INFORMATION LAW AND AUTOMATED GOVERNANCE

Keynote address at the Information Law Conference

Institute of Directors, 24 April 2023

Lord Sales, Justice of the UK Supreme Court[*]

A. INTRODUCTION

Government is becoming increasingly automated, and for good reason in terms of speed, efficiency and cost. But this trend brings problems and challenges as well.

Perhaps the best known algorithmic tool so far to be used in UK public services is the system initially developed for grading A-levels in 2020.[1] By applying an algorithm within parameters set by Ofqual, the exams regulator, exam boards generated grades for each student using a range of data including historic performance at the student's school and rankings of students provided by their teachers. But there are many other examples.

The Public Law Project recently published its 'Tracking Automated Government Register',[2] a database recording details of various algorithms used by the executive to make or inform decisions in a range of policy areas. It has identified over 40 such algorithms across the public sector.

For example, the Home Office and Border Force use a tool called 'Border Risk and Targeting Capability' to identify risks at the border of the United Kingdom. These include risks of fraud, crime, and illegal migration.

[*] I am indebted to Raphael Hogarth for his invaluable assistance in preparing this address.

[1] A tribunal case, relating to a Freedom of Information Act request for certain information relating to the operation of that algorithm, is ongoing. See the First-tier Tribunal's decision, https://informationrights.decisions.tribunals.gov.uk/DBFiles/Decision/i3045/047%20130622%20DECISION-Dismissed.pdf (EA/2021/0234), which is understood to be subject to appeal to the Upper Tribunal.

[2] https://trackautomatedgovernment.shinyapps.io/register/.

According to the Home Office, this tool "*ingests large volumes of data*" from various data streams, then "*analyses, structures, and matches the data so that Border Force officers have the information they need to find patterns and generate intelligence. Much of this data analysis and processing is automated [...]*." [3]

The Department for Work and Pension's "*Advance Payments Fraud Machine Learning Tool"* has been used to *"analys[e] information from historical fraud cases to predict which cases are likely to be fraudulent in the future. Cases scored as potentially fraudulent by the model are flagged to caseworkers, who then prioritise the review and processing of such cases accordingly.*"[4]

Such technology is being used at a local level, too. Durham Constabulary has reported on the use of a 'Harm Assessment Risk Tool' (HART) to predict how likely an individual is to commit a violent or non-violent offence over the next two years. The tool based its predictions on 34 pieces of data, including relating to the person's past criminal history and personal characteristics such as age, gender, and postcode.[5]

Policing applications of artificial intelligence have been considered by the Courts. ***R (Bridges) v Chief Constable of South Wales*** concerned the use of automated facial recognition technology by the South Wales Police, to compare faces captured by CCTV cameras to images on police watchlists.

According to a press release by ViewSonic Corporation, it has partnered with a school academy in Wolverhampton to deploy an artificially intelligent analysis tool in the classroom, which is said to be capable of "*analyzing students' attentiveness [and focus] through articulated human pose estimation*".[6]

[3] https://hodigital.blog.gov.uk/2022/01/31/analysing-data-to-identify-risks-at-the-uk-border/

[4] https://www.nao.org.uk/wp-content/uploads/2022/07/Report-on-Accounts-Department-for-Work-and-Pensions.pdf

[5] M. Oswald, J. Grace, S. Urwin & G.C. Barnes, 'Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality', (2018) 27:2 Information & Communications Technology Law, 223-250.

[6] https://www.prnewswire.com/in/news-releases/viewsonic-s-myviewboard-sens-brings-uk-s-first-ai-powered-classroom-to-smestow-academy-896798309.html.

It is difficult to say how widespread the use of automation technology has become in public administration. The range is clear from these examples. The trend is international.[7]

B. BENEFITS AND RISKS

It is inevitable that the use of such technologies by the state will become more widespread in view of the potential benefits. Four stand out.

The first is efficiency. Automation of decisions speeds them up. It also reduces the labour costs of decision-making. That can be good news for the citizen wanting a decision in their case. It is good for the public body's accounts, which is a powerful incentive in an era of stretched public resources.

Secondly, an algorithm may be able to make decisions far more accurately than a human decision-maker. Research on the HART system in Durham found that artificial intelligence was "*more accurate than experienced police officers in predicting the likelihood of criminals reoffending*".[8]

Thirdly, automated processes can be programmed to leave a good audit trail to allow post hoc review of decisions. If systems are properly constructed, that could facilitate the giving of reasons to the citizen, allowing them - and reviewers in the case of dispute - to understand how power has been exercised in their case.

Fourthly, through accuracy, automation of decision-making is capable of promoting the rule of law. One aspect of this is the elimination of capriciousness through the consistent application of rules. Where the volume of decisions is very large, like in the immigration or social welfare

[7] G. Misuraca and C. Van Noordt, 'AI Watch - Artificial Intelligence in public services' (2020) EUR 30255 EN, Publications Office of the European Union, Luxembourg, identified 230 initiatives using artificial intelligence in public services across Europe; this was only a sample <https://publications.jrc.ec.europa.eu/repository/handle/JRC120399>/.

[8] https://www.durham.police.uk/News/News-Articles/2022/January/AI-can-predict-reoffending-university-study-finds.aspx

contexts, a human decision-maker would not be able to check their reasoning against the reasoning of all past decisions to make sure that they are being consistent. Humans are also liable to make decisions based upon their subjective will or whim, even if only subconsciously.

In line with the objective of the law itself, automation of discretionary decision-making converts acts of will into the operation of a code of rules.

That is all good in theory. Yet it would be naïve to suppose that the spread of automated - extending increasingly into artificially intelligent - technology through government will necessarily bring us closer to an ideal of administrative decision-making. The risks of automation are as important as the potential benefits. There is no avoiding a risk-benefit analysis highly attuned to each specific context.

There are three salient problems. The first is the opacity of automated decision-making. In an influential article, Jenna Burrell has written that this opacity takes three forms.[9]

There is *intentional* opacity, where the workings of an algorithm are deliberately concealed by its developers in order to protect their intellectual property. There may be "*opacity as technical illiteracy*". Understanding algorithms requires specialised skills, which most people do not have. Finally, there is also *intrinsic opacity* – an opacity that "*arises from the characteristics of machine learning algorithms and the scale required to apply them usefully*". Particularly in the case of artificial intelligence and machine learning systems, the workings of the technology can elude even the computer scientists who created it.[10]

[9] J. Burrell, 'How the machine 'thinks': Understanding opacity in machine learning algorithms' (2016) 3:1, Big Data & Society.

[10] Simon Chesterman gives the example of AlphaGo, a computer programme developed by Google to play the complex board game 'Go'. The programmer could not explain how it devised the strategies it used to defeat a grand master in 2016: S. Chesterman, *We, the Robots? Regulating Artificial Intelligence and the Limits of the Law* (CUP, 2021), 65.

Opacity is the reverse of the audit trail point in terms of transparency to the citizen and to a reviewer.

In the case of ***Wisconsin v Loomis*** the Supreme Court of Wisconsin was called upon to decide whether it had been lawful for a Court to sentence Mr Loomis for criminal offences with the benefit of a 'risk assessment' produced by a controversial algorithm which uses historic data about other individuals to make inferences about which offenders pose the highest risk of recidivism.[11] One justice commented that the opacity of the algorithm, which prevented the court from understanding how it worked, was a significant problem.

A second risk is the potential for substantively bad decisions of various types, which is to say mistakes. This is the reverse of the accuracy point. Mistakes could arise from:

a. Systems that misunderstand, or use defective methodology to analyse, the information inputted into them; or

b. Systems insufficiently flexible or sensitive to accommodate themselves to the messy facts of individual cases.

A 2021 UK Supreme Court case, ***HMRC v Tooth***,[12] illustrates the problem. A taxpayer had entered a loss item not in the box of a form in which HMRC said it should have been entered, but rather in a 'white space' included in the form to ensure that the taxpayer was not constrained by the limitations of boxes.[13] The Revenue argued that the resultant inaccuracy in one part of the form, when read in isolation, should be regarded as a "*deliberate inaccuracy in a document given to [HMRC]*",[14] the relevant statutory test. It relied on the fact that "*the return would be read upon receipt at the Revenue by a computer rather than, initially at least, by a sentient, literate, human being. Computers, it was said, do not do contextual interpretation, but look at each part of, or box in, the return separately*".[15] The Supreme Court rejected that argument. A document written in the English language did not, the Court

[11] 881 N.W.2d 749 (Wis. 2016).
[12] [2021] UKSC 17, [2021] 1 WLR 2811.
[13] §51.
[14] Section 118(7) Taxes Management Act 1970.
[15] §49.

found, have a different meaning because it is read by a computer. The more realistic analysis was simply that the computer had made a mistake.

Mistakes by machines have also been responsible for litigation against the government in Australia. A system nicknamed 'robo-debt' has been used to calculate and collect debts owed on account of over-payment of welfare benefits. The system would average out certain income data held by the tax authority and use this to calculate what individuals should have been paid. The system would then raise a debt if this process indicated that there had been an over-payment. As a result of the averaging process, many debts were wrongly issued. This led to a test case,[16] a class action[17] and a Royal Commission inquiry.[18]

Some bad outcomes may be harder to spot and more systemic than such individual mistakes. There are forms of algorithmic 'bias', a term which covers the conceptually distinct questions of whether an algorithm treats different groups differently or whether it does so unjustly.[19] The root concern is that the algorithm operates so as to *unfairly* or *unjustly* prejudice a group with some characteristic, particularly a protected characteristic such as a particular ethnicity or gender.

Injustice may enter into a system in various ways. Systems may absorb and reproduce the legally or morally objectionable prejudice of their creators. In 'supervised learning' systems, a human operator feeds the algorithm with example cases, telling the algorithm which decision ought to be made in each case, allowing it to learn how to weight factors in making future decisions. Prejudice may have infected the labelling practice of the humans who fed the model.[20]

[16] ***Amato v The Commonwealth of Australia***, VID611/2019. See consent order at https://www.comcourts.gov.au/file/Federal/P/VID611/2019/3859485/event/30114114/document/1513665.

[17] ***Prygodicz v Commonwealth of Australia*** (No 2) [2021] FCA 634. See approval of settlement at https://gordonlegal.com.au/media/1365/prygodicz-v-commonwealth-of-australia-no-2-2021-fca-634-11-june-2021.pdf.

[18] See the inquiry website: https://robodebt.royalcommission.gov.au/.

[19] J. Susskind, *Future Politics* (OUP, 2020), 280.

[20] Qing and Lim give the hypothetical example of technology that is trained to grade examination papers. There is research indicating that human essay markers have "*prejudices on the linguistic choices of students which signify membership in demographic groups*"; accordingly they argue that "*automatic essay grading models might then be trained on a dataset of essays with the corresponding scores assigned by such human essay graders, thus incorporating the biases of the*

Machine learning cannot, absent human intervention, tell the difference between a link between variables that should be learnt and applied, and a link between variables that should be condemned and ignored. It is no solution to force models to close their eyes to protected characteristic data, because the model may in practice rely on these characteristics by relying on other information associated with them.

Datasets created in other ways to train automated models may also reflect latent prejudice, for instance because they are statistically imbalanced. Jamie Susskind gives the examples of algorithms to identify human faces which struggle to recognise the faces of non-white people if they are trained using majority-white faces, or voice recognition algorithms that will not 'hear' women's voices if they are predominantly trained on men's.[21] In complex systems, the route to a discriminatory outcome can be circuitous and poorly signed.

The third risk is more abstract. The individual's interaction with the state may be devalued or degraded by its mechanisation. While automation removes the caprice of a human decision-maker, it also removes capacity for empathy and respect.

One of the purposes of administrative law is to protect the dignity of the individual in his or her interaction with the state. Sabino Cassese writes of administrative law having a tripolar nature.[22] The "*original, traditional pole*" is that of "*public power*": administrative law as an "*instrument for securing obedience*". The second is the "*social*" pole of administrative law: the use of that instrument to enable society to realise its collective goals, and not merely those of the ruling elite. The third component is the "*liberal*" pole, which defines the limits of public power. This operates from the perspective of the citizen. This aspect of administrative law is rooted in individual dignity.

Accordingly, it may be a requirement of procedural fairness that, in certain circumstances, a public body should hold an oral hearing at which an individual affected by a decision can make

humans into the models": Qing and Lim, 'A legal framework for artificial intelligence fairness reporting' [2022] CLJ 610.

[21] J. Susskind, *Future Politics* (OUP, 2020), 282.

[22] S. Cassese, '"Le Droit Tout Puissant et Unique de la Société": Paradoxes of Administrative Law' (2010) 22:1, European Review of Public Law.

representations. Lord Hoffmann said that one of the two purposes of a fair hearing is the avoidance of the "*sense of injustice*" which the person who is the subject of the decision will otherwise feel.[23] In *Osborn's* case[24] Lord Reed cited Jeremy Waldron's statement that applying a law to an individual embodies a "*crucial dignitarian idea*", of "*respecting the dignity of those to whom the norms are applied as beings capable of explaining themselves".*[25]

Not every administrative decision touches the dignity of the individual so as to engage this level of procedural protection. For instance, automated systems for enforcement of parking regulations depend on photos of licence plates and offending behaviour.[26] It is doubtful that it will be an affront to the dignity of the citizen to take a photograph of their number plate and send them a fine for parking on a double yellow line, at any rate if they are given the option – rarely taken – to challenge the decision and offer special extenuating reasons for consideration.[27] The impact of dignitarian considerations on systems design is a topic awaiting more detailed academic and judicial exploration.

C. TENSIONS IN REGULATING AUTOMATED DECISION-MAKING

The challenge is how to regulate the use of automation in decision-making in a way that allows administrators to pursue the benefits while protecting the public from the risks from opacity, mistakes, inflexibility, injustice, and debasement of the state's respect for the individual.

There are two overarching tensions or thematic challenges inherent in that exercise. The first is a tension between maximising efficiency through automation and preserving flexibility and empathy through human involvement. The second is the tension between two different models of legal regulation: administrative law and the law of data protection.

[23] ***Secretary of State for the Home Department v (AF (No 3)*** [2009] UKHL 28; [2010] 2 AC 269, para 72.
[24] ***Osborn v Parole Board*** [2013] UKSC 61, [2014] AC 1115.
[25] J. Waldron, 'How Law Protects Dignity' [2012] CLJ 200, 210.
[26] M. Bovens and S. Zouridis, 'From Street-Level to System-Level Bureaucracies: How Information and Communication Technology Is Transforming Administrative Discretion and Constitutional Control' (2002) 62:2 Public Administration Review, 174-184.
[27] See the Civil Enforcement of Parking Contraventions (England) Representations and Appeals Regulations 2007, Regulation 4(2)(b)(ii).

A spectrum of human involvement

The first tension is that the greater the human involvement to avoid certain risks of automation, the less the efficiency gains will be. The greater the human involvement in decision-making, the lower the risk of rigidity, inflexibility and the 'computer says no' type of mistake, of the kind that occurred in ***Tooth*** and the robo-debt cases. However, greater human involvement will also make decision-making more time-consuming and expensive. There is a spectrum: human involvement is not 'all or nothing'.

The artificial intelligence literature discusses various levels of human involvement, sometimes labelled as "human in the loop", "human over the loop", and "human out of the loop".

"Human-in-the-loop" decision-making means human decision-making supported by the automated system. This is a structure that gives lawyers comfort. The contribution of the model is to provide information, but the locus of decision-making remains the human mind.

The court in the ***Loomis*** case laid heavy emphasis on the need to retain the sentencing judge 'in the loop'.[28] The opacity of the algorithm was not unfair to Mr Loomis, because once the report was before the ultimate decision-maker, Mr Loomis could "*review and challenge the risk scores*" produced by the technology before the judge.[29] The Court was also presented with some evidence, albeit disputed by the developers of the system, that the tool was more likely to be inaccurate in respect of black defendants than white ones, to their detriment.[30] This worry, too, could be overcome if sentencing courts were properly informed about the limitations of the technology and the existence of evidence suggesting a racial skew, in order to "*better assess the accuracy of the assessment and the appropriate weight*" to give to the fruits of the automated analysis.[31] Loomis's sentence was ultimately upheld on the basis that the automated risk assessment was "*not determinative*" in deciding the nature or severity of the sentence: the judge, in the end, exercised judgment.[32]

[28] §35.
[29] §53.
[30] §63.
[31] §66.
[32] §109.

"Human over the loop" is a model whereby a human oversees the operation of the automated system, and intervenes where necessary, but not by default.

This is the model contemplated by Article 22 of the General Data Protection Regulation ("**GDPR**"), and Article 22 of the UK GDPR, at least in respect of some automated decision-making. That article says that the data subject "*shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her*", unless the decision is (a) necessary for contractual purposes, (b) authorised by member state law which lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, or (c) based on the data subject's explicit consent. In the case of solely automated decision-making grounded in contract or consent, the GDPR states that those "suitable measures" to safeguard the data subject's rights and interests must include "*at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision*".

That is, the Regulation contemplates human involvement not *ab initio*, but human intervention on request.

Another example discussed in the literature is decision-making about student finance in Sweden. Though some aspects of decision-making by the Swedish National Board of Student Finance are only partly automated, some of their decisions are fully automated, such as decisions on loan re-payments. One article describes how, "*[w]hile it is the system that 'makes' decisions*", officers of the Board are *"obliged by law to take responsibility for them and to communicate the decisions to the customers by editing the default formulation and signing it*".[33] Humans are also in control of the appeal process. The authors commend the Swedish system as one "*which puts a strong emphasis on compliance with national legislation, officers' ethical codes, and publishing of the rules*", demonstrating how "*a carefully designed system integrating automation with human responsibility can realise many benefits, while remaining sensitive to the values expressed in the rule of law.*"

[33] M. Zalnieriute, L. Bennett Moses, and G. Williams, 'The Rule of Law and Automation of Government Decision-Making' (2019) 82(3) Modern Law Review 425.

It would be possible to design a system where humans sat still further over the loop. For example, one could have a system whereby, if an individual were dissatisfied with the result produced by an automated system, they would have the opportunity to fill out a more involved and detailed form, which would allow the system itself to undertake a more thorough review. The case could then be passed on to a human if the individual remained aggrieved. A staged system of this kind, which uses automated technology effectively as a triage tool, may create disincentives to disproportionate use of the public authority's resources, while identifying the truly problematic cases for resolution.

A third model is "human-out-of-the-loop", whereby a process operates with no human involvement at all, or very minimal human involvement. Perhaps it will one day be the case that systems are sophisticated enough to operate in this way, without an unacceptable risk of error. At present, this seems remote.

Different models of legal regulation

The second tension is between the administrative law and data protection regimes of regulation.

Administrative decision-making is traditionally regulated by the Administrative Court applying principles of administrative law in judicial review. This has distinctive features.

One is the supervisory nature of the Administrative Court's jurisdiction. The Court reviews whether a public authority's decision is so unreasonable as to be unlawful, but does not ask itself whether the decision was wrong – nor even usually in domestic law challenges whether it was disproportionate. There is ordinarily no oral evidence. There are sometimes experts, though this is the exception rather than the rule. There is ordinarily no order for disclosure and inspection of documents. The Administrative Court is capable of determining disputed facts, but the starting point is that the Court will consider the factual material before the decision-maker, at the time of the impugned decision.

Secondly, there is a large body of developed doctrine, based on a significant volume of cases at appellate level and commentary in textbooks.

But for automated decision-making by public authorities, many of the most obviously relevant legal obligations arise instead from the data protection framework. Those include the duties under Article 22 of the UK GDPR concerning automated individual decision-making and compliance with the data protection principles under Article 5, including in particular the requirement of fairness under Article 5(1)(a) and the lawful bases for processing under Article 6.

This regime can be contrasted with the administrative law regime:

(1) The body with primary practical authority in the data protection field is not a Court, but a regulator. The Information Commissioner publishes guidance about how to comply with the UK GDPR and the Data Protection Act 2018. The Commissioner's office engages with data controllers to guide them towards compliance, in a way that a court cannot. If they do not comply, it has regulatory tools to secure compliance, which include investigatory powers (such as the issue of 'information notices'[34]).

(2) Data protection matters do still come before the courts, in a variety of different forums. There are sometimes data protection grounds in judicial review, and there are also civil actions founded upon data protection breaches. However, where the Information Commissioner has taken action, the primary means of putting the issues before a Court is to appeal the Commissioner's decision to the First-tier Tribunal.[35] That is a very different jurisdiction from the Administrative Court. Its jurisdiction is not supervisory: it undertakes a full merits review,[36] though it does not start entirely afresh and must pay careful attention to the reasons of the Commissioner.[37] Unlike the Administrative Court, it is accustomed to hearing oral evidence, including of a technical nature, and making findings as to disputed matters of fact.

[34] S.142 Data Protection Act 2018.

[35] S.162 Data Protection Act 2018.

[36] See ***Central London Community Healthcare NHS Trust v Information Commissioner*** [2013] UKUT 0551, §50.

[37] See ***R (Hope and Glory Public House Limited) v City of Westminster Magistrates' Court*** [2011] EWCA Civ 316 , [2011] 3 All ER 579, §45 *per* Toulson LJ (as he then was), approved by the Supreme Court in ***Hesham Ali (Iraq) v Secretary of State for the Home Department*** [2016] UKSC 60, [2016] 1 WLR 4799, §45 *per* Lord Reed.

(3) The law applied is EU-derived, and so – unlike domestic administrative law – is suffused with the proportionality principle, which is a general principle of EU law.[38]

(4) There is, overall, less doctrine or law in the data protection field than there is in administrative law. There are far fewer reported cases, as the first port of call is the regulator rather than the court, and much of the litigation that does take places goes before a tribunal that cannot set binding precedent. In the case of Article 22 in particular, it had no analogue in the data protection regimes which pre-dated the GDPR coming into effect in 2018. Controllers are, therefore, more reliant on the regulator's guidance.

There are advantages to both models of regulation. The administrative law model promotes certainty and predictability: the law is 'hard' rather than 'soft', laid down by courts rather than the executive or a regulator. Yet there are procedural advantages of the data protection model, too: its ability to deal nimbly with facts in a technical area, regulation by decision-makers with significant subject-specific expertise, and a style of legal regulation that can allow for relatively informal resolution of issues, without needing to commission pages of legal advice from expensive counsel, or trouble the Court.

It remains to be seen what the respective contributions of these jurisdictions will be in regulating administrative decision-making. It is to be hoped, though, that they can usefully learn from one another, rather than pulling in different directions.

They do share some substantive concepts, and the development of each body of substantive law can inform the other. For example,

(1) The UK GDPR requires that controllers process data fairly;[39] administrative law requires that public bodies act fairly. The common ground between these two concepts

[38] See recital (4): "*[t]he right to the protection of personal data is not an absolute right; it must be considered in relation to its function in society and be balanced against other fundamental rights, in accordance with the principle of proportionality.*" See further ***Ittihadieh v 5-11 Cheyne Gardens*** [2017] EWCA Civ 121, [2018] QB 256 at §§95-103 per Lewison LJ.
[39] Article 5 (1)(a).

should not be overstated: the focus of administrative law is procedural fairness,[40] whereas fairness in a data protection context is a wider concept, requiring a balancing exercise between different interests.[41] Nevertheless, at least in respect of procedural fairness, administrative law can supply much learning on what fairness requires.

(2) Administrative law requires that a decision-maker "*ask himself the right question and take reasonable steps to acquaint himself with the relevant information to enable him to answer it correctly*";[42] Christopher Knight has written of how obligations to conduct a data protection impact assessment under data protection law could inform the content of that duty,[43] albeit that the duties themselves remain distinct (the Divisional Court in the ***Rwanda*** litigation rejected an argument that the home secretary's decisions under the immigration rules could be rendered unlawful by breaches of data protection legislation;[44] this is subject to an application for permission to appeal[45]).

(3) I have already mentioned Article 22 of the UK GDPR, on decisions based on solely automated processing of personal data. This requires that domestic law lays down "*suitable measures to safeguard the data subject's rights and freedoms and legitimate interests*" in cases where such decisions are made. Some of these measures are laid down by statute,[46] which requires a controller in certain cases to consider and comply with any request for reconsideration of a solely automated decision. Administrative law may assist in specifying how a controller should undertake that process. It may even assist in specifying suitable safeguards by analogy in cases where the statutory process does not apply, including in cases involving private sector controllers.

[40] ***R (Gallaher Group Ltd) v Competition and Markets Authority*** [2018] UKSC 25, [2019] AC 96, §41.

[41] ***Johnson v Medical Defence Union Ltd (No 2)*** [2007] EWCA Civ 262, [2008] Bus LR 503, §141 *per* Arden LJ: "*the very word "fairness" suggests a balancing of interests. In this case the interests to be taken into account would be those of the data subject and the data user, and perhaps, in an appropriate case, any other data subject affected by the operation in question.*"

[42] ***Secretary of State for Education and Science v Tameside Metropolitan Borough Council*** [1977] AC 1014, 1065B, *per* Lord Diplock

[43] C. Knight, 'Automated Decision-making and Judicial Review' (2020) 25:1, Judicial Review 21-27.

[44] ***R (AAA and others) v Secretary of State or the Home Department*** [2022] EWHC 3230 (Admin), §143. Appeal proceedings in relation to this case are ongoing.

[45] [2023] EWCA Civ 266, §53.

[46] S.14 Data Protection Act 2018.

There may also be some procedural innovation or flexibility required in the Administrative Court, if it begins to hear more challenges raising automated decision-making issues, and data protection issues in particular. I have spoken before of how the Administrative Court may need to become comfortable with expert evidence and the determination of contested facts in such claims, in a way that feels alien in the judicial review context.[47]

D. THE ADOPTION AND DESIGN OF AUTOMATED SYSTEMS

I turn now to consider how the law – in particular, the common law – will apply to regulate:

(1) First, the adoption and design of automated systems by public bodies; and

(2) Secondly, the operation of those systems in individual cases.

Vires

The first and fundamental question a public body will need to ask itself when considering adoption is: do I have the legal power to make decisions in this way at all?

With 'human in the loop' systems, this may not be a significant issue.

But with 'human out of the loop' systems, or systems where the human sits over the loop, it may be argued that the nature of the decision-making has indeed changed, so as to raise a question as to whether the public body has the legal power to proceed in this way, to use a dehumanised system.[48] If the nature of the decision has changed, it is important to identify the source of the power.

The law of data protection adverts to this question, but does not answer it in clear terms. Under Article 22 of UK GDPR decisions "*based solely on automated processing*" may be taken if there is consent or a contractual requirement. Contract will not generally apply in cases involving public

[47] Lord Sales, 'Algorithms, Artificial Intelligence and the Law' [2020] Judicial Review 46.

[48] C. Harlow and R. Rawlings, 'Proceduralism and Automation' in Fisher, Young and King (eds), *The Foundations and Future of Public Law* (OUP, 2020), arguing that computerisation is "*apt to change the nature of an administrative process, translating public administration from a person-based service to a dehumanised system*".

bodies undertaking public functions. As for consent, Recital (43) to the GDPR states that "*consent should not provide a valid legal ground for the processing of personal data in a specific case where there is a clear imbalance between the data subject and the controller, in particular where the controller is a public authority*". Public authorities are thus likely to be left relying on the third possible ground, namely that the decision is "*is required or authorised by domestic law which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests*". This leaves important questions open.

In ***Khan Properties Ltd v HMRC***[49] the question was whether HMRC's penalty of £100, issued to the taxpayer for late filing of a return, was validly assessed. The First-tier Tribunal found that the penalty had been issued automatically by a computer.[50]

The Tribunal noted that, in social security legislation, Parliament had expressly provided that "*any decision, determination or assessment falling to be made [...] by the Secretary of State*" under certain enactments "*may be made or issued not only by an officer of his acting under his authority but also [...] by a computer for whose operation such an officer is responsible*".[51] There was no equivalent provision in the tax context, where the requirement was that a decision be taken by an officer.[52] The Tribunal drew the conclusion that an automated decision was not permissible.[53] Parliament had to change the law.[54]

The Australian case of ***Pintarich v Deputy Commissioner of Taxation***, about automatically generated correspondence from the tax authorities, is in the same vein.[55] A decision intimated by such correspondence was held not be a decision at all, so it was open to a human officer to make a different, less generous decision binding on the taxpayer later on. The Court held that, by a majority, that in order for there to be a relevant binding decision, there had to be a "*mental process of reaching a conclusion and an objective manifestation of that conclusion*".[56] Kerr J, in dissent, thought that the concept of a decision had to be kept up to date, to cover an automated decision.

[49] [2017] UKFTT 830 (TC).
[50] §27.
[51] S.2(1)(a) Social Security Act 1998.
[52] S.100(1) Taxes Management Act 1970.
[53] §23.
[54] S.103 Finance Act 2020: *"Anything capable of being done by an officer of Revenue and Customs by virtue of a function conferred by or under an enactment relating to taxation may be done by HMRC (whether by means involving the use of a computer or otherwise"*.
[55] [2018] FCAFC 793.
[56] §140.

It would be preferable for Parliament expressly to confront the issue of automation in decision-making by public authorities, but given fast-moving technology and the constraints on Parliamentary time, this may not be realistic. The courts will have to decide on the basis of the existing legislative texts. They will have to make a judgment in the light of the doctrine that an Act of Parliament is taken to be 'always speaking'. The Supreme Court's recent judgment in the ***News Corp*** case, concerning application of VAT law, addresses this doctrine in the context of technological change from print to digital newspapers.[57] The Court ultimately held that the word "*newspapers*" in the legislation did not embrace the digital editions, because the technological change between a newspaper and a digital edition was so radical as to make them things of a different type. Applying the always speaking doctrine to statutory powers to work out the proper role of automation, therefore, may require some careful consideration of what the nature of the decisions envisaged by Parliament was when conferring powers on public bodies.

Another relevant concept is the boundary between lawful and unlawful delegation of authority. Administrative law imposes a general rule that "*a statutory power of decision-making must be exercised by the person on whom the power has been conferred*".[58] Yet public policy considerations also dictate that some forms of delegation are necessary in practice and are lawful. A well-known example is the 'Carltona' principle, that "*a decision made on behalf of a minister by one of his officials is constitutionally the decision of the minister himself*".[59] This is justified, the Courts have held, on the basis that "*the administration of government in this country the functions which are given to ministers […] are functions so multifarious that no minister could ever personally attend to them*".[60] The case law on delegation may assist in delineating the line between using technology to discharge a statutory function, and abdicating responsibility to that technology.

A third concept of potential relevance is the doctrine of implied powers. A public authority has an implied power to do what is necessary for, reasonably incidental to or consequential upon the

[57] ***News Corp UK & Ireland Ltd v Commissioners for His Majesty's Revenue and Customs*** [2023] UKSC 7.

[58] ***Shahid v Scottish Ministers*** [2015] UKSC 58, [2016] AC 429, §68 *per* Lord Reed.

[59] ***R (King) v Secretary of State for Justice*** [2015] UKSC 54, [2016] AC 384, §49 *per* Lord Reed.

[60] ***Carltona Ltd v Comrs of Works*** [1943] 2 All ER 560, 563 *per* Lord Greene MR.

performance of its functions.[61] It may be argued that automation of public administration can be authorised under that doctrine. Ultimately, much is likely to depend on the precise scheme that comes before the Court.

Considerations

If a public body does have the power to adopt an automated system, that does not mean that the system is *ipso facto* lawful. The adoption of the system will be subject to other duties, arising under

(i) administrative law,

(ii) human rights law,

(iii) equality law, and

(iv) the law of data protection.

These bodies of law can overlap, as is shown by the ***Bridges*** case on police use of facial recognition technology. The Court of Appeal held that rights to respect for private life under Article 8 of the ECHR had been infringed and there had been a failure to comply with the public sector equality duty in setting up the system and in establishing safeguards in relation to its operation. The police's discretion in relation to which individuals would be targeted and in which locations the technology would be used was too unconstrained. The public sector equality duty had been breached because the system had been established without proper regard to the need to eliminate discrimination. The police force had not sought to check that the software was free from racial bias. The opacity of the system due to the developer's interest in maintaining their commercial confidentiality was held to be no excuse.

So in this case we see the opacity of the system posing problems for a public authority. Public authorities have duties to acquaint themselves with relevant information before taking their decisions, and opacity may prevent this. How can public authorities satisfy themselves that the systems they adopt are appropriate and lawful?

[61] ***Commissioner of the Independent Commission of Investigations v Police Federation*** [2020] UKPC 11, §29.

One possible solution is for public authorities to seek disclosure, on confidential terms, of key information – including the make-up of training datasets, and possibly relevant information about how those data are converted into predictions – from their contractors. The government's recently published algorithmic transparency recording standard sets out a range of information that a public authority ought to collect when using "*algorithmic tools that either have a significant influence on a decision-making process with direct or indirect public effect, or directly interact with the general public*".[62] These include training datasets, testing datasets, and information about the system itself. The standard also suggests that a data protection impact assessment, an algorithmic impact assessment, an ethical impact assessment and equality impact assessment may be appropriate.

However, pulling all of that together will not always be straightforward in practice. Contractors may be loath to share that information directly with public authorities, even at the adoption stage, and even on confidential terms. Public authorities will often lack the resources or expertise to understand the information anyway.

I have previously suggested that Parliament may need to set up an expert commission staffed by coding technicians, with lawyers and ethicists to assist them, which is given access to commercially sensitive code on strict condition that its confidentiality is protected. If that body is expert and widely trusted in industry, then this would go a long way to resolving the difficulties of direct disclosure to the public authority, particularly if public authorities were able to apply to the commission for an opinion or certification of a system, before adopting it.[63]

A tailored legal framework, as well as new institutional architecture, may be required to make that work.

[62] Central Digital and Data Office and Centre for Data Ethics and Innovation, 'Guidance for organisations using the Algorithmic Transparency Recording Standard', 5 January 2023 https://www.gov.uk/government/publications/guidance-for-organisations-using-the-algorithmic-transparency-recording-standard.

[63] Lord Sales, 'Algorithms, Artificial Intelligence and the Law' [2020] Judicial Review 46. Rebecca Williams, in 'Rethinking Administrative Law for Algorithmic Decision Making' (2022) 42 OJLS 468, suggests a "*form of regulatory verification*", whereby the verifying entity would have access to commercially sensitive data, as does e.g. the patent office, to carry out necessary checks, in something akin to a "kitemark" system. This is not dissimilar from the Commission proposal.

Private companies producing machine learning models are themselves subject to several obligations of their own under data protection law. Their processing must be fair and accurate in accordance with the UK GDPR. They must also implement appropriate technical and organisational measures to take into account the risks to the rights and freedoms of data subjects and implement the data protection principles effectively, on which the Information Commissioner publishes guidance relating specifically to AI and discrimination.[64] There are also obligations to conduct data protection impact assessments. However, these obligations do not specifically require the disclosure of the metrics of fairness, or any de-biasing methods, used.[65] To address this, it has been proposed that there should be a framework for AI fairness reporting, whereby companies must disclose the fairness of their machine learning models against certain statistical metrics on a 'comply or explain' basis.[66]

Bridges also provides an example of the need for procedural flexibility. The Court considered an expert report from a professor of computer science, specialising in machine learning, which explained that the accuracy of a facial recognition system depends to a considerable extent on the training dataset. The Court also considered a witness statement from a senior officer of the company that had developed the relevant software, explaining that the precise makeup of the training data was commercially confidential but that the training dataset contained "*roughly equal quantities of male and female faces*", and a "*wide spectrum of different ethnicities*".[67] The Court had to engage with the detail of this expert evidence, and evidence of fact, in a way going beyond what might ordinarily be expected in a claim for judicial review.

[64] https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/guidance-on-ai-and-data-protection/what-do-we-need-to-do-to-ensure-lawfulness-fairness-and-transparency-in-ai-systems/

[65] Article 35(7) of the GDPR states that a Data Protection Impact Assessment should include "*a description of the envisaged processing operations and the purposes of the processing*", "*an assessment of the necessity and proportionality of the processing*", "*an assessment of the risks to the rights and freedoms of data subjects*" and the measures envisaged to "*address the risks*" and "*demonstrate compliance with this Regulation*".

[66] Qing and Lim, 'A legal framework for artificial intelligence fairness reporting' [2022] CLJ 610.

[67] §198.

A third point to emerge from the case is that one cannot assume too readily that human involvement always mitigates the risk of adopting an algorithmic system. If human involvement is not properly structured, then it can have the effect of augmenting the risk. In ***Bridges***, human decisions about the deployment of the technology were found to be too discretionary and insufficiently governed by rules. There is also always the risk that, if humans are not brought into the loop in the right way, they may fall victim to 'automation bias', whereby they too readily trust the output of the algorithm, failing to apply their own judgement.

E. THE OPERATION OF AUTOMATED SYSTEMS

Finally, I make four observations about the duties that apply to operating an algorithmic system, and the possibility of challenge in individual cases.

First, there is obviously some artificiality in distinguishing between system-level duties and duties that apply in individual cases, when it comes to operating algorithms. This is because the operation of an algorithm is more deterministic than the exercise of human discretion: if the system gets something wrong in an individual case, then that can only arise from a problem with the system itself.

This could, in time, require some doctrinal development. The Supreme Court recently handed down a judgment on the issue of when a *policy* or *statement of practice* issued by a public body will be unlawful.[68] Yet algorithms are not the same as policies. In some respects, they are more flexible: they are capable of taking into account millions of factors, that would never be specified in a policy document, and rendering a profoundly subtle decision. In other ways, they are inflexible by comparison to policies, because they produce a completely determined outcome. In this respect, algorithms are more akin to law itself. The Courts will need to give careful thought to when an algorithm, or an artificially intelligent system, is rendered unlawful by the decisions it 'takes', or will take, in individual cases. Much is likely to depend on exactly how and where the human is situated in decision-making.

[68] ***R (A) v SSHD*** [2021] UKSC 37, [2021] 1 WLR 3931.

Secondly, it is desirable in principle that there should be some iteration between the overall design of a system, and the experience of operating that system in practice – including any successful challenges by individuals to the operation of the system in their case as happens with human systems.

In an ideal world, that process of learning could itself be automated. However, the technological challenge of doing so presently appears immense. Decisions by courts, tribunals or regulators are likely to be insufficient in quantity as a dataset to allow for effective automated learning. There therefore remains an important role for humans, even those out of the decision-making loop: to ensure that the technology itself reflects the development of the law as it applies to individual cases.

Thirdly, just as transparency *to the authority* is necessary in order to ensure public bodies can discharge their *duties*, so transparency *to the individual* is necessary to ensure that they can vindicate their *rights*. Because automated decision-making in an individual case is likely to involve processing the individual's personal data, the law of data protection immediately assists with securing that transparency.

> 1) There is the general duty of transparency of processing under Article 5(1)(a) of the UK GDPR.
>
> 2) There is also a duty under Article 13(1) to inform the data subject of the purposes of processing for which the data are intended at the time of collecting the data.
>
> 3) Article 13(2)(f) provides that the controller must, when obtaining the data subject's data, inform the data subject of the existence of any automated decision-making including "*meaningful information about the logic involved*", as well as the "*significant and the envisaged consequences of such processing*"

– although the meaning of 'meaningful' remains hotly contested in the academic literature.[69]

4) Article 14 imposes similar duties to provide this information where the controller obtains the data from a source other than the data subject, albeit subject to certain exceptions.

5) Section 14 of the Data Protection Act 2018 requires the controller to notify the data subject in writing if certain decisions producing legal or other significant effects in their case have been taken based solely on automated processing.

It remains to be seen whether those obligations suffice for effective challenge. Certainly, some jurisdictions have gone further by legislation. French law, for example, specifies that an individual decision taken on the basis of algorithmic processing must include an explicit statement to that effect, as well as information about the parameters of the algorithm, their weighting and the contribution made by the algorithmic processing to decision-making in "*intelligible form*".[70]

Even in the absence of further legislative intervention, there may be some room for fruitful interaction between the requirements of administrative law and the transparency requirements of data protection law. The administrative case law has much to say about the circumstances in which procedural fairness requires disclosure of background information by the decision-maker to the individual, and what will constitute adequate reasons to support a public law decision.

Fourthly, I have no doubt that the substantive law that applies to administrative decisions will be able to cope with the spread of algorithm decision-making and AI. But it will not always be easy to work out how it should be applied.

[69] See the range of positions referenced in R. Williams, 'Rethinking Administrative Law for Algorithmic Decision Making' (2022) 42 OJLS 468.

[70] *loi pour une République numérique*, LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique, Article 4 and Decret no 2017-330 du 14 mars 2017, Article 1.

Several of the legal concepts we use when analysing administrative decision-making direct the Court's attention, in some way, to the mind of the decision-maker. Did the decision-maker take into account a relevant consideration? Did they have an improper motive? Were they acting in bad faith? Mental concepts such as these sit uneasily with automated decision-making. Given the nature of machine learning tools, the law of relevant considerations is a particularly challenging example. Human beings typically make decisions by some process of reasoning. Machine learning, by contrast, relies on statistical inferences: this variable is statistically associated with that outcome. As has been observed, when a system bases its 'decision' on a particular input factor, "*it is not identifying that factor as relevant to the decision it is making in the way that a human would, but merely recognising that, statistically, that factor often correlates with the relevant outcome*".[71] This will force the common law to confront, perhaps more squarely than it has done before, what 'relevance' truly means, and why it matters.

A related issue is the nature of discrimination. The Equality Act 2010 lays down concepts of direct and indirect discrimination. Direct discrimination, in broad terms, is treating a person less favourably than another *because of* their protected characteristic. If done in the exercise of a public function, this is *ipso facto* unlawful.[72] Indirect discrimination focuses on the impact of a policy or practice that is, on its face, neutral as between different groups in society. If such a practice has a *disparate impact* on a protected group, it may be unlawful, but only if it cannot be justified.[73]

A system may be instructed by its programmers to achieve some *neutral* objective – like accurately predicting recidivism, as in the ***Loomis*** case. However, if the operation of artificial intelligence finds that certain protected characteristic data (like sex, ethnicity or some other protected characteristic) can be used as part of an algorithm that accurately predicts that variable, then it will make decisions in individual cases, at least partly, "*because of*" that characteristic. At a system level, this looks like something neutral on its face, that should be capable of justification. Yet at an individual level, the system has made a decision *because of* a protected characteristic. That looks more like direct discrimination: a difference in treatment that is incapable of justification.

[71] R. Williams, 'Rethinking Administrative Law for Algorithmic Decision Making' (2022) 42 OJLS 468.

[72] Ss.13, 29 Equality Act 2010.

[73] S.19 Equality Act 2010.

There are difficult judgements to be made about what, in the operation of an automated system, is truly antithetical to the values protected by Parliament in statute and given effect in the common law by the Courts. The advent of sophisticated information technology in administration will require not just the insight and expertise of those lawyers who understand administrative law, but those who understand information law too.