A W O

Analysis

Effective protection against AI harms

Alex Lawrence-Archer & Ravi Naik

JULY 2023

TABLE OF CONTENTS

July 23

A. Executive summary

1. We were instructed to consider the extent to which key regulations[1] provide effective protection against AI harms as exemplified by three hypothetical scenarios in the fields of employment, financial services, and the public sector. These scenarios are designed to represent plausible near-term deployments of AI technology that could impact ordinary people.

2. The likelihood of AI harms being prevented or redressed depends on:

   a) The existence of regulations – whether sector-specific or cross-cutting – which prevent a harmful AI tool from being used or require decision-makers to consider and address the harms that might arise;

   b) The presence of regulators with the powers and resources to enforce those regulatory requirements;

   c) A private right to redress for individuals who suffer harm, and accessible forums through which to enforce them; and

   d) Mandated, meaningful and in-context transparency to ensure individuals become aware of and can evidence how they have been harmed, as a precursor to obtaining redress.

3. Ideally all would be present. If multiple elements are missing or weak, this indicates that harms arising from the use of AI tools ('AI harms') are unlikely to be effectively prevented or redressed.

4. The Scenarios demonstrate the central importance of cross-cutting laws that relate to the use of data and decision-making generally: the UK GDPR and the Equality Act 2010. Equally, they demonstrate common gaps in protection, including:

   a) The lack of legally mandated, meaningful, and in-context transparency that would alert individuals to the possible harm they face and allow them to evidence it;

   b) Gaps in regulation due to a lack of resources and access to information for some regulators, combined with a lack of enforcement powers or a low (relative to peer regulators) use of the powers that do exist; and

[1] Given the general-purpose application of AI tools it is naturally not possible to consider every conceivable application of law. We have focused in our analysis on the kinds of regulation mentioned in the Government's AI White Paper, such as the UK GDPR, Equality Act, Financial Conduct Authority Rules etc.

c) The need to enforce GDPR (and in some cases, Equality Act) rights through the civil courts, which is lengthy, expensive, uncertain and off-putting for most ordinary people.

5. Factors specific to certain sectors complicate the picture. For example, individuals are better protected in highly regulated sectors such as financial services. However, we conclude that collectively the Scenarios demonstrate that there are significant gaps in the effective protection from AI harms in the current regulatory regime.

6. It is notable that in some areas the Government's proposed reforms to the UK GDPR in the Data Protection and Digital Information Bill (No 2)[2] are set to further weaken the level of protection from AI harms, for example by loosening requirements to assess risky data processing prospectively, by weakening individuals' access to data rights, and by expanding the circumstances in which significant decisions may be taken by solely automated means.

[2] See further at https://www.awo.agency/blog/the-data-reform-bill-uncertainty-and-missed-opportunities/

Summary Table: Layers of Protection from AI Harms

Are there legal requirements that the decision-maker must consider in advance?	**Is it likely that a regulator would prevent the AI harm through enforcement of those requirements?**	**Would the individual be able to find out about and evidence the harm?**	**Is there a legal right to redress for the harm?**	**Is it practical for individuals to enforce any legal rights to redress?**
Scenario 1 (Employment)				
Limited: The UK GDPR and Equality Act impose some requirements, but these do not address all the harms in the scenario or fundamentally prevent the tool from being used.	Unlikely: relies on enforcement by the ICO and the EHRC, both of which are limited in the information available to them, their powers and enforcement approach, and their resources.	Low/medium: Some additional protections from ERA in relation to statements of pay.	Medium: GDPR and Equality Act give rise to causes of action for some harms in the Scenario (but not those relating to general working conditions). Additionally, some harms covered by the Employment Rights Act where an employee is dismissed.	Impractical: requirement to bring a civil claim for GDPR breaches. Employment Tribunal for ERA and Equality Act breaches. But this relies on having a protected characteristic and/or employment status, and does not protect against diminished working conditions.
Scenario 2 (Biometric Mortgage Assessment)				
Medium: both Cross-cutting (GDPR and Equality Act) and Sector-Specific FCA Rules are relevant to the tool, suggesting it may not be permissible to implement it in the way described.	Medium: reason to believe FCA is a more effective *ex ante* regulator, as it is focused on one sector and has strong enforcement powers Super-complaints may also bring issues to the FCA's attention	Poor: it would be especially difficult for an individual to identify the harm in this scenario given the opacity of the algorithmic logic, even taking into account GDPR transparency rights.	Good: as well as GDPR and Equality Act causes of action, able to seek redress under FCA rules.	Practical: Financial Services Ombudsman provides free-of-charge resolution with need for legal representation
Scenario 3 (DWP Chatbot)				
Low: the UK GDPR likely does not rule out the use of the tool. Further, any additional guidance for public bodies on the use of AI is non-binding and compliance with the guidance is not monitored.	Very unlikely: relies solely on enforcement by the ICO, which takes a light-touch approach to regulating public bodies, which arguably reduces incentives for compliance.	Poor: relies on non-binding guidance on the part of the DWP and GDPR transparency, which does not require explanations of automated decisions *in situ*.	Medium: beyond GDPR rights, voluntary DWP maladministration scheme and rights to appeal benefit decisions But the DWP scheme may not fully compensate consequential losses.	Practical: appeal from DWP scheme plus option to appeal to Parliamentary Ombudsman.

B. Introduction and approach

7. We were instructed to consider key regulations applicable in three hypothetical scenarios in which the use of algorithmic tools might cause harm (‘**AI harm(s)**’) to individuals (the ‘**Scenarios**’). For each scenario, we have considered a range of factors which might contribute to providing effective protection for individuals against AI harms:

    a) Laws and regulations applying to the body(ies) responsible for the creation and use of the algorithmic tool i.e. is there regulation that would prevent this harm from arising? If so, is that regulation sufficiently-well enforced?

    b) Laws and regulations that allow individuals to obtain redress for the harm i.e. would an individual know they had been harmed and be able to evidence it? If so, does the law give them (i) a right to redress and (ii) a venue in which that right can be enforced on a basis that is realistic for ordinary people?

    c) Laws requiring transparency are not of themselves capable of providing redress for AI harms. We nevertheless include consideration of them since it is not practically possible for an individual to obtain redress without knowledge or evidence of harm, and a lack of transparency as to how algorithmic tools are used is a barrier to respect for individuals’ rights[3].

8. The Scenarios show that AI harms may engage law and regulation that is either *specific* to a sector or field of activity (such as employment law), or *cross-cutting* in the sense that it deals with the use of data or fairness in decision-making (the UK GDPR[4] and Equality Act 2010). We have subdivided our analysis on this basis in order to demonstrate the strengths and weaknesses of cross-cutting regulation for AI harms.

9. We have considered the law as it would apply if the current draft[5] of the Data Protection and Digital Information Bill (‘**DPDI Bill**’) becomes law. We assume that all the acts described take place in England and Wales after 1 August 2023.

C. Cross-Cutting regulation: The UK GDPR

10. The UK GDPR regulates all processing of personal data provided its territorial scope is engaged[6]. Processing is defined very broadly (Article 4(2)) as:

[3] See e.g. Gryz, J. and Rojszczak, M. (2021). *Black box algorithms and the rights of individuals: no easy solution to the “explainability” problem*. Internet Policy Review, 10(2). https://policyreview.info/articles/analysis/black-box-algorithms-and-rights-individuals-no-easy-solution-explainability

[4] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (Retained EU Law)

[5] https://commonslibrary.parliament.uk/research-briefings/cbp-9746/ Published March 2023

[6] See Article 3 for details but it will be engaged for all the Scenarios.

> *"any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction."*

11. Personal data is (Article 4(1):

> *"any information relating to an identified or identifiable natural person (data subject); an identifiable natural person is one who can be identified, directly or indirectly […]"*

12. The concept of identification has been clarified by courts to include where an individual is capable of being 'singled out' from data[7]. That is, where an individual's 'identity' (such as their name) is unknown, but data still allows decisions to be made about them specifically, the individual is identifiable. Opinions and inferences about a person – regardless of the level of certainty or accuracy associated with them – are personal data relating to that person[8].

13. The broad scope of the UK GDPR and the way in which it has been interpreted is a significant strength in relation to AI harms. Each of the Scenarios involves the processing of personal data by one or more 'controllers' – that is, the entity which determines the purposes and means of the processing. In these scenarios, the controller is the decision-maker using the algorithmic tool.

14. The UK GDPR creates obligations on the part of the relevant controller, including the following most relevant to the Scenarios:

<u>Lawful, fair and transparent processing Article 5(1)(a) and 6</u>

15. Processing of personal data must have a valid legal basis, such as it being necessary for the performance of a contract, benefiting from the consent of data subjects, or necessary for a legitimate interest pursued by the data controller. It must also be transparent to data subjects (not least through compliance with Articles 13-15 UK GDPR) and fair. However, the requirement for processing to be fair does *not* mean that substantive decision outcomes where processing is involved must be objectively fair. All

[7] Breyer Case C-582/14 (CJEU)

[8] See e.g. Wachter and MIttelstadt (2018) *A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI,* Columbia Business Law Review 2019/2. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3248829

that is required is that the *processing* be fair in the sense of being justifiable and within the reasonable expectations of data subjects (i.e. not unduly hidden or misleading)[9].

16. We address the lawful bases that might be relied upon for each Scenario in the analysis below.

Provision of information to data subjects Articles 13/14

17. Linked to the general requirement of transparency, data controllers are required by these Articles to provide a range of information to data subjects about processing concerning them. This is often done through 'privacy notices' which state (among other things) the legal basis relied on, the purposes of processing etc. It is important to note that the requirements are not absolute; some information need only be provided to the extent necessary to *"ensure fair and transparent processing"*. Further, where data has not been obtained from a data subject[10], such as inferences drawn through the use of algorithmic tools, information on processing need not be provided where to do so would involve disproportionate effort.

Restrictions on the processing of special category data (Article 9)

18. There is a general prohibition on the processing certain categories of personal data ('**special category data**' or '**SCD**'). Namely processing:

> *"revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation"*

19. The precise wording of Article 9 is important, as different types of SCD attract slightly different treatment, which we address where relevant to the Scenarios.

20. Some exemptions are available to the general prohibition, but if none is available any processing of SCD will be unlawful. We address which might be relied upon for each Scenario in the analysis below.

Accurate processing Article 5(1)(d)

21. This article requires that processing be:

[9] See ICO Guidance for further detail: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-protection-principles/a-guide-to-the-data-protection-principles/the-principles/lawfulness-fairness-and-transparency/

[10] Directly observed data is 'obtained' from data subjects for these purposes, but inferences drawn from such data are not.

> *"accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay."*

22. Accuracy is a relative standard, since what is a '*reasonable step*' depends on the circumstances. Where processing is more consequential or risky, a higher standard of accuracy will be required[11].

Safeguards and prohibitions for significant automated decisions Articles 22A-C

23. These Articles (inserted by the DPDI Bill) govern decisions which are both (i) solely automated (meaning there is no meaningful human involvement in them) and (ii) significant (meaning they have a legal effect on someone or a similarly significant effect).

24. Article 22B prohibits solely automated significant decisions where they are based in whole or in part on the processing of SCD. Article 22C mandates a range of safeguards that data controllers are required to put in place for other solely automated significant decisions, such as providing general information about them, and allowing data subjects to make representations or obtain human intervention.

Data protection by design and default Article 25

25. Controllers are required to put in place measures in their processing that ensure processing is lawful, fair accurate and so on. Given its link to other requirements of the UK GDPR this provision will rarely lead to an infringement or right of action *on its own*, but is part of the framework against which the ICO is likely to judge a controller, should the ICO proactively investigate their processing.

Assessment of high-risk processing Article 35[12]

26. A controller must prepare AHRP where processing *"is likely to result in a high risk to the rights and freedoms of [individuals]."* Weakened from the better-known data protection impact assessment by the DPDI Bill, this merely requires a 'summary' of the processing and an assessment of risks for data subjects and how they can be mitigated. This is a purely internal document; it need not be shared with the ICO or with data subjects (per changes introduced by the DPDI Bill).

27. The UK GDPR also creates a range of data subject rights[13], including:

    a) **The right to be informed** (Article 15): data subjects are entitled to copies of their personal data from controllers processing that data, as well as information about

[11] See also ICO Guidance: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-protection-principles/a-guide-to-the-data-protection-principles/the-principles/accuracy/

[12] As amended by the DPDI Bill.

[13] Although note that Article 12 – as amended by the DPDI Bill – sets a new lower bar for data controllers to reject the exercise of these rights; they are not absolute.

the processing. This is requested through what is generally known as a 'data subject access request'.

b) **The right to object** (Article 21(1)): data subjects may object to processing where the controller's legal basis is its legitimate interests or a task of a public authority. The objection must be based on factors specific to the data subject and may be overridden (i.e. the controller may continue processing) if there are '*compelling legitimate grounds*' to do so.

c) **The right to rectifications** (Article 16) where data are inaccurate or out of date, data subjects have a right to have them corrected or updated by the relevant controller.[14]

28. Where the UK GDPR has been infringed, it provides – alongside the Data Protection Act 2018 ('**DPA 2018**') – for any affected data subject to lodge a complaint with the 'supervisory authority' which in the UK is the Information Commissioner's Office (**ICO**')[15] (Article 77) and to bring a civil claim (Article 79 and §167-8 DPA 2018). Practical oversight and enforcement of the requirements of the UK GDPR are discussed at section E below.

i. Limitations on transparency in the UK GDPR

29. A range of UK GDPR provisions mandate the provision of information to data subjects[16] about how their data is being processed. This goes some way to helping individuals know when AI harms have occurred, understand them, and evidence them, but the obligations on controllers and benefits for individuals are limited.

Provision of information

30. Articles 5(1)(a), 13, 14 and 22C require only the provision of *general* information to data *subjects* as opposed to specific information to any specific data subject[17]. In practice these obligations are often discharged by a high-level privacy notice that discusses in general terms (albeit at great length) the kinds of processing that a controller engages in, without tailoring this to an individual, or attempting to make the information accessible or user-friendly.

[14] It is important to note that the DPDI Bill slightly lowers the threshold at which data subject rights in the UK GDPR may be refused by controllers. As it has not yet been implemented it is difficult to assess precisely how this will manifest, but it seems clear that the level of protection provided by these rights will be somewhat lower for individuals under the UK's new regime.

[15] When the DPDI Bill passes the ICO will be reconstituted as the Information Commission. We refer to it as the ICO for convenience.

[16] As to information required to be provided to the ICO, see section E.

[17] Notably Article 22C does not require a data subject to be informed of a significant solely automated decision *at the time that decision is made*.

31. By way of example two major high street mortgage lenders' privacy notices run to 16 and 13 densely-packed pages[18] respectively and contain vague and conditional statements such as:

> *"We must have a legal basis (lawful reason) to process your personal data. In most cases, the legal basis will be one of the following [list of 5 separate lawful bases]"*
>
> *"We will keep your information confidential but we may share it with third parties (who also have to keep it secure and confidential) in the following circumstances [1.5 page list of 21 different categories of potential recipient]"*

32. The Department for Work and Pensions' *Personal Information Charter*[19] runs to 14 pages and includes the statement:

> *"DWP is developing new digital services all the time. If any new services involve automated decision-making, we will tell you about this when the decision is made."*

Explanations for algorithmic decisions

33. Articles 13 and 14 make only limited provision for the explanation of algorithmic decisions, providing for the provision (to the extent necessary to ensure fair and transparent processing) of:

> *"the existence of automated decision-making, including profiling, which is subject to the requirement to provide safeguards under Article 22C and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject"*

34. Whilst the words 'at least in those cases' may suggest that 'meaningful information' about the logic of algorithmic decisions may be required in other cases, in practice there is little case law establishing when such meaningful information is required, if the rules on significant solely automated decisions are not engaged.

[18] Each with a Fleisch-Kincaid score in the 50s indicating they are 'hard to read'.

[19] https://www.gov.uk/government/organisations/department-for-work-pensions/about/personal-information-charter#automated-decision-making

35. Further, the meaning of 'meaningful information' is unclear, and the general obligation to provide information can qualified where to do so would be impossible or involve disproportionate effort.

36. In summary therefore the UK GDPR does not entitle data subjects to a *full* explanation of an algorithmic decision. This might present a significant barrier to substantiating a challenge to such a decision, especially given the opaque nature of machine learning driven tools, since it may be very difficult to show that a specific factor (such as a protected characteristic, or special category data) has played a role[20].

Accessing personal data

37. Accessing the underlying data might be a further important means by which an individual could evidence an AI harm, but it too is a limited right. The right to access copies of one's personal data is expressed in Article 15 to be limited such that it "*shall not adversely affect the rights and freedoms of others*." This rather open-ended qualification has been interpreted by data controllers as exempting them from the requirement where it presents even a slight risk, particularly in contexts where they have strong incentives not to provide data[21].

38. At a practical level, obtaining access under Article 15 is not always straightforward or fast. Controllers may take up to 3 months to (lawfully) respond, and the most recent available figures show that around 40% of over 10,000 complaints received by the ICO in one year related to Article 15. Changes to Article 15 by the DPDI Bill amend the basis on which the right of access can be refused to make it harder for data subjects to get access to their data.

ii. ICO Guidance on explaining AI decisions

39. The ICO has published guidance[22] for organisations as to how they should explain decisions made using AI to those affected. It is framed as elaborating on the transparency and data access requirements in the UK GDPR, but also sets out further principles, such as (among many others) the need to '*consider the context'* in which a decision is made, and to *'[t]hink about how to build and implement your AI system in a way that […] fosters the physical, emotional and mental integrity of affected individuals.'*

[20] The absence of a right to an explanation has been explored further in academic commentary: Wachter et al (2017) *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation,* International Data Privacy Law 7(17): https://academic.oup.com/idpl/article/7/2/76/3860948 It is also notable that locating the 'decision' caught by GDPR provisions is not always straightforward, although this does not affect our analysis of the Scenarios. See Binns and Veale (2021) *Is that your final decision? Multi-stage profiling, selective effects, and Article 22 of the GDPR,* International Data Privacy Law 11(4): https://academic.oup.com/idpl/article/11/4/319/6403925

[21] See e.g. reporting relevant to contract workers: https://www.workerinfoexchange.org/wie-report-managed-by-bots

[22] https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/explaining-decisions-made-with-artificial-intelligence/

40. The guidance Is extensive and detailed, encouraging data controllers to design systems which offer meaningful and useful explanations of AI decisions *in context* to those affected. Were the guidance followed to the letter in all cases (and in the three Scenarios), this would overcome many of the limitations on transparency in the UK GDPR. But the guidance has no legal status and therefore does not create enforceable requirements that would aid individuals in identifying and evidencing the AI harms in the Scenarios. Whilst laudable, in our view it goes beyond what the law requires and is therefore of limited relevance to a legal analysis of effective protection against the harms in the Scenarios.

iii. Summary

41. The UK GDPR does not entitle individuals to an explanation of specific algorithmic decisions that have affected them at the time those decisions are made. Nor does it entitle them to scrutinise in full the algorithmic tool used to make the decision, and there may even be limits on the extent of the individual's *own* personal information that they can retrieve to understand a decision.

42. At best, the UK GDPR is likely to provide clues to motivated individuals that would act as the beginning of a process to finding out that an AI harm had occurred, understanding and evidencing it.

D. Cross-cutting regulation: The Equality Act 2010

43. The Equality Act 2010 ('**EA**', §4-12) regulates discrimination on the basis of a specified range of protected characteristics:

- age;
- disability;
- gender reassignment;
- marriage and civil partnership;
- pregnancy and maternity;
- race;
- religion or belief;
- sex;
- sexual orientation.

44. The relevant type of discrimination for two of the three Scenarios is indirect discrimination, which takes place when a person (A) applies to another (B) a 'provision, criterion or practice' which:

a) A also applies to people other than B, including those who do not share B's protected characteristic;

b) Puts B – or those who share B's protected characteristic – at a 'particular disadvantage' compared to those who do not share it; and which

c) Cannot be justified as a proportionate means of achieving a legitimate aim. (s. 19 EA).

45. Discrimination arising from a disability (s.15 EA) is also relevant to the Scenarios and operates in a similar way to indirect discrimination for the purposes of the Scenarios.

46. The EA only makes indirect discrimination unlawful in certain circumstances including, relevantly to this analysis (i) where a person provides a service to the public or a section of the public (for payment or not) (s.29 EA); and (ii) In offering employment, the terms of employment, or the terms of contract work (§39 and 41 EA).

47. Breaches of the EA give rise to a private right of action for the individual(s) affected.

E. Regulatory enforcement of the UK GDPR by the ICO

i. Mandate and powers

48. The ICO is "responsible for monitoring the application of [the UK GDPR]" (Article 51) and is further mandated:

> *"to secure an appropriate level of protection for personal data, having regard to the interests of data subjects, controllers and others and matters of general public interest," and*
>
> *"to promote public trust and confidence in the processing of personal data."*[23]

49. The ICO has a range of legal powers which it may use to secure compliance with the UK GDPR by controllers including:

a) The power to compel information (§142-145 DPA 2018);

b) The power to enter premises as part of an assessment (§146 and 154 DPA 2018);

[23] s.120A DPA 2018 as amended by the DPDI Bill

c) Requiring controllers to take steps to remedy infringements of the UK GDPR (s.149 DPA 2018); and

d) Issuing fines to controllers for up to £17.5m or 4% of worldwide turnover, whichever is higher (Article 83)[24].

50. The ICO thus has extensive statutory enforcement powers which it may use to enforce the UK GDPR.

ii. Sources of information about compliance

51. As well as carrying out investigations of its own volition (which may include using its information-gathering powers), the ICO must investigate complaints by data subjects (s.165 DPA 2018), which might lead to enforcement action.

52. However, it is important to note that there is no formal requirement for data controllers to share information with the ICO on a regular basis. Controllers are required to report 'data breaches' to the ICO (Article 33 UK GDPR) but this has a limited meaning:

> *"[the] accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data."*

53. That is, there is requirement for a controller to report on wider infringements of the requirements of the UK GDPR, such as lawfulness and transparency of processing.

54. Article 36 of the UK GDPR[25] provides that data controllers *may* consult the ICO where they assess that their processing is likely to result in a high risk to data subjects' rights and freedoms. This is a notable change from the previous position, which *required* controllers to do this. It is likely that over time this will reduce the amount of information available to the ICO about the kinds of processing taking place which may lead to AI harms, and therefore reduce its ability to proactively regulate to mitigate those harms. The ICO's central role in supporting other regulators with a responsibility for regulating AI harms (confirmed in the Government's AI White Paper) means that this diminution in understanding of how controllers are using technology in practice could affect regulatory capacity across the economy.

iii. Approach to enforcement

55. In practice there is some evidence that the ICO uses its 'harder' powers of enforcement in relation to breaches of the UK GDPR less frequently than other data protection

[24] Note that these are administrative fines; the ICO cannot order the payment of compensation by a controller to a data subject.

[25] As amended by the DPDI Bill.

regulators[26]. In 21/22 (the most recent year for which an annual report is available), saw fines of just £183,000 issued[27]. Whilst a full assessment of the ICO's approach to regulation is outside the scope of this analysis, academic commentators have described the ICO's level of enforcement as 'low'[28].

56. In relation to enforcing against public bodies specifically, in 2022, the Information Commissioner stated:

> *"In central Government, fines create a 'money-go-round', moving funds from one department to the Treasury and then to the consolidated account. It's not effective and can have the opposite effect to what we want."*[29]

57. Relatedly, the ICO's regulatory strategy published in 2022 states:

> *"For public sector organisations, our approach also means that any financial penalty may be reduced or replaced by a public reprimand."*[30]

58. That is, the ICO will rarely if ever consider levying a fine on a public body for a breach of the UK GDPR or DPA 2018. This approach was exemplified by the recent decision by the ICO to issue only a 'reprimand' – a non-binding published statement – to Thames Valley Police for an incident in which they had released identifying details of a witness to the very criminal group against whom that witness was to give evidence.

59. There is no objectively correct level of fines for the ICO to issue, but even commercially-focused legal commentators have remarked on a stark contrast between the ICO and – for example – the Spanish regulator, which 50 times as many GDPR fines in the same period as the ICO and described the ICO as focussing on the 'carrot rather than the stick' in its enforcement[31]. New, stronger requirements for the ICO to take into account 'the desirability of promoting innovation' (s.120B DPA 2018 inserted by the DPDI Bill) may deepen this tendency towards lighter-touch regulation.

[26] As opposed to breaches of the Privacy and Electronic Communications Regulations 2006 which regulate specific aspects of direct marketing such as unsolicited phone calls.

[27] https://ico.org.uk/media/about-the-ico/documents/4021039/ico-annual-report-2021-22.pdf note that the figure quoted in the report was later reduced following a decision by the ICO to agree a c.90% reduction in the fine to the Cabinet Office: https://ico.org.uk/about-the-ico/media-centre/news-and-blogs/2022/11/ico-and-cabinet-office-reach-agreement-on-new-year-honours-data-breach-fine/

[28] Erdos, D (2022) *Towards Effective Supervisory Oversight? Analysing UK Regulatory Enforcement of Data Protection and Electronic Privacy Rights and the Government's Statutory Reform Plans* University of Cambridge Faculty of Law Research Paper No. 16/2022: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4284602

[29] https://ico.org.uk/about-the-ico/media-centre/news-and-blogs/2022/11/how-the-ico-enforces-a-new-strategic-approach-to-regulatory-action/

[30] https://ico.org.uk/media/about-the-ico/policies-and-procedures/4022320/regulatory-posture-document-post-ico25.pdf

[31] https://www.dacbeachcroft.com/en/gb/articles/2023/may/five-years-of-gdpr-standing-the-test-of-time/

iv. Resource constraints and areas of focus

60. The UK GDPR is a law of general – indeed extraordinarily wide – application. Given the way personal data is defined, it has been described as a 'law of everything'[32]. Conversely, the ICO has finite financial and human resources. In its regulatory strategy, it explicitly states:

> *"We have finite resources and we are unable to look into every matter raised with us."*[33]

61. This issue was also addressed in the ICO's response to the Government's consultation on the AI White Paper[34].

62. Given the lack of a formal requirement for data controllers to report on their compliance to the ICO and its finite resources, this means that in practice the ICO's enforcement action will not cover every sector or all types of UK GDPR infringement. It will be focused on particular areas chosen by the regulator[35]. As well as assessing general factors such as the level of harm, the ICO chooses specific fields of activity in which to focus its investigations and enforcement. For example in the current period it is focussing on:

- Children's privacy;
- Impact of technology on vulnerable groups;
- Deprivation; and
- Personal safety.[36]

63. The level of protection offered by the UK GDPR as enforced by the ICO (as opposed to by an individual) therefore depends on whether an AI harm occurs in a sector which the ICO has elected to prioritise for investigation and enforcement.

64. The ICO is also required by the DPA 2018 to publish codes of practice clarifying the application of the UK GDPR in a number of specific areas and must prepare further codes of practice if required to by the Secretary of State (s.124A DPA 2018 inserted by the DPDI Bill). No extant or proposed codes cover the matters in the Scenarios.

65. The ICO also publishes a range of non-statutory guidance on data protection designed to help controllers comply with the UK GDPR. This includes guidance on the use of

[32] Purtova (2018) *The law of everything. Broad concept of personal data and future of EU data protection law,* Journal of Global Information Technology Management: https://www.tandfonline.com/doi/full/10.1080/17579961.2018.1452176

[33] https://ico.org.uk/media/about-the-ico/policies-and-procedures/4022320/regulatory-posture-document-post-ico25.pdf

[34] https://ico.org.uk/media/about-the-ico/consultation-responses/4024792/ico-response-ai-white-paper-20230304.pdf at [25].

[35] The ICO's Regulatory Strategy 2022 explicitly recognises this.

[36] https://ico.org.uk/about-the-ico/our-information/our-strategies-and-plans/ico25-strategic-plan/annual-action-plan-october-2022-october-2023/safeguard-and-empower-the-public/

algorithmic tools[37] which may offer some additional protection against AI harms, although this would require a data controller to have found and implemented the guidance, and be highly motivated to proactively comply with it.

v. Summary

66. The ICO has strong powers in theory to enforce the UK GDPR, but

a) By comparison to other data protection regulators, it makes limited use of those enforcement powers, especially against public bodies, which may reduce incentives for compliance;

b) It relies for information about compliance on proactive investigations and complaints rather than compulsory auditing or reporting by data controllers; and

c) It has an extremely broad remit, which means it is compelled to focus its finite resources on a few specific fields of activity.

67. In this light, the level of effective protection provided by *ex-ante* regulation by the ICO for any specific AI harm is low to medium at best. That is, the data protection regulatory regime is not likely to identify and provide effective protection against AI harms as a general rule, despite the cross-cutting nature of the UK GDPR and its relevance to those harms. Rather, in many cases AI harms will manifest despite the work of the regulator, leaving individuals to seek redress themselves.

F. Regulatory enforcement of the Equality Act 2010 by the Equality and Human Rights Commission

i. Mandate and powers

68. The Equality and Human Rights Commission (‘**EHRC**’) is tasked by the Equality Act 2006 (‘**EA 2006**’ as amended by the EA) with:

> *“encouraging and supporting the development of a society in which—*
>
> *(a)people’s ability to achieve their potential is not limited by prejudice or discrimination,*
> *(b)there is respect for and protection of each individual’s human rights*
> *(c)there is respect for the dignity and worth of each individual,*
> *(d)each individual has an equal opportunity to participate in society, and*

[37] https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/

> *Ithere is mutual respect between groups based on understanding and valuing of diversity and on shared respect for equality and human rights.”*[38]

69. As well as having general powers to undertake research and publish guidance (§13-19 EA 2006), the EHRC is empowered to:

    a) Carry out investigations into whether someone as breached equality law (s.20) including a power to compel information (s.20 and Sch 2);

    b) Enter into binding agreements with entities regarding compliance (s.23);

    c) Issue a notice requiring that person to prepare an action plan or recommending action to be taken, where a breach of equality law has been found (s.21); and

    d) Institute or intervene in judicial review proceedings (s.30)[39].

70. These enforcement powers are weaker than those given to the ICO. For example the EHRC cannot directly require a person to take steps to comply with the EA or fine entities. Rather, it must instead apply to the court to secure compliance with an action plan (s.22) or for an injunction to restrain an unlawful act (s.24).

ii. Sources of information about compliance

71. There are no mandatory reporting or audit requirements for private organisations in relation to equality law[40]. Like the ICO, the EHRC must rely on information obtained of its own volition to support a decision to take enforcement action. In fact, the EHRC has even fewer sources of information to draw on, since there is no formal mechanism by which individuals can bring complaints about infringements of the EA to the EHRC, and organisations are not required to report on infringements – even in the same limited sense as the requirement to report data breaches under Article 33 UK GDPR.

iii. Approach to enforcement

72. A full assessment of the EHRC’s work as a regulator is beyond the scope of this analysis. However there is evidence that the EHRC makes relatively sparing use of the limited enforcement powers it has. In 2019 a Select Committee Report[41] stated:

[38] s.3 EA 2006

[39] The EHRC may also intervene in other proceedings under this section.

[40] Public authorities do have reporting requirements as part of the Public Sector Equality Duty in s.149 of the EA, but these do not arise in relation to any of the Scenarios.

[41] https://publications.parliament.uk/pa/cm201719/cmselect/cmwomeq/1470/147002.htm

> *"As an organisation [the EHRC] must overcome its timidity and be bolder in using the existing powers that only it has."*

73. Since that report, enforcement action by the EHRC does not appear to have significantly increased. The most recent legal intervention listed on its site is from 2020[42] and the most recent investigation was launched in 2021[43]. The EHRC's *Human Rights Legal Cases* page has not been updated since 2016[44], and the most recent agreement under s.23 listed is from 2018[45].

iv. Resource constraints and areas of focus

74. Much like the UK GDPR, the provisions of the EA may apply in a wide variety of domains. The EHRC is a cross-cutting regulator, not a sector-specific one. This means it in theory is responsible for enforcing compliance with equality law across the whole of society despite being a relatively small organisation[46]. Much like the ICO it is not realistic to expect the EHRC to procure full compliance with equality law in every area. It chooses areas of focus through periodic strategic plans. Its current priorities are (emphasis added):

- equality in a changing workplace
- equality for children and young people
- upholding rights and equality in health and social care
- addressing the equality and human rights impact of digital services and artificial intelligence
- fostering good relations and promoting respect between groups
- ensuring an effective legal framework to protect equality and human rights[47].

75. The EHRC has emphasised this challenge in its response to the Government's AI White Paper, stating:

> *"While the Commission is committed to the regulation of AI under both equality and human rights law, these additional duties are outside our current business plan commitments and unfunded. The Government must invest in the Commission and other regulators to ensure that the regulatory community can*

[42] https://www.equalityhumanrights.com/en/legal-casework/legal-cases
[43] https://www.equalityhumanrights.com/en/inquiries-and-investigations/inquiry-challenging-decisions-about-adult-social-care
[44] https://www.equalityhumanrights.com/en/legal-casework/human-rights-legal-cases
[45] https://www.equalityhumanrights.com/en/legal-casework/enforcement-work
[46] In the most recent year for which data is available the EHRC had a budget of just over £18m: https://www.equalityhumanrights.com/en/what-we-do/business-plan-2021-2022
[47] https://www.equalityhumanrights.com/en/our-litigation-and-enforcement-policy/how-we-decide-whether-use-our-powers

> *build the capacity and expertise needed to support safe, responsible and ethical innovation in AI."*[48]

76. Much like the ICO, the EHRC publishes guidance on a range of issues relevant to its duties. There is no current guidance directly relevant to the Scenarios[49], although the EHRC's current strategic plan indicates an intention to carry out work broadly on the issues raised by them.

v. Summary

77. Much as in the field of data protection, the broad applicability of equality law and the availability of (at least some) enforcement powers for the EHRC belie a relatively weak enforcement environment. The EHRC:

    a) Has relatively limited enforcement powers, especially vis-à-vis private entities;

    b) Uses those limited enforcement powers sparingly and rarely;

    c) Relies on its own investigations to uncover lack of compliance with the law; and

    d) Is obliged to focus on only some areas given its broad remit and finite resources.

78. Enforcement of the EA by the EHRC therefore provides only a limited degree of protection from AI harms. It should not be expected that proactive enforcement by the EHRC will identify and prevent AI harms as a general rule. They will manifest in many circumstances, leaving individuals with the burden to identify them and seek redress.

G. Practicality of enforcing civil claims under the UK GDPR and Equality Act

79. Both the EA and UK GDPR give rise to private rights of action for compensation, injunctive or declaratory relief against the decision-maker which has infringed the relevant provision(s) (§167-9 DPA 2018 and Part 9 EA). That is, an individual affected by an AI harm which engages either or both the EA or UK GDPR can sue the entity responsible for redress, assuming they are aware of and can evidence the harm. This appears to provide an strong level of protection, but the practicalities of bringing such claims are a crucial consideration in whether this protection is in fact effective.

80. Where discrimination takes place in the context of employment (or contract work), proceedings may be brought in the Employment Tribunal (see section 0ii below). But for

[48] https://www.equalityhumanrights.com/en/file/43396/download at [45]
[49] There is guidance on the use of AI in public services, but Scenario 3 does not raise issues under the EA.

discrimination in the context of service provision, and for all claims under the UK GDPR, proceedings must be brought in the civil courts – usually the county court (s.114 EA and §167-9 DPA 2018). This is far from straightforward. An individual bringing proceedings in the county court faces barriers including:

a) The complexity of the process and possible need for legal advice;

b) The difficulty of funding that legal advice; and

c) The risk of being made to pay the other side's costs if unsuccessful ('adverse costs risk').

81. The county court has three 'tracks' to which cases may be allocated, the small claims track, the fast track, and the multi-track[50]. The main factor determining allocation is the value of the claim. Claims for smaller amounts (below £10,000) are likely (though not guaranteed) allocated to the small claims track of the county court. Where the claimant is seeking something other than money – e.g. a court order for compliance with legislation, this is more likely to be allocated to the 'fast track'.

82. Claimants seeking to enforce rights to redress under the UK GDPR and EA are likely to use either the small claims or fast track. In rare cases, claims may be allocated to the Media and Communications List of the King's Bench Division of the High Court.

i. Complexity and need for legal advice

83. The small claims track is in theory designed for individuals to use without legal representation. However in practice even a claim on the small claims track is very challenging for an ordinary person. The fast track and High Court even more so. This is particularly acute where complex or novel issues arise, which the Scenarios present.

84. Claimants may very well therefore need legal representation to effectively enforce rights to redress through the courts.

ii. Paying for legal advice

85. Legal aid is not generally available for claims under the UK GDPR (unless the infringement is by a public authority), but may be available for claims for discrimination under the EA. However this is means-tested and the threshold for eligibility is very low, meaning it will not be a practical option for most claimants even in the types of case where it is theoretically available.

[50] https://www.gov.uk/government/publications/small-claims-track-fast-track-and-multi-track-ex305-and-ex306

86. Where the claim is under the EA, support with legal costs may be available from the EHRC under §28-9 EA 2006. This will not be made available in all cases however, and only 14 instances of such support being made available are listed on the EHRC's website.

87. Failing such support, a claimant would need to fund their own legal fees which may be very substantial – especially in the High Court – and might have to be funded through a damages-based agreement (i.e. using any compensation achieved to pay legal fees at the end of a case). Even a damages-based agreement will not be an option where the remedy sought focuses on non-monetary redress as opposed to financial compensation.

88. Legal fees are in addition to court fees which will be payable in any event and may run into the high hundreds or even low thousands of pounds for one claim, depending on its value.

iii. Adverse costs

89. There is very limited (though not zero) adverse costs risk on the small claims track. But on the other tracks of the county court and in the High Court, there could be a very real risk of being made to pay a significant proportion of the other side's legal costs if unsuccessful. For many this could make bringing a claim unrealistic[51].

iv. Time to resolution

90. Recent statistics show that the average time for a small claim to reach trial is approximately 1 year, while claims on other tracks take significantly longer[14].

v. Complaints to the ICO: *not* a right to redress

91. As mentioned in section E above, the ICO is required to facilitate the making of – and respond to – complaints by data subjects about infringements of the UK GDPR. This might at first appear to provide an alternative route by which rights to redress can be enforced, but it is not. The ICO's obligations in relation to complaints are limited (s.165(4) DPA):

> *"If the Commissioner receives a complaint under subsection (2), the Commissioner must—*
>
> > *(a)take appropriate steps to respond to the complaint,*

[51] See e.g. Warby J in *Lloyd v Google* [2018] EWHC 2599 (QB) at [29]: *"The claim is being funded by Therium Litigation Funding IC ("Therium"), an investment vehicle associated with and advised by Therium Capital Management Limited. Therium has engaged to provide funding in up to three tranches, the first and second being of £5 million each, and the third of £5.5 million."*

> *(b)inform the complainant of the outcome of the complaint,*
>
> *(c)inform the complainant of the rights under section 166, and*
>
> *(d)if asked to do so by the complainant, provide the complainant with further information about how to pursue the complaint."*

92. Case law[52] has clarified that the 'appropriate steps' requirement is itself extremely limited in scope. It is a purely procedural requirement and not a substantive one. That is, the response provided by the ICO is not held to any objective standard and need not provide any vindication to a data subject, even one who is right about how their rights have been infringed. Indeed, the ICO is not even required to inform the data subject of *whether* their rights have or have not been infringed. All that is required is that the ICO investigate the complaint in some way. In practice this may mean as little as forwarding it to a policy team investigating the broad issue raised, and informing the complainant that it has done so.

93. The right to complain to tle ICO is therefore of very limited practical utility to data subjects[53]. As well as not being a means of obtaining redress, it may not even assist in building a case against a decision-maker with a view to bringing a civil claim.

vi. Summary

94. The clarity of legal rights to redress – including financial compensation – for AI harms through cross-cutting legislation is undermined by the fact that enforcing those rights is impractical in all but the strongest cases, or in the unlikely event that the claimant is very wealthy, since:

    a) Many claimants will require expensive legal representation, which could either prevent a claim entirely or significantly reduce compensation received;

    b) Adverse costs risks could make bringing a claim unrealistic; and

    c) Resolution through the courts is time-consuming and slow.

[52] *R (Delo) v Wise Payments Ltd* [2022] EWHC 3046 (Admin) and *Killock and Veale v The Information Commissioner [2021] UKUT 299 (AAC)*

[53] Commentators have noted that the 'overwhelming majority' of complaints to the ICO result in 'no further action': Erdos (2022) *Towards Effective Supervisory Oversight? Analysing UK Regulatory Enforcement of Data Protection and Electronic Privacy Rights and the Government's Statutory Reform Plans,* University of Cambridge Faculty of Law Research Paper No. 16/2022: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4284602

H. Scenario 1: Employment and recruitment

A supermarket chain operates an automated system to allocate shifts to warehouse workers and set variable levels of pay. The automated system uses AI to analyse data on the previous availability and productivity of individual workers, but does not account for the reasons behind particular absences or dips in productivity. In some instances this may result in some workers being given few or no shifts, and in others it may lead to contract termination. This system is also used to make inferences about the potential performance of prospective employees as part of a pre-hire sifting process, which uses postcode data from CVs to link job applicants to risk profiles created on existing workers.

95. This scenario discloses three principal types of harm, relating to both existing and prospective employees:

a) Existing employees may suffer detriment – either a reduction in pay or offered hours or termination of employment – for previous low availability or productivity in ways that are unfair, for example by being linked to protected characteristics such as disability.

b) Prospective employees may be less likely to be offered employment based on inferences made about them, which may be inaccurate or unfair.

c) Employees' working conditions may be significantly worse as a result of the automated system, with progressively more stretching auto-generated productivity targets and always-on monitoring creating an atmosphere of stress, high-pressure and a lack of dignity.

96. We assume that the people working in the warehouse are on zero-hours contracts of employment (as opposed to having only worker status). Where individuals are workers but not employees, the section below on unfair dismissal under the Employment Rights Act 1996 ('**ERA**') will not be applicable. We also assume that there are no relevant collective agreements in place for the supermarket chain's warehouse workers.

97. There is 'sector-specific' regulation (and a dispute resolution mechanism) applicable to this Scenario, but no 'regulator' of employment as such. As in Scenario 2 (and, in respect of the UK GDPR, Scenario 3), the Equality Act 2010 ('**EA**') and UK GDPR are relevant cross-cutting regulations applicable to this Scenario 1.

Is there processing of special category data?

98. This Scenario raises the prospect that profiling of warehouse employees may be carried out which *correlates* to employees' health. For example, dips in productivity or absences from work might be due to health issues.

99. The UK GDPR makes special provision for processing of personal data '*concerning health*' (Article 9(1)). In our view it is doubtful that the processing in this Scenario is *'concerning'* health even if, for some individuals, productivity or absence scores *correlate* to health issues. It does not appear to be the intention of the supermarket chain to obtain or record data concerning health, and for the substantial majority of data subjects, the profiling will not in fact concern their health, merely their availability and productivity (see also para 167).

100. However we have noted below the additional considerations which would apply, were it successfully argued that the use of the automated variable shifts/pay system does constitute the processing of data concerning the health of warehouse employees.

Are decisions to terminate automated?

101. The Scenario contemplates that profiling by the automated system which 'leads to' termination of employment contracts. To determine warehouse employees' rights, it will be important in practice to establish whether such decisions are 'solely automated' for the purposes of the UK GDPR. Article 22A(1) provides that a decision is solely automated if there is no '*meaningful human involvement*' in it. For this analysis we assume that supermarket chain managers follow recommendations from the automated system very closely, effectively as a 'tick-box' exercise, meaning that decisions to terminate are solely automated. However it is important to note that greater human manager involvement such as taking other factors into account or overriding the automated system's recommendation in an appreciable number of cases, would render the decisions *not* solely automated and outside the scope of Articles 22B and 22C UK GDPR.

i. ***Ex ante* regulation**

Sector-specific requirements

102. There are no sector-specific ex ante regulatory requirements that would *prevent* the supermarket chain in this Scenario from introducing the automated variable shifts/pay system[54].

Cross-cutting requirements

Assessment of High Risk Processing and DPDD

103. An AHRP will be required under Article 35 UK GDPR (as amended by the DPDI Bill) as this would be high-risk processing. However it only requires a relatively high-level summary of the processing and is not required to be reviewed externally, so offers only limited practical protection against this kind of harm. It should in theory address how the

[54] It is assumed that the system does not permit rates of pay to fall below the national minimum wage (s.1 National Minimum Wage Act 1998), which would be unlawful.

supermarket chain's use of the automated variable shifts/pay system complies with the principle of DPDD.

104. It is notable that there is no requirement in the UK GDPR to consult with representatives of employees when carrying out an AHRP[55], further increasing the risk that the supermarket chain would fail to recognise the risks of harm in this Scenario, or fail to put in place sufficiently robust mitigations against them.

ICO Enforcement

105. As set out in section H(ii) below, the harms in this Scenario may constitute breaches a number of rights of individuals under the UK GDPR. To the extent the ICO becomes aware of such breaches – such as through data subject complaints, submissions from civil society, or its own investigations – it would be open to the ICO to take enforcement action to prevent the harms from occurring again in future.

106. However as set out in section E, the scale of the ICO's mandate relative to its resources means that proactive enforcement by the ICO of rights under the UK GDPR provides only limited protection against the risks in this Scenario from materialising.

The Equality Act 2010

107. Where the automated variable shifts/pay system treats individuals with protected characteristics (such as disability) less favourably, this may amount to discrimination (see section H(ii) below).

108. The EHRC's regulatory work might address the risks in this Scenario in a way that alerts the supermarket chain the potential breaches of the EA, but as discussed in section F the EHRC's lack of any enforcement powers, and remit relative to resources, means that this provides very limited ex ante protection in this Scenario.

109. **In sum**, ex ante regulation covering the harms in this Scenario is limited to that in the UK GDPR and Equality Act. The effectiveness of these requirements in preventing this type of harm is severely limited by the roles played by the ICO and EHRC respectively.

ii. ***Ex post* redress**

How would an individual know about and evidence harm?

Requirement for a statement of particulars of employment

110. Both workers[56] and employees are entitled to a statement of particulars of employment from the supermarket, chain. This must include *"the scale or rate of remuneration or the method of calculating remuneration"* (s.1(4)(a) ERA). It must also include information on

[55] This is a change from the previous position which imposed a requirement to do so 'where appropriate'.

[56] This is one requirement of the ERA which applies to workers who lack employee status. We assume that any workers in the Scenario started working for the supermarket chain after 1 January 2021.

"whether or not [hours of work] may be variable, and if they may be how they vary or how that variation is to be determined" (s.1(4)(c)(iii) ERA).

111. So the supermarket chain would certainly have to provide the employees with at least *some* information on how the automated system used past absence and productivity data to set rates of pay and shifts allocated[57]. However, we are not aware of any case law dealing with the application of s.1(4) to this method of varying pay rates and shifts offered that clarifies the extent of information required.

112. An employer might argue that it is sufficient merely to state that past absence and productivity are taken into account. Conversely, it could be argued that the references to the '*method'* of calculating pay and '*how'* variation in hours is determined require detailed explanations of precisely how past absence and productivity factor into the calculations. It is unlikely that the statement of particulars would be required to detail any hypothetical ways in which the automated system might be *unfair*, such as by drawing attention to the potential for correlation between periods of absence and disability.

113. Unless and until tested by case law, these provisions therefore provide only relatively limited transparency, alerting warehouse employees that a system is used which takes past absence and productivity into account when setting pay and offering shifts.

Provision of reasons for dismissal

114. Where an employee[58] has been in continuous employment for two years with the supermarket chain, they are entitled to a "*a written statement giving particulars of the reasons for the employee's dismissal"* (s.92 ERA). As with s.1(4) this has not been tested in the context of this kind of technology, so it is possible that the supermarket chain can comply by stating that the employee's productivity or absence rate was too low or high (respectively), without necessarily providing details of how the automated system made those calculations.

GDPR Transparency and data access

115. Articles 13 to 15 would apply, which should *in theory* ensure:

a) Individuals are aware that processing – including automated decision-making – is taking place, and including information about the logic of that decision-making.

[57] It might also be argued that the common law requirement of trust and confidence in the employment contract requires the supermarket chain to explain decisions on shift allocation and pay rates (see e.g. *Commerzbank AG v Keen* [2006] EWCA Civ 1536 §44) but it is unlikely that this provides any additional transparency beyond s.1(4) ERA.

[58] Not applicable to those with only worker status.

b) Individuals can obtain access to their personal data, which may be an important means of discovering and evidencing how an individual has been profiled by the productivity algorithm.

116. However, as set out in section C(i), there are significant limitations on the rights to transparency and data access in practice. Of particular relevance is the limitation of the right of access so as not to *“adversely affect the rights and freedoms of others.”* Controllers in the gig economy have used these qualifications in the GDPR as a way to limit the data and explanations they provide to workers about algorithmic systems that determine pay, allocate shifts, and monitor productivity.

117. By way of example a controller might be able to comply with a vague statement to current workers such as “We use a range of personal data about you including your past shifts, location, and what you’ve told us about your preferences to generate the best match between an individual worker and an available shift. We never make solely automated decisions about you which have a legal or similarly significant effect”. This might alert workers to some degree of productivity and availability profiling, but not reveal how it could unfairly discriminate against e.g. disabled workers. Were a worker to exercise his/her right of access, he/she might receive his productivity and availability scores, but no explanation or underlying calculations as to why they were low. There is extensive evidence of this kind of approach from controllers from within the gig economy in particular[59]. The lawfulness of such approaches is debatable, but there is no case law requiring controllers to go beyond them, and they represent common practice.

118. In other jurisdictions, arguments by workers that the UK GDPR requires fuller explanations of this type of algorithmic system have been successful[60] but this has not been tested in the English courts.

119. Motivated and informed warehouse workers (current and prospective) would therefore likely be able to establish (e.g. from reading a privacy notice) that some kind of profiling is taking place. But would likely find it very challenging to understand the full workings of that profiling and – by extension – that the profiling could result in unfairness, creating a real barrier to obtaining redress.

Article 22C Transparency

120. Where the automated variable shifts/pay system terminates an employee’s contract, this will be a significant, solely automated decision (Article 22A(1) UK GDPR) since the

[59] https://www.workerinfoexchange.org/wie-report-managed-by-bots
[60] https://protect-eu.mimecast.com/s/W8lJCg5BQCp6m2miEbjUA?domain=uitspraken.rechtspraak.nl (in Dutch) and https://protect-eu.mimecast.com/s/h7J1Cj2GOiPE3q3Sn9RDP?domain=uitspraken.rechtspraak.nl (in Dutch)

termination of an employment contract is a legal effect. The same is true of the automated sifting out of individuals at the hiring stage.

121. Unlike under the UK's previous data protection regime, this solely automated and significant decision would not be prohibited, since it does not involve the processing of SCD (Article 22B UK GDPR as amended by the DPDI Bill). However, the UK GDPR imposes transparency obligations about the decision.

122. Article 22C requires safeguards to be in place, including a requirement to *"provide the data subject with information about decisions* [which are solely automated and significant]*".* As this requirement is new, there is no case law or guidance on what information is required. It might be assumed that the obligation goes beyond the requirements of Articles 13 and 14 (otherwise it would be redundant) but this is not supported by specific case law[61].

123. It is also notable that the requirement is a general one – to provide information about '*decisions*' rather than each decision. It may well be that this can be discharged through a general, up-front privacy notice to employees, rather than by way of specific notification of *each* significant solely automated decision made about an employee (or prospective employee) *at the time it is made* – that is, at the time that the decision to hire or terminate is made.

124. It is not clear that this additional transparency requirement would apply to decisions by the automated system to vary pay rates or shift allocations, since it is unclear whether such decisions have 'legal or similarly significant' effect on current workers/employees. Under a zero-hours contract, a worker has no right to be offered any specific number of hours. A change in the number of hours offered in any one week therefore by definition does not affect the worker's legal rights. The same is true where a contract does not entitle a worker to a specific rate of pay (provided it is above the minimum wage) albeit this may be an unusual arrangement.

125. For that reason there is a strong argument that a change in hours offered - or rate of pay – does not have a legal or similarly significant effect on a worker. At the margins, this might be different where a worker has very consistently been offered a large number of hours and this precipitously drops off as a result of a solely automated decision, but there is no guidance on this point, and it has not been tested in case law. At a minimum, it should not be assumed that Article 22C applies to these decisions. That is,

[61] And the costs of and barriers to satellite litigation to establish this may make it redundant in practice.

transparency requirements regarding decision-making for changes to shifts and pay rates are limited to those in Articles 13-15 UK GDPR.

126. **In sum**, both employment-specific and cross-cutting regulation impose transparency and data access requirements around the use of the automated variable shifts/pay system. This is particularly the case where the system is responsible for a decision to terminate a contract of employment (especially where the employee has two continuous years' service) or sift someone out of a hiring process.

127. *However*, it is likely that these transparency requirements will only alert (prospective) warehouse workers to the use of the automated system, with (at best) a high-level explanation of how it works. They are unlikely to have access to a detailed explanation of its algorithmic logic, or to notification at the point of any specific automated decisions affecting them. This would present substantial challenges in *evidencing* specific harms or breaches as a precursor to obtaining *ex post* redress.

Sector-specific rights to redress

Unfair dismissal

128. Employees with two years' continuous service for the supermarket chain have a right not to be unfairly dismissed (s.94 ERA). The Scenario contemplates the automated system leading to terminations of employment for either low availability (i.e. failure to take up offered shifts) or low productivity.

129. A dismissal can only be fair if it is based on capability, conduct, redundancy, or legal requirement (s.94(2) ERA). An automated dismissal for failure to take up offered shifts would likely therefore be unfair.

130. Where an automated dismissal was for low productivity (i.e. 'capability' of the warehouse employee), it may well nonetheless be unfair due to procedural unfairness[62], depending on the process used by the supermarket chain in making use of the employee's automatically generated profile.

131. Where a dismissal meets the definition of 'solely automated' for UK GDPR purposes, it will almost certainly be procedurally unfair for the purposes of the ERA, since no disciplinary procedure is followed at all. Even if there is some limited human involvement in the dismissal, it may still be procedurally unfair if it significantly varies from the ACAS Code of Practice[63].

[62] *Polkey v AE Dayton Services Ltd* [1988] AC 344

[63] https://www.acas.org.uk/acas-code-of-practice-on-disciplinary-and-grievance-procedures

132. A successful claim for unfair dismissal would entitle a warehouse employee to compensation (§118-127 ERA)[64]. The law on the amount of compensation is complex but it typically consists of a 'basic' award calculated by reference to years of service, plus a 'compensatory' award for consequential losses (such as lost income whilst looking for a new job), such that a successful unfair dismissal claim goes much of the way to redressing the harm caused by the unfair dismissal. However, the warehouse employee would not be entitled to compensation for distress unless there is a discrimination element to the dismissal (see para 138 below).

133. Where a dismissal is unfair *only* for procedural reasons (i.e. if the supermarket chain would have dismissed the employee for low productivity anyway, following a proper process), any compensation for the unfair dismissal will be significantly reduced[65].

Practical considerations for *ex post* redress in employment law

134. Disputes about unfair dismissal (or a failure to provide particulars of employment or a reason for dismissal, where applicable) are adjudicated by the Employment Tribunal (s.111 ERA and §1-2 Employment Tribunals Act 1996).

135. Use of the Employment Tribunal is free, and adverse costs risk is limited[66]. However, whilst the Employment Tribunal is in theory designed to be capable of use by unrepresented claimants, in practice the majority of claimants are represented[67]. As a corollary to the low risk of adverse costs, a warehouse employee claimant's own legal costs would only be recoverable in the Employment Tribunal where the supermarket chain acted unreasonably in defending the claim (i.e. in very limited circumstances).

136. In practice this means there are some real barriers to ordinary people bringing Employment Tribunal claims in this Scenario, since the majority of claimants will need to find either free legal representation or meet costs through damages-based agreements.

137. The law of unfair dismissal only deals with *one* of the harms raised by the Scenario, and that only for employees with more than two years' continuous service for the supermarket chain. Employment law does not provide protection for employees against the decline in their working conditions or increased stress or indignity that may result

[64] In rare cases the Employment Tribunal may order that an employee be reinstated (§113-116 ERA). This is more likely where an employee can return to the same or a similar role with a different team, which may be possible assuming the supermarket chain has multiple locations.

[65] *Polkey*

[66] An adverse costs order may only be made against a claimant where they or their representative have acted unreasonably or brought a claim with no prospect of success (Rule 76, Employment Tribunals (Constitution and Rules of Procedure) Regulations 2013.

[67] In the most recent quarter for which statistics are available, 67% of claimants in the Employment Tribunal were represented.

from having their productivity constantly automatically monitored under the threat of reduced working hours or pay.

Unlawful discrimination

Disability discrimination for existing employees and workers

138. As set out at section D, The Equality Act 2010 ('**EA**') protects against discrimination on the basis of a number of 'protected characteristics', of which 'disability' is relevant to this Scenario. Periods of low availability (i.e. failure to take up offered shifts) or low productivity may arise due to a warehouse employee's disability. Reducing pay or shifts offered, or terminating a disabled employee's employment could therefore amount to discrimination:

a) s.15 EA provides that unfavourable treatment "*because of something arising in consequence of B's disability*" (where the employer knows – or should know – about the disability) is discriminatory. Reduction of a disabled employee's hours or pay (or termination of their employment contract) due to low availability or productivity caused by their disability could meet this definition and therefore be unlawful when read with s.39 EA (or s.41 for workers without employment status).

b) s.19 EA provides that application of "*a provision, criterion or practice*" which puts disabled warehouse employees "*at a particular disadvantage*" compared to non-disabled ones is indirectly discriminatory. The supermarket chain's *practice* of reducing shifts/pay and terminating contracts for low availability could meet this definition[68] and therefore be unlawful when read with s.39 EA (or s.41 for workers without employment status).[69]

139. In both cases, the supermarket chain would not be liable if it can show that its conduct is a '*proportionate means of achieving a legitimate aim'*. The supermarket chain's desire to run an efficient warehouse with acceptable levels of productivity is a legitimate aim. Whether automatic reductions in pay or shifts (or termination) for a disabled warehouse employee is *proportionate* to that aim will be case specific. However our view is that the automatic changes described in the Scenario would – without any further safeguards – likely be unlawful disability discrimination to the extent they are applied to disabled employees.

140. A successful claim for disability discrimination entitles the claimant to compensation. This includes consequential losses (such as lost income while seeking other

[68] Note that these tests are looser than the test for whether there is processing of special category data under Article 9 UK GDPR.

[69] For job applicants, the Scenario only foresees profiling on the basis of postcode. This would not be related to any protected characteristic of an applicant, so there would be no discrimination under the EA.

employment), injury to feelings, and can include aggravated damages. The average total award for disability discrimination in 2021/22 was £26,172[70].

Practical considerations for bringing claims for disability discrimination

141. Claims for disability discrimination by employees and workers may be brought in the Employment Tribunal (s.120 EA), as to which see para134 above.

142. Therefore where the automated system unfairly treats an employee with a disability, cross-cutting regulation provides a clear cause of action and a forum in which to bring it (the Employment Tribunal) which is preferable to bringing a court claim. However this is only relevant to one of the three harms raised by the Scenario.

GDPR rights to redress

Lawful basis for the processing

143. The processing by the automated variable shifts/pay system could be carried out in reliance on either Article 6(1)(b) (necessary for performance of – or for steps taken prior to entering into[71] – the employment contract) or 6(1)(f) (necessary in the legitimate interests of the supermarket chain)[72].

144. Whilst 'necessary' in the context of Articles 6(1)(b) and (f) does not mean 'strictly necessary', there is guidance and case law to suggest that the processing should be more than merely useful, and that less intrusive means ought to have been considered[73]. The legitimate interests basis also requires a balancing of the supermarket chain's interests against those of existing and prospective employees as data subjects. In this Scenario those data subjects' interests would weigh fairly heavily due to the risk of contract termination and even discrimination as a result of the processing, as well as the negative impact on general working conditions and employee dignity. There may be further arguments against the necessity of the processing of applicants data to the extent that there is no actual connection between postcode and likely availability/productivity, since such processing would not be helpful in sifting applicants and therefore could not be 'necessary'.

145. Warehouse employees and applicants therefore have a number of ways to challenge the lawfulness of the processing involved in the automated system by arguing that it lacks a valid legal basis under Article 6 UK GDPR. Although there is no English case

[70] https://www.dacbeachcroft.com/en/gb/articles/2023/january/employment-tribunals-statistics-published/

[71] The full wording relevant to applicants is that the processing must be necessary *"in order to take steps at the request of the data subject prior to entering into a contract."* In this case the steps would be consideration of the job application, which is at the request of the data subject/applicant.

[72] We do not consider that consent would be a viable lawful basis given the power imbalance between the supermarket chain and its current and prospective employees. Articles 4 and 7 UK GDPR require *inter alia* that consent be 'freely given' and capable of being withdrawn. Guidance strongly indicates that consent will not generally be valid in an employment relationship: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/lawful-basis/consent/when-is-consent-appropriate/

[73] https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines-art_6-1-b-adopted_after_public_consultation_en.pdf and *Rigas*

law directly analogous to this situation, we consider that such challenges would have a reasonable prospect of success.

The right to object

146. Where the supermarket relies on its legitimate interests, individual data subjects may object to the processing under Article 21(1) UK GDPR where there are factors specific to them justifying that objection. Unless the supermarket chain can give compelling legitimate grounds to continue the profiling of an employee or applicant who objects, the profiling would need to stop (or be prevented from happening in the first place). However, in a zero-hours contract context, the utility of an individual right to object is clearly very low, since the supermarket chain can simply respond by ceasing processing and either not considering an application or ceasing to offer shifts to the objecting data subject.

Processing of special category data

147. As stated above, we do not consider that the system described in the Scenario is processing special category data. However, if it *were* shown to be doing so, this would require explicit consent from warehouse employees or applicants (Article 9(2)(a)) since such processing is prohibited unless an exemption applies, and none of the other exemptions in Article 9 could apply in this Scenario.

148. Therefore if employees or applicants can surmount the hurdle of demonstrating that SCD is being processed, this would make the processing unlawful (assuming they do not explicitly consent to it).

Accuracy of processing in relation to applicants

149. The Scenario states only that job applicants are linked by postcode to current employees (presumably those from postcodes with similar socio-economic profiles). That is, inferences are made about the likely availability and productivity of prospective applicants based on their coming from similar areas to existing employees.

150. We are not told whether applicant postcode is in fact a strong and reliable predictor of productivity and availability. There seems to be at least a reasonable chance that it does not predict productivity or availability, or does so only very weakly. If that is true, then the profiling of applicants by postcode would likely be inaccurate in breach of Article 6(1)(d). Accuracy is a relative standard, depending on the context and consequences of processing. Where the ability to be considered for employment is on the line, a weak or unreliable inference about availability/productivity based postcode on would likely not meet the standard of accuracy required by the UK GDPR and would therefore be unlawful.

Fairness of processing (existing employees)

As stated in section C, the concept of fairness does not extend to imposing requirements of procedural fairness or fair decision outcomes. Rather it is related to the reasonable expectations of data subjects about how their data will be processed. In our view it would be difficult to argue that assessing employees' past availability and productivity and making decisions on those bases is so unexpected as to be unfair for the purposes of Article 5(1)(a).

Fairness of processing (applicants)

151. For applicants however, the lack of an apparent rational connection between one's postcode and likely availability and productivity levels could be enough to render the processing unfair. It is well arguable that it would be completely outside applicants' expectations that inferences about their productivity would be drawn from such an apparently unrelated characteristic, making the processing unfair and unlawful under Article 5(1)(a).

Automated decision-making

152. A termination of employment by the automated system would be a solely automated and significant decision[74] (Article 22A(1)). The ICO has also stated that "*e-recruiting practices without human intervention*" will constitute decisions with 'legal or similarly significant effect' for the purposes of Articles 22A-C UK GDPR[75]. The use of inferences generated by the automated system for pre-hire sifting meets this definition, meaning both current and prospective workers/employees are subject to solely automated significant decisions by the automated system.

153. If the processing of the automated system is established to be 'concerning' health[76], then it may be that these automated decisions are *"based entirely or partly on special categories of personal data referred to in Article 9(1)"* (Article 22B) and therefore unlawful[77]. This may be easier to demonstrate for an existing employee with an actual health condition, but could also apply to an applicant, because a postcode-based inference about productivity would be – at least in part – an inference that the applicant is more likely to have a health condition[78].

[74] Given our assumption about the minimal level of human involvement.

[75] https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/individual-rights/automated-decision-making-and-profiling/what-does-the-uk-gdpr-say-about-automated-decision-making-and-profiling/#id3

[76] Which, as stated above, we do not think is likely.

[77] Article 22B exempts some decisions from this prohibition, but that exemption requires the decision to be *both* '*necessary for entering into or performing a contract*' and '*necessary for reasons of substantial public interest*. Whilst the first limb of this exemption may be met, the second is clearly not.

[78] Note that it is not relevant that the automated system is trained using some data subjects' data (existing employees) and then used to make decisions about others (applicants). This is because any inference made about an applicant – even if it draws on training data from other data subjects – will be personal data *of the applicant*. It is also notable that Article 22B only requires that the decision be 'based entirely or partly on [SCD]' – it does not specify that it needs to be based on the SCD of any particular data subject.

154. If – as seems more likely – the automated decisions do *not* concern SCD, then Article 22C merely requires safeguards to be in place, including the provision of information about automated decisions and the ability for employees or applicants to seek human review of the decision[79].

Causation, loss and compensation for GDPR breaches

155. We identify a number of ways in which (prospective) warehouse employees might be able to show that the profiling of current and prospective warehouse employees is unlawful, including that it:

a) Lacks a valid legal basis (or exemption from the prohibition in Article 9);

b) Is inaccurate and/or unfair (for prospective applicants); or

c) Involves prohibited solely automated significant decisions.

156. Whilst there may be strong arguments that the processing is unlawful, claimants would only be entitled to financial compensation[80] where it can be shown that the breach of the UK GDPR *caused* loss (which may include distress). So for example:

a) If an applicant has an inaccurate inference made about them based on their postcode, they would need to prove that *but for* that inaccuracy, they would have got the job in order to claim financial compensation for the breach (beyond any distress).

b) If an employee is terminated as a result of a prohibited automated decision (Article 22B) (or a permissible one but without safeguards in place, Article 22C), then they would need to prove that *but for* the infringement they would have been kept on in employment in order to claim compensation for lost earnings. If the supermarket chain can show that any manager intervening would have made the same decision to terminate, only compensation for distress (to the extent distress occasioned *by the decision being automated* can be proven) will be available.

157. As well as the difficulties in enforcing data rights detailed at section G, it must be borne in mind that no compensation is available for the mere fact of a breach of the GDPR. In this Scenario, a (prospective) employee might be able to show some distress caused by

[79] It is notable that whilst the DPDI Bill requires guidance and codes of practice to be published by the ICO on some matters, it does not require any specific guidance to be prepared on the safeguards required for permitted significant and solely automated decisions.

[80] Claimants would also be entitled to injunctive relief – i.e. a court order to stop any unlawful processing – and in some cases declaratory relief, i.e. court confirmation that the processing is unlawful.

the GDPR breaches, but the level of compensation would be low (i.e. in the low hundreds of pounds[81]).

158. **In sum**, the GDPR on its face appears to provide protection in a number of ways against the harms in this Scenario. Unlike employment and equalities regulation, the GDPR offers potential routes to challenge the processing itself, and therefore the full range of risks which it poses, including to general working conditions.

159. But even if a (prospective) employee can overcome challenges in (i) finding out about the supermarket chain's breach(es), and (ii) obtaining enough evidence of it/them, there remain significant obstacles to enforcing their GDPR rights through the courts (as discussed in section G above). Even where the prospects of obtaining a court order to stop the processing are good, this may be of very limited value to employees in an economically insecure position who might have their hours reduced or contracts terminated (particularly if they only have worker status or have less than 2 years' continuous service). The burden of securing compliance with regulation of direct relevance to the risks is placed on those who can ill-afford to bear it.

iii. <u>Conclusion: level of effective protection from AI Harm(s) in this Scenario</u>

160. Of the three, this Scenario highlights an area in which effective protection from AI harms is the **weakest**. Whilst there is some sector-specific regulation, it only makes certain practices unlawful relating to individuals with employment status and protected characteristics, failing to address the totality of the harm envisaged by the Scenario. A lack of a sector-specific regulation (and consequent reliance on the ICO and EHRC which have limited resources, information, and enforcement capacity) compound this, as does a lack of effective transparency, which would likely prevent individuals from understanding that they might have been unfairly treated by the supermarket chain's use of this algorithm.

[81] We assume that no recognised psychiatric injury such as clinical stress or depression has been caused by the unlawful processing.

Are there legal requirements that the decision-maker must consider in advance?	Is it likely that a regulator would prevent the AI harm through enforcement of those requirements?	Would the individual be able to find out about and evidence the harm?	Is there a legal right to redress for the harm?	Is it practical for individuals to enforce any legal rights to redress?
Scenario 1 (Employment)				
Limited: The UK GDPR and Equality Act impose some requirements, but these do not address all the harms in the scenario or fundamentally prevent the tool from being used.	Unlikely: relies on enforcement by the ICO and the EHRC, both of which are limited in the information available to them, their powers and enforcement approach, and their resources.	Low/medium: Some additional protections from ERA in relation to statements of pay.	Medium: GDPR and Equality Act give rise to causes of action for some harms in the Scenario (but not those relating to general working conditions). Additionally, some harms covered by the Employment Rights Act where an employee is dismissed.	Impractical: requirement to bring a civil claim for GDPR breaches. Employment Tribunal for ERA and Equality Act breaches. But this relies on having a protected characteristic and/or employment status, and does not protect against diminished working conditions.

I. Scenario 2: financial services

As part of a mortgage application process, a lender requires applicants to undertake a video interview. The recording of these interviews are analysed with biometric technologies that use AI to infer characteristics to create a risk score. Applicants have a range of prosody and accents. Some have speech impairments due to disability or another condition and some applicants may be vulnerable, for example by reason of reduced mental capacity. This information is considered by the lender as part of their due diligence for the mortgage application. The lender ultimately chooses to deny many mortgage applications, with the risk score as part of the matrix. These decisions are made by human staff members, but the use of the risk scores provided by the biometric software is influential. The human staff members are only provided with the risk score, rather than any underlying information about how the score was calculated.

161. This Scenario discloses one principal kind of harm, that individuals might receive a low biometric risk score, reducing their access to credit, due to patterns of speech or facial expression that have no (apparent) rational connection to their creditworthiness. The Scenario indicates that low scores may arise due to:

- Regional or ethnic accents;
- Speech impairments (which may constitute disabilities); and/or
- Other impairments which may constitute disability, such as neurodivergence.

162. The source of the low score will have an impact on the level of protection afforded by regulation, as set out below.

163. There is 'sector-specific' regulation applicable to this Scenario, which is overseen by a dedicated regulator, the Financial Conduct Authority ('**FCA**'). Separately, there is also a dedicated sector-specific dispute resolution mechanism, the Financial Ombudsman Service ('**FOS**'), which resolves disputes by applying sector-specific regulation (s.226 of the Financial Services and Markets Act 2000, '**FSMA**').

164. As in Scenario 1 (and, in respect of the UK GDPR, Scenario 3), the Equality Act 2010 ('**EA**') and UK GDPR are relevant cross-cutting regulations applicable to this Scenario.

Does the technology work?

165. The Scenario does not state whether or not the biometric risk score is accurate. That is, it is not known whether the biometric risk score is a reliable and robust indicator of future creditworthiness which it would make sense for the lender to take into account in deciding whether or not to provide a mortgage. For this analysis, we assume that the biometric risk score *has not* been demonstrated to provide a reliable and robust indicator

of creditworthiness. Where this assumption is important to the legal analysis, we highlight it below.

Is there processing of special category data?

166. The UK GDPR makes special provision for processing of certain categories of personal data in Article 9. Most relevant for this Scenario, Article 9 prohibits (unless a valid exemption applies):

a) Processing of biometric data *for the purpose of* uniquely identifying a natural person[82];

b) Processing of personal data *revealing* racial or ethnic origin; and

c) Processing of data *concerning* health.

167. This raises the prospect that the creation of the biometric risk score may be processing of special category data engaging Article 9 ('**SCD**'), but this is not straightforward and depends on a very close reading of Article 9:

167.1. Where a biometric risk score is at least in part determined by characteristics related to racial or ethnic origin (e.g. a strong ethnic accent), then this may well be said to be processing 'revealing' those special characteristics. This is likely the strongest of the three arguments, but even this is not certain. The EDPB has stated83 that processing such as video surveillance, which plainly 'reveals' individuals ethnic origin, does not always engage Article 9, implying that there is a requirement that controller have a degree of intent to capture or reveal such information.

167.2. Where a risk score is at least in part determined by characteristics relating to a health condition (such as a speech impairment or neurodivergence) then it could be said to be processing 'concerning health'. Both 'concerning' and 'health' require some unpacking:

- The concept of 'concerning' is (confusingly) further defined in Article 4(15) as meaning 'related to' physical or mental health, and '*reveal[ing] information about […] health status*').

[82] Article 4(14) further clarifies this as meaning *"personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data."*
[83] Guidelines 3/109 at §62

- The word 'health' is not further defined, but it presumably imports a relationship to a generally recognised medical condition (though it would not be necessary for that condition to have been formally diagnosed)[84].

Thus if the generation of a low biometric score reveals that someone has a diagnosable physical or mental health condition, this could engage Article 9. However it is not the lender's *intention* to identify or record health characteristics specifically. By the same logic as that for 'revealing' ethnic origin, Article 9 might not be engaged (though there is no direct guidance on this point).

167.3. A 'biometric risk score' might be thought self-evidently to be '*processing of biometric data'* for the purposes of Article 9. But the Article 9 definition goes on to provide that the processing must be '*for the purpose of uniquely identifying'*. Whilst any risk score will be linked to a specific applicant, it is not clear that this type of biometric processing – where the aim is to classify or analyse a person, rather than uniquely identify them – meets the definition[85].

168. The most that can be said is that for some applicants, the creation of the risk score may constitute processing engaging Article 9, although this might be very difficult to prove. We set out below the additional considerations where Article 9 is engaged, but it should be borne in mind that it is far from certain that an applicant would succeed in persuading the ICO or a court that the application process involved the processing of SCD. Further, for many applicants the question will not arise, such as where a low score is due to a regional rather than an ethnic accent.

Are decisions to refuse credit automated?

169. The Scenario states that humans take the final decision to deny credit to applicants, and that the biometric risk score is '*part of the matrix'* of factors in that decision (albeit an influential part). The UK GDPR makes special provision for decisions which are *'based solely on automated processing*'. This definition is only met if there is '*no meaningful human involvement'* in a decision; that is, mere 'rubber-stamping' of decisions will not prevent a decision from being solely automated. In this Scenario it seems that whilst the

[84] For the avoidance of doubt it is not relevant whether any such health condition reaches the threshold of a disability within the meaning of the Equality Act 2010.

[85] Guidance on this is somewhat contradictory. The ICO has stated "The individual does not have to be identified for this data to become biometric data - it is the type of processing that matters" in its guidance on the use of live facial recognition in public places published June 2021 [https://ico.org.uk/media/2619985/ico-opinion-the-use-of-lfr-in-public-places-20210618.pdf]. However the European Data Protection Board has stated in Guidelines 3/2019 at §76 that the purpose of unique identification is a crucial part of the definition. Legal commentators in England have argued that the UK GDPR does *not* provide special protection for biometric classification as opposed to identification: *Independent legal review of the governance of biometric data in England and Wales,* Matthew Ryder KC, June 2022 [https://www.adalovelaceinstitute.org/wp-content/uploads/2022/06/The-Ryder-Review-Independent-legal-review-of-the-governance-of-biometric-data-in-England-and-Wales-Ada-Lovelace-Institute-June-2022.pdf]

biometric risk score is important, there is meaningful human involvement in the decision to refuse credit, since other factors are taken into account by the human decision-maker (Article 22A(1) as amended by the DPDI Bill)[86].

i. **Ex ante regulation: sector-specific requirements**

Principle 6 and The Consumer Duty

170. The provision of mortgages is a regulated activity under FSMA (s.22 FSMA). Part 9A, Chapter 1 FSMA empowers the FCA to make 'general rules' about regulated activities, which the lender in this Scenario must follow. FCA rules are extremely extensive, and include specific rules on mortgage lending[87]. However the most relevant rules for this Scenario is the relatively new 'Consumer Duty', which replaces the FCA's Principle 6 (also known as 'Treating Customers Fairly).

171. The new Consumer Duty, which comes into force 31 July 2023, introduces a 'Consumer Principle', which would require the lender in this Scenario to "*act to deliver good outcomes for retail customers* [i.e. including the mortgage applicants in the Scenario]*".* It includes cross-cutting rules requiring firms to act in good faith towards retail customers, avoid causing foreseeable harm to retail customers, and enable and support retail customers to pursue their financial objectives.

172. One of the these cross-cutting rules includes acting in "*good faith by designing products or services to support the objectives and needs of customers in the target market*". The FCA provide "examples of not acting in good faith", including:

> *"Using algorithms, including machine learning or artificial intelligence, within products or services in ways that could lead to consumer harm. This might apply where algorithms embed or amplify bias and lead to outcomes that are systematically worse for some groups of customers, unless differences in outcome can be justified objectively."*[88]

173. The lender in this scenario is therefore subject to a regulatory requirement not to apply the biometric risk score in this way, since it can lead to unfair outcomes which cannot be justified objectively.

[86] It might be possible to argue that there are multiple 'decisions' in the process leading up to the decision to refuse credit. For example the decision to apply the biometric test, or the decision to set an applicant's score. However it is very unlikely that any of these decisions would engage Articles 22A-C, since they do not '*produce a legal effect*' or '*have a similarly significant effect*' on the data subject. To the extent that Articles 22A-D are engaged, see the discussion of their operation in the context of Scenario 1.

[87] The Mortgage and Home Finance Conduct of Business rules: https://www.handbook.fca.org.uk/handbook/MCOB/1.pdf

[88] Non-Handbook guidance for firms on the Consumer Duty at 5.11 https://www.fca.org.uk/publication/finalised-guidance/fg22-5.pdf

Vulnerable customers

174. The FCA has promulgated (on the same legal basis as the Consumer Duty above) separate guidance on the fair treatment of 'vulnerable' customers[89] (the '**Vulnerable Customers Guidance**'). A vulnerable customer is someone who:

> *"due to their personal circumstances, is especially susceptible to harm, particularly when a firm is not acting with appropriate levels of care."*

175. In this scenario, individuals with learning difficulties or other mental disabilities or difference might achieve low biometric risk scores for credit. They would also be vulnerable customers of the lender, to whom it owes special duties under the Vulnerable Customers Guidance, including most notably:

a) Asking [itself] what types of harm or disadvantage the lenders customers may be vulnerable to (e.g. being unfairly denied credit due to a low biometric score);

b) Considering the potential negative impacts of a product or service on vulnerable consumers and designing products and services to avoid potential harmful impacts; and

c) Considering how to communicate with vulnerable consumers; where possible the lender should offer multiple channels so vulnerable consumers have a choice.

176. It is clear therefore that implementing the biometric credit score in the way described without any safeguards for vulnerable customers would be a breach of the FCA rules in the form of the Vulnerable Customers Guidance.

Enforcement by the FCA

177. A full assessment of the FCA's enforcement powers and regulatory activity is not possible within the scope of this analysis, but the FCA has very significant regulatory powers where regulated firms do not follow FCA rules, including the Consumer Duty. For the lender in this Scenario this could include:

- Withdrawing its authorisation making it unlawful for it to carry on business;
- Issuing fines against the lender or individuals within it; and
- Seeking injunctions from the court in support of enforcement.

178. s.384 FSMA also provides that where if it is demonstrated that the use of the biometric risk score was in breach of the Consumer Duty and the applicant suffered loss by being

[89] https://www.fca.org.uk/publication/finalised-guidance/fg21-1.pdf

refused the mortgage, then the FCA may require the lender to pay to the applicant '*such amount as appears to the FCA to be just*'.

179. The FCA has a clear sector focus and is significantly larger than the ICO as a cross-sector regulator[90], giving it a more realistic task in enforcing its rules. Financial regulation also requires regulated firms – such as the lender in this scenario – to report on their activities directly to the FCA. This gives the FCA an important source of information in assessing compliance and identifying where there is a risk of AI harms materialising. This is in stark contrast to the limitations on the sources of information available to the ICO and EHRC (see sections E and F above).

Super-complaints

180. FSMA provides for designated consumer bodies to bring 'super-complaints' to the FCA (rather than to the FOS) where a feature of a market *"is, or appears to be, significantly damaging the interests of consumers."* Consumer bodies must be designated by the Treasury and the criteria for designation are strict[91]. At the time of writing only two bodies – Which? and Citizens Advice – are so designated. The FCA *must* respond with reasons to a valid super-complaint within 90 days (s.234E).

181. Although super-complaints are relatively rare, there is no reason why the harm caused by the biometric risk score in this Scenario could not be the subject of one, demonstrating an additional route by which the harm might be brought to the FCA's attention (and potentially a route to collective redress for AI harms that does not depend on individuals taking action) and prevented or ameliorated through regulatory action.

182. **In sum** whilst not every firm will comply with FCA rules at all times, the existence of a sector-focused regulator which is clearly alive to the risks presented by automated decision-making means that there is a better level of ex ante regulatory management of the harms in this Scenario than Scenarios 1 and 3.

ii. Ex ante regulation: cross-cutting requirements

UK GDPR: Assessment of High Risk Processing and DPDD

183. An AHRP will be required under Article 35 UK GDPR (as amended by the DPDI Bill) as this would be high-risk processing. However it only requires a relatively high-level summary of the processing and is not required to be reviewed externally, so offers only limited practical protection against this kind of harm.

[90] As a crude comparison the FCA's budget is approximately 10 times the size of the ICOs. See https://ico.org.uk/media/about-the-ico/documents/4021039/ico-annual-report-2021-22.pdf and https://www.fca.org.uk/publications/business-plans/2022-23

[91] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/200454/guidance_for_super_complainants_120313.pdf

184. The lack of an obvious connection between speech patterns and creditworthiness would make it more challenging for the lender to demonstrate compliance with the principle of data protection by design and default, in theory monitored by the ICO. It is arguable that because of the serious impact of a mortgage refusal and the difficulty in understanding the reason for a low score, it cannot be '*necessary'* to use this as part of the mortgage application process.

ICO Enforcement

185. As set out in section I(iii) below, the harms in this Scenario may be unlawful and/or constitute breaches of a number of rights of individuals under the UK GDPR. To the extent the ICO becomes aware of such breaches, or issues with ex ante requirements such as AHRP and DPDD, it would be open to the ICO to take enforcement action to prevent the harms from occurring again in future.

186. However as set out in section E, the scale of the ICO's mandate relative to its resources means that proactive enforcement by the ICO of rights under the UK GDPR provides only limited protection against the risks in this Scenario from materialising.

Enforcement of the Equality Act 2010 by the EHRC

187. Where the biometric scoring algorithm treats individuals less favourably in connection with a protected characteristic (such as disability or race), this may amount to discrimination (see section I(ii) below). A refusal by the lender to make reasonable adjustments would also be a breach of the EA.

188. The EHRC's regulatory work might address the risks in this Scenario in a way that alerts the lender to the potential breaches of the EA, but as discussed in section F, the EHRC's lack of strong enforcement powers, and remit relative to resources, means that this provides very limited ex ante protection in this Scenario.

189. **In sum**, while cross-cutting regulation provides limited ex ante protection, this is one area in which sector-specific regulation – and the associated powers of the regulator – do have a reasonable prospect of preventing the harm in this Scenario from arising.

iii. Ex post redress

How would an individual know about and evidence harm? Transparency aspects of the Consumer Duty and Vulnerable Customers Guidance

190. The Consumer Duty extends to "providing effective communications that customers can understand." Similarly, the Vulnerable Customers Guidance requires that the lender "ensure[s] all communications and information about products and services are understandable for consumers in their target market and customer base."

191. It could be argued that these rules include providing mortgage applicants with detailed information about the biometric risk score and its role in the credit application. However this is not certain. The lender might argue that the FCA rules relate to transparency about the products offered (e.g. interest, repayment terms etc.) rather than transparency about the lender's business processes. It is at least arguable that FCA rules require transparency about the use of the biometric score, but we are not aware of any enforcement action analogous to this scenario.

<u>GDPR Transparency and data access</u>

192. Articles 13 to 15 would apply, which should *in theory* ensure:

a) Individuals are aware that processing is taking place, including at least some information about the logic of that decision-making.

b) Individuals can obtain access to their personal data, which may be an important means of discovering that a biometric credit risk score has been applied, and that it was generated as a result of the video interview.

193. However, as set out in section C(i) above, there are significant limitations in practice to the rights to transparency and access in the UK GDPR. As in Scenario 1, the lender in this Scenario might well use qualifications to data subject rights to limit the data and explanations provided to applicants about the biometric risk score, arguing that to provide full information would endanger its intellectual property, or increase the risk of fraud.

194. Thus even where an applicant discovers that a risk score has been applied to them as part of their application, they would not necessarily be able to work out that it was developed through biometric analysis of their speech patterns or facial expressions, much less determine that the biometric analysis systematically underscores applicants with certain accents or impairments.

195. A further challenge in applying GDPR (or indeed FCA rules) transparency requirements is that the lender itself may not understand how the biometric risk scoring works. The lender may not be in a position to explain to an applicant *how* their speech patterns have produced a low score, even if it wanted to. This is more likely where the lender assigns biometric scores using technology sourced from a third party technology provider.

196. This creates a major barrier to **both** ex ante regulation **and** ex post redress, since it will be unattractive to a regulator or claimant to proceed against the lender without at least a reasonable, evidence-based belief that the biometric risk score is treating people unfairly in the way described in the Scenario. If the lender did not understand its own

scoring approach, this may constitute a breach in itself of the Consumer Duty, though this would not overcome the practical barrier to evidencing the AI harms of the biometric scoring system itself. It might also be argued to be a breach of Article 5(1)(a) of the UK GDPR, but bringing free-standing proceedings about transparency issues alone, before any substantive challenge can be considered, would be very expensive and unattractive for an individual affected by the decision with suspicions about its fairness.

197. As noted above we do not consider that this Scenario involves solely automated decision-making. If it did, limited additional transparency requirements might apply under Article 22C UK GDPR, although we do not consider that these would not materially improve the prospects of an applicant discovering and evidencing the harm caused by the biometric scoring system, as Article 22C does not require explanations of decisions in context or at the point when a decision is made.

198. **In sum,** in the absence of clear sector-specific requirements for algorithmic transparency (as opposed to transparency about the features of financial products), the GDPR at first appears to require information to be provided to applicants about the biometric risk score. In practice however, an ordinary person in this Scenario would face very significant barriers in knowing the role the risk score had played in their application, and understanding the ways in which it might be discriminatory (especially given that the lender itself might not understand how the risk score works).

Sector-specific rights to redress

<u>The Financial Ombudsman Service</u>

199. Where the Customer Duty or Vulnerable Customers Guidance have been breached, a mortgage applicant in this Scenario may complain to the FOS[92], which will determine an outcome that it considers to be *"fair and reasonable in all the circumstances of the case"* (s.228). This may include financial compensation, including for non-monetary loss, and/or a direction to take "such steps in relation to the [mortgage applicant] as the ombudsman considers just and appropriate*"*. In this Scenario that could for example include reconsidering the mortgage application on a fair basis.

200. The applicant would need to complain to the lender first, and normally would need to complain to the FOS within six months of receiving a final response from the lender.

201. The FOS is free-to-use and complainants do not require legal representation. There is no risk of being made to pay the lender's costs if unsuccessful. An initial response from a case handler is 'typically' provided within 90 days[93]. Whilst the time to resolve a

[92] s.226 FSMA and para 2.3 of the FCA Handbook on Dispute Resolution

[93] https://www.financial-ombudsman.org.uk/consumers/expect/how-long-it-takes#:~:text=Typically%2C%20this%20part%20of%20our,case%20handler%20as%20things%20progress.

complaint where a formal ruling from an ombudsman is required is less certain, complainants are in no worse a position on timing than if they had to bring a court claim.

202. **In sum** there is an effective and accessible practical mechanism for applicants who find out about the harm in this Scenario to vindicate their rights of redress arising from sector-specific regulation.

Cross-cutting rights to redress: discrimination

Indirect discrimination

203. As set out at section D, The Equality Act 2010 ('**EA**') protects against discrimination on the basis of a number of 'protected characteristics', of which 'disability' and 'race' are relevant to this Scenario, which indicates that some low biometric risk scores arise from speech impairments, neurodivergence and particular accents where they are associated with race[94].

204. Applying lower biometric risk scores in this way would likely constitute disability discrimination under the EA. s.19 EA provides that application of "*a provision, criterion or practice*" which puts potential customers who are disabled or from particular racial backgrounds "*at a particular disadvantage*" compared to those without the protected characteristics is indirectly discriminatory.

205. The lender's *practice* of setting risk scores using the video interview seems likely to meet this description[95] and therefore be unlawful when read with s.29 EA, as the lender is offering a service to the public.

206. The lender would not be liable if it can show that the use of the biometric risk score is a '*proportionate means of achieving a legitimate aim'*. We have assumed that the risk score is *not* a reliable indicator of creditworthiness, meaning it cannot possibly meet this test, since it does not help the lender achieve anything useful. However, *even if the biometric risk score was accurate on average*, it might still fail this test. Whilst the lender has a legitimate aim in lending credit to the appropriately credit-worthy people and at the right rates, its scoring system is applied in a blanket way without any safeguards for those with protected characteristics, such as to have a face-to-face interview. Further, the consequences of credit refusal are serious, meaning the criteria put in place by the lender needs a strong justification. Without any further safeguards, we consider that it would be difficult for the application of the biometric score to be a proportionate means

[94] Note that if the tool systematically disfavours individuals with accents which are *not* related to race – such as UK regional accents or class-based accents – then indirect discrimination could not arise, since neither regional origin within the UK nor socio-economic status are protected characteristics.

[95] Note that this test is looser than the test for whether there is processing 'concerning' health data under Article 9 UK GDPR.

of achieving the lender's legitimate aim, but it would be fact-specific and is difficult to predict with certainty.

207. A successful claim for indirect discrimination in the provision of a service under the EA entitles the claimant to compensation. This includes consequential losses (which might include losses due to a house purchase falling through), injury to feelings, and can include aggravated damages. Any financial loss in this Scenario would be fact-specific. Damages for injury to feelings would likely be in the low £,000s at most due to its one-off nature and taking place in private[96].

Right to reasonable adjustments

208. §20 and 27 EA impose on the lender a duty to make '*reasonable adjustments*' to avoid those with the protected characteristic of disability from being put "*at a substantial disadvantage*". This duty would arise in this Scenario – a reasonable adjustment might for example be for a human to conduct the interview. This gives disabled applicants a legal means of avoiding the harm caused by the biometric risk score, assuming they are aware of it in advance.

Practical considerations for bringing claims for discrimination under the EA

209. Claims for indirect discrimination in this Scenario must be brought in the civil courts (normally the county court) (s.114(1)(a) EA). This is far from straightforward, as set out in section G above.

210. **In sum** for those unfairly treated by the biometric risk score *with a protected characteristic*, whilst there is a clear right of action under the EA, enforcing that right in the county court would be uncertain, risky, costly, and daunting, leaving a gap in effective protection.

211. Moreover many applicants may be treated unfairly *without* having a recognised protected characteristic (e.g. if low scores result from regional White British accents), and they would be unprotected by the EA.

Cross-cutting rights to redress: the UK GDPR

Lawful basis for the processing

212. The Scenario does not state the legal basis on which the lender relies for the processing to create and use the biometric risk scores. Conceivably, the lender might seek to rely on:

a) Article 6(1)(a): consent from the applicant;

[96] See guidance from the EHRC: https://www.equalityhumanrights.com/sites/default/files/quantification-of-claims-guidance.pdf

b) Article 6(1)(b): the processing is necessary for the entry into a contract (i.e. the mortgage); or

c) Article 6(1)(e): the processing is necessary for the lender's legitimate interests in assessing creditworthiness.

213. The fact that the biometric risk score is not reliable makes it very difficult for Articles 6(1)(b) or (e) to apply. For processing to be 'necessary' it must be rationally connected to either the entry into/performance of the contract, or to the lender's business interests. Since the score does not reflect creditworthiness, there is no rational connection, and the processing cannot be 'necessary' for either the contract or the lender's interests[97].

214. This would force the lender to rely on obtaining applicants' consent to the preparation and use of the biometric risk score. Valid consent is a high standard under the GDPR. It must be freely given, specific and informed. For consent to the biometric risk score to be informed, detailed information about it would be required, perhaps going beyond the transparency requirements of Articles 13-14. This might be very challenging for the lender (especially where it does not understand the biometric risk score itself).

215. The GDPR states at Article 7(4) that:

> *"When assessing whether consent is freely given, utmost account shall be taken of whether, inter alia, the performance of a contract, including the provision of a service, is conditional on consent to the processing of personal data that is not necessary."*

216. So any consent would have to be truly optional. That is, applicants should be able to withhold consent to the biometric scoring but still be fairly considered for the mortgage, giving them a route to avoid the harm.

217. Applicants in this Scenario therefore have a number of ways to challenge the lawfulness of the processing involved in the biometric risk scoring by arguing that it lacks a valid legal basis under Article 6 UK GDPR. Although there is no English case law directly analogous to this situation, we consider that such challenges would have a reasonable prospect of success.

[97] If the biometric risk score *were* accurate, the lender might have a better chance of relying on these two bases. However, there would still be some arguments that the use of the biometric risk score is not 'necessary' bur merely 'useful' or 'convenient', and that applicants' interests in avoiding discriminatory outcomes outweigh those of the lender. So there would still be scope to challenge the lawful basis for the processing.

Processing of special category data

218. For some applicants, the generation of the risk score may involve processing of SCD (see para 166 above). Where this is the case, the processing would require explicit consent from mortgage applicants (Article 9(2)(a)) since such processing is prohibited unless an exemption applies, and none of the other exemptions in Article 9 could apply in this Scenario.

219. Thus if applicants can surmount the hurdle of demonstrating that SCD is being processed, this would make the generation of the risk score unlawful (or provide a route to avoid it by withholding consent).

Accuracy of processing

220. As stated above, it appears that the biometric risk score is not a reliable indicator of creditworthiness. Given how the score is used (i.e. it is interpreted as an indicator of creditworthiness by the lender), the processing would likely be inaccurate in breach of Article 6(1)(d). Accuracy is a relative standard, depending on the context and consequences of processing. The consequences in this Scenario are serious and so such an irrational method of generating risk scores would not meet the level of accuracy required by the UK GDPR and would be unlawful.

Fairness of processing

221. The lack of an apparent rational connection between one's mode of speech and creditworthiness could be enough to render the processing unfair, especially if little information is provided to clarify that this is what the biometric risk score does. It is well arguable that it would be completely outside applicants' expectations that they would be scored in this way, making the processing unfair and unlawful under Article 5(1)(a).

Causation, loss and compensation for GDPR breaches

222. We identify a number of ways in which mortgage applicants might be able to show that the profiling of mortgage applicants is unlawful, including that it:

a) Lacks a valid legal basis;

b) Violates the prohibition in Article 9; and/or

c) Is inaccurate and/or unfair.

223. Whilst there may be strong arguments that the processing is unlawful, applicants would only be entitled to financial compensation[98] where it can be shown that the breach of the UK GDPR *caused* loss (which may include distress). So an applicant would need to

[98] Claimants would also be entitled to injunctive relief – i.e. a court order to stop any unlawful processing – and in some cases declaratory relief, i.e. court confirmation that the processing is unlawful.

show that they would have been granted the mortgage *but for* the irrationally-low risk score, and that the inability to get the mortgage resulted in loss. In this Scenario, a mortgage applicant might be able to show some distress caused by the GDPR breaches, but the level of compensation would be low (i.e. in the low hundreds of pounds[99]).

224. **In sum**, the GDPR on its face appears to provide protection in a number of ways against the harms in this Scenario.

225. But even if a mortgage applicant can overcome challenges in (i) finding out about the risk score and its role in the mortgage decision (ii) understanding how it might be discriminatory, and (iii) obtaining enough evidence of this, there remain significant obstacles to enforcing their GDPR rights, leaving a major gap in effective protection.

iv. Conclusion: level of effective protection from AI Harm(s) in this Scenario

226. Current law provides the best (though far from complete) effective protection from the AI harms in this Scenario. Protection from cross-cutting regulation is mixed: there are strong protections where unfair treatment by the biometric score is related to particular characteristics like race, health and disability. Whilst this is welcome, other arbitrary and unfair discrimination – such as on the basis of regional or class-linked accents – is much less well-protected.

227. The key layer of protection however is sector-specific. It provides for sector-specific rules (the FCA Rules), a sector-specific regulator to monitor compliance with those rules (the FCA) and a sector-specific mechanism for redress which is accessible. The stumbling block that prevents effective protection from being complete is transparency. This Scenario is a good illustration of how complex and opaque algorithmic logic challenges GDPR (and other) rules regarding transparency. It would likely be rather difficult for the lender itself or even the regulator to identify systematic patterns of unfairness in the biometric scoring algorithm, let alone for an individual to do this as part of seeking redress for an unfair lending decision.

[99] We assume that no recognised psychiatric injury such as clinical stress or depression has been caused by the unlawful processing.

Are there legal requirements that the decision-maker must consider in advance?	**Is it likely that a regulator would prevent the AI harm through enforcement of those requirements?**	**Would the individual be able to find out about and evidence the harm?**	**Is there a legal right to redress for the harm?**	**Is it practical for individuals to enforce any legal rights to redress?**
Scenario 2 (Mortgage Assessment)				
Medium: both Cross-cutting (GDPR and Equality Act) and Sector-Specific FCA Rules are relevant to the tool, suggesting it may not be permissible to implement it in the way described.	Medium: reason to believe FCA is a more effective *ex ante* regulator, as it is focused on one sector and has strong enforcement powers Super-complaints may also bring issues to the FCA's attention	Poor: it would be especially difficult for an individual to identify the harm in this scenario given the opacity of the algorithmic logic, even taking into account GDPR transparency rights.	Good: as well as GDPR and Equality Act causes of action, able to seek redress under FCA rules.	Practical: Financial Services Ombudsman provides free-of-charge resolution with need for legal representation

J. Scenario 3: Universal credit

> *The Department for Work and Pensions (DWP) has introduced a new AI chatbot service to provide advice to recipients and potential recipients of universal credit (UC), replacing phone-based customer service as the main point of benefits enquiries. The AI chatbot is deployed by the DWP itself, but it is based on a large language model developed by a US-based technology firm, which has been licensed by an intermediary firm based in the UK and fine-tuned by this firm using data from DWP customer service logs, augmented with structured information about DWP eligibility rules appended to the LLM's prompt Individuals can use the chatbot to query the status of their UC payments, and their assessed eligibility for payments.*
>
> *The information provided by the chatbot is not always accurate, and it can provide incorrect advice that sometimes leads to individuals failing to claim benefits to which they would be entitled. The chatbot is also agentic: it can carry out certain actions such as amending data held about an individual on the basis of information shared with it in conversation, which can lead to an individual's assessed eligibility for payments changing.*

228. The scenario contemplates two principal potential harms which regulation might be expected to guard against:

a) Individuals may be provided with incorrect information, causing them to fail to claim or underclaim UC to which they are entitled.

b) Individuals' records may be automatically updated by the chatbot, with the potential for updates to be inaccurate, negatively affecting benefits entitlement.

229. As with Scenarios 1 and 2, the cross-cutting requirements of the UK GDPR apply to this Scenario because both the provision of personalised advice and the amendment of individuals' DWP records involve the processing of personal data.

230. There is no 'regulator' as such for this Scenario, since the data controller/decision-maker is a public body. However there are public-sector specific policies and legal principles which may offer ex ante protection against harms or ex post scope for individuals to seek redress.

i. Ex ante regulation: sector-specific

231. A wide range of **non-binding** guidelines apply to the procurement and use of AI tools by public bodies, as well as to the use of data in the implementation of new services[100]:

Guidelines for AI procurement[101]

232. These cross-departmental guidelines encourage those procuring new AI tools in government to take a range of sensible steps to ensure the tools are fit for purpose. These include having an appropriately-skilled team, considering questions of bias, and carrying out an 'AI Impact Assessment' including assessing risks of inaccuracy in results.

Guide to using AI in the public sector[102]

233. This collection of guidance, published in 2019, includes content on "understanding AI ethics and safety", which advises readers to (among other things) "guarantee as much as possible the safety, accuracy, reliability, security, and robustness of its product".

Digital Service Standard[103]

234. This policy encourages public bodies providing digital services to (notably) "make sure the service helps the user to do the thing they need to do as simply as possible - so that people succeed first time" (Principle 4).

Data Ethics Framework[104]

235. This guidance aims to "help public servants understand ethical considerations, [and] address these within their projects", including "*Understanding [the] unintended consequences of your project"* and *"ensuring that the data for the project is accurate [and] representative".*

Government Transparency Standard[105]

236. Through this initiative, the Government has both articulated a level of transparency to which public bodies using algorithmic decision-making should aspire, and sought to centralise reports from public bodies in compliance with the standard. In theory, the DWP's chatbot would be an example of an algorithmic tool which should be self-reported according to the standard, with the (some tools have begun to be reported against the standard, but this appears to be limited to date).

Guidance to civil servants on use of generative AI[106]

[100] A Crown Commercial Services AI framework appears to be on hold: https://www.crowncommercial.gov.uk/agreements/RM6201

[101] https://www.gov.uk/government/publications/guidelines-for-ai-procurement/guidelines-for-ai-procurement

[102] https://www.gov.uk/government/collections/a-guide-to-using-artificial-intelligence-in-the-public-sector

[103] https://www.gov.uk/service-manual/service-standard

[104] https://www.gov.uk/government/publications/data-ethics-framework

[105] https://www.gov.uk/government/collections/algorithmic-transparency-recording-standard-hub

[106] https://www.gov.uk/government/publications/guidance-to-civil-servants-on-use-of-generative-ai/guidance-to-civil-servants-on-use-of-generative-ai

237. This primarily guides civil servants in the use of generative AI tools to do their work – as opposed to making generative AI tools available to the public – but could have some relevance to the Scenario.

The Digital, Data and Technology Profession Playbook[107]

238. This guidance for civil servants working on digital and data issues is extensive and includes some content regarding the testing of products which could in theory help to identify the errors in the Chatbot.

239. **Taken together** this guidance, whilst somewhat generic, constitutes sensible advice which, if followed, should alert public bodies to the risk of the harms in this Scenario. However this guidance is entirely voluntary, and we are not aware of any one body within Government mandated (much less formally empowered) to ensure compliance with it.

240. A full analysis of the effectiveness of this system of voluntary guidance across Government is beyond the scope of this analysis, but a range of examples of public sector data-driven projects that have received significant public backlash[108] indicate that it is not infrequently ignored, even in high-risk contexts, therefore providing very limited ex ante prevention of the kinds of harms in this Scenario.

241. Note that the Public Sector Equality Duty under the EA would apply to the introduction of the Chatbot (which the DWP might discharge by carrying out an Equality Impact Assessment). However the Scenario does not indicate that the Chatbot disfavours individuals with protected characteristics, thus any impact assessment would not highlight issues in complying with the Equality Duty and it would therefore not prevent the Chatbot from being introduced.

ii. Ex ante regulation: cross-cutting

The UK GDPR

Assessment of High Risk Processing and DPDD

242. An AHRP will be required under Article 35 UK GDPR (as amended by the DPDI Bill), since the use of a chatbot in this context would almost certainly be high-risk processing. However, Article 35 only requires a relatively high-level summary of the processing (in contrast to pre-DPDIB requirements which were more prescriptive). It should in theory address how the DWP complies with the principle of DPDD in the use of the Chatbot.

[107] https://www.gov.uk/government/publications/the-digital-data-and-technology-playbook/the-digital-data-and-technology-playbook

[108] See e.g. https://en.wikipedia.org/wiki/2020_United_Kingdom_school_exam_grading_controversy and https://www.theguardian.com/uk-news/2020/aug/04/home-office-to-scrap-racist-algorithm-for-uk-visa-applicants among others

Enforcement by the ICO

243. The ICO is empowered to enforce the principle of DPDD, as well as other breaches of the GDPR in the use of the chatbot, such as the potential individual rights breaches set out at section below. However, as set out in section E, the scale of the ICO's remit relatively to its resources and – arguably – its approach to enforcement, means that proactive enforcement by the regulator of GDPR requirements offers limited practical protection against the kinds of risks in this Scenario. Of particular relevance to this Scenario is the fact that it is very unlikely that the ICO would ever do more than issue a 'reprimand' to DWP, which arguably reduces incentives for the DWP to comply with *ex ante* requirements.

iii. Ex post redress

How would an individual know about and evidence harm?

244. As set out above, were the DWP to follow all voluntary government guidance on the use of AI tools to the letter, comprehensive information about the chatbot would at least have been made available online. Comprehensible explanations of its logic and limitations would also have been made available to UC claimants at the point of use, in line with digital service design standards. However these are not binding transparency requirements, so we turn to those which derive from cross-cutting regulation.

Retention of records

245. Since the processing involved in the operation of the Chatbot is high-risk, Article 30A UK GDPR (inserted by the DPDI Bill) requires DWP to retain '*appropriate records*' of its processing of personal data. This obligation is limited to recording what processing took place. Thus the records of advice provided or any changes to a person's file would need to be retained. But the duty to keep records would not extend to DWP being able to re-run any automated analysis or decision-making (for example if advice had been generated using an earlier model of the Chatbot), which could present a barrier to effective transparency.

GDPR transparency and data access

246. Articles 13 to 15 would apply, which should *in theory* ensure:

a) Individuals are aware that processing – including automated decision-making – is taking place when the Chatbot gives advice and updates records, including information about the logic of that decision-making (for updates to records, which constitute profiling).

b) Individuals can obtain access to their personal data, which may be an important means of discovering and evidencing a mistake in advice or inaccurate updating of records by the Chatbot.

247. However as set out at section C(i) above there are significant practical limitations to transparency and access under the UK GDPR. Even were an ordinary person to take the time to read the DWP's information charter and privacy notice, it is unlikely to be obvious to them how the Chatbot works, or that there are risks of it making mistakes. Indeed paradoxically as the Chatbot introduces mistakes into the system in a systematic way, it may give the impression of objectively rational and unchallengeable information by reason of being automated.

248. Further, as with the controllers in the other scenarios, the DWP might well resist or limit any requests for copies of data on the basis that to release full information about the chatbot could increase the potential for fraud.

Transparency for automated decision-making

249. Where the Chatbot updates an individual's record in a way that affects benefits entitlement, this would likely amount to an automated decision covered by Articles 22A-C UK GDPR, since it has 'legal or similarly significant effect' and the Scenario indicates it is solely automated (i.e. lacking any '*meaningful human involvement*' Article 22A(1)).

250. Assuming the decision is not *prohibited* (see section J(iii) below), Article 22C requires it to be accompanied by safeguards, including a requirement to *"provide the data subject with information about decisions* [which are solely automated and significant]*".* As this requirement is new, there is no case law or guidance on what information is required. It might be assumed that the obligation goes beyond the requirements of Articles 13 and 14 (otherwise it would be redundant) but this is not assured.

251. It is also notable that the requirement in Article 22C is a general one – to provide information about '*decisions*' rather than each decision. It may well be that this can be discharged through a general, up-front privacy notice provided by DWP to UC claimants, rather than by way of specific notification of *each* significant solely automated decision made about a UC claimant *at the time it is made* – that is, at the time that the chatbot updates the claimant's record.

252. **In sum** the UK GDPR would require DWP to make information available to UC claimants about the processing involved in the Chatbot's operation, including automated decisions such as updating records. It's likely that many UC claimants would not be *actually* aware of the processing involved in the Chatbot, and even those that are would not necessarily have an explanation enabling them to easily identify the potential for errors in its advice.

253. More realistically, where it becomes apparent to a UC claimant or their advisors that something has gone awry with their entitlement, UK GDPR transparency requirements

will at best create enough of a trail for them to follow, by which they stand some chance of identifying the Chatbot's errors.

Sector-specific rights to redress

No legal right to compensation

254. The provision of inaccurate information – or the inaccurate updating of claimant information – by the chatbot would likely not give rise to a legal cause of action against DWP:

a) Judicial review would not be available as a remedy for incorrect advice, since the provision of advice is not a 'decision', and judicial review is only available where there is a decision which can be challenged (and, if it was unlawful, required to be remade)[109].

b) The tort of misfeasance in public office requires an element of 'malice' or bad faith[110], which is not present in this Scenario.

c) Traditionally the courts have been extremely reluctant to impost a duty of care in the tort of negligence on public bodies.[111] Whilst some recent decisions[112] suggest that it *may* now be possible to formulate a duty of care in negligence in certain circumstances, this has not been conclusively established in any circumstances analogous to this Scenario. Therefore bringing a claim in negligence against the DWP in this Scenario would not be a realistic *ex post* means of redress for an ordinary person as it would be far too uncertain and costly.

The DWP's Customer Charter and Special Payments Scheme

255. The DWP undertakes in its Customer Charter[113] to *"provide you with the correct decision, information or payment."*

256. DWP operates its own complaints procedure[114] through which an individual could complain about being provided with incorrect advice by the Chatbot, or about their

[109] Where the Chatbot updates an individual's records, this could amount to a decision which could be subject to judicial review. If the update was clearly erroneous, any challenge would likely be successful, since it would be irrational to update someone's record with incorrect information. However, given the availability of other more realistic routes of redress, we do not discuss judicial review in detail, since it is expensive and involves significant legal barriers (including being barred where the individual has an adequate alternative remedy, as they do here through the DWP's maladministration scheme), which would make it a disproportionate means of simply having an inaccurate record corrected.

[110] See e.g. *Young v Chief Constable of Warwickshire* (2020) EWHC 308 (QB)

[111] *Home Office v Mohammed* [2011] 1 WLR 2862; *Gorringe v Calderdale MBC* [2004] UKHL 15 applying the test in *Caparo v Dickman plc* [1990] 2 AC

[112] *Poole Borough Council v GN* [2019] 2 WLR 1478

[113] https://www.gov.uk/government/publications/our-customer-charter/our-customer-charter

[114] https://www.gov.uk/government/organisations/department-for-work-pensions/about/complaints-procedure#before-you-make-a-complaint

records being updated inaccurately, causing them to underclaim universal credit (‘**UC**’), each of which would constitute ‘maladministration’[115].

257. Assuming an individual realises they received incorrect advice or had their records changed inaccurately, demonstrating maladministration through a complaint should be relatively straightforward[116]. The DWP has published guidance[117] on payments in cases of maladministration which provides:

> *“Individuals should not be disadvantaged as a result of maladministration:*
>
> *it is not necessary for an individual to request consideration of a special payment. The appropriateness of making a payment should be routinely considered in any attempt to rectify departmental maladministration, which may have resulted in a customer (or a third party) experiencing injustice and/or hardship*
>
> *the purpose of the special payment scheme is, wherever possible, to return the individual to the position they would have been in but for the maladministration. If this cannot be achieved the aim is to provide redress that is reasonable and proportionate in light of the individual circumstances of the case.” (at 4.2, emphasis added)*

258. That is, where the poor advice or inaccurate record led to the individual underclaiming UC, backdated payments should be available as an outcome of the complaint. This would presumably be backdated to the point at which the claimant first sought advice or can otherwise demonstrate they would have applied for UC had they had the correct advice or records but for the Chatbot.

259. The guidance also provides that compensation for consequential financial losses may be available (though this will not generally include professional fees), as well as minimal payments for distress suffered (between £50 and £500).

[115] Note that this is a *separate process* to challenging a substantive benefits decision, which is done through the ‘mandatory reconsideration’ process with a right of appeal (within a time limit) to the Social Security and Child Support Tribunal (s.30 and Sch 6 of the Tribunals Courts and Enforcement Act 2007). It is of course possible that an individual provided whose record was incorrectly updated might appeal the substantive decision, but we focus here on obtaining redress for the erroneous advice or inaccurate change of records, rather than resolving the substantive forward-looking entitlement to UC which would presumably be straightforward once an error was identified. See https://www.gov.uk/mandatory-reconsideration

[116] This will be made more easy where the UC claimant has successfully exercised their right of access under Article 15.

[117] https://www.gov.uk/government/publications/compensation-for-poor-service-a-guide-for-dwp-staff/financial-redress-for-maladministration-staff-guide#introduction

260. If the DWP does not follow this policy, the individual may appeal to the Independent Case Examiner[118], and failing that, to the Parliamentary and Health Service Ombudsman[119].

Practical considerations for ex post redress from the DWP

261. Complaining to the DWP (and if necessary, appealing to the ICE and/or PHSO) is free of charge, does not carry adverse costs risks, and does not require professional advice or representation (assuming that the individual has already established that the Chatbot made an error in their case, which may not be straightforward).

262. **In sum**, although there is no *legal* right to redress for errors made by the Chatbot in this Scenario, in practice redress can be sought on a sector-specific basis for the DWP's maladministration, and this is supported by multiple layers of relatively accessible appeal. Assuming an individual somehow becomes aware of the Chatbot's error – and the conditionality of this cannot be overstated – there are reasonably good prospects of rectifying the harm caused.

Cross-cutting rights to redress[120]

Accuracy of processing

263. When the Chatbot provides personalised advice, it does so by processing the enquiring UC claimant's personal data. The presentation of that advice to the UC claimant is also an act of processing of their personal data. More obviously, the amendment of a UC claimant's record is also an act of processing of their personal data. All of these acts of processing are required to be sufficiently accurate, given the circumstances, in accordance with Article 5(1)(d) UK GDPR.

264. The consequences of inaccurate processing in this Scenario potentially grave. Vulnerable claimants may end up underclaiming or being denied UC to which they are entitled, affecting their health and opportunities significantly. In severe cases, perhaps endangering their life. In this sense, for at least some individuals the processing results in consequences that have legal or similarly significant effect.

265. The standard of accuracy therefore required by the UK GDPR is high. Any erroneous advice causing a UC claimant to underclaim, and any erroneous update causing a

[118] Which operates as an independent officer under the remit of the Parliamentary Ombudsman, established by the Parliamentary Commissioner Act 1967. See https://www.gov.uk/government/publications/how-to-take-a-complaint-to-the-independent-case-examiner/how-to-bring-a-complaint-to-the-independent-case-examiner

[119] Established by the 1967 Act https://www.ombudsman.org.uk/

[120] We do not consider the fundamental lawfulness of the processing since it has an obvious lawful basis – including where SCD is processed – and in any event challenging the processing itself as unlawful would not provide redress since it is the *erroneous* advice or updating of records that has caused detriment to the UC claimant.

claimant to lose entitlements would certainly be inaccurate and unlawful processing under the UK GDPR.

266. Inaccurate processing would give the UC claimant a cause of action, including the right to compensation for damage caused by the breach, which in this Scenario would include any UC underclaimed, consequential losses and potentially a small amount of compensation for distress (i.e. similar to the compensation likely available from the DWP's maladministration payments scheme).

Right to rectification

267. Where the Chatbot has inaccurately updated records, the individual would have a right to obtain rectification of the data under Article 16 of the UK GDPR.

Article 22 automated decision-making

268. Where the Chatbot updates an individual's record in a way that affects benefits entitlement, this would likely amount to an automated decision covered by Articles 22A-C UK GDPR, since it has 'legal or similarly significant effect' and is solely automated (Article 22A(1)).

269. In many cases, this would be prohibited, since:

a) The decision would be based '*entirely or partly*' on SCD (most obviously health data, which will be relevant to many aspects of UC entitlement, Article 22B(1)); and

b) We are not aware of any specific enabling legislation of which we are which requires or authorises solely automated significant decisions by the DWP for this purpose (Article 22B(3)).

270. Where the decision is unlawful on the basis of Article 22B, this breach would give rise to a cause of action for any damage suffered[121]. This would include underclaimed UC where it can be shown that if a human had been involved, a different decision would have been reached. That is, a human would not have made the erroneous update to the UC claimant's record, which should be relatively easy to demonstrate if the Chatbot has made a clear error.

271. For some individuals, no SCD will be involved, meaning that the solely automated significant decision is lawful, but must be accompanied by safeguards, such as informing individuals of how the Chatbot may update their records (presumably going beyond the requirements of Articles 13 and 14) and giving them a right to request human

[121] Article 82 and s.168DPA entitle a data subject to "compensation from the controller or processor for the damage suffered [due to an infringement of the UK GDPR]" which includes distress.

intervention. This provides no cause of action in relation to any loss suffered due to the inaccurate updating of records, but may help individuals (if they are well-advised) to *identify* the error or rectify it before it has a significant impact on them.

272. **In sum**, the GDPR in theory provides individuals with causes of action to have losses caused by the Chatbot's errors, assuming they are identified by the individual and can be evidenced. Rules on automated decision-making will prohibit these kinds of decisions in some cases, but in many others will only make provision for transparency around the decision and a right to human review.

273. In practice, the enforcement of such causes of action will be very challenging for most individuals for the reasons set out in G and, since the causes of action do not provide for compensation beyond what is available through the DWP's maladministration special payments scheme, it is unlikely that a UC claimant affected by this scenario would seek to enforce their GDPR rights in this way.

iv. Conclusion: level of effective protection from AI Harm(s) in this Scenario

274. This Scenario further demonstrates the limitations of cross-cutting regulation in that the Equality Act 2010 has no relevance since protected characteristics are not directly engaged, and the GDPR, whilst it provides important legal rights to redress, would only be enforceable through the civil courts, which is impractical for many. Perhaps unusually, important protection is provided through the existence of an entirely voluntary (albeit well-established and long-lasting) DWP maladministration scheme, which provides a practical route to redress for incorrect advice or records changes.

275. Again, transparency presents an important stumbling block to effective protection. Neither the UK GDPR nor voluntary public sector guidance ensure that individuals affected would come to find out that they'd been given incorrect advice. Indeed the use of an AI tool may make ordinary people even *less* likely to question the advice they receive on eligibility.

Are there legal requirements that the decision-maker must consider in advance?	Is it likely that a regulator would prevent the AI harm through enforcement of those requirements?	Would the individual be able to find out about and evidence the harm?	Is there a legal right to redress for the harm?	Is it practical for individuals to enforce any legal rights to redress?
Scenario 3 (DWP Chatbot)				
Low: the UK GDPR likely does not rule out the use of the tool. Further, any additional guidance for public bodies on the use of AI is non-binding and compliance with the guidance is not monitored.	Very unlikely: relies solely on enforcement by the ICO, which takes a light-touch approach to regulating public bodies, which arguably reduces incentives for compliance.	Poor: relies on non-binding guidance on the part of the DWP and GDPR transparency, which does not require explanations of automated decisions *in situ*.	Medium: beyond GDPR rights, voluntary DWP maladministration scheme and rights to appeal benefit decisions But the DWP scheme may not fully compensate consequential losses.	Practical: appeal from DWP scheme plus option to appeal to Parliamentary Ombudsman.

K. Conclusion: gaps in effective protection from AI harms

276. Our analysis of the Scenarios demonstrates that there are gaps in the effective protection from AI harms in the current regulatory regime.

277. The potential for algorithmic tools to be applied in almost any domain means that AI harms engage a complex patchwork of regulation[122], both cross-cutting and sector specific. In that sense, it is true that much AI regulation already exists. The requirements of the UK GDPR and EA have broad application across the Scenarios (albeit that the EA is only engaged where the person harmed has a protected characteristic). Indeed, much of the processing involved in the decisions described in the scenarios would be unlawful under the UK GDPR, raising the prospect of enforcement by the ICO or private causes of action by individuals harmed. And there is sector-specific regulation applicable to Scenarios 1 and 2.

278. A focus on the letter of the law would be misguided, however. Effective protection is only provided where (i) regulators are empowered, resourced, and motivated to enforce the law proactively, and (ii) where it is also realistic and practical for individuals to enforce their rights to redress through the court system. It is through this lens that the gaps in effective protection can be most clearly seen, as indicated on the Summary Table at section A above. Principally these gaps are:

a) It is not realistic to expect the ICO and EHRC as cross-cutting regulators to enforce the UK GDPR and EA with a completeness that will reliably protect against AI harms. They do not have sufficient powers, resources, or sources of information, and cooperation between regulators is not assured[123]. And they have not always made full use of the powers they have. There will be greater protection against AI harms occurring in sectors with specific regulators like in Scenario 2, but this may be the exception rather than the rule.

b) Legal rights to redress are only meaningful if they can be enforced in accessible forums. Where specific redress forums such as ombudsmen exist – as in Scenarios 2 and 3 – individuals have a better chance of enforcing their rights at a realistic cost and risk level. Scenario 1 shows that where individuals harmed by

[122] This patchwork is itself a barrier to effective protection, since it is challenging to navigate even for professionals, let alone the ordinary individuals subject to AI harms.

[123] It is notable that the EHRC is not a member of the Digital Regulation Cooperation Forum, a principal mechanism for AI regulatory cooperation cited in the Government's AI White Paper.

algorithmic tools are left to fall back on the UK GDPR and the civil courts, this creates serious barriers to effective protection.

c) Crucially, to enforce rights to redress, individuals must know about and be able to evidence AI harms. Regulation fails to ensure this in any of the scenarios, with no right to an explanation of complex and opaque algorithmic decisions. This leaves real questions about whether an individual would in practice get redress, even in Scenario 2, where the regulatory and redress environment is otherwise strongest.

279. A further gap in effective protection is only implicit in the scenarios, in that it is entirely absent from the UK's regulatory regime. Our analysis focuses on the obligations or rights against the decision-makers *using* the algorithmic tools. There are no regulatory requirements that bind the *developers* or *sellers* of those tools that would require them to – for example – consider the risks of their tools making inaccurate or biased decisions. This is in notable contrast to some other areas of UK regulation, such as that covering medical devices or civil aviation.