*false indicators of causality, typically arising when an extraneous variable that affects two other variables is omitted*". See Imad A Moosa, 'Blaming Suicide on NASA and Divorce on Margarine: The Hazard of Using Cointegration to Derive Inference on Spurious Correlation' (2017) 49 Applied Economics 1483.

[112] See Vigen (n 34). See also 'Beware Spurious Correlations' [2015] *Harvard Business Review* <https://hbr.org/2015/06/beware-spurious-correlations>.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

system is based on erroneous underlying assumptions about society that may not be easily verifiable – a point discussed further below.

Even basic algorithmic systems that do not rely on data-driven models or multiple datasets can contain errors which may lead to disastrous consequences. The example of the UK's Post Office scandal, which involved reliance on a faulty software called Horizon to carry out transaction, accounting and stocktaking tasks, is a testimony to this.[113] Due to bugs in the software, the amounts reported sometimes indicated shortfalls of thousands of pounds, which the Post Office took at face value and – rather than considering that the system was flawed – used as a basis to prosecute postmasters, some of whom ended up in prison following convictions for false accounting and theft.[114] Accordingly, insufficient attention to errors in the software's programming led to great human tragedy, which was only redressed two decades after the facts, when some of the falsely convicted postmasters already served many years in prison, lost their marriages or died.[115]

In sum, just as mistakes can occur when human beings make decisions, so can mistakes occur when human beings design, develop and use algorithmic tools to inform or make decisions. Yet due to their *systemic* nature, the mistakes that occur in the context of algorithmic systems can lead to *systemic* types of harm, a point I will come back to in Chapter 4.

2.2.3 *Risk of Bias*

Another concern typically associated with human action and decision-making is 'bias'. As human subjects, we are not isolated entities, but inherently influenced by our societal surroundings. Our education, upbringing, culture, social circles, language and societal roles shape our thoughts, opinions and values. While we can try to detach ourselves from these influences and seek an 'objective' point of view to consider facts and events, our inherent positionality in the world – including our motivations, ambitions and relationships – means that we inevitably look at the world from a certain angle. Our biases are not necessarily problematic, as they enable us to draw on our knowledge and experience to make sense of the world around us.[116] However, bias can also be fallacious and cloud our rational judgment,

[113] 'Post Office Scandal: What the Horizon Saga Is All About' *BBC News* (22 July 2021) <www.bbc.com/news/business-56718036>.

[114] Mitchell Clark, 'Bad Software Sent Postal Workers to Jail, Because No One Wanted to Admit It Could Be Wrong' (*The Verge*, 23 April 2021) <www.theverge.com/2021/4/23/22399721/uk-post-office-software-bug-criminal-convictions-overturned>.

[115] Zoe Darling, 'More than 30 Victims of Post Office IT Scandal Died without Justice' (*The Justice Gap*, 15 February 2022) <www.thejusticegap.com/more-than-30-victims-of-post-office-it-scandal-died-without-justice/>.

[116] See also Endre Begby, 'The Epistemology of Prejudice' (2013) 2 Thought: A Journal of Philosophy 90; Thomas L Griffiths, 'Understanding Human Intelligence through Human Limitations' (2020) 24 Trends in Cognitive Sciences 873.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

as well as having a discriminatory effect on the people we interact with or make decisions about.

One of the proclaimed benefits of algorithmic regulation is the fact that – unlike human beings – algorithmic systems do not suffer from human biases and achieve more objective decision-making, hence diminishing the risk of unjust discrimination.[117] After all, algorithmic systems lack moral agency and are hence not inherently motivated by their own self-interests or by prejudices in the way human beings are. Yet the myth that algorithmic systems generate purely objective outcomes has by now been busted, as it is widely acknowledged that they can reflect the prejudices and cognitive biases of their makers and users. Bias can also manifest itself through the data fed into the system.[118] If a facial recognition system is only trained on datasets showing pictures of white men, the system will not be able to recognise women and people of colour as accurately.[119] Accordingly, the outcomes of algorithmic systems, and the biases they reflect, hinge on the human decisions that lie at their origin. The fact that a broad range of societal domains, and hence also the knowledge about or data collected from these domains, are still plagued by inequalities and historic discriminations, renders the unchecked use of algorithmic systems liable to perpetuate and even exacerbate unjust bias – at scale.

Consider the by now well-known example of Amazon's data-driven algorithmic system built to evaluate the CVs of incoming applicants. The choice of the system's designers to rely on data from previous (successful) applicants, which were primarily white males, to train the algorithm's model resulted in the algorithm assigning lower scores to incoming female CVs, and hence to discriminate applicants based on their gender.[120] The outcome of this design choice was unlikely deliberate, yet had an adverse impact on numerous individuals, as the non-representative dataset led to unlawful discrimination.[121] One can also raise questions about the fact that most

[117] See, e.g., David Moschella, 'Machines Are Less Biased than People' (*Verdict*, 12 November 2019) <www.verdict.co.uk/ai-and-bias/>.

[118] Harini Suresh, 'The Problem with "Biased Data"' (*Medium*, 26 April 2019) <https://medium.com/@harinisuresh/the-problem-with-biased-data-5700005e514c>; Frederik J Zuiderveen Borgesius, 'Strengthening Legal Protection against Discrimination by Algorithms and Artificial Intelligence' [2020] The International Journal of Human Rights 1; Eirini Ntoutsi and others, 'Bias in Data-Driven Artificial Intelligence Systems – An Introductory Survey' (2020) 10 WIREs Data Mining and Knowledge Discovery e1356.

[119] Joy Buolamwini and Timnit Gebru, 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification', in *Proceedings of Machine Learning Research* (2018) <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>; Timnit Gebru, 'Race and Gender' in Markus D Dubber, Frank Pasquale and Sunit Das (eds), *The Oxford Handbook of Ethics of AI* (Oxford University Press 2020); Ntoutsi and others (n 118).

[120] Jeffrey Dastin, 'Amazon Scraps Secret AI Recruiting Tool That Showed Bias against Women' *Reuters* (San Francisco, 10 October 2018) <www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>.

[121] Consider also the research undertaken by journalists from Bavarian Broadcasting who examined an algorithmic application developed by a Munich-based start-up that analyses videos from job applicants based on their tone of voice, language, gestures and facial expressions, in

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

algorithmic voice assistants, such as Amazon's Alexa, Apple's Siri or Microsoft's Cortana, typically have a female name and voice by default, thereby – consciously or not – risking to reinforce the societal paradigm of female subservience.[122]

Evidently, the risk of biased decision-making is also present without reliance on algorithmic regulation. It has therefore been argued by some that, since the decision-making process of algorithmic systems can at least be rendered transparent by 'printing out and examining the program' so as to 'de-bias' it (unlike the thought processes of biased public officials) algorithmic regulation remains the better option.[123] Without taking a stance on which option is 'best', it must be pointed out that the prejudices embedded in algorithmic systems are not always visible or easily perceivable, for instance when they arise from reliance on proxies that are only indirectly discriminating, or from biased assumptions underpinning the system's design choices or the selection of the data. Moreover, as I noted in Section 2.1, the decision-making process of algorithmic systems using deep-learning techniques are not intelligible, hence making it challenging to assess their potentially biased character. Finally, the significant difference in terms of the decision-making's *scale* should be pointed out. Unlike a public official handling individual cases, in just a few seconds a biased algorithmic system can inform or adopt impactful administrative acts for thousands or even millions of citizens. Accordingly, those behind the proverbial wheel of the algorithmic system – its designers, developers and users – carry an important responsibility in this respect.

2.2.4 *Opacity*

One of the reasons it can be difficult to detect and mitigate the risk of errors or bias in the context of algorithmic systems is the fact that they can suffer from a

order to make the application process "*faster, but also more objective and fair*". The application's evaluation of candidates was found to be easily swayed by appearances in the video, such as the fact that an applicant wears glasses or a headscarf, or the presence of a painting or bookshelf against the wall, or the lighting quality of the video – all of which should in principle be irrelevant for a job applicant's evaluation. See Bavarian Broadcasting, 'Objective or Biased' (*BR*), 2021 <https://interaktiv.br.de/ki-bewerbung/en/>.

[122] Caitlin Chin and Mishaela Robison, 'How AI Bots and Voice Assistants Reinforce Gender Bias' (*Brookings*, 23 November 2020) <www.brookings.edu/research/how-ai-bots-and-voice-assistants-reinforce-gender-bias/>. As a response to such criticism, many AI-enabled voice assistants are now also equipped with male voices.

[123] See, for instance, the contention by Boden that

> many people – for instance, those who are female, working class, Jewish, disabled, or black – encounter unspoken (and often unconscious) prejudice in their dealings with official or professional bodies. An AI welfare advisor, for example, would not be prejudiced against such clients unless its data and inferential rules were biased in the relevant ways. A program could, of course, be written so as to embody its programmer's prejudices, but the program can be printed out and examined, whereas social attitudes cannot.

See Margaret A Boden, 'The Social Impact of Artificial Intelligence' in Ray Kurzweil (ed), *The Age of Intelligent Machines* (Cambridge University Press 1990) 451.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

transparency deficit. As pointed out in Section 2.1.3, certain data-driven systems are also denoted as 'black-box' systems, given the unexplainable nature of their inner decision-making processes.[124] This problem is less manifest with knowledge-driven systems, which are typically more interpretable. The notion that algorithmic systems are opaque should thus be nuanced, especially when making general statements. Broadly speaking, a lack of transparency in the context of algorithmic systems can manifest itself in at least three non-exclusive ways, concerning, respectively: (1) the fact that an algorithmic system is used, (2) the way in which it is used and (3) the way in which it works.

First, public authorities that rely on algorithmic systems to inform or adopt administrative acts can omit to communicate, or choose to obscure, the fact that they deploy such systems. Given the system's digital and hence potentially frictionless nature, individuals may not always be aware of the fact that they are being subjected to a (partially) automated decision-making process.[125] This need not be problematic, especially if the decisions made by the system do not affect people in any meaningful way. If a Flemish-speaking public official uses an algorithmic translation tool to draft an email in French rather than asking a colleague for help, this is unlikely to affect anyone's rights or interests, and unlikely needs to be rendered explicit. If, however, that public official outsources the decision of whether a citizen should receive social welfare benefits to an algorithmic system, the omission to communicate about this is a different matter, since this information is valuable to a citizen wanting to challenge the decision, particularly if she suspects error or bias. I will come back to this issue in subsequent chapters when analysing the impact of algorithmic regulation on the rule of law. Suffice it to state here that this type of opacity is human-chosen, and does not depend on the technical specificities of the system.[126]

Second, there can be opacity around the way in which an algorithmic system is used. This concerns, for instance, opacity around the type of data fed into the system, what the system was optimised for, which assumptions underly the system's design, and how the outcomes of the system are used by the public authority that

[124] See, e.g., Frank Pasquale, *The Black Box Society: The Secret Algorithms that Control Money and Information* (Harvard University Press 2015).

[125] Think of remote facial recognition systems that might scan our faces without us being aware of this, but also of online psychographic targeting that can be used to manipulate us into buying certain products or believing disinformation. See also Kate Crawford, *Atlas of AI* (Yale University Press, 2021), 109.

[126] Furthermore, it can also be noted that algorithmic systems can be deliberately used to pretend they are a human being rather than a machine, in light of their ability to be programmed so as to mimic human behaviour – for instance in the form of a chatbot. When not transparently communicated about, such opaque use of algorithms can adversely affect interests such as privacy, autonomy and dignity. See, e.g., Catelijne Muller, 'The Impact of Artificial Intelligence on Human Rights, Democracy and the Rule of Law' (Council of Europe 2020) CAHAI(2020)06-fin <www.coe.int/en/web/artificial-intelligence/cahai>.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

deploys it. These value-laden choices are rarely made transparent, which strengthens the idea that those choices do not exist. That idea is mistaken, yet it risks overshadowing the fact that the systems' developers and deployers, who are usually already in a position of power, can retain power precisely through the non-contestable configuration of these systems.[127]

Third, there can be opacity around the algorithmic system's inner workings. This is where the black-box problem truly comes in.[128] For certain data-driven systems, it cannot be explained how their internal decision-making processes work, given the high level of complexity of the model's functions. This renders it difficult to evaluate whether the processes followed by the system are based on robust and fair assumptions and comply with existing legislation.[129] However, even when such inscrutable systems are used, transparency regarding the two other points above remains possible, and can already enhance the possibility to exercise oversight over the system. Information can in any case be provided about the way in which the system was designed and developed, what it was optimised for, which input was selected, which techniques were used, and how the system was tested for potential bias or inaccuracies. It should hence be ensured that this third type of opacity is not used as a pretence to also maintain *human-chosen* opacity to avoid public scrutiny.[130] Moreover, to address the challenges of this third type of opacity, a research domain has developed itself around the notion of 'explainable AI', aimed at rendering such models more intelligible and at enabling developers and deployers of these systems to gain more insights into their internal processes.[131]

Finally, it has been claimed that the choice between using a black box model and an interpretable model often comes down to making a trade-off between accuracy and interpretability since, despite their opaque nature, in some situations deep learning models can generate more accurate results.[132] However, this juxtaposition is not always accurate. Researchers have shown that, in some situations, a similar level of accuracy can also be reached with more interpretable models, and that this binary choice can hence be a false (and potentially misleading) dichotomy.[133]

[127] See in this regard also Steven Feldstein, *The Rise of Digital Repression: How Technology Is Reshaping Power, Politics, and Resistance* (Oxford University Press 2021).
[128] Bathaee (n 43).
[129] Jenna Burrell, 'How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms' (2016) 3 Big Data & Society 1.
[130] Crawford (n 125) 12.
[131] See in this regard, e.g., Andrea Vedaldi and others, *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning* (Springer 2019); Pantelis Linardatos, Vasilis Papastefanopoulos and Sotiris Kotsiantis, 'Explainable AI: A Review of Machine Learning Interpretability Methods' (2021) 23 Entropy 18.
[132] See for instance High-Level Expert Group on AI, 'Ethics Guidelines for Trustworthy AI' (n 85) 18.
[133] Rudin and Radin (n 48) 7.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

2.2.5 *Dependency on Data and Proxies*

Algorithmic systems are highly dependent on data. For knowledge-driven systems, this dependency may be less pronounced, as the system's functioning often relies on a pre-articulated model of reality that is codified into the system, yet this model ultimately also consists of machine-readable data based on which the system can reason. For data-driven systems this dependency is far more explicit, as they rely on (very large) datasets to derive a model in the first place. This also implies that the quality of the system's output strongly hinges on the quality and accuracy of the data it is fed during the training, testing and use phase. Public authorities' reliance on such systems in policy-making can be accompanied by the belief that, if only we have enough data about a certain phenomenon, we can use data-driven techniques to make normative decisions. This has also been referred to as *dataism* or "*a belief in data as the enabler of a better, more effective, and objective society*".[134] While the term dataism sounds almost spiritual, its definition actually corresponds rather well to so-called *evidence-based* approaches of public decision-making, grounded in 'data' and 'science',[135] which often comes down to statistical analysis.

While this is of course a lofty goal, it is important to keep in mind that "*data are not the facts themselves, but rather traces or marks of these facts*".[136] Moreover, the facts that data represents "*are about a phenomenon of interest, which is chosen by an observer from a number of different possibilities*",[137] and hence cannot be said to reflect an Archimedean perspective. Likewise, the design and development of algorithmic models rely on "*socially derived perceptions and understandings, not fixed universal, physical laws*".[138] This implies the potential existence of a gap between the data and model on the one hand, and the reality that they represent on the other hand.[139] Accordingly, when developing a model or drawing inferences from a dataset, this gap needs to be kept in mind. Moreover, even when this gap is

[134] Heather Broomfield and Lisa Reutter, 'In Search of the Citizen in the Datafication of Public Administration' (2022) 9 Big Data & Society 1, 3. See also Jose van Dijck, 'Datafication, Dataism and Dataveillance: Big Data between Scientific Paradigm and Ideology' (2014) 12 Surveillance & Society 197.

[135] See, e.g., Gary Banks, 'Evidence-Based Policy Making: What Is It? How Do We Get It?' (2009) SSRN ANU Public Lecture Series, Productivity Commission <https://papers.ssrn.com/abstract=1616460>.

[136] Wolfgang Pietsch, *Big Data* (1st edn, Cambridge University Press 2021) 11.

[137] ibid.

[138] Kyle Eischen, 'Opening the "Black Box" of Software: The Micro-Foundations of Informational Technologies, Practices and Environments' (2003) 6 Information, Communication & Society 57, 61. See also Katherine Fink, 'Opening the Government's Black Boxes: Freedom of Information and Algorithmic Accountability' (2018) 21 Information, Communication & Society 1453, 1454.

[139] A common aphorism in statistics that acknowledges this limitation states that 'all models are wrong', even if 'some are useful'. See also George EP Box, 'Science and Statistics' (1976) 71 Journal of the American Statistical Association 791.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

minimal, and the dataset can be said to provide an adequate representation of reality, one must consider that a precise reproduction of reality may not always be desirable from a normative point of view. As explained above, in many societal domains, historic inequalities persist, which will also be reflected in the data collected about that domain, necessitating a cautious approach when using data about how things *are* to make predictions, recommendations or decisions about how things *should be*.

Unfortunately, the push for technocratic governance, which "*assumes that complex societal problems can be deconstructed into neatly defined, structured and well-scoped problems that can be solved algorithmically and in which political realities play no role*",[140] sometimes leads to a conflation of the positive and the normative. This can also be referred to as the *is–ought fallacy*.[141] Let me clarify with an example. Consider the use of an algorithmic system by a public authority to assist in hiring new officials at a Ministry.[142] It is one thing for this system to help assess whether a formal eligibility criterion for the job (for instance, having a specific degree) has been met. In this case, the system is not asked to help determine which grounds should render a job candidate eligible or right for the job, but merely to assess whether the already explicitly defined prerequisite is attained. It is, however, another thing for the system to help determine, for instance based on data of past hired public officials and how well they perform, which candidate should be interviewed.

The first algorithmic task can be said to belong to the positive realm. Someone already took a normative decision by deciding that one must have a specific degree to be eligible for the job, and the algorithm is merely deployed to peruse data to determine whether this is the case. The second task, however, belongs to the normative realm. The ministry here outsources the normative decision of what makes a public official 'right for the job' to the algorithmic system, which will in turn rely on the optimisation function and the data it was fed. The underlying normative grounds of the algorithmic recommendation may not be transparent here (perhaps the fact that they have much prior work experience or, to revisit the above example of Amazon,[143] the fact that they are male), because it is not necessarily known which factors were flagged by the algorithmic system as relevant, yet it comes down to a normative decision nonetheless. As I discussed

[140] Marijn Janssen and George Kuk, 'The Challenges and Limits of Big Data Algorithms in Technocratic Governance' (2016) 33 Government Information Quarterly 371, 372.

[141] The articulation of this fallacy is 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.

[142] I have taken this example from annex III of the AI Act mentioned above, which classifies such systems as high risk.

[143] See *supra*, Section 2.2.3.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

elsewhere,[144] when humans seek to understand the best approach to deal with a problem, they already – often implicitly – have an idea of what the ideal outcome would be, based on their values and preferences. While algorithmic systems can help determine what the best course of action might be given a value X, they will never be able to determine the value that humans should strive for *as such*. Thinking otherwise is a naive approach at best, and leads to a dangerous discharge of responsibility at worst.

In addition, it should be borne in mind that not everything can easily or fully be captured by data. Complex social phenomena are not readily translatable to quantifiable metrics, and hence typically require the mediation of indicators, metrics and proxies.[145] Since hardly any social phenomenon can be entirely reduced to metrics and indicators, something risks getting lost when translating information about such phenomena into a format that can be algorithmically computed and analysed.[146] This intangible information deficit is also problematic in the context of algorithmic regulation, which is dependent on the quality of these indicators and proxies, and on the soundness of the assumptions underlying their use.[147] Moreover, it places the official or authority responsible for the identification and collection of these indicators in a position of power (albeit typically a hidden one) since they can not only frame the problem that needs analysis, but also the norm or ideal that the system should seek to optimise, as well as indications that may constitute a deviation of the norm.

Consider the example of an algorithmic system used by a public authority to predict which individuals might commit fraud, and hence where it should focus its limited resources. While a person's 'propensity to commit fraud' is difficult to quantify, there are some elements that will typically be used to provide an indication thereof, such as possible previous convictions of fraud, missing documents, unusual transactions or past complaints. These elements are more easily quantifiable and data-fiable and could thus be used as a proxy for 'the propensity to commit fraud', which is what the public authority is ultimately after. Crucially, however, the chosen proxies are not always reflective of the sought-after phenomenon: this will depend

[144] 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) 12.

[145] Sally Engle Merry, 'Measuring the World: Indicators, Human Rights, and Global Governance' (2011) 52 Current Anthropology S83; Viktor Mayer-Schönberger and Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work, and Think* (Houghton Mifflin Harcourt 2013).

[146] See also Geoffrey C Bowker and Susan Leigh Star, 'Building Information Infrastructures for Social Worlds – The Role of Classifications and Standards' in Toru Ishida (ed), *Community Computing and Support Systems: Social Interaction in Networked Communities* (Springer 1998); Luke Stark, 'Algorithmic Psychometrics and the Scalable Subject' (2018) 48 Social Studies of Science 204.

[147] Rachel Thomas and David Uminsky, 'The Problem with Metrics Is a Fundamental Problem for AI' [2020] Ethics of Data Science Conference 2020 <http://arxiv.org/abs/2002.08512>.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

on the soundness of the assumptions made by the system's developers.[148] Furthermore, even if the indicators are well-chosen (by being relevant to analyse the phenomenon in question) they remain mere indicators: in practice, a citizen might very well have missing documents and make unusual transactions, yet nevertheless not in the slightest be prone to commit fraud. In short, the human condition cannot simply be reduced to numerical utility functions.[149] It must therefore be ensured that such a reductionist approach to human beings – and to social phenomena more generally – does not ignore relevant and essential aspects of their humanity for the sake of speed and efficiency. This is especially important when the proxies and indicators are not based on actual behaviour or facts about the citizen in question, but on correlations that were identified with other individuals and groups.[150]

Finally, it must be noted that the data analysed by data-driven systems typically concerns information about individualisable human beings, in which case it is considered to be 'personal data'.[151] The personalisation of administrative acts essentially relies on public authorities' ability to collect data about citizens, and to use such data to profile them and draw inferences about their character, preferences or behaviour.[152] As the combination of different data-sets might yield new possibilities for analysis and insights, authorities are incentivised to both collect more data and

[148] Furthermore, certain proxies may be relevant in theory, but are illegal to take into account in practice given that they can lead to unjust discrimination. A hypothetical study might indicate, for instance, that over the past fifty years women were overall less creditworthy than men (without necessarily explaining the historical reasons for this). While, on that basis, banks could choose to make the assumption that sex is a valid indicator of someone's creditworthiness, they are in principle not allowed to take this factor into account in their evaluation, since sex is a prohibited discrimination ground.

[149] Foreword by Danielle Allen, x, in Hannah Arendt, *The Human Condition* (University of Chicago Press (2019) 1958).

[150] See in this regard also Sandra Wachter and Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) 2 Columbia Business Law Review 494; Eduard Fosch-Villaronga and others, 'A Little Bird Told Me Your Gender: Gender Inferences in Social Media' (2021) 58 Information Processing & Management 102541.

[151] Under EU General Data Protection Regulation ("GDPR"), such information is broadly defined as "*any information that relates to an identified or identifiable living individual*". See European Parliament and Council Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) 2016.

[152] Such predictions not only rely on the data of the individual that is being assessed, but also on the data of many other individuals, and how their traits correlate. On algorithmic profiling, see, e.g., Mireille Hildebrandt and Bert-Jaap Koops, 'The Challenges of Ambient Law and Legal Protection in the Profiling Era' (2010) 73 Modern Law Review 428; Stefanie Hänold, 'Profiling and Automated Decision-Making: Legal Implications and Shortcomings' in Marcelo Corrales, Mark Fenwick and Nikolaus Forgó (eds), *Robotics, AI and the Future of Law* (Springer 2018). See also Omri Ben-Shahar, 'Data Pollution' (2019), 11 Journal of Legal Analysis 104 and Salomé Viljoen, 'A Relational Theory of Data Governance' (2021) 131 The Yale Law Journal 573.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

keep it stored for a longer time, in case an opportunity arises to use it in another context.[153] Clearly, this creates tensions with the fundamental rights to privacy and data protection,[154] as the incentive to gather and analyse data from as many individuals as possible de facto leads to mass surveillance.[155] As I discuss further below,[156] the EU has legislation in place to protect individuals when their personal data is processed. However, this does not take away the fact that mass personal data-collection is taking place on a daily basis, which increases the asymmetry of information and hence also the asymmetry of power between government and citizen. It goes beyond this book to provide a thorough description of the impact of algorithmic regulation on the rights to privacy and data protection, yet it must be highlighted as an important factor when considering its societal aspects, since this impact considerably influences society's shape and direction.[157]

2.2.6 *Impact on Human Agency*

Human agency denotes the capacity of human beings to act in a particular situation or environment, and is typically associated with the notion that they can do so intentionally and autonomously.[158] Agency is traditionally also linked to responsibility since, generally speaking, one can only be deemed responsible for situations in which one has a certain level of agency.[159] Since the very raison d'être of algorithmic systems is to take over tasks from human beings to make their lives 'easier' or 'better',

[153] For a commercial actor, these insights could focus on the way in which a certain service or product can best be commercially marketed based on individuals' preferences. For a law enforcer, these insights could focus on the physical places in which most crimes occur, and where police resources should hence be prioritised. See also Crawford (n 125) 95.

[154] These rights are respectively protected by Articles 7 and 8 of the Charter of Fundamental Rights of the European Union, as well as by secondary legislation – including, for instance, the abovementioned GDPR.

[155] See Karen Yeung, 'Five Fears about Mass Predictive Personalization in an Age of Surveillance Capitalism' (2018) 8 International Data Privacy Law 258; Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (1st edn, PublicAffairs 2019).

[156] See *infra*, Section 5.3, focusing particularly on the GDPR and the LED.

[157] See also Paul De Hert and others, 'Legal Safeguards for Privacy and Data Protection in Ambient Intelligence' (2009) 13 Personal and Ubiquitous Computing 435; Bart Custers, 'Data Dilemmas in the Information Society: Introduction and Overview' in Bart Custers and others (eds), *Discrimination and Privacy in the Information Society*, vol 3 (Springer Berlin Heidelberg 2013) <http://link.springer.com/10.1007/978-3-642-30487-3_1>; Julie E Cohen, 'Turning Privacy Inside Out' (2019) 20 Theoretical Inquiries in Law 1.

[158] See, e.g., Markus Schlosser, 'Agency' in Edward N Zalta (ed), *The Stanford Encyclopedia of Philosophy* (Winter 2019, Stanford University 2019) <https://plato.stanford.edu/archives/win2019/entries/agency/>.

[159] See, for instance, Aristoteles, *Nicomachean Ethics* (Terence Irwin tr, Hackett 1999), section III.1 1109b30. See also Maureen Sie, 'Self-Knowledge and the Minimal Conditions of Responsibility: A Traffic-Participation View on Human (Moral) Agency' (2014) 48 The Journal of Value Inquiry 271; Mark Balaguer, *Free Will* (MIT Press 2014).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

their impact on human agency is a given, as humans essentially outsource their capacity of action to the system.[160] As long as they have agency regarding the act of outsourcing, and regarding the consequences attached thereto, this need not be an issue. Conversely, the loss of such agency might also hamper their sense of responsibility.[161] The link between reliance on algorithmic systems and human responsibility – and in particular, the risk of a negative correlation between both – has extensively been discussed in scholarship.[162] I will not attempt to reproduce this discussion here, but will merely make some observations on how the deployment of algorithmic systems can impact the agency of the deployers of algorithmic systems in the public sector – an environment that is typically marked by hierarchical relationships.[163]

For this purpose, let me recall an experiment carried out by Stanley Milgram in the 1960s, to assess how obediently people act under 'authority', as this experiment likewise involved the mediation of a machine.[164] Individuals were asked to administer increasingly high electric shocks to a volunteer, whenever that volunteer answered a question erroneously. While the shocks were fake, the grim results of his experiment indicated that there was a high rate of individuals who, sitting behind a machine and faced with the choice to obey to authority or refrain from hurting another person (even upon that person's specific request to stop), all too often opted for the former.[165] Milgram analysed the results of the – multiple variations of his – experiment and drew a number of conclusions that are also of relevance when considering the societal aspects of algorithmic systems.

First of all, Milgram observed that "*distance, time and physical barriers neutralise the moral sense*".[166] The further away the individual was from the volunteer

[160] See in this regard also European Group on Ethics in Science and New Technologies (EGE), 'Statement on Artificial Intelligence, Robotics and "Autonomous" Systems', March 2018, <https://op.europa.eu/en/publication-detail/-/publication/dfebe62e-4ce9-11e8-be1d-01aa75ed71a1>, 9.

[161] See, e.g., Gerben Meynen, 'Why Mental Disorders Can Diminish Responsibility: Proposing a Theoretical Framework' in Bert Musschenga and Anton van Harskamp (eds), *What Makes Us Moral? On the Capacities and Conditions for Being Moral* (Springer Netherlands 2013).

[162] See, for instance, Jaana Hallamaa and Taina Kalliokoski, 'How AI Systems Challenge the Conditions of Moral Agency?' in Matthias Rauterberg (ed), *Culture and Computing* (Springer International Publishing 2020). See also Elena Popa, 'Human Goals Are Constitutive of Agency in Artificial Intelligence (AI)' (2021) 34 Philosophy & Technology 1731.

[163] I discuss the organisational environment of public decision-making in more detail in Section 2.3.

[164] Stanley Milgram, *Obedience to Authority: An Experimental View* (Harper Perennial 2009).

[165] Milgram also refers to the obedience to authority organised through bureaucratic organisation in Nazi Germany, and the actions of Adolf Eichmann who – pursuant to Hannah Arendt's account – banalised the evil he committed by stating that he was simply obeying orders. This 'thoughtlessness' and banalisation of evil is discussed by Arendt in Hannah Arendt, *Eichmann in Jerusalem: A Report on the Banality of Evil* (Viking Press 1963); Hannah Arendt, *Lectures on Kant's Political Philosophy* (University of Chicago Press 1982).

[166] Milgram (n 164) 157.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

subjected to the shock, the higher the obedience rate, despite the moral concern.[167] In the context of algorithmic regulation, there is typically a physical distance between the person responsible for the system and the individual subjected thereto. Indeed, the system is meant to take over tasks from human beings, providing them with the possibility to monitor these tasks from a distance. However, as Milgram noted, this physical distance also facilitates an emotional distance from the individual subjected to the system, and hence from responsibility in case the individual is adversely impacted by the system's outcomes.

Second, Milgram also explains that the individuals who participated in his experiment automatically adopted a number of internal mechanisms or 'buffers' to cope with the tension they faced in an ethically unclear situation, and to divest themselves of moral responsibility. One of those mechanisms concerns deference to a hierarchically higher authority that imposed the decision. In the context of algorithmic regulation, one can point to the fact that algorithmic systems often tend to have a sense of authority, due to their typically superior computational capabilities as well as their aura of objectivity, since they are based on technical rules and mathematical functions rather than on 'biased' human decisions.[168] Reference can also be made to the known risk of automation bias, or the propensity for human beings to favour suggestions made by automated systems and to ignore contradictory information made without automation, even if correct.[169] In a more banal sense, the fact that the public authority's hierarchy decided that decisions should henceforth be informed or taken through algorithmic systems can also constitute an act of authority through which public officials can divest responsibility. It inevitably constraints their individual agency, thus potentially making them feel less responsible for the system's problematic outcomes.

Third, Milgram describes "*the tendency of the individual to become so absorbed in the narrow technical aspects of the task that he loses sight of its broader consequences*".[170] The fact that the problematic act becomes fragmented (the official is no longer the sole person behind the act and is no longer directly faced with its consequences, but there is a chain of actions in between, mediated by technology) likewise facilitates the act's execution.[171] Similarly, when it comes to algorithmic

[167] See ibid. Milgram explicitly refers to the set-up of the experiment and the role that technology played therein: "*While technology has augmented man's will by allowing him the means for the remote destruction of others, evolution has not had a chance to build exhibitors against these remote forms of aggression to parallel those powerful inhibitors that are so plentiful and abundant in face-to-face confrontations.*"

[168] See also Linda J Skitka, Kathleen L Mosier and Mark Burdick, 'Does Automation Bias Decision-Making?' (1999) 51 International Journal of Human-Computer Studies 991.

[169] Mary Cummings, 'Automation Bias in Intelligent Time Critical Decision Support Systems' [2014] AIAA 1st Intelligent Systems Technical Conference.

[170] Milgram (n 164) 7.

[171] Milgram calls this a dangerously typical situation in complex societies: "*it is psychologically easy to ignore responsibility when one is only an intermediate link in a chain of evil action but is far from the final consequences of the action. Even Eichmann was sickened when he toured the*

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

regulation, it can be noted that algorithmic systems are often composed of different components that interact with each other within a broader network or chain, which further alienates the system's designer or deployer from its consequences and facilitates the evasion of responsibility. This has also been discussed in scholarship as the difficulty of the *many hands* problem[172] which, in the context of algorithmic systems, is only intensified by the opacity surrounding the different types of (interacting) conducts and systems.[173]

Finally, one can also point out Milgram's warning that individuals tend to dehumanise the persons affected by their action by attributing impersonal qualities to those persons, and thereby making it easier to cope with the role of 'hurter'. In the context of algorithmic regulation, the affected person is typically an unknown plurality of citizens, reduced to numeric abstractions and data-points in the system and hence also de-humanised in a more literal sense. All of these elements individually as well as collectively should be carefully taken into consideration when outsourcing impactful decisions about human beings to algorithmic systems. With this in mind, let us now explore how algorithmic regulation is relied upon by public authorities.

2.3 ALGORITHMIC REGULATION IN THE PUBLIC SECTOR

Previously, I described the building blocks of algorithmic regulation, namely algorithmic systems, and I discussed their technical and societal aspects. Since this book is concerned with the use of algorithmic regulation by *public authorities of the executive branch of power*, in this section I conduct a closer examination of how these authorities operate. A better understanding of the bureaucratic environment in which administrative acts are adopted, and of the inherent challenges associated with such an environment, is an important prerequisite to examine the role that algorithmic regulation can play therein. Therefore, in this section, I start by setting out how public authorities are organised and describe the key features of their bureaucratic environment (Section 2.3.1), as well as the pitfalls associated therewith (Section 2.3.2). Subsequently, I discuss the role of administrative discretion, and its ability to mitigate some of bureaucracy's pitfalls (Section 2.3.3). I then examine the history of public authorities' reliance on algorithmic systems (which is not a new

concentration camps, but to participate in mass murder he had only to sit at a desk and shuffle papers." See ibid 11.

[172] 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 Jennifer Cobbe, Michael Veale, and Jatinder Singh, 'Understanding Accountability in Algorithmic Supply Chains' (2023), Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, 1186–97.

[173] See in this regard Nathalie A Smuha, 'Beyond the Individual: Governing AI's Societal Harm' (2021) 10 Internet Policy Review 10–11.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

phenomenon) and discuss how the logic of such systems resonates with the logic of bureaucracy (Section 2.3.4). Finally, I conclude with an assessment of the reasons underlying public authorities' embrace of algorithmic regulation (Section 2.3.5).

A caveat must, however, be made. While this section provides a broad overview of the operations and use of algorithmic regulation by public authorities, such overview should not be generalised for all EU Member States. Each Member State has its own legal and administrative traditions, and its own pace of technological uptake. This also means that the internal practices and cultures of their public authorities, including their deployment of algorithmic regulation, vary significantly.[174] Even within the same Member State, notable differences in technological uptake exist when comparing one public authority with another. Let me therefore stress that my observations on the use of algorithmic regulation by public authorities are general in nature rather than representing a meticulous description of the situation in every public authority in the EU, given the large diversity of such uptake.

2.3.1 *The Organisation of Public Authorities*

The executive branch of power is responsible for the execution of laws, and its competences are often characterised as having a residual character, encompassing "*anything that is neither legislative nor judicial*".[175] Historically, the executive has been in charge of law enforcement, warfare and national security, as well as tax collection. Yet from the late nineteenth century onwards, with the rise of the welfare state, its competences significantly expanded, covering also the governance of welfare programmes, economic state interventions, and the regulation of societal risks – such as the risks raised by new technologies to public health or safety.

Increasingly, the general laws and policies adopted by the legislative branch of power required more sophisticated implementation techniques, whereby the executive branch became competent to adopt a growing range of administrative acts, both of individual and general application, often necessitating more specific knowledge or technical expertise.[176] To exercise these more specialised functions, a large

[174] See in this regard also European Commission. Joint Research Centre, *AI Watch, Artificial Intelligence for the Public Sector: Report of the "3rd Peer Learning Workshop on the Use and Impact of AI in Public Services", 24 June 2021* (Publications Office 2021) <https://data.europa.eu/doi/10.2760/162795>.

[175] Robert Schütze, *European Constitutional Law* (Cambridge University Press 2012) 221.

[176] See Sean Gailmard and John W Patty, 'Formal Models of Bureaucracy' (2012) 15 Annual Review of Political Science 353. Note that the role and place of experts in bureaucracy and in policy-making more generally (including reliance on external experts when it concerns highly specialised scientific domains) is much debated – especially in light of the greater call for public participation in policy-making by non-experts – but will not be discussed further in this book.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

organisation of expert administrators was built, denoted as bureaucracy.[177] As stated by Francesca Bignami, the nineteenth century can be described as one "*of great optimism in the ability of public servants and the bureaucratic form of organisation to pursue the common good and advance the interests of society as a whole*".[178] Public officials working at public authorities are (largely) unelected, but they in principle operate under the supervision and control of the government's (largely) elected members.[179]

Bureaucracy has most famously been conceptualised by Max Weber in his work *Economy and Society*, as part of his broader reflections on modernity and the birth of the rational state.[180] Weber's conceptualisation of bureaucracy should be seen as an 'ideal type' or 'analytical concept' rather than as a factual description of how public authorities function.[181] He examined how the bureaucratic organisation of public administration was able to legitimise social control through legal-rational means, as opposed to more traditional or feudal forms of domination.[182] While not devoid of criticism,[183] Weber's work is still a cornerstone of public

[177] This does not mean that bureaucratic forms of state organisation did not exist before. See in this regard Peter Crooks and Timothy H Parsons (eds), *Empires and Bureaucracy in World History: From Late Antiquity to the Twentieth Century* (Cambridge University Press 2016).

[178] Francesca Bignami, 'From Expert Administration to Accountability Network: A New Paradigm for Comparative Administrative Law' (2011) 59 American Journal of Comparative Law 859, 862.

[179] Accordingly, the expansion of the executive's tasks and the concomitant establishment of public administration also marked the start of a distinction between the 'political' component of the executive ('political executive') and the 'administrative' component of the executive ('administrative executive' or 'bureaucracy'). This has raised concerns around the 'unaccountable' nature of civil servants (especially those working in executive agencies, who typically work more independent from politically elected officials), despite the public power they wield and the impact their decisions can have on citizens. However, elections are but one manner of organising accountability, and a rich body of administrative law – together with broader constitutional principles – has ensured various accountability mechanisms for public administrators. See, e.g., Andrew B Whitford, 'Decentralization and Political Control of the Bureaucracy' (2002) 14 Journal of Theoretical Politics 167; Anya Bernstein and Cristina Rodriguez, 'The Accountable Bureaucrat' (2023) 132 Yale Law Journal 1600. See also, e.g., Gillian E Metzger, 'Legislatures, Executives, and Political Control of Government' in Peter Cane and others (eds), *The Oxford Handbook of Comparative Administrative Law* (Oxford University Press 2020) 697. See also Denis J Galligan, 'Public Administration and the Tendency to Authoritarianism' in András Sajó (ed), *Out of and into Authoritarian Law* (Brill – Nijhoff 2002) 193.

[180] Note that this work was published after Weber's death, by his wife Marianne. See also Glynn Cochrane, *Max Weber's Vision for Bureaucracy: A Casualty of World War I* (Springer International Publishing 2018) 68.

[181] See, e.g., Lorenz, Meijer and Schuppan (n 26) 73. See also Johan P Olsen, 'Maybe It Is Time to Rediscover Bureaucracy' (2006) 16 Journal of Public Administration Research and Theory 1, 2.

[182] Chris Muellerleile and Susan L Robertson, 'Digital Weberianism: Bureaucracy, Information, and the Techno-Rationality of Neoliberal Capitalism' (2018) 25 Indiana Journal of Global Legal Studies 187, 192.

[183] See, for instance, the discussion by Harro Höpfl of the many inconsistencies in Weber's conceptualisation of bureaucracy, drawing on similar critiques from other scholars in 'Post-

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

administration research[184] and proves helpful in discerning how public authorities operate today,[185] since "*bureaucracies continue to be pillars of public service provision*".[186]

Several features have been distinguished as characteristic of modern bureaucracy.[187] While these features are grouped differently by various commentators,[188] the most notable are:

- Efficiency: the optimisation of working methods and resources in order to achieve the desired goals in the most efficient manner;
- Legality: the legitimation of administrative action through formalised rules, standards and procedures;
- Rationality: the execution of tasks based on expert knowledge and reason, and the gathering of further knowledge to advance the organisation's goals and create an institutionalised memory;[189]
- Objectivity and impartiality: the application of laws and procedures to all subjects equally, without privileging certain individuals or groups;
- Impersonality: the execution of tasks driven by impersonal official obligations rather than by personal interests or emotions – or *sine ira et studio*, as Weber put it;[190]

bureaucracy and Weber's "Modern" Bureaucrat' (2006) 19 Journal of Organizational Change Management 8. See also the critique that Weber's conceptualisation was hardly original, and largely drawn from earlier work by Hegel, in Cochrane (n 180) 72 and following.

[184] See Erin L Borry and Tina Kempin Reuter, 'Humanizing Bureaucracy: Applying the Human Rights-Based Approach to Weber's Bureaucracy' (2022) 5 Perspectives on Public Management and Governance 164, 164.

[185] While some scholars claimed that public administration is no longer organised bureaucratically, but rather in terms of market structure or network structure, these claims have been rebutted and incorporated into an understanding of modern public organisation as carrying mixed features which include but are not limited to those structures. See also Olsen (n 181); Höpfl (n 183).

[186] 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. See also Olsen (n 181).

[187] It should be stressed that there does not exist a uniform conceptualisation of bureaucracy. Note in this regard also Claude Lefort's statement that "*Bureaucracy appears as a phenomenon that everyone talks about, feels and experiences, but which resists conceptualization*", in Claude Lefort, 'What Is Bureaucracy?' (1974) 22 Telos: Critical Theory of the Contemporary 31.

[188] See, e.g., Höpfl (n 183); Olsen (n 181); Lorenz, Meijer and Schuppan (n 26); Borry and Reuter (n 184) 165.

[189] Rationality has been defined and used by scholars in many different ways. As regards the use of the term by Weber, Kalberg, for instance, identified four related but distinct types of rationality. See Stephen Kalberg, 'Max Weber's Types of Rationality: Cornerstones for the Analysis of Rationalization Processes in History' (1980) 85 The American Journal of Sociology 1145.

[190] Meaning 'without hatred or passion'. See Max Weber, *Economy and Society (1968)* (Keith Tribe tr, Harvard University Press 2019) 353.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

- Hierarchy: the establishment of relationships of super- and subordination, including the importance of loyalty and obedience to hierarchical authority;
- Specialisation: the organisation of tasks based on specific jurisdictional and competence areas, driven by a functional division of labour.

Note that, beyond this set of formal or procedural features, bureaucracy has also been described more substantively as "*an expression of cultural values and a form of governing with intrinsic value*", whereby "*administration is based on the rule of law, due process, codes of appropriate behavior, and a system of rationally debatable reasons*".[191] Public officials' obedience to authority is therefore seen as the other side of the same rule of law-coin, as it legitimises their actions and ensures their accountability to society at large. Moreover, their duty to act in the public interest[192] renders them "*guardians of constitutional principles, the law, and professional standards*.[193] Evidently, this duty also raises questions on the relationship between, on the one hand, the 'public' interests or values that public authorities should advance and, on the other hand, the 'individual' interests or values that private persons might hold dear – especially in a pluralistic society.[194]

Over the last century, the organisation of public authorities has known several waves of reform, inspired by neoclassical and private management ideas.[195] This has led to the privatisation of services and an embrace of market competition, whereby citizens have become "*a collection of customers with a commercial rather than a political relationship to government, and legitimacy is based on substantive performance and cost efficiency*".[196] Despite these reforms, the core features of bureaucracy were retained, and most particularly the focus on procedural rationality and the efficient execution of tasks.[197] Indeed, as noted by Galligan, "*the premise of efficient administration in implementing the policies of government remains the dominant paradigm*".[198] In what follows, I will therefore be using the notion of bureaucracy and public administration interchangeably.

2.3.2 *The Pitfalls of Bureaucracy*

A lot has been written about the downsides of bureaucracy, and it goes well beyond the scope of this book to reconstruct those rich debates here. I therefore limit myself

[191] Olsen (n 181) 3.
[192] See in this regard also Barry Bozeman, *Public Values and Public Interest: Counterbalancing Economic Individualism* (Georgetown University Press 2007).
[193] Olsen (n 181) 3.
[194] See in this regard also Sebastiaan P Tijsterman and Patrick Overeem, 'Escaping the Iron Cage: Weber and Hegel on Bureaucracy and Freedom' (2008) 30 Administrative Theory & Praxis 71.
[195] See Kennedy (n 106) 210.
[196] Olsen (n 181) 6.
[197] Muellerleile and Robertson (n 182).
[198] Galligan, 'Public Administration and the Tendency to Authoritarianism' (n 179) 192.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

here to a brief analysis of how some of the core features of public administration can run counter to the normative ideals of liberal democracy, as this will be relevant for my subsequent examination of the impact of algorithmic regulation on the rule of law. Such analysis also enables a better understanding of the challenges that are *inherent* to bureaucracy, as opposed to the challenges that are raised or exacerbated by algorithmic regulation in particular.

Despite the benefits that Weber identified with the bureaucratic form of organisation, he also acknowledged it might pose certain threats to substantive values such as individual freedom, and the risk of a concentration of power.[199] For instance, in related work,[200] he noted that an overly instrumental conception of rationality can result in a *stahlhartes Gehäuse* or *iron cage* in which people are subjected to rules and procedures without ensuring that the substantive values these rules are supposed to serve are actually achieved.[201]

This echoes an oft-made distinction between two types of rationality, namely procedural and substantive rationality.[202] As explained by Muellerleile and Robertson, *substantive rationality* is the value-laden framework that people draw on to determine their actions in a particular situation, typically based on their individual ethical perspective.[203] Conversely, *procedural rationality* draws on scientific and economic calculation, and is based on a depersonalised set of codified rules, laws and regulations, which is typical in bureaucratic organisations. When the execution of actions becomes a mere instrumental undertaking, without regard for their practical impact and the extent to which they advance the goals and values they should secure, procedural rationality can undermine substantive rationality, to the detriment of the values intrinsic thereto.

This risk has been raised by various scholars. In particular, it has been pointed out that the impersonal nature of bureaucracy (or the 'rule of nobody', as bureaucracy was called by Hannah Arendt[204]) can lead to a 'disinterestedness' that erodes public officials' sense of responsibility and morality, which can turn into amorality[205]

[199] Borry and Reuter (n 184) 165. See also Kari Palonen, 'Max Weber's Reconceptualization of Freedom and Foundations' (1999) 27 Political Theory 523; Terry Maley, 'Max Weber and the Iron Cage of Technology' (2004) 24 Bulletin of Science, Technology & Society 69.

[200] Max Weber, *The Protestant Ethic and the Spirit of Capitalism (1905)* (Stephen Kalberg tr, Blackwell 2002).

[201] See also Peter Baehr, 'The "Iron Cage" and the "Shell as Hard as Steel": Parsons, Weber, and the Stahlhartes Gehäuse Metaphor in the Protestant Ethic and the Spirit of Capitalism' (2001) 40 History and Theory 153; Tijsterman and Overeem (n 194); Peeters and Widlak (n 186).

[202] See, e.g., Arre Zuurmond, *De Infocratie: Een Theoretische en Empirische Heroriëntatie Op Weber's Idealtype in Het Informatietijdperk* (Phaedrus 1994) 4.

[203] Muellerleile and Robertson (n 182) 195.

[204] Hannah Arendt, *On Violence* (Harcourt Brace and Company 1970).

[205] See in this regard, for instance, the critique formulated by MacIntyre in his seminal work 'After Virtue' (*After Virtue: A Study in Moral Theory (1981)* (Bloomsbury Academic 2015). See also Ron Beadle and Geoff Moore, 'MacIntyre on Virtue and Organization' (2006) 27 Organization Studies 323; Matthew Sinnicks, 'Leadership after Virtue: MacIntyre's Critique of Management

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

(or worse, immorality). One of the most fervent critiques of rational-legal bureaucracy has been formulated by Zygmunt Bauman in his *Modernity and the Holocaust*.[206] According to Bauman, the features of modern bureaucracy, particularly the emphasis on efficiency and procedural rationality, and the specialisation of tasks which risks obliterating the overall result of one's actions, can undermine the possibility for moral action by individuals working for the state.[207] Bauman called attention to the 'important role' that bureaucratic culture played in the Holocaust (also commented on by Hannah Arendt in her characterisation of Eichmann as an amoral bureaucrat[208]), and characterised modern bureaucracy as 'a moral sleeping pill'.[209] As summarised by du Gay, "*for Bauman, the essence of bureaucratic structure and process' is the dissociation of 'instrumental rational criteria' from 'moral evaluations' of the ends they serve*",[210] leading to the threat of moral discharge and the dehumanisation[211] of those subjected to the impersonally applied procedures.[212]

It has also been argued that, precisely because of its emphasis on procedural rationality and rule-following, the organisation of public administration carries an inherent risk of authoritarianism, and a neglect of the rights of individuals.[213] As, for instance, noted by Galligan, "*administrative bureaucracies are naturally governed by procedural rigidity and a disregard for individualized differences; efficiency and self-interest prevail over fairness, and secrecy militates against explanation and justification*".[214] Accordingly, public administration can be considered as more than a mere 'neutral' instrument to carry out public policies. Instead, by virtue of its inherent features, it can rather be seen as a mode of organisation that can *aggravate* potential authoritarian and illiberal elements already contained in the laws it implements.[215]

Reconsidered' (2018) 147 Journal of Business Ethics 735. See also See also Camilla Stivers, 'Rule by Nobody: Bureaucratic Neutrality as Secular Theodicy' (2015) 37 Administrative Theory & Praxis 242; Hannah Spector, 'Bureaucratization, Education and the Meanings of Responsibility' (2018) 48 Curriculum Inquiry 503.

[206] Zygmunt Bauman, *Modernity and the Holocaust* (Polity Press 1989). See also Smuha, 'The Human Condition in an Algorithmized World' (n 144) 35.

[207] Paul du Gay, *In Praise of Bureaucracy: Weber, Organization, Ethics* (SAGE Publications 2000), chapter 2 in particular.

[208] Arendt, *Eichmann in Jerusalem: A Report on the Banality of Evil* (n 165).

[209] Bauman (n 206) 26.

[210] Du Gay (n 207).

[211] See also Borry and Reuter (n 184). Reference can also be made to Section 2.2.6, where I discussed similar concerns voiced by Stanley Milgram.

[212] While these statements might seem condemning, it should be noted that both Arendt and Bauman have nuanced their critique of bureaucracy by conducting more in-depth investigations into its merits and pitfalls, which goes beyond the space I can allocate to this subject in this book.

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

[214] Galligan, 'Public Administration and the Tendency to Authoritarianism' (n 179) 187.

[215] ibid. It should be noted that more nuanced stances have been taken in this regard. For instance, Claude Lefort, another philosopher who opposed totalitarianism and engaged in a critical examination of the role of bureaucracy in the political sphere, drew on Weber to reject

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

Others have taken up bureaucracy's defence, arguing that bureaucratic objectivity "*entails a trained capacity to treat people as 'individual' cases, i.e. apart from status and ascription, so that the partialities of patronage and the dangers of corruption might be avoided*".[216] Moreover, one could also consider bureaucracy's "*instituted blindness to inherited differences in status and prestige*" as "*a source of democratic equalization*".[217] Du Gay also nuances the above, by pointing out that the distinction between formal and substantive rationality has not been deployed by Weber and that it may be artificial since, in practice, a distinction between 'means versus ends' is not always straightforward.[218] Rules can, after all, be designed with the specific aim of protecting and enhancing substantive rights and values, and can contribute to the institutionalisation of public accountability.

Furthermore, it has also been argued that an over-emphasis on *substantive* rationality might overlook the fact that there exists a plurality of (often conflicting) values.[219] Bureaucracy can allow the state to remain 'neutral' and maintain its legitimacy in a pluralistic society, precisely because it focuses on procedural rationality and efficiency, and thereby enables "*the expression and protection of a broader range of conflicting values held to be important by human beings*".[220] While the theme of state neutrality has been widely debated,[221] it should be noted that Weber considered value conflict to be an inherent feature of modernity,[222] and he was hopeful that democratic institutions, through an agonistic political process, could

the claim that "*the development of bureaucracies must affect the nature of a political and economic regime, no matter how necessary they might seem once certain conditions are fulfilled*". He noted that,

> On the contrary, Weber claims that the numerical importance of this form of organization does not in any way determine its relation to power. The proof is that the state bureaucracy accommodates itself to diverse regimes – as demonstrated by France, where the state bureaucracy has remained remarkably stable. The proof lies also in the fact that during war, the bureaucratic staff of a conquered country is used by the foreign power, and continues to carry out its administrative tasks. In principle, bureaucracy is indifferent to the interests and values of a political system, i.e., it is an organ at the service of rulers located somewhere between the rulers and those who are ruled.

See Lefort (n 187). At the same time, the indifference of bureaucracy to the values of a political system does not exclude the fact that its mode of organisation can enhance the excesses of authoritarian governance approaches, as argued by Galligan.

[216] Du Gay (n 207).

[217] ibid.

[218] ibid.

[219] This plurality of values – and the fact that these values can be irreconcilable – was already acknowledged by Weber too. According to him, this irrevocability was furthered by the broader disenchantment of modernity, and the erosion of the role of religion and more traditional sources of moral authority as unifying factor. See in this regard 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, 27.

[220] ibid 25.

[221] See in this regard also Stivers (n 205).

[222] Spicer (n 219) 32.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

address these conflicts. Consequently, he believed in the power of democratic control to mitigate the pitfalls of the 'iron cage', and insisted on the need to make space for politics and democratic oversight amidst and alongside bureaucracy.[223]

Interestingly, the two views I just outlined on bureaucracy's merits and pitfalls foreshadow an important tension that is inherently part of society's reliance on legal rules and procedures to exert social control, which will be closely examined in this book. On the one hand, we require abstract rules of general application, to ensure the law's impartiality ('procedural rationality'). On the other hand, we also require that the application of those general rules results in individual justice,[224] in light of the concrete particularities of each person and situation, and the way in which different values and interests matter to them ('substantive rationality'). Based on the discussion above, it can be concluded that a combination of both is warranted, albeit not always evident, and that political oversight based on a democratic process is an important mechanism to foster this.

In addition to external oversight, there are other mechanisms that can curb the risks evoked above. One such mechanism is ensuring the 'internal morality' of public authorities. Notably, "*Weber did not see bureaucrats, particularly civil servants, simply as mindless automatons but believed them perfectly capable of ethically principled conduct within their own proper sphere of action*".[225] He also emphasised their ethical duties and their need for 'moral discipline'.[226] Indeed, normativity can also play a role *inside* public authorities, by incorporating substantive principles within the organisation of bureaucracy and requiring public officials to adhere thereto.[227] As I shall discuss extensively in Chapter 3, these substantive principles can be distilled from the more general commitment to the rule of law, and from other core tenets of liberal democracy.[228] Another mechanism concerns administrative discretion, which is prevalent in public authorities' day-to-day organisation and can be used to soften the rigidity of procedural rules.[229] Since discretion will be an important theme in this book, especially when discussing its relationship with algorithmic regulation, let me dissect this concept more closely in the next section.

[223] See Max Weber, *Weber: Political Writings* (Peter Lassman and Ronald Speirs eds, Cambridge University Press 1994) 222. See also Palonen (n 199) 532; Spicer (n 219) 35.

[224] See in this regard also Reuben Binns, 'Human Judgment in Algorithmic Loops: Individual Justice and Automated Decision-Making' (2022) 16 Regulation & Governance 197, which will also be discussed *infra*, in Section 4.2.

[225] Spicer (n 219) 34.

[226] ibid.

[227] Denis James Galligan, 'Senses of Discretion', in his *Discretionary Powers: A Legal Study of Official Discretion* (Oxford University Press 1990) 5.

[228] Typically, these principles are concretised through the branch of administrative law.

[229] Galligan, 'Discretionary Powers in the Legal Order' (n 213); 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).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

2.3.3 *Administrative Acts in between Rules and Discretion*

In essence, discretion provides public authorities with a certain level of autonomy when making a decision amongst a variety of options, based on their assessment and judgment.[230] It can be defined as "*a power which leaves an administrative authority some degree of latitude as regards the decision to be taken, enabling it to choose from among several legally admissible decisions the one which it finds to be the most appropriate*".[231] When exercising discretion, public authorities still do so within the confines of the law, which led Dworkin to state that "*discretion, like the hole in a doughnut, does not exist except as an area left open by a surrounding belt of restriction. It is therefore a relative concept. It always makes sense to ask, 'Discretion under which standards?'*"[232] Unsurprisingly, the expansion of the modern state and the increase in government functions, including the increase of legislation that needs to be implemented and applied by the executive, also led to an expansion of discretionary power.

Discretion can arise in different circumstances. The most common one relates to the tension I indicated above, between the general nature of legal rules on the one hand, and the need for their individual application on the other hand. When determining a desirable policy outcome or goal, it is impossible for the legislator to determine in advance all the particular situations that may arise, and to provide precise instructions to the executive as to how it should act in each of these situations. Public authorities therefore typically have some discretion as to how precisely they will implement and enforce generally applicable rules and policies. In this sense, discretion is a side-effect of organising social control based on a system of legal rules.

However, discretion is more than a side-effect, since it can also be seen as "*a positive way of conferring powers where it is important that officials have more freedom as to the way they are to be exercised than a detailed set of rules might allow*".[233] In some situations, applying a general rule to advance a just cause, without the possibility of tempering its concrete impact in potentially unforeseen or unanticipated circumstances, can actually lead to injustice. On the one hand, securing

[230] Galligan notes that the emphasis in discretion should not lay on the fact that public authorities have *autonomy*, but rather that they have scope for judgment and personal assessment – which also implies the need to rely on reason. See 'Senses of Discretion' (n 227) 8.

[231] This definition is drawn from the Council of Europe. See 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', 1980.

[232] See Ronald Dworkin, *Taking Rights Seriously* (Harvard University Press 1978) 31. See also Tony Evans and John Harris, 'Street-Level Bureaucracy, Social Work and the (Exaggerated) Death of Discretion' (2004) 34 British Journal of Social Work 871, 881; Mireille Hildebrandt, 'Algorithmic Regulation and the Rule of Law' (2018) 376 Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 20170355, 5.

[233] Galligan, 'Senses of Discretion' (n 227) 2.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

public services for millions of citizens naturally requires a certain level of organisation and systematisation. On the other hand, any system will also inevitably overlook the uniqueness of the individual cases that it systematises.

This dilemma was also discussed by Emmanuel Levinas, famous for his emphasis on the primacy of ethics in human relationships and the responsibility we have for the 'other'.[234] He considered it inevitable that any system seeking to ensure justice for the many can become dehumanising precisely by approaching individuals in a general rather than an individualised manner.[235] Indeed, within such a system, "*the other is no longer the unique person offering himself to the compassion of my responsibility, but an individual within a logical order or a citizen of a state in which institutions, general laws, and judges are both possible and necessary*".[236] This leads to the danger that the other is "*extinguished in the system of universal laws*".[237] While this danger does not imply that all systematisation of public services should be rejected, it does mean that the system needs to be continually corrected and perfected against its own harshness.[238] This permanent correction of the system, however, cannot be expected to come from another *system*. Instead, Levinas believes it can be found in acts of 'little goodness', which he juxtaposes against a systematised 'Goodness'.[239] The little goodness is "*a goodness outside of every system, every*

[234] See also Emmanuel Levinas, *Autrement qu'être ou au-delà de l'essence* (11th edn, Le Livre de Poche (2019) 1974); Emmanuel Levinas, *Le temps et l'autre* (11th edn, Presses Universitaires de France (2014) 1979).

[235] Emmanuel Levinas, *Is It Righteous to Be?: Interviews with Emmanuel Levinas* (Jill Robbins ed, Stanford University Press 2001). See also Luc Anckaert, 'Ethics of Responsibility and Ambiguity of Politics in Levinas's Philosophy' (2020) 97 Problemos 61.

[236] Lévinas, *Is It Righteous to Be?* (n 235) 116.

[237] ibid.

[238] ibid 206–07. 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) 230.

[239] In this regard, Levinas draws on the novel *Life and Fate* by Vasily Grosman (written in 1959 but only published in 1980). *Life and Fate* details events during the Second World War and provides (comparative) perspectives about the totalising regimes of Nazism and Stalinism. A few characters in the novel – especially Ikonnikov, described as a 'holy fool' – showcase "isolated acts of senseless kindness", which stand in stark opposition to the great totalitarian visions of "the Good". Levinas recounts the significance thereof as follows:

> Grossman's eight hundred pages offer a complete spectacle of desolation and dehumanization … Yet within that decomposition of human relations, within that sociology of misery, goodness persists. There is a long monologue where Ikonnikov – the character who expresses the ideas of the author – casts doubt upon all social sermonizing, that is, upon all reasonable organization with an ideology, with plans … Every attempt to organize humanity fails. The only thing that remains undying is the goodness of everyday, ongoing life. Ikonnikov calls that 'little act of goodness' … This 'little goodness' is the sole positive thing … [I]t is a goodness outside of every system, every religion, every social organization.

See Levinas, *Is It Righteous to Be?* (n 235) 89. See also Michael L Morgan, *The Cambridge Introduction to Emmanuel Levinas* (Cambridge University Press 2011) 23.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

religion, every social organization".[240] It is "*a correction of the impersonality of a system that both tries to realize justice but also disregards the invisible tears of people who, despite all their efforts, fall outside of this whole*".[241]

In the context of public authorities' application of the law, it is this little goodness, this case-by-case correction to ensure justice in individual situations where the application of general laws might lead to injustice, that can be likened to the role of discretion. Indeed, discretion enables public officials to exercise a 'little goodness' that softens the hard edges of a generalised legal system. This is precisely why hardship clauses are sometimes incorporated in government policies, allowing public authorities to provide relief from the law's application or to deviate from its implementation based on the specific individual circumstances.[242] In this sense, discretion can counter the excesses of procedural rationality, by ensuring that rules are not applied overly rigidly, but with due regard to the ends they aim to serve, based on substantive rationality. It hence provides space to make trade-offs between different interests and values in a particular situation, without the need for the legislator to always anticipate those situations.

Discretion can arise due to other factors too. Certain risks that the legislator seeks to prevent or mitigate arise in situations of inherent complexity or uncertainty, requiring more comprehensive assessments by experts to identify the best course and timing of action. For instance, in the area of technology regulation, the legislator typically sets out a legal framework with general safety norms that should be met, but relies on public authorities and their risk analyses to develop more specific standards, procedures and guidelines. Furthermore, the fact that public authorities inevitably have limited resources at their disposal means they must optimise these resources in a way that allows them to reach their goals in the most efficient manner by setting certain priorities. Discretion has therefore been "*long identified as necessary for administrative authorities to operate effectively within all modern legal systems*".[243]

Note that the role of discretion applies not only to the implementation of legislation at the national level, but also to the implementation of European law. Indeed, when the EU adopts new legislation, certain legal provisions may explicitly or implicitly enable Member Sates' public authorities to exercise some discretion as

[240] Levinas, *Is It Righteous to Be?* (n 235) 89.

[241] Anckaert (n 238) 230.

[242] For instance, the Dutch tax authorities allow citizens to apply for hardship relief in tax cases "*when the law has a consequence that was unintended*" and "*which the legislator would have been able to prevent if it had anticipated such consequence*" (as detailed on the Dutch government's website <www.government.nl/topics/paying-taxes/applying-for-hardship-relief-in-tax-cases>.

[243] Karen Yeung and Lee A Bygrave, 'Demystifying the Modernized European Data Protection Regime: Cross-Disciplinary Insights from Legal and Regulatory Governance Scholarship' (2022) 16 Regulation & Governance 137, 148; Galligan, 'Discretionary Powers in the Legal Order' (n 213).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

regards the ways in which this legislation will be implemented at the national level. In this context, discretion hence also serves as a tool that allows individual Member States to make trade-offs between different values and interests in line with their national traditions, as long as they remain within the confines set out by the EU legislator and by primary EU law.[244]

Depending on the context in which public power is exercised, the executive's discretion can be more or less extensive. Let me concretise this with an example. Consider the legal rule in the area of Belgian migration law, which grants a migrant in Belgium the possibility to apply for a residence permit "*if exceptional circumstances justify the submission of this application in Belgium rather than abroad*".[245] In this case, discretion is rather vast: the law itself does not define a list of 'exceptional' circumstances, so it is up to the officials implementing the law to interpret this term, and to decide whether such circumstances are present. The relevant public authority (the Immigration Office, acting under the supervision of the 'political' executive) typically issues policy guidelines that set out which procedures will be followed in the implementation of this rule. On the government's website, one can, for instance, read that "*a long stay in Belgium, or integration into Belgian society, is not, in itself, an exceptional circumstance justifying an application for a residence permit in Belgium*".[246] This statement or policy reflects the executive's discretionary choice of implementing the law in this way rather than in another, as it is not an explicit part of the text adopted by the legislator. Had a more migration-friendly government been elected, this policy might, for instance, have been different.

Contrast this with the Belgian legal rule that for every child, regardless of income level, parents receive a standard sum of childcare benefits.[247] In this situation, the general rule to be applied is rather clear, and discretion is near inexistent: if an individual has a child, it is clear that the public official in charge needs to allocate that person the standard sum. However, even in this context, discretion is not *entirely* absent. One can still conceive that the responsible public authority may adopt guidelines or procedures that set out which type of evidence is accepted to prove the existence of a child, and which process must be followed to apply for such benefits if they are not allocated automatically. Accordingly, even within the scope of a single administrative act – the allocation of childcare benefits – public

[244] See Koen Lenaerts, Ignace Maselis and Kathleen Gutman, *EU Procedural Law* (Janek Tomasz Nowak ed, Oxford University Press 2015) 141.

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

[246] See the website of the Belgian Immigration Office: https://dofi.ibz.be/en/themes/third-country-nationals/residence-permit-articles-9-9bis-9ter/exceptional-circumstances.

[247] While these procedures are slightly different in Belgium's different regions, in Flanders this is governed by the Flemish Decree of 27 April 2018 (*Decreet tot regeling van de toelagen in het kader van het gezinsbeleid*), accessible at www.ejustice.just.fgov.be/mopdf/2018/07/31_1.pdf#Page140.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

authorities can have discretionary competences over some of the act's aspects, and bound competences over other aspects.[248]

As regards the adoption of administrative decisions, Galligan identifies three elements to the decision-making process, each of which allows for the exercise of discretion: (1) finding facts, (2) setting standards and (3) applying the standards to the facts.[249] Discretion is often associated with the second element (requiring the interpretation of vague or ambiguous legislation, or the creation of more specific standards based on broad legislation), and third element (requiring an element of judgment and assessment, even if the way in which the standard must be interpreted is clear). However, even the first element already implies some discretion (e.g. how will evidence regarding the facts be gathered and assessed?). Accordingly, drawing on the example above, discretion not only arises as regards the final decision (e.g. the allocation of childcare benefits) but also in various intermediary steps (e.g. how can the existence of a child be proven, how is the application process for benefits organised, how is the allocation calculated, and so on).

To summarise, public authorities and officials exercise their functions and take administrative acts based on a set of rules, yet they always have a varying level of discretion at their disposal which enables them to carry out their tasks in a way that is, ideally, both efficient *and* just. Of course, this does not mean that discretion is always exercised in a sound manner. The autonomy afforded by discretion also opens the door for deviations of the law and potential abuses. This is why, in line with the principles of the rule of law that I will discuss in Chapter 3, public authorities typically also adopt guidelines or procedures that set out how public officials should exercise discretion, thereby enhancing the predictability and consistency of the law's application across various departments within the organisation. Moreover, democratic oversight and judicial review play an important role in ensuring that discretion – whether to apply, interpret or deviate from the general law – is used commensurably with the rule of law and other constitutional values and principles.

With this conception in mind of how public authorities carry out their functions and adopt administrative acts, in between rules and discretion, let me now examine the role that algorithmic systems can play in this context.

2.3.4 *From Bureaucracy to Algocracy*

The processing of information and the adoption of administrative acts based on such information is a core task of public authorities. As discussed above, the aspiration to do so efficiently, objectively and rationally already underpinned the nineteenth-century model of administrative bureaucracy and remained an important aim of

[248] Steven Van Garsse (ed), *Handboek Bestuursrecht* (Politeia 2016) 41.
[249] Galligan, 'Senses of Discretion' (n 227) 9.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

public administrations ever since.[250] It is therefore no surprise that, as soon as affordable, algorithmic regulation became part and parcel of public authorities' working methods, aimed at rendering their information processing activities more efficient.

In this regard, Zuurmond highlights that "*bureaucracy and informatisation seem to go hand in hand*".[251] Indeed, bureaucracy relies on the execution of administrative tasks based on expert knowledge and information, which in turn requires the collection of data – including data about legal subjects. According to Peeters and Widlak: "*as state tasks expanded, especially in welfare states, so did the number of registrations and their importance. Knowing your citizens has never been more important as when you try to decide who is eligible to student grants, social security, health care, social housing, or pensions*".[252] Besides the collection of information, the need for the speedy processing of such information also set in motion a broader process of digitalisation. When public authorities started to deploy computer systems in the twentieth century, they realised that the information they sought to process first had to be converted from an analogue format to computer-readable code ('digitisation'). Only then, analytical and decision-making processes could be transformed to the digital realm too ('digitalisation').

Gradually, the image of bureaucrats sitting behind a stuffy pile of papers thus transformed into an image of bureaucrats sitting behind large computer screens. In early 2002, Bovens and Zouridis pointed out that "*window clerks are being replaced by Web sites, and advanced information and expert systems are taking over the role of case managers and adjudicating officers. Instead of noisy, disorganized decision-making factories populated by fickle officials, many of these executive agencies are fast becoming quiet information refineries, in which nearly all decisions are pre-programmed by algorithms and digital decision trees*".[253] The uptake of algorithmic regulation in public administration is hence nothing new. Yet over time, public authorities "*experienced several leaps of technological innovation*",[254] and the systems they relied upon became ever more sophisticated. Today, algorithmic regulation is not only based on basic decision trees, but also on more complex knowledge-driven systems and increasingly on data-driven systems too. Accordingly, a distinctive growth can be perceived both as regards the scale on which algorithmic regulation

[250] See Höpfl (n 183); Gailmard and Patty (n 176).

[251] Zuurmond (n 202) 2.

[252] Peeters and Widlak (n 186) 176.

[253] 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, 175.

[254] 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) 302.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

is used in the public sector, as well as the importance and impact of the acts that are being automated.[255]

Evidently, the introduction of algorithmic regulation also impacted the way in which public officials take decisions. Instead of making judgment calls regarding the application of the law to concrete situations, their role became increasingly focused on filling in electronic forms and templates, based on which an algorithm can compute certain outcomes. Bovens and Zouridis conceptualised this as a turn from 'street-level' bureaucracy (a term coined by Michael Lipsky to denote public officials interacting with and taking decisions about citizens)[256] to 'screen-level' bureaucracy. As the uptake of algorithmic regulation increased, and the computers assigned to public officials became a networked digital infrastructure that connects databases and processes across public authorities, 'screen-level' bureaucracy developed further into 'system-level' bureaucracy.[257] The advent of the internet and the Internet of Things also played an important role in this regard, as it progressively enabled citizens to directly provide information to public authorities (knowingly or not), thereby facilitating the collection of their data.

Today, reliance on algorithmic regulation can hence be called *systemic*, as it underpins the functioning of public authorities at large. Algorithmic systems can be used for a myriad of functions within public authorities, from the automated translation of text and the filtering of incoming mail, to the formulation of replies to citizen questions and the adoption of administrative acts. This does not mean that the adoption of administrative acts is necessarily *entirely* automatised. As noted in Section 2.1.1, algorithmic systems can be used for the *adoption* of administrative acts (decision-making sensu stricto), yet in most public authorities these systems are still primarily used to *inform* or *recommend* administrative acts. This also implies that, today, the collection, processing, analysis and assessment of data by public authorities, which lies at the heart of their decision-making processes, is primarily carried out by algorithms.

Initially, it was thought that the embrace of algorithmic systems would alter the nature of bureaucracy, or even lead to its collapse by giving rise to a radically new 'post-bureaucratic order'.[258] However, as already hinted at above and as noted by various scholars, "*rather than less bureaucracy, we seem to experience its propagation and expansion at every turn*".[259] Instead of being dissipated, the 'original' bureaucracy merely turned into an 'algorithmic' or 'digital' bureaucracy. Muellerleile and Robertson note that "*the digital bureaucracy is a world of data in motion, given*

[255] See also Caroline Lequesne-Roth, 'Livre blanc: La digitalisation du service public – Pour une éthique numérique inclusive' (Observatoire de l'éthique publique 2021).
[256] Michael Lipsky, *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services* (Russell Sage Foundation 1980).
[257] Bovens and Zouridis (n 253) 178.
[258] Höpfl (n 183).
[259] Muellerleile and Robertson (n 182) 187.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

direction and shape by new kinds of digital infrastructures – from codes to algorithms to platforms, whose digital footprint replaces the material archive, and whose experts are the new data scientists".[260] Lorenz and others even speak of an 'algocracy', arguing that, "*whereas the bureaucracy denotes the exercise of power through the office, the algocracy shows that power is exercised through algorithms*".[261]

The adoption of algorithmic systems by public authorities hence did not simply come down to the introduction of a new tool amongst others to rationalise public decision-making. It also fundamentally altered the organisation of public administrations[262] and brought along several side-effects. In Chapter 4, I will carry out an extensive analysis of the impact of those effects on one particular societal interest, namely the rule of law. In this section, already foreshadowing this analysis, I will only mention four general consequences of the uptake of algorithmic systems in the public sector.

First, one can observe that public authorities have increasingly formalised aspects of their decision-making processes, as such formalisation is necessary to express information and rules through computer-readable binary code.[263] As Bovens and Zouridis point out, "*a conditionally programmed legal framework will lend itself much easier to ICT applications than a goal-oriented legal framework*".[264]

Second, this algorithmisation also led to a reduction of discretion at the level of individual public officials. As previously described, algorithms rely on input and instructions to deliver certain outcomes, and are able to do so at scale. To benefit from these efficiencies of scale, public authorities hence seek to routinise and centralise processes.[265] Accordingly, "*many decisions are no longer made at the street level by the worker handling the case; rather, they have been programmed into the computer in the design of the software*".[266] Indeed, decisions are increasingly guided by algorithmic systems and databases,[267] whereby public officials are often "*no longer involved in handling individual cases, but direct their focus toward system development and maintenance, toward optimizing information processes, and toward creating links between systems in various organizations*".[268] This does not mean that administrative discretion has dissolved. Instead, the introduction of algorithmic regulation and the routinisation of decisions has pushed this discretion higher up the value chain, to the level of the designers and developers of the algorithmic

[260] ibid 190.
[261] Lorenz, Meijer and Schuppan (n 26) 72.
[262] See also Kennedy (n 106).
[263] Bovens and Zouridis (n 253) 178.
[264] ibid 181.
[265] See also Kennedy (n 106) 231.
[266] Bovens and Zouridis (n 253) 177.
[267] Lorenz, Meijer and Schuppan (n 26) 72.
[268] Bovens and Zouridis (n 253) 178.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

systems. I will come back to this point in the following chapters, given the implications this has on the rule of law.[269]

A third effect of the public sector's algorithmisation is the extensive process of quantification and datafication this brought along, particularly given the more recent uptake of data-driven systems.[270] Today, public authorities have more data than ever at their disposal to inform their policy and decision-making processes. Furthermore, the computational abilities of algorithmic systems allow them to process and analyse such data at unprecedented speed.[271] This also reinforces the bureaucratic tendency towards rationalisation, under the heading of 'evidence-based' decision-making. As Lorenz and others describe, the introduction of algorithmic systems "*enables government organizations to quantify the uncertainty inherent in decision-making processes on the basis of data analysis by expressing it as probability and to thereby further rationalize this process: even though there is no full certainty about a situation, a more rational choice can be made based on probabilities*".[272]

Finally, along with this rationalisation and datafication one can also discern a thoroughgoing formalisation of citizen registration and classification.[273] Algorithmic regulation typically requires that natural and legal persons are classified according to traits that are relevant for the regulation's application, based on the various datapoints that have been collected about them. This classification can be inserted into the system manually, but it can also be generated through a data-driven system programmed to identify patterns and, on that basis, classify citizens into various categories, and evaluate or score them.[274] Since these classifications are used to inform administrative acts, their contours are not without substantive consequences.

As noted by Peeters and Widlak, "*classifications are by their very nature contested, because they are abstractions and simplifications of a complex social reality that highlight certain elements of that reality while ignoring others*".[275] However, while humans know that concepts and classifications are mere social constructs, and that they only represent a selective and partial aspect of that which is being classified, this cannot be said of algorithmic systems. For instance, my mother is not just a 'mother', but also a 'wife', a 'daughter', a 'colleague', a 'friend', a 'consumer', a 'reader', a

[269] See *infra*, particularly Section 4.1.3.

[270] 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, 3.

[271] The gathering of data is also strongly promoted by international organisations such as the OECD, as it is believed that "*intelligent data usage offers a myriad of possibilities to fundamentally transform public sector activities, how services are designed, delivered and monitored*". See OECD, *The Path to Becoming a Data-Driven Public Sector* (OECD 2019) <www.oecd-ilibrary.org/governance/the-path-to-becoming-a-data-driven-public-sector_059814a7-en> 9.

[272] Lorenz, Meijer and Schuppan (n 26) 72.

[273] See also Hänold (n 152); Broomfield and Reutter (n 134).

[274] See also Lina Dencik and others, 'Data Scores as Governance: Investigating Uses of Citizen Scoring in Public Services' (Data Justice Lab, Cardiff University 2018).

[275] Peeters and Widlak (n 186) 176.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

'Belgian', a 'woman' and more. Depending on the particular context, one or more aspects of an individual will be particularly focused on, for example to determine the applicability of certain legal rules, even if that aspect is but one part of a more comprehensive picture.[276] It is therefore possible, yet always partial, to place people and things into a category, even if we do so on a daily basis.[277]

These conventional concepts, limited as they may be to reflect the richness of reality, are how we structure our world, and – by extension – our legal system. Yet as long as we find ourselves in an intersubjective environment, we can draw attention to the limitations of these concepts, contest them, explain why a certain categorisation is erroneous, provide nuance, or ask for additional options or categories given that we share the concept's *meaning*.[278] Such explanation and contestation is not possible when classifications are made by an algorithmic system. Furthermore, the registration of information into databases also tends to reverse the burden of proof when seeking to correct such information. Indeed, "*once something is registered, it is considered 'true' according to the principles of formal bureaucracy*".[279] And, as discussed above, the larger the database, the more chance that datapoints are erroneously registered or classified, which can evidently affect administrative acts that are based on such erroneous information.

This hints to the fact that algorithmic regulation not only generates benefits, but that it can also reinforce some of the pitfalls we already encountered with bureaucracy. I will revisit this issue expansively in further chapters. At this stage, suffice it to conclude that the logic of bureaucracy and the logic of algorithmic systems (and particularly the efficiency-oriented informatisation, standardisation, and rationalisation they bring along) seem rather aligned. While the uptake of algorithmic regulation does impact the way in which public officials perform their tasks, the core features of bureaucratic organisation that underpin those tasks appear to have remained in place.

Before concluding this chapter on algorithmic regulation, let me briefly make explicit what I have thus far mostly implied, namely the benefits that public authorities aspire to materialise when adopting algorithmic regulation.

2.3.5 *Rationale of Algorithmic Regulation*

"*What if more poor families were to receive the state benefits they need? What if more patients' lives can be saved? What if more children could be protected from situations*

[276] Consider, for instance, also the categorisation of behaviours that are 'legal' and 'illegal', or distinctions between people that are 'single', 'married' or 'divorced'. See in this regard also Larry Alexander, 'Scalar Properties, Binary Judgments' (2008) 25 Journal of Applied Philosophy 85.
[277] Bowker and Star (n 146).
[278] Smuha, 'The Human Condition in an Algorithmized World' (n 144) 31.
[279] Peeters and Widlak (n 186) 176.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

of abuse? What if more people can be rescued from imminent flooding? What if more terrorist attacks could be prevented?"[280] The OECD does not beat around the bush regarding the opportunities it sees in the adoption of algorithmic systems. As of yet, evidence of the concrete benefits of algorithmic regulation is scarce, particularly as regards data-driven systems, given the relative novelty of their scaled use.[281] As noted by van Noordt and Misuraca, "*many of the benefits ascribed to AI for the public sector are not always based on empirical data, but often rely only on assumptions*", rendering it difficult to validate and assess the actual benefits of the adoption of algorithmic regulation in public authorities.[282] In what follows, I discuss the aspired or expected benefits they generate, without taking the materialisation of such benefits for granted.

First, as already pointed out, the automated nature of algorithmic regulation means that its computations can be carried out on a vast scale (even population-wide), thereby enabling mass decision-making.[283] As Schartum notes,

> Of course, it is possible to imagine many more office buildings where thousands of men and women would do all the detailed processing of individual cases that are processed today by computers, but this alternative is not very realistic: Modern taxation systems, national social insurance schemes and management of many other welfare programs would not be feasible without the use of computers and the algorithmic law that is integrated in the software.[284]

Accordingly, by processing vast amounts of data in a short amount of time, algorithmic systems may increase the efficiency of processes, as the speed of their computations supersedes that of human decision-making. These efficiencies are also believed significantly to reduce costs given that "*cognitive technologies could free up hundreds of millions of public sector worker hours*".[285]

Second, the ability to peruse a large amount of data in a short amount of time also means algorithmic systems can help optimise decision-making processes.[286] Public authorities possess more data than ever before, yet this easily leads to an excess of

[280] 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.

[281] Colin van Noordt and Gianluca Misuraca, 'Artificial Intelligence for the Public Sector: Results of Landscaping the Use of AI in Government across the European Union' [2022] 39 Government Information Quarterly 101714.

[282] ibid.

[283] See also Fleur Johns, 'Governance by Data' (2021) 17 Annual Review of Law and Social Science 53.

[284] Schartum (n 254) 301.

[285] Peter Viechnicki and William D Eggers, 'How Much Time and Money Can AI Save Government?' (Deloitte Center for Government Insights 2017).

[286] See Svenja Falk, *Digital Government: Leveraging Innovation to Improve Public Sector Performance and Outcomes for Citizens* (Springer Berlin Heidelberg 2016).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

information when resources to analyse such data are limited. As Michèle Finck puts it, "*in an age of informational overload through the continuous generation of ever more data, computational learning may become the only means of making sense of data the quantity of which exceeds the capacities of human cognition*".[287] As discussed above,[288] data-driven systems are indeed able to identify patterns in data that public officials may not be able to see, and can on that basis provide recommendations that could in principle enable public authorities to reach their goals more effectively. Consider the example of an algorithmic system used by the French tax authorities to identify non-declared swimming pools based on Google maps images, thereby recuperating about 10 million euros of non-paid taxes in a short amount of time.[289]

Third, algorithmic regulation is sometimes introduced with the explicit aim of reducing the risk of partial, erroneous, arbitrary, or biased decision-making, and to decrease the risk of corruption.[290] By codifying the application of rules designed to address these risks into algorithmic systems, an informational architecture is created that can prevent public officials to deviate from the law *by design*.[291] Furthermore, the automated application of legal rules could also lead to their more consistent application, as the same rule will no longer be applied by different public officials who might have their own interpretation or judgment.[292] It has been argued that, in this way, algorithmic regulation might not only enhance the law's effectiveness, but also public authority's legitimacy.[293]

Fourth, algorithmic systems could take over tasks that are highly repetitive and intellectually unstimulating for public officials, or tasks that are dangerous and would

[287] Michèle Finck, 'Automated Decision-Making and Administrative Law' in Peter Cane and others (eds), *The Oxford Handbook of Comparative Administrative Law* (Oxford University Press 2020) 659.

[288] See *supra* Section 2.1.3.

[289] 'La détection par intelligence artificielle de piscines non déclarées va être généralisée en France' *Le Monde.fr* (29 August 2022) <www.lemonde.fr/pixels/article/2022/08/29/experimentee-dans-neuf-departements-la-detection-de-piscines-non-declarees-par-intelligence-artificielle-va-etre-generalisee_6139439_4408996.html>.

[290] See for instance Nils Köbis, Christopher Starke and Iyad Rahwan, 'The Promise and Perils of Using Artificial Intelligence to Fight Corruption' [2022] Nature Machine Intelligence <www.nature.com/articles/s42256-022-00489-1>.

[291] See the discussion in this regard in Roger Brownsword, 'Technological Management and the Rule of Law' (2016) 8 Law, Innovation and Technology 100; Karen Yeung, 'Can We Employ Design-Based Regulation While Avoiding Brave New World?' (2011) 3 Law, Innovation and Technology 1.

[292] See also Thomas J Barth and Eddy Arnold, 'Artificial Intelligence and Administrative Discretion: Implications for Public Administration' (1999) 29 The American Review of Public Administration 332; Justin B Bullock, 'Artificial Intelligence, Discretion, and Bureaucracy' (2019) 49 The American Review of Public Administration 751.

[293] Irina Pencheva, Marc Esteve and Slava Jankin Mikhaylov, 'Big Data and AI – A Transformational Shift for Government: So, What Next for Research?' (2020) 35 Public Policy and Administration 24, 28.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

unduly expose human beings to risks,[294] such as cleaning a nuclear site.[295] Reports that highlight the advantages of such systems typically emphasise that this frees up valuable time which public officials can spend on safer or more interesting tasks, or for better interactions with citizens in need, rather than replacing their jobs.[296]

Fifth, algorithmic regulation, especially when based on data-driven systems, can also enable the personalisation of administrative acts at lower cost, thereby reconciling the massive scale of decisions that public authorities must take, with the need for individual tailoring.[297] More generally, the algorithmisation of the public sector and establishment of well-kept interoperable databases should avoid the need for citizens to provide the same data multiple times to different authorities,[298] and might enable public authorities to provide their services proactively based on the information they have, without the need for a citizen to request a service, let alone the need to physically visit the authority's office.[299]

Importantly, all of these benefits must be considered against a background of increasing pressure that public officials face to 'do more with less'. Their responsibilities keep on increasing, while public spending is significantly being cut. As early as 2018, a survey conducted by McKinsey revealed that "*43 percent of all public sector transformation efforts over the past five years have had cost reduction as a core goal*".[300] The global recession and the war in Ukraine arguably only magnified this financial pressure. At the same time, I must underline that the aim to cut back public spending by adopting technology is by no means a new development. Already

[294] See, e.g., Kai-Fu Lee, 'AI's Real Impact? Freeing Us from the Tyranny of Repetitive Tasks' [2019] *Wired* <www.wired.co.uk/article/artificial-intelligence-repetitive-tasks>.

[295] H2020, 'The Robots Are Coming to Clean up Our Nuclear Sites' (*CORDIS – European Commission*), 2019. <https://cordis.europa.eu/article/id/358596-the-robots-are-coming-to-clean-up-our-nuclear-sites>. See also Rob Spencer, 'Got a Dirty, Dangerous, Dull Job? Let a Robot Do It and Keep Your Workers Safe' (2002) 20 Robotics World 14.

[296] As discussed further below, the introduction of algorithmic systems, however, often stems from a desire to cut costs – including personnel costs. In the abovementioned example of the French swimming-pool detection system, trade unions in fact expressed their concerns around the system's use, fearing that it will be used to avoid recruiting new officials in a context of a continuous decline in staff since several years. See 'La détection par intelligence artificielle de piscines non déclarées va être généralisée en France' (n 289).

[297] See Pascal D König, 'Dissecting the Algorithmic Leviathan: On the Socio-Political Anatomy of Algorithmic Governance' (2020) 33 Philosophy & Technology 467, 470. See also Horst Eidenmueller, 'Why Personalized Law?' (Social Science Research Network 2021) SSRN Scholarly Paper ID 3969934 <https://papers.ssrn.com/abstract=3969934>.

[298] See, e.g., OECD, *The Path to Becoming a Data-Driven Public Sector* (n 271). See also Schartum (n 254) 322.

[299] See, e.g., Regina Sirendi and Kuldar Taveter, 'Bringing Service Design Thinking into the Public Sector to Create Proactive and User-Friendly Public Services' in Fiona Fui-Hoon Nah and Chuan-Hoo Tan (eds), *HCI in Business, Government, and Organizations: Information Systems* (Springer International Publishing 2016).

[300] Tera Allas, Roland Dillon and Vasudha Gupta, 'A Smarter Approach to Cost Reduction in the Public Sector', McKinsey (8 June 2018) <www.mckinsey.com/industries/public-and-social-sector/our-insights/a-smarter-approach-to-cost-reduction-in-the-public-sector>.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

in the 1980s, under the influence of the New Public Management movement (NPM) which espoused the integration of private sector management ideas into public administration,[301] algorithmic systems were increasingly embedded in public processes. NPM reinforced the ideal of efficiency and imbued it with private sector tools such as market-based mechanisms, performance indicators, outsourcing and procurement, customer-service orientation, but also budget cuts and performance management of staff. The adoption of digital technology was seen as an important part thereof.[302] While NPM has been subjected to criticism,[303] the push towards performance indicators and cost-savings remained, and similar ideas also permeated the subsequent 'digital-era governance' and 'e-governance' movements, which more explicitly focused on efficiency improvements through automated data analysis and electronic platforms.[304]

Furthermore, a strong push for the uptake of algorithmic systems in the public sector also came from the European Commission. In 2009, EU Member States signed the *Malmo Declaration on eGovernment*,[305] which the Commission implemented through the *European eGovernment Action Plan 2011–2015*[306] and the *eGovernment Action Plan 2016–2020*,[307] each underlining the need to accelerate the digital transformation of governments. The *Tallinn eGovernment Declaration* in 2017 added an important impetus for Member States and the Commission to continue investing in the modernisation of the public sector, seen as indispensable to increase "*the transparency, responsiveness, reliability, and integrity of public governance*".[308]

Most recently, with its Coordinated Plan on AI in 2019[309] and 2021,[310] the European Commission urged all EU Member States to adopt their own national

[301] Broomfield and Reutter (n 134) 3.
[302] Kennedy (n 106) 211.
[303] See also P Dunleavy, 'New Public Management Is Dead – Long Live Digital-Era Governance' (2005) 16 Journal of Public Administration Research and Theory 467.
[304] OECD, 'A Data-Driven Public Sector' (n 280) 9.
[305] 'Ministerial Declaration on EGovernment, Approved Unanimously in Malmö, Sweden, on 18 November 2009', Malmö, 2009, <www.mt.ro/web14/documente/date-deschise/reglementari/Ministerial-declaration-on-egovernment_Malmo_2009.pdf>.
[306] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: The European eGovernment Action Plan 2011–2015 – Harnessing ICT to promote smart, sustainable & innovative government 2010 (COM(2010) 743 final).
[307] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: EU eGovernment Action Plan 2016–2020 – Accelerating the digital transformation of government 2016 (COM(2016) 179 final).
[308] 'Tallinn Declaration on EGovernment at the Ministerial Meeting during Estonian Presidency of the Council of the EU on 6 October 2017', Talinn, 2017, <https://digital-strategy.ec.europa.eu/en/news/ministerial-declaration-egovernment-tallinn-declaration>.
[309] European Commission, 'Coordinated Plan on Artificial Intelligence', Brussels, 7.12.2018, COM(2018) 795 final.
[310] European Commission, 'Coordinated Plan on Artificial Intelligence: 2021 Review. Fostering a European Approach to Artificial Intelligence' Annex to the Communication from the

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

AI strategy, including a plan for the technology's adoption in the public sector.[311] The EU also explicitly finances the uptake of algorithmic regulation by Member States through the recovery and resilience facility that was established to meet the financial challenges raised by the Covid-19 pandemic,[312] and more indirectly through other EU budgets.[313] In sum, Member States have reasons enough to adopt algorithmic regulation, and it can only be expected that this trend will be accelerated in the years to come.

A caveat should be made, as the benefits aspired by algorithmic systems are not automatically achieved.[314] After all, the *actual* benefits of these systems entirely depend on how they are designed, developed and used, and whether they take into account the risks set out in previous sections. It is not because a system is developed or deployed with good intentions, that it is also developed and deployed in a good manner, with due attention to unintended consequences or problematic uses later on. Moreover, even when benefits are achieved, this does not mean they actually benefit *all*. Often, those who already find themselves in a beneficial position will be best placed to reap those benefits, whereas those who are in a vulnerable position may not necessarily be better off,[315] a point that I will revisit later on.

In addition, the capacities of algorithmic systems, and particularly systems defined as 'AI', are sometimes oversold,[316] and hyperbolic statements about their benefits, followed by disappointing results, are not uncommon.[317] Yet once significant investments have been made to establish an algorithmic regulation project, sunken costs and path dependencies render it difficult to abandon it, even if the output is not up

Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions COM(2021) 205 final 46.

[311] It can be noted that these strategies often present the digital transformation as enabling 'progress' towards an ever-better human condition, a view that bears strong affinities with progressivism. See in this regard also Smuha, 'The Human Condition in an Algorithmized World' (n 144) 26.

[312] Regulation (EU) 2021/241 of the European Parliament and of the Council of 12 February 2021 Establishing the Recovery and Resilience Facility.

[313] Consider, for instance, the many AI projects funded through the H2020 and HorizonEurope research programmes of the European Union.

[314] Smuha, 'The Human Condition in an Algorithmized World' (n 144) 10.

[315] See also Sahajveer Baweja and Swapnil Singh, 'Beginning of Artificial Intelligence, End of Human Rights' (*LSE Human Rights*, 16 July 2020) <https://blogs.lse.ac.uk/humanrights/2020/07/16/beginning-of-artificial-intelligence-end-of-human-rights/>.

[316] Kate Crawford also describes the phenomenon of so-called 'Potemkin AI', whereby a product is sold as an autonomous system for marketing purposes, but in fact primarily relies on human labour behind the scenes, often in very dire labour circumstances. See Crawford (n 125) 65.

[317] A recent example is the use of algorithmic systems to help counter the Covid-19 pandemic. Soon after Covid-19 broke out, numerous tech developers enthusiastically started designing and deploying AI systems with great expectations of how these could be used against the virus. However, the results were disappointing, and AI was not able to deliver its promise. See in this regard Will Douglas Heaven, 'Hundreds of AI Tools Have Been Built to Catch Covid. None of Them Helped' (*MIT Technology Review*, 30 July 2021) <www.technologyreview.com/2021/07/30/1030329/machine-learning-ai-failed-covid-hospital-diagnosis-pandemic/>.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

to standard. Finally, it is precisely the desire to reap the benefits of this technology, often with the explicit aim to increase individual and societal welfare, that can blind public authorities to its risks and that eventually gives cause to concerns.

2.4 CONCLUDING REMARKS

In the sections above, I delineated the concept of algorithmic regulation for the purpose of this book by discussing its technical and societal aspects, and by examining how public authorities rely thereon within their broader organisational environment. I started by describing algorithmic systems as the building blocks of algorithmic regulation, and explained that these are essentially comprised of *input*, algorithmic *instructions* and *output*. I distinguished knowledge-driven systems from data-driven systems and discussed some of the differences in both approaches, whilst simultaneously cautioning against their strict distinction, since algorithmic systems can rely on a mixed approach. Moreover, regardless of the underlying approach and of whether they are considered 'intelligent' enough to fall under the AI umbrella, the impact of these systems on individual, collective and societal interests can be significant, especially when used by public authorities. Therefore, I decided to focus my analysis on all algorithmic systems that are used by public authorities to inform or take administrative acts, regardless of their underlying approach.

Since algorithmic systems are not isolated entities but part of a broader environment, I complemented their technical description with a discussion of some of their societal characteristics, which cannot be seen as separate therefrom. It is only by clarifying the underlying socio-technical infrastructure in which these systems are embedded that the mutual influence of algorithmic systems and society upon each other can be made more visible. I therefore emphasised the need to examine the implicit and explicit human choices that underlie the systems' design and use – and hence their affordances – as well as the power relationships that shape these choices. The extent to which algorithmic systems are opaque, for instance, relies not merely on the technique underlying their functioning, but also on how developers and deployers communicate about the normative choices they made throughout the systems' design and use.

I also highlighted several risks associated with the use of algorithmic regulation, including human errors or mistakes, (unintended) bias and discriminatory outcomes, the impact on human agency, and emphasised the wide-ranging effects these risks can have given not only the systems' opacity but also the scale of their deployment. In addition, I discussed the systems' dependency on data, and the importance of keeping in mind that not all social phenomena can easily be captured by quantifiable metrics. This broader picture of the societal aspects of algorithmic systems, which also highlights their function as *regulatory* tools broadly speaking, is essential to understand the concrete effects they can have on their environment.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005

Finally, I examined how algorithmic regulation is used by, and influences the organisation of, public authorities. I first discussed how public authorities function, and highlighted their bureaucratic environment, observing that the logic of bureaucracy is in certain aspects very similar to the logic of algorithmic systems. Both 'systems' are underpinned by a drive towards procedural rationality and efficiency – to the potential detriment of substantial rationality. I also emphasised the important role of discretion in public decision-making – particularly the way in which it can counter some of the excesses of procedural rationality – and the need for public authorities to have both political oversight over their actions and an 'internal morality' to ensure they execute their tasks both efficiently *and* justly.

Across this chapter, I provided various illustrations of how algorithmic systems can be used in the public sector for a diversity of tasks, including managing public welfare programmes, conducting criminal investigations, evaluating asylum applications or assessing tax fraud. While most applications today are still primarily focused on *informing* rather than *adopting* administrative acts, the trend of ever-increased reliance on this technology provides strong indications that in the next decade, the automated adoption of administrative acts will become part and parcel of public administration. In this book, I will therefore analyse how the rule of law can be impacted by algorithmic regulation, conceptualised as the reliance on algorithmic systems either to *inform* or to *adopt* administrative acts. Having clarified what this practice entails, I can now move on to a conceptualisation of the rule of law.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 13 Nov 2026 at 18:24:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009427500.005