Public Law Project

Public Law Project response to the AI White Paper consultation

JUNE 2023

Summary

I. PLP welcomes the introduction of AI regulation through this White Paper, we agree that regulation must be introduced to allow for the full benefit of AI to be realised, whilst protecting individuals and communities from AI-related harms. Our work in this area focusses on ensuring that public authorities operate AI and automated decision-making (ADM) systems transparently, reliably, lawfully, fairly, in a non-discriminatory way, and with adequate authorisation in law for their use. However, we are concerned about the Government's lack of consideration within this White Paper of the specific risks and regulatory requirements posed by the use of AI and automated decision-making (ADM) by public authorities.

II. PLP's research into automation and digitalisation has shown that public authorities are increasingly using AI and ADM in a wide range of areas (such as immigration, welfare benefits, policing and prisons) that have the potential to impact significantly on the lives of our beneficiaries. Underregulating this area of AI usage and will impede the state's ability to effectively prevent, or even mitigate, the very risks it identifies.

III. The 'pro-innovation' focus of the proposed AI regulation has led to an unnecessarily 'light-touch' approach to the regulation of public authority use of AI and ADM. In addition, the iterative approach set out reveals that the Government are willing to accept that AI may "cause and amplify discrimination that results in, for example, unfairness in the justice system [or] risks to our privacy and human dignity, potentially harming our fundamental liberties." [1] PLP urges the Government to take a more hands-on approach now, and introduce regulation on a statutory footing to ensure it has the necessary bite to effectively prevent AI-related harms.

IV. Opacity is an inherent challenge in understanding and assessing the operation of AI by public authorities. Disclosure must be improved, but simply putting information on the use of AI into the public domain will not achieve meaningful transparency alone. Information must be provided to the public in ways that are concise, intelligible, and easily accessible.

V. To improve transparency, the Government must introduce mandatory disclosure requirements for public authorities who use AI to make decisions partially or entirely. PLP recognises the potential value of the Algorithmic Transparency Recording Standard (ATRS) in improving levels of transparency, but this potential cannot be realised unless the submission of reports becomes compulsory. AI transparency could also be improved through the introduction of the requirement to notify individuals who are subject to decisions made or supported by AI, alongside the publication of the rules and reason for processing.

VI. Currently, the ability to seek redress for AI-related harms is limited by the lack of available information on where such systems are being used and how they operate. Where individuals and communities are not aware of the existence of AI in decision-making processes, it is not possible for them to even consider contesting decisions or processes on this basis. Our research shows that a comprehensive legal and regulatory framework is still a work in progress. Although not created with AI in mind, some existing legal frameworks contain a number of crucial, albeit imperfect,

[1] Department for Science, Innovation and Technology (DSIT) and Office for Artificial Intelligence, *A pro-innovation approach to AI regulation* (29 March 2023) https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper,, paragraph 25.

safeguards which can be interpreted to regulate its use, but remain fragmented.

VII. To improve routes to contest or seek redress, the focus should first be on fortifying existing safeguards and ensuring clarity and coherence between existing laws. PLP is concerned about the lack of bite judicial review is having in practice in this context, and the direction of legislative reforms and the effect this will have on individuals and communities subject to AI-related harms. One way such routes could be improved is through a specialist regulator and forum for complaints relating to public authority use of AI.

VIII. PLP agrees that the revised principles proposed in the White Paper are steering AI regulation in the right direction, but that the principles in and of themselves will not go far enough to cover the risks posed by AI technologies. Principles provide structure to the development of regulation but are inadequate due to existing gaps in the regulatory and governance framework. To properly address the risks posed by AI technologies the Government must introduce new cohesive legislation.

IX. PLP invites the Government to introduce new substantive legal obligations to ensure that public sector use of AI takes due precaution of the associated risks. The light-touch, 'test and learn' approach's primary policy goal is to promote private sector innovation,[2] but it remains to be seen why proper regulation of public sector use of AI would impede private sector innovation. Individuals are put at great risk by taking an entirely reactive 'hurt first, fix later' approach to regulating public sector use of AI.

X. PLP therefore supports the introduction of a statutory duty for regulators but invites the Government to go further and introduce new substantive obligations for public authorities using AI and ADM systems. We encourage the Government to pay mind to precautionary examples of AI regulation from other jurisdictions in order that it might prevent uncertain risks before they are realised.

XI. PLP acknowledges the importance of central oversight. Without this, the regulation of AI will remain piecemeal, with different approaches existing and developing in different sectors. However, we do not believe that the revised cross-sector principles and central functions will effectively address the risk posed by AI used in public decision-making and administration. The White Paper itself acknowledges the limitations to the "current patchwork of regulation".[3] The proposed regulatory framework will inherit the gaps in existing UK regulation and therefore we do not envisage the central functions to adequately cover all high impact areas and uses of AI.

XII. PLP recently convened a joint statement titled, 'Key principles for an alternative AI White Paper',[4] in which we and 29 other individuals and civil society organisations call for a specialist regulator to ensure people can seek redress when things go wrong. This new regulator needs to be adequately resourced, have oversight of the entire AI landscape and have the right tools to enforce the regulatory regime, including powers to proactively audit public ADM tools and their operation across different AI applications and systems.

[2] DSIT (n 1), paragraphs 8-18.
[3] DSIT (n 1), para 70, page 42.
[4] Key principles for an alternative AI White Paper (June 2023) https://publiclawproject.org.uk/content/uploads/2023/06/AI-alternative-white-paper-in-template.pdf.

XIII. To fairly and effectively allocate legal responsibility for AI use across the public sector, specific obligations that require adherence from public authorities developing, deploying, and operating AI should be introduced via statute. The legal framework should seek to bring coherence and clarity by building upon and working with existing data protection safeguards and our public law framework.

Introduction

1. Public Law Project (PLP) is an independent national legal charity founded in 1990 with the aim of improving access to public law remedies for marginalised individuals. Our vision is a world in which the state acts fairly and lawfully. Our mission is to improve public decision-making, empower people to understand and apply public law, and increase access to justice. We deliver our mission through five programmes: litigation, research, advocacy, communications, and training.

2. One of PLP's five priority areas for 2022-25 is ensuring that government use of new technologies is transparent and fair. Given our specific expertise, this consultation response focuses mostly on AI in the context of ADM systems used by public bodies.

3. We are not opposed in principle to government use of ADM systems; we recognise their potential benefits in terms of improving the accuracy and quality of decision-making, as well as efficiency and cost-saving. To realise the potential benefits of this rapidly expanding practice and to allow AI to influence the lives of PLP's beneficiaries positively, these systems must operate fairly, lawfully and in a non-discriminatory way.

4. PLP welcomes the opportunity to respond to this consultation. There are a growing number of instances in which AI is causing harm. Without regulatory intervention, instances of harm will only increase. For that reason, we support the establishment of a new framework to bring clarity and coherence to the regulation of AI.

5. Our central concern is that the White Paper's almost complete failure to pay attention to the use of AI in the public sector fundamentally undermines the regulatory framework it proposes. The significant omission will impede the framework's ability to effectively prevent, or even mitigate, the risks it has identified.[5]

6. The use of AI, specifically within ADM systems used by public authorities, is an important matter which requires serious thought and deliberation. Yet, the White Paper essentially glosses over the specific uses, risks, and need for regulation. Section 3.2 of the White Paper sets out the principles-based regulatory framework and identifies six characteristics of the regime.[6] As first signalled in the policy paper of July 2022,[7] "pro-innovation" continues to be the central driving force, with the Government's

[5] We use the term 'public sector organisation' (interchangeably with 'public body and 'public authority) in line with the broad definition under section 6(3)(b) of the Human Rights Act 1998 (HRA), under which a 'public authority' includes organisations whose functions are of a public nature.
[6] DSIT (n 1), page 21.
[7] Secretary of State for Digital, Culture, Media and Sport, 'Establishing a pro-innovation approach to regulating AI' (20 July 2022) https://www.gov.uk/government/publications/establishing-a-pro-innovation-approach-to-regulating-ai/establishing-a-pro-innovation-approach-to-regulating-ai-policy-statement.

aspiration to "make the UK one of the top places in the world for AI companies" a central focus. Whilst PLP does not wish to challenge this in relation to the private sector, we are concerned that it is unnecessary 'light-touch' approach to the regulation of public authority use of AI and ADM.

7. PLP's research into automation and digitalisation has shown that public authorities are increasingly using ADM in a wide range of high impact areas, *inter alia*, immigration, welfare benefits, policing and prisons, and education. To date, we have gathered more than forty examples of public ADM systems through our investigative research, full details of which are publicly available in our 'Tracking Automated Government (TAG) Register'.[8] The TAG register is reflective of only a small amount of ADM systems used by public authorities, as opacity is an inherent challenge of working in this area. Of the tools we do know about, we have identified key risks and challenges in terms of the risk of discrimination, unlawfulness and unfairness.[9]

8. In November 2022, PLP submitted written evidence to the Science and Technology Committee (STC) inquiry on AI governance in the UK.[10] We drew on ADM tools detailed in the TAG Register to illustrate that governance has been ineffective in relation to public sector use of AI and ADM to the extent that is has not succeeded in securing an adequate level of transparency; accountability; and protection against unlawfulness and unfairness, including ensuring privacy and data rights and protection against discrimination.

9. In 2022, PLP ran two roundtables on regulating government use of AI, with over twenty participants, including civil society organisations, grassroots organisations, academics, and individuals affected by ADM tools. Our response draws on the key themes arising from the roundtable discussions, as well as our research.

10. It is disappointing to see that the AI regulation White Paper has not given due attention to the growing use of AI in government, particularly because of the parallels between the tools currently being used in the UK and those that have been successfully challenged – and their use stopped - in other jurisdictions. Legal challenges have been successful in stopping the use of the Netherlands System Risk Indication (SyRI), the Australian Online Compliance Intervention System (RoboDebt), and the American Michigan Integrated Automated System, finding their operation unlawful. The UK Government must learn from the failures of other governments and ensure that its adoption and use of AI in administrative systems is fair, lawful and non-discriminatory. One way of achieving this is to implement effective regulation of AI in this context, which would allow for individuals to feel the intended benefits of introducing automation into public administration whilst sparing them the unintended effects that have been realised elsewhere.

11. The Government has chosen not to use the White Paper to introduce new legislative requirements to govern the use of AI. The non-statutory nature of the proposed AI regulation means its implementation relies heavily on existing legal and regulatory frameworks. The effectiveness of this approach, against a backdrop of weakening existing data protection and our human rights framework, is uncertain. PLP endorses the concerns of the Ada Lovelace Institute, that the proposed changes to

[8] Public Law Project, Tracking Automated Government 'TAG' Register (9 February 2023) http://trackautomatedgovernment.org.uk/.
[9] See 'Unequal impact(s)' column of the 'Public Law Project, Tracking Automated Government 'TAG' Register (9 February 2023) http://trackautomatedgovernment.org.uk/' for more detail.
[10] Public Law Project's submission of written evidence to the House of Commons Science and Technology Committee's Governance of artificial intelligence (AI) inquiry (13 December 2022) https://committees.parliament.uk/writtenevidence/113825/pdf/.

other legislative frameworks is simultaneously "undercutting the new AI regulatory framework before it is even in place".[11]

12. The Data Protection and Digital Information Bill (No.2) (hereafter the 'DPDI Bill'), if passed in its current form, will water down current data protections. Additionally, threats to leave the European Convention on Human Rights – and legislation which disapplies parts of the Human Rights Act – are putting our human rights framework at risk. Against this backdrop, it is vital that the proposed AI regulation is legally enforceable.

13. The White Paper states that "[r]egulators will lead the implementation of the framework, for example by issuing guidance on best practice for adherence to these principles",[12] and therefore it is concerning that the DPDI Bill will potentially reduce the independence of the ICO by giving the Secretary of State the power to approve or reject novel guidance issued by the regulator. It is difficult to see how existing regulators will carry out their expected role if Ministers are to have such powers.

14. A more 'hands-on' regulatory approach to the use of AI in the public sector, specifically in relation to ADM systems, would not stymie aspirations for broader AI innovation or place additional burdens on businesses. These new governance tools sit closely to the decision-making point, and thus entail greater displacement of human discretion than rounds of innovation that are desired in the private sector. Ultimately, the AI regulation White Paper pays inadequate regard to the specific uses, risks, and need to regulate AI in the public sector.

15. Due to our organisational focus and expertise on public sector use of AI and ADM, we confine our comments to Qs 1-10, 17-18, and L1-L2.

Question 1

Do you agree that requiring organisations to make it clear when they are using AI would improve transparency?

16. PLP agrees that requiring organisations to make it clear when they are using AI would be a helpful first step in improving transparency. However, being 'clear' about the use of AI alone will not be enough to improve transparency.

17. Opacity is an inherent challenge in understanding and assessing the operation of AI, specifically within public sector use of ADM systems. At present, most uses of AI by public authorities have been uncovered through resource-intensive research, including the submission of requests under the Freedom of Information Act 2000 (FOIA). The proactive disclosure from organisations of their use of AI would improve the extent to which individuals, organisations and legal practitioners are aware of the role of AI in administrative decisions.

18. Simply putting information on the use of AI into the public domain will not achieve meaningful transparency. We echo the concerns of Ananny and Crawford that disclosure (as opposed to proactive

[11] Ada Lovelace Institute, 'Policy briefing – the UK government's White Paper on AI regulation' (March 2023) https://www.adalovelaceinstitute.org/wp-content/uploads/2023/03/Briefing-on-AI-White-Paper-March-2023-Ada-Lovelace-Institute.pdf.
[12] DSIT (n 1), para 49, page 26.

transparency) can place "a tremendous burden on individuals to seek out information about a system, to interpret that information, and determine its significance".[13] Instead, the public must be informed of the use of AI by public authorities alongside the notification of decisions. In other words, the public authority must be *proactively* clear about the presence of AI or automation within any stage of a decision-making process. Transparency will not be meaningfully achieved if individuals need to know what to ask for and where to look to obtain the relevant information. There is an asymmetry in power between data controllers, AI and ADM systems operators, and those subject to decisions. AI and ADM systems are involved in decision-making that has significant and material impacts on people's lives, especially because these systems are being used by public authorities such as the Home Office and Department for Work and Pensions. Ultimately, the disclosure of information about the use of AI and ADM systems must be proactively and prospectively led by the public authority using the system.

19. It is also vital for this information to be provided in a way that is concise, transparent, intelligible, easily accessible, and in clear and plain language. In practical terms, this information could include, *inter alia*, the existence of the system, operating details, the (unequal) impacts on those with protected characteristics, specific groups or communities and the role of AI in the decision recommendation, or final decision.

Question 2

Are there other measures we could require of organisations to improve AI transparency?

20. There are three measures to improve AI transparency that the White Paper unfortunately does not consider. They are:

    (a) Introduce specific compulsory transparency requirements for public sector use of AI and ADM systems,

    (b) Make the submission of reports under ATRS mandatory for public authorities and increase the level of operational details required,

    (c) Introduce a requirement for public authorities to notify individuals when automation is used to reach a decision, similar to that under the Canadian Directive on Automated Decision Making (DADM),[14] and France's Loi pour une Republique Numérique (Law for a Digital Republic) 2016.[15]

21. First, specific transparency requirements for public sector use of AI and ADM systems must be compulsory. The General Data Protection Regulation (GDPR), as implemented by the Data Protection Act 2018, already requires transparency in relation to automation in government. Under Art. 5(1)(a) GDPR personal data must be "processed lawfully, fairly and in a *transparent* manner in relation to the data subject".[16] Under Article 22, in combination with Articles 12, 13, and 14 GDPR (governing transparency of information), data subjects can request information to find out if they are being subjected to solely automated decision-making. However, the transparency requirements under

[13] Ananny, M. and Crawford, K., *Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability,* New Media and Society, 2016, 7 https://journals.sagepub.com/doi/10.1177/1461444816676645
[14] Government of Canada, Direction on Automated Decision-Making (1 April 2019), available at: https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592.
[15] Law No. 2016-1321 of October 7, 2016, for a Digital Republic.
[16] Emphasis added.

existing legal frameworks do not go far enough in securing meaningful transparency. In our experience, it is often difficult to find out about the existence of AI and ADM systems, let alone obtain information about their operation – both in general and in application to a specific individual.[17]

22. In its response to the consultation 'Data: a new direction', the Government stated that it "does not intend to take forward legislative change at this time", despite widespread support for compulsory transparency reporting,[18] and, indeed, the DPDI Bill does not include any such requirements. It is for that reason we believe a specific statutory requirement for transparency of public authority use of AI and ADM systems would be the best first step to ensuring meaningful transparency and greater public understanding of the use of AI. It is disappointing therefore, that the Government have chosen not to use the White Paper to introduce legislative requirements, despite acknowledging not only the importance of transparency in increasing public trust but also that the lack of transparency may lead to AI users "breaking laws, infringing rights [and] causing harm".[19] Our response to Qs 7 and 8 detail our position on the introduction of statutory duties related to public authority use of AI.

23. Second, we recognise the potential of the Central Digital and Data Office (CDDO) and Centre for Data Ethics and Innovation's (CDEI) 'Algorithmic Transparency Recording Standard' (ATRS) to increase transparency of the public sector's use of algorithms. However, its potential to deliver a minimum viable standard of transparency is and will remain limited if engagement with the ATRS is not made compulsory. At present, it is not mandatory for public sector organisations to engage with the standard or produce reports, and this does not appear likely to change in the near future. PLP's TAG Register should not be the most comprehensive account of government use of AI and ADM. [20] As set out in paragraph 7, the TAG Register is reflective of only a small amount of ADM systems used by public authorities and the onus must be placed on public authorities to disclose information of this detail proactively and prospectively.

24. Moreover, even if it were placed on a statutory footing, the ATRS does not ask for sufficient operational detail to ensure meaningful transparency. The purpose of transparency bears on its meaning in the context of ADM. The Berkman Klein Center conducted an analysis of 36 prominent AI policy documents to identify thematic trends in ethical standards.[21] They found there is convergence around a requirement for systems to be designed and implemented to allow for human oversight through the "translation of their operation into intelligible outputs". In other words, transparency requires the 'translation' of an operation undertaken by an ADM system into something that the average person can understand. Without this, there can be no democratic consensus-building or accountability. Another plank of meaningful transparency is the ability to test explanations of what an algorithmic tool is doing. In our view, meaningful transparency requires that people lacking specific technical expertise—i.e., the vast majority of us—are able to understand and test how an algorithmic

[17] See PLP's submission to the Science and Technology Committee inquiry on AI governance in the UK for more information, available at: https://committees.parliament.uk/work/6986/governance-of-artificial-intelligence-ai/

[18] The government proposed introducing compulsory transparency reporting on the use of algorithms in decision-making for public sector bodies and "the majority of respondents agreed with this proposal, with respondents outlining they believe it would improve public trust", Section 4.4. of Department for Digital, Culture, Media and Sport (DCMS), *Data: a new direction – government response to consultation* (17 June 2022), https://www.gov.uk/government/consultations/data-a-new-direction/outcome/data-a-new-direction-government-response-to-consultation.

[19] DSIT (n 1), section 3.2.3, page 29.

[20] 'Public Law Project, Tracking Automated Government 'TAG' Register (9 February 2023) http://trackautomatedgovernment.org.uk/.

[21] Jessica Fjeld, and Achten, Nele and Hilligoss, Hannah and Nagy, Adam and Srikumar, Madhulika, Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI (January 15, 2020). Berkman Klein Center Research Publication No. 2020-1 (15 January 2020), available at https://ssrn.com/abstract=3518482 or http://dx.doi.org/10.2139/ssrn.3518482.

tool works. Against this interpretation of transparency, the ATRS falls short. Neither Tier 1 nor Tier 2 of the ATRS requires sufficient operational details for individuals to properly understand the decision-making process to which they are subjected. At Tier 1, organisations are asked to explain 'how the tool works', but nowhere is there a reference to any criteria or rules used by simpler algorithmic tools. At Tier 2, a 'technical specification' is requested, but this appears to mean nothing more than a brief descriptor of the type of system used, e.g., 'deep neural network'.

25. We propose the following measures for improving transparency under the ATRS:

    (a) At Tier 1, the high-level explanation of how the algorithm works should include rules or criteria used by the algorithm.

    (b) At Tier 2, there should be sufficient detail for an individual whose rights may be affected to fully understand how the process works. For example, we invite the Government to consider whether a link to an 'executable version' (EV) should be a requirement. PLP has written in detail about disclosure of EVs as a means to secure meaningful transparency.[22] Based on our definition, an EV of a model is one that allows someone with access to it to: (1) change the inputs or assumptions of the model; (2) run the model; and (3) see the outputs. In our view, the salience of an EV is that it allows someone to see and use the 'front-end' of the decision-making tool. It does not offer access to the 'back-end', and it is only a copy – it does not allow a third party to make changes to the system actually used by the decision-maker. Where an EV is disclosed, a third party or affected individual can run the EV on a range of different inputs and generate their own counterfactual explanations e.g., they would be able to say, "If I earned more, my application would have been successful".

    (c) Arguably, an EV is—for transparency purposes—the closest equivalent to a written policy because it allows an ordinary person to understand and test explanations of how discretionary power is to be exercised in a given case. For an example of an EV, the investigative newsroom Lighthouse Reports, based in the Netherlands, were able to obtain sufficient information to create an EV by reconstructing the algorithm previously used by Dutch municipalities in an attempt to stop welfare fraud.[23] The algorithm profiles citizens on social assistance benefits and categorises them into risk groups, with specific target groups being rendered as potential fraudsters.[24] Lighthouse Reports' EV allows people to generate their own risk score, understand how those scores are calculated, and what characteristics and behaviours are taken into account.

    (d) Alongside impact assessments, Tier 2 should include other ongoing monitoring and evaluation work conducted by government departments, such as data on the impact and efficacy of their algorithmic tools.

26. Third, we urge the UK to draw inspiration from the compulsory transparency regimes found in Canada under the Directive on Automated Decision Making (DADM),[25] and France's Loi pour une Republique

[22] Mia Leslie and Tatiana Kazim, 'Executable versions: an argument for compulsory disclosure (part 1)' (The Digital Constitutionalist, 03 August 2022), available at: https://digi-con.org/executable-versions-part-one/; Mia Leslie and Tatiana Kazim, 'Executable versions: an argument for compulsory disclosure (part 2)' (*The Digital Constitutionalist*, 03 November 2022), available at: https://digi-con.org/executable-versions-part-two/.

[23] FNV, Lighthouse Reports, Argos, and NRC news' reconstruction of the Dutch 'Fraud Scorecard' algorithm (14 July 2022), available at: https://www-fnv-nl.translate.goog/nieuwsbericht/sectornieuws/uitkeringsgerechtigden/2022/07/ben-jij-door-je-gemeente-mogelijk-als-fraudeur-aan?_x_tr_sl=auto&_x_tr_tl=en&_x_tr_hl=en-US&_x_tr_pto=wapp

[24] Gabriel Geiger and others, 'Junk Science Underpins Fraud Scores' (*Lighthouse Reports*, 22 June 2022), available at: https://www.lighthousereports.nl/investigation/junk-science-underpins-fraud-scores/.

[25] Government of Canada, Directive on Automated Decision-Making (1 April 2019), available at: https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592.

Numérique (Law for a Digital Republic) 2016.[26]

27. The DADM requires federal institutions to disclose information about ADM systems used to recommend or make administrative decisions about an individual. There are three disclosure requirements that apply to all ADM systems. At all levels, operators are required to i) disclose the components of the ADM system, ii) disclose the source code—subject to certain exemptions—and iii) document decisions made by the ADM systems for monitoring and reporting. Under the Canadian regime, the level of transparency depends on the level of expected impact on individuals, communities and ecosystems, which is assessed using an Algorithmic Impact Assessment tool (AIA).[27] The AIA is a mandatory risk assessment questionnaire that determines the impact level of an automated decision-system. Some of the criticisms of UK's ATRS detailed in paragraphs 23-25 could be addressed if it was amended in line with, or included a risk assessment section similar to, the Canadian AIA.

28. Furthermore, when a higher impact system is identified under the AIA more detailed disclosure is required. Notice must be given to individuals subject to ADM and a meaningful explanation provided to affected individuals as to how and why the decision was made in this way, in addition to disclosure of information about how the components work; how the algorithm supports the administrative decision; results of any reviews or audits; and a description of the training data, or a link to the anonymised training data if this data is publicly available.

29. Similarly, the French regime requires those implementing ADM systems to provide notice to those subject to decisions that a decision is made or supported by an algorithm. But the regime goes further and includes a requirement to publish the rules defining the algorithmic processing, the main characteristics of its implementation, and the purpose of such processing. Further still, if requested by the person concerned, the implementing authority must also disclose the extent to which the algorithm contributed to the decision-making process, the data processed by the system (including its sources), and the processing criteria and their weighting.

30. If the Government aspires for its AI regulation to be "word leading", we encourage it to first bring AI transparency requirements in line with those under the Canadian and French regimes at minimum. The adoption of an impact and risk assessment, and corresponding transparency requirements, like those under the Canadian DADM would allow for proportionate and meaningful transparency that relates to the assessed impact of public sector use of AI and ADM systems. Additionally, the Government should take inspiration from both regimes and use forthcoming AI regulation to require public authorities to provide notice to individuals who are subject to decisions made or supported by an algorithm. This would be a positive step towards securing meaningful and proactive AI transparency in the public sector, as called for in response to Q1.

Question 3

Do you agree that current routes to contest or get redress for AI-related harms are adequate?

31. We do not agree.

[26] Law No. 2016-1321 of October 7, 2016, for a Digital Republic.

[27] Directive on Automated Decision-Making, Algorithmic Impact Assessment https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html

32. It is important to first note that the ability to seek redress for harms caused by AI systems is necessarily predicated on the knowledge of the existence of such systems. As articulated in our response to Qs 1 and 2, current transparency of the use of AI and ADM systems by public authorities is insufficient – both in its practical realisation and the legal and regulatory frameworks that govern it. Transparency is a necessary starting point for evaluating AI technologies in order to contest or redress related issues or harms, and without transparency there cannot be adequate routes to contest or seek redress for AI-related harms. [28] Where individuals and communities are not aware of the existence of AI in decision-making processes, it is not possible for them to even consider contesting decisions or processes on this basis, frustrating any efforts to contest or seek redress for AI-related harms at the first hurdle.

33. The contestation of AI-related harms and redress requires accountability from the public authorities developing and deploying AI and ADM systems. Accountability goes beyond transparency in that it requires adequate avenues to challenge, together with effective enforcement mechanisms and the possibility of sanctions. Interpretations of accountability differ, but we find Mashaw's 'six inquiries' a helpful tool in understanding and assessing accountability regimes. Mashaw's position is that in any accountability relationship, we should be able to specify at least six important things: [29]

    (a) *Who* is liable or accountable

    (b) *To whom* are they liable or accountable

    (c) *What* they are liable to be called to account for

    (d) *Through what processes* accountability is to be assured

    (e) *By what standards* the putatively accountable behaviour is to be judged

    (f) What the potential *effects* are of finding those standards have been breached

34. Our research shows that a comprehensive legal and regulatory framework is still a work in progress and that the current legal and regulatory requirements are fragmented at best. Although not created with AI in mind, some existing legal frameworks contain a number of crucial safeguards which can be interpreted to regulate its use. For instance, AI is subject to the provisions of the UK GDPR; public law doctrines developed through the common law, such as the duty to give reasons; the Equality Act 2010 (EA 2010); the Human Rights Act 1998 (HRA 1998); and the European Convention on Human Rights (ECHR), particularly Articles 8 and 14. However, the few cases in which individuals and organisations overcome the transparency hurdle and seek to contest or get redress for AI-related harms are being settled before they get to court. At the time of writing, contesting decisions made or supported by AI and ADM systems is one of the most uncertain areas of public law because there are no clear judicial principles on how existing legal and regulatory frameworks apply when seeking accountability for AI related harms.

35. In what follows, we give some additional detail on the requirements of the DPA 2018 and UK GDPR, and the EA 2010 and how they can currently be used to contest or get redress for AI-related harms.

36. Under the DPA 2018 and UK GDPR:
    (a) Data subject access requests (Article 15 of the UK GPDR) allow individuals to check the accuracy of their personal data, learn more about how their data is being used and with whom their data is

[28] See Justice and Home Affairs Committee, 'Technology rules? The advent of new technologies in the justice system' (30 March 2022) https://committees.parliament.uk/publications/9453/documents/163029/default/.
[29] Jerry Louis Mashaw, Accountability and Institutional Design: Some Thoughts on the Grammar of Governance (1 November 2006), 118, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=924879.

being shared, and obtain a copy of the data held about them.

(b) Article 22 of the UK GPDR provides that a data subject shall have the right "not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her". Article 22 is capable of placing a meaningful limit on the deployment of AI within ADM systems.

(c) The requirement to undertake a Data Protection Impact Assessment (DPIA) is provided for under Section 64 of the DPA 2018. DPIAs are an important tool for guarding against some of the risks posed by ADM systems, including discrimination, by helping with the identification and minimisation of such risks before they arise. DPIAs that fail to sufficiently assess the risks to the rights and freedoms of data subjects, and therefore do not take the measures envisaged to address the risks, breach the requirement under Section 64.30

(d) The requirement that data processing is lawful, fair and transparent is found in Article 5(1)(a) of the UK GDPR. Harmful and rights-violating uses of AI can be challenged under this provision, alongside Articles 6 and 9(2) which set out conditions for lawful processing of data and personal data, and Article 14, which requires data subjects to be informed about how their personal data is used.

(e) Where these provisions are contravened, individuals are able to contest the breach and seek redress. Under Articles 77-79, 80 and 82, the GDPR provides mechanisms for individuals to challenge the breach of their privacy rights and harm caused by the collection, processing, storage, or transmission of personal data. These provisions allow individuals to seek redress and compensation against organisations for harms caused by AI systems.

37. Whilst providing vital safeguards, the routes to contest or obtain redress for AI-related harms through the DPA 2018 and UK GDPR are limited. Article 22 could be made more effective through clarification of its key terms – especially "a decision based solely on automated processing" – to ensure that it has broad practical application. Article 22, properly defined, should prohibit de facto solely automated decision-making where, due to automation bias[31] or for any other reason, the human official is merely rubber-stamping a score, rating or categorisation determined by the computer. It should require meaningful human oversight, rather than a token gesture. Whilst it is not legally mandatory to publish a DPIA, it can help to foster trust in public authorities handling and processing personal data. It demonstrates accountability and transparency and allows for any unequal impacts of the processing to be publicly assessed and scrutinised. However, there is a significant lack of transparency of DPIAs, of the 42 tools in the TAG Register, only 45.2% have publicly available government assessments of their impact on the protected characteristics of individuals.[32] Overall, it is uncertain how long these safeguards will persist if the DPDI Bill passes in its current form.

38. The EA 2010: Meaningfully performed,[33] the Public Sector Equality Duty (PSED) as set out in section 149 of the EA 2010 helps public bodies consider the impact of policies and decisions on people who share a protected characteristic. It is the prerogative of the decision-maker to decide the process to ensure

[30] *R (Bridges) v Chief Constable of the South Wales Police [2019] EWHC 2341 (Admin)* [153]

[31] See, for example, L.J. Skitka and others, 'Does automation bias decision-making?' (1999) 51 International Journal of Human-Computer Studies 991; Tatiana Kazim and Joe Tomlinson (2023) Automation Bias and the Principles of Judicial Review, Judicial Review https://www.tandfonline.com/doi/epdf/10.1080/10854681.2023.2189405?needAccess=true&role=button.

[32] Figure relates to both Equality Impact Assessments and Data Protection Impact Assessments, 'Public Law Project, Tracking Automated Government 'TAG' Register (9 February 2023) http://trackautomatedgovernment.org.uk/.

[33] The PSED is often not meaningfully performed. The substance of the duty is often replaced by the prioritisation of form and/or procedure of discharging the duty.

compliance with the PSED. In practice public bodies often carry out Equality Impact Assessments (EIAs) to fulfil their obligations under section 149. The PSED also applies to any AI systems that public bodies are already using or that others may be developing or using on their behalf. Where the PSED is breached, a claim for judicial review may be made to challenge the failure of the public authority to comply with its duties. The 2020 Committee on Standards in Public Life report, *Artificial Intelligence and Public Standards*,[34] looked at the risks and opportunities for public standards posed by AI, and found that data bias could cause AI to produce decisions and policy outcomes that are discriminatory, which may breach the EA 2010.[35] Whilst contributors to the review believed the PSED is the "single best tool available" for dealing with data bias if used correctly,[36] they said there was uncertainty about how the EA 2010 applies to ADM in practice. There is currently no sufficiently detailed guidance for public bodies using AI on how to comply with anti-discrimination law.

39. Whilst public authorities should ensure that the results of EIAs are published as soon as possible after the date of the decision, our experiences have demonstrated a reluctance from many public authorities to proactively publish or disclose (in response to requests submitted under the FOIA 2000) EIAs. Without access to EIAs it is difficult for affected individuals to adequately assess whether the public authority have taken the necessary steps to discharge the PSED. As stated above in relation to DPIAs, of the 42 tools in the TAG Register, only 45.2% have publicly available government assessments of their impact on the protected characteristics of individuals.[37] Lack of transparency surrounding EIAs can limit the potential of the EA 2010 to allow for contestation and redress for AI-related harms.

40. Ultimately, PLP does not agree that the that current routes to contest or get redress for AI-related harms caused by public authorities are adequate. As set out above, there are shortfalls in the existing non-AI specific legal frameworks, particularly the lack of transparency of both the use of AI and ADM systems, and of relevant DPIAs and EIAs.

41. Drawing again on Mashaw's inquiries, without the disclosure of information about the method of decision-making it is difficult for affected individuals to identify *who* is liable or accountable for the AI-related harms they suffer, and exactly *what* they are liable to be called to account for. Without the disclosure of assessments of the impact of policies and processes, it is not possible to know on a systemic level *what* the public authority are liable for when AI systems cause harm. Without clear, comprehensive, and unified frameworks for challenging AI-related harms it is not possible to know exactly *through what* processes accountability is to be assured or *by what standards* the putatively accountable behaviour is to be judged. And finally, without the development of judicial principles through caselaw, the public cannot know what the potential *effects* are of finding those standards have been breached. Without certainty in relation to these inquiries, the accountability regime for AI-related harm will remain inadequate.

Question 4

How could current routes to contest or seek redress for AI-related harms be improved, if at all?

[34] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/868284/Web_Version_AI_and_Public_Standards.PDF

[35] https://www.libertyhumanrights.org.uk/issue/policing-by-machine

[36] https://www.equalityhumanrights.com/en/advice-and-guidance/public-sector-equality-duty

[37] Figure relates to both Equality Impact Assessments and Data Protection Impact Assessments, 'Public Law Project, Tracking Automated Government 'TAG' Register (9 February 2023) http://trackautomatedgovernment.org.uk/.

42. To improve routes to contest or seek redress, the focus should first be on fortifying existing safeguards and ensuring clarity and coherence between existing laws. PLP is concerned about the lack of bite judicial review is having in practice in this context, and the direction of legislative reforms and the effect this will have on individuals and communities subject to AI-related harms.

43. Public bodies are responsible and accountable for the lawfulness of their decision-making whether involving AI and ADM systems or not. As emphasised by Cobbe, they are required to meet public law's "standards when using [AI and] ADM just as with human decision-making, and… an unlawful decision made by or with the assistance of ADM should be dealt with by reviewers as it would had a similarly unlawful decision been taken by a human".[38] However, it remains unclear how public law will apply to AI and ADM systems used in the public sector. This is in large part due to the current tendency for legal challenges to be settled before reaching court.

44. The DPDI Bill is, at the time of writing, making its way through Parliament. We have set out our detailed position on the DPDI Bill in our second reading briefing,[39] and oral evidence session.[40] In the context of AI regulation, our concern is that if the Bill passes in its current form it would weaken the above important data protection rights and safeguards, making it more difficult for people to know how their data is being used, how decisions about them are being made, and weakening requirements on those who process data to consider the rights and interests of those their actions will affect. Personal data is fed into ADM systems and the DPDI Bill would mean sweeping changes to data protection law, including both the DPA 2018 and the UK GDPR - changes which prioritise economic growth and innovation over rights-protection, transparency and accountability. While the Bill does not outright remove any of the current protections in data protection law, it weakens many of them to the extent that they will struggle to achieve their original purposes.

45. At the very least, the safeguards under the DPA 2018 and UK GDPR must be protected. The watering down of requirements on public authorities to consider the rights and interests of those their actions will affect will allow for the same, if not more, AI-related harms to occur without the corresponding restrictions and routes for challenge.

46. If AI regulation is to make an improvement on existing frameworks, it will be necessary to ensure that there are quick and effective ways of enforcing existing rights. At PLP's roundtables, better enforcement mechanisms were seen as a more urgent priority than new digital rights. One way to do this could be through a specialist regulator and forum for complaints relating to public authority use of AI. However, roundtable attendees were concerned that a specialist forum may not be accessible for affected individuals.

47. Another option could be sector- or system-specific avenues for redress. For example, if welfare benefits are suspended through the use of an automated system, an affected individual should be specifically informed that an automated system was used in the decision-making process and there should be a dedicated and adequately resourced complaints procedure if they suspect unfairness as a result of AI in a particular government process.

---

[38] Jennifer Cobbe, 'Administrative Law and the Machines of Government: Judicial Review of Automated Public-Sector Decision-Making', 7 https://core.ac.uk/reader/334953720.; In another common law jurisdiction, the Australian Government's best practice principles for ADM emphasise that decisions made by or with the assistance of ADM must comply with administrative law (Australian Government 'Automated Assistance in Administrative Decision-Making: Better Practice Guide'(2007), p ix. Available at https://www.oaic.gov.au/images/documents/migrated/migrated/betterpracticeguide.pdf.

[39] https://publiclawproject.org.uk/content/uploads/2023/04/PLP-Briefing-DPDI-Bill-No.2-Second-Reading-Final-1.pdf

[40] See from 15:54, Data Protection and Digital Information (No. 2) Bill Committee, Oral evidence session (10 May) available at: https://parliamentlive.tv/event/index/12fd22dd-4fed-4d01-90a9-354c1dae034a?in=15:54:12.

Question 5

Do you agree that, when implemented effectively, the revised cross-sectoral principles will cover the risks posed by AI technologies?

Question 6

What, if anything, is missing from the revised principles?

48. We will address Qs 5 and 6 together. PLP agrees that the principles are steering AI regulation in the right direction, but that the principles in and of themselves will not go far enough to cover the risks posed by AI technologies.

49. We support Lord Clement Jones' view that the revised principles amount "to toothless exhortation by sectoral regulators to follow ethical principles".[41] Introducing 'AI regulation' means addressing the issues and risks posed by AI technologies. Numerous examples of non-statutory guidance exist from regulators, governmental and non-governmental bodies, both domestically and internationally.[42] These initiatives have provided structure and demonstrated cohesion across early-stage public debate on how we want to regulate and govern AI. Yet, the AI regulation White Paper adds little to this ecosystem. The revised principles are effectively instructions to regulators about the outcomes they ought to be working towards when AI is used in the areas for which they are responsible.

50. The principles are underpowered and, as articulated by the Ada Lovelace Institute, the proposed framework "inherits gaps in existing UK regulation, and therefore risks leaving prominent AI harms unaddressed".[43] The White Paper itself acknowledges that some regulators have warned they may "lack the statutory basis to consider the application of the principles".[44]

51. To adequately address the risks posed by AI technologies the Government must introduce new legislation (addressed in full under Qs 7 and 8).

Question 7

Do you agree that introducing a statutory duty on regulators to have due regard to the principles would clarify and strengthen regulators' mandates to implement our

[41] Lord Clement Jones, 'The long-awaited AI Governance White Paper falls far short of what is needed' (3 April 2023) https://www.politicshome.com/thehouse/article/longawaited-ai-governance-white-paper-falls-far-short-needed.
[42] 'Guidance on AI and data protection', Information Commissioner's Office https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/guidance-on-ai-and-data-protection/; 'Guidelines for AI procurement. Office for AI (June 2020) https://www.gov.uk/government/publications/guidelines-for-ai-procurement/guidelines-for-ai-procurement; 'Artificial intelligence in public services', Equality and Human Rights Commission (September 2022) https://equalityhumanrights.com/en/advice-and-guidance/artificial-intelligence-public-services; Organisation for Economic Co-operation and Development, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449 (22/05/2019) https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449; Council of Europe, Ad hoc Committee on Artificial Intelligence (CAHAI) (December 2021) https://rm.coe.int/cahai-2021-09rev-elements/1680a6d90d.
[43] Ada Lovelace Institute, 'Policy briefing – the UK government's White Paper on AI regulation' (March 2023) https://www.adalovelaceinstitute.org/wp-content/uploads/2023/03/Briefing-on-AI-White-Paper-March-2023-Ada-Lovelace-Institute.pdf.
[44] DSIT (n 1), page 36.

principles while retaining a flexible approach to implementation?

52. PLP welcomes statutory intervention to strengthen regulators' mandates to implement the principles. However, we are concerned for two reasons that the White Paper does not adequately do so. Firstly, the Government has not committed to introducing the proposed statutory duty. It has only committed to assess whether doing so is necessary after AI users have tested the existing regulatory framework. Secondly, the White Paper has not adequately considered public sector use of AI, which presents a different regulatory challenge. Therefore, PLP invites the Government to introduce new substantive obligations to ensure that public sector use of AI takes due precaution for the risks associated.

53. The stated aim of the White Paper is to address the risk of AI in a proportionate manner to ensure the continued growth and development of AI in the UK.[45] We do not oppose this aim, but we do consider the risks posed by AI use in the public sector to be different to those posed by private sector use and the potential for regulation of public bodies to stifle innovation is lower. Accordingly, PLP considers that the White Paper's proportionate approach requires that regulation of AI use in the public sector be more precautionary than in the private sector and we demonstrate that the White Paper has not given due consideration to public sector use of AI.

54. If, at some point in the future, the Government's assessment is that statutory intervention is necessary, the proposed statutory duty would not introduce any new substantive obligations for regulators or for public bodies. Regulators are not to be provided with any new enforcement powers. The duty is merely 'to have due regard' to the general ethical principles set out in the White Paper. Regulators are expected to enforce the ethical principles to the extent that it is already possible, but they are expected to do so more proactively because of the new statutory duty.[46]

55. The Government anticipates that there will be instances of AI being used in contravention of the principles but not in contravention of any law.[47] Accordingly, such instances may be free from interference by regulators. The Government argue that such instances are a necessary price to pay in a trade-off for becoming "the jurisdiction with the most pro-innovation approach."[48] The primary policy goal of the light-touch 'test and learn' approach is to promote private sector innovation,[49] but it remains to be seen why light-touch regulation of public sector AI use would foster private sector innovation. The White Paper has not set out why possible contraventions of the principles are justified in the public sector, where the risk of harm to ordinary people is significant.

56. The statutory duty would affect how public sector – as well as private sector - use of AI is regulated. The 10+ regulators with digital technologies in their direct remit (who will be subjected to the duty) regulate AI use in both the private and the public sector. Therefore, the statutory duty aims to strengthen regulators in how they guard individuals against AI risks including those arising from public sector use. But the White Paper does not introduce any new obligation on public bodies that would prevent such harms occurring: the statutory duty does not apply to the public sector users of AI and ADM systems. In fact, the Government is actively watering down its existing obligations that promote the principles through legislation such as DPDI Bill.[50] Therefore, the Government will be in the contradictory position of instructing regulators to apply its ethical principles on AI use, whilst not

[45] *UK Artificial Intelligence Regulation Impact Assessment* (23rd March 2023), paragraph 61.
[46] DSIT (n 1), paragraph 67.
[47] *UK Artificial Intelligence Regulation Impact Assessment* (23rd March 2023), paragraph 100.
[48] *UK Artificial Intelligence Regulation Impact Assessment* (23rd March 2023), paragraph 66.
[49] DSIT (n 1), paragraphs 8-18
[50] For further information on this point, see our second reading briefing on the Bill: https://publiclawproject.org.uk/content/uploads/2023/04/PLP-Briefing-DPDI-Bill-No.2-Second-Reading-Final-1.pdf

specifically accepting an obligation to do the same.

57. This could lead to a situation where regulators are unable to enforce the principles against harmful use of AI by public bodies. For example, the Information Commissioner's Office (ICO) will play a significant role in promoting the AI transparency principle as it enforces data protection legislation and the Freedom of Information Act (FOIA). The statutory duty will require the ICO to be more pro-active in promoting transparent use of AI through FOIA, but public authorities using AI will not themselves be under any new transparency obligation. If the public sector use of AI is novel, and falls outside of existing regulations, then an individual harmed by it will have no means of enforcing the proposed principles against the relevant body. This situation will be allowed to occur because the statutory duty only applies to regulators. Novel situations are likely because the existing patchwork of regulation was developed before the wide-spread use of AI systems by government and are therefore poorly tailored to its use.[51] We therefore recommend that the Government reconsider its decision not to introduce any new obligations for public authorities using AI.

58. Such instances and their harmful effects have been justified in the White Paper as test cases, from which the Government will learn. The potential harmful effects of public sector use of AI are significant and is often deployed on such a scale that the inherent risks of a decision are amplified. It has been highlighted that even if a system is almost completely accurate, the scale of public sector use means that thousands can still be harmed by a mistake.[52] Consider, for example, the abandoned A-level results algorithm used to decide grades during the COVID-19 pandemic. Ofqual tested the algorithm's accuracy by applying it to students from the three years preceding implementation. In 51 of 55 subjects being assessed, the algorithm correctly predicted actual attained grades in over 90% of cases.[53] But when the algorithm was used to determine such a significant outcome for so many people, even a small proportion of inaccurate predictions led to the large-scale amplification of unfairness. We submit that public authorities should only be permitted to run such risks where there is sound justification for doing so.

59. The justification offered by the White Paper for permitting test cases is the promotion of private sector innovation. PLP considers this justification to have been offered without specific consideration of the different challenges posed by public sector use of AI. The economic modelling cited in support of the test and learn approach argues that the risk from regulation to private sector innovation arises primarily from compliance costs, levels of prohibition on private firms, and levels of public trust in AI.[54] It remains to be seen why regulating public sector use would introduce compliance costs for private sector firms; and why prohibitions in the public sector would increase private sector costs. Furthermore, given that innovation is encouraged by high levels of public trust, heed must be paid to the fact that public sector use of AI enjoys a lower level of public confidence than AI use in any other sector.[55] The Government's decision not to create a statutory obligation to apply the principles may damage already low public trust and, should the assumptions of the economic model hold, this may stifle innovation. Accordingly, the potential benefit of a test and learn approach is less in the public

[51] Council of Europe, Ad hoc Committee on Artificial Intelligence (CAHAI) (December 2021) https://rm.coe.int/cahai-2021-09rev-elements/1680a6d90d at paragraph 82.
[52] Council of Europe, Ad hoc Committee on Artificial Intelligence (CAHAI) (December 2021) https://rm.coe.int/cahai-2021-09rev-elements/1680a6d90d at paragraph 13.
[53] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/909368/6656-1_Awarding_GCSE__AS__A_level__advanced_extension_awards_and_extended_project_qualifications_in_summer_2020_-_interim_report.pdf at pages 76 - 81.
[54] Evidence to support the Analysis of Impacts for AI Governance (2023).
[55] Trust in AI – A Five Country Study, KPMG (2021) https://kpmg.com/au/en/home/insights/2021/03/artificial-intelligence-five-country-study.html at page 20.

sector than it is in the private sector. A proportionate approach would recognise that further consideration must be given to the specific challenges posed by public sector use of AI.

60. PLP recognises that the White Paper aims to harness the economic opportunity of promoting private sector AI innovation. We submit, however, that this does not provide the basis for a test and learn approach for public sector use. Individuals are put at great risk by taking an entirely reactive 'hurt first, fix later' approach to regulating public sector use of AI. The statutory intervention is required to adequately mitigate against this risk, and it is not clear that there would be an economic harm in doing so. Therefore, PLP supports the introduction of a statutory duty for regulators but invites the Government to go further and introduce new substantive obligations for public authorities using AI and ADM systems.

Question 8

Is there an alternative statutory intervention that would be more effective?

61. PLP considers that further statutory intervention is required to promote safe public authority use of AI. We recommend that the Government adopt a precautionary approach to regulating AI in the public sector.

62. In our response to Q7, we have set out that the proposal to 'wait and see' if statutory intervention is required accepts a significant risk of harm to individuals and appears to have been arrived at without adequate, or indeed any, consideration of public sector use of AI. The risk of harm from public sector use of automated decision making is significant. The White Paper recognises this, stating that *"without Government action, AI could cause and amplify discrimination that results in, for example, unfairness in the justice system [or] pose risks to our privacy and human dignity, potentially harming our fundamental liberties"*.[56]

63. We consider statutory intervention necessary to prevent harms before they occur. However, the full nature of potential future AI harms is unknown. The White Paper highlights that *"given the pace at which AI technologies and risks emerge, […] we know that there is no time to waste*."[57] As articulated in the introduction to this response, AI risks are more than hypothetical. There have been successful legal challenges in the Netherlands, Australia, and USA and PLP's TAG Register highlights the presence of similar systems in the UK.[58] The UK must learn from the failures of other governments and introduce effective regulation to guard the public against uncertain risks. Such risks require caution; uncertainty cannot justify inaction. It is not enough for the Government to merely scan the horizon for harms that have already emerged from public sector AI.

64. Statutory intervention is required before, not after, the full extent of AI-related harm in the public sector is realised. PLP considers that to do so, the Government must learn from other instances where it has regulated to minimise uncertain risks. In such circumstances, international agreements and domestic policy have taken a precautionary approach *"in advance of scientific proof, or in the face of fundamental ignorance of possible consequences"*.[59] The UK Government has, for example, adopted a precautionary approach to regulation in the context of uncertainty as to the risks posed by beef on the

[56] DSIT (n 1), paragraph 25.
[57] DSIT (n 1), paragraph 124.
[58] Public Law Project, Tracking Automated Government 'TAG' Register (9 February 2023) http://trackautomatedgovernment.org.uk/.
[59] J. Cameron and T. O'Riordan *'Interpreting the Precautionary Principle'* (2013).

bone during the 1997 BSE epidemic,[60] and environmental regulation.[61] PLP therefore recommends that an analogous approach to regulation be taken for public sector use of AI, and endorses the view expressed by Professor Joe Tomlinson and Jack Maxwell that uncertainties warrant a precautionary approach to the development and use of automated systems in government.[62]

65. A precautionary approach to regulating public sector use of AI is being taken in other jurisdictions. PLP encourages the Government to learn from such approaches and highlights three approaches in particular: the proposed EU AI Act, and Canadian and French domestic public sector AI Governance regimes.

66. The proposed EU AI Act would introduce a centralised regulatory framework for AI. To mirror this proposal, the Government considered a 'do maximum' option. PLP recognises the potential benefits of such an approach, as highlighted in the White Paper Impact Assessment. [63] Individuals receive the highest direct benefit where they are protected by new legislative requirements. More stringent regulation of public sector use of AI can reduce the risks and harms posed to individuals and achieve greater trust in AI use. [64] We note that the Government did not consider introducing new legislative requirements exclusively for the public sector to achieve these benefits. Rather, it appears to have understood statutory intervention as all or nothing.

67. PLP does, however, agree with some of the concerns held by Government regarding the EU approach. PLP agrees that AI should be defined according to its use, rather than by its form. At the roundtables we held in 2022, participants expressed widespread concern regarding the EU AI Act's attempt to place AI systems into discrete risk categories. Producing a satisfactory definition of 'high risk' is difficult because AI technologies are fast developing, and new emergent tools may fall outside of the definition, which will be slow to be amended. We therefore agree that regulating the uses of AI provides necessary flexibility in the Government's approach.

68. Accordingly, PLP recommends that the Government learn from the precautionary approach to public sector AI regulation that is being taken in the EU, whilst retaining an approach which targets the function and not the form of AI. We consider that further consideration must be given to new substantive obligations being placed on public authorities to promote adherence to the principles. For example, as we have set out in our answer to Q2, we have particular concern that public use of AI is not transparent and that there is no effective mechanism for ensuring transparency. Therefore, we have recommended a further statutory intervention to require that the Algorithmic Transparency Recording Standard is complied with and to improve the operational detail of information being released.

69. In considering how it might move to a more thorough approach to public sector regulation, the Government ought to consider alternatives to the draft EU AI Act. PLP urges the UK to learn lessons from positive examples of AI governance, particularly the French Loi Pour une Republique Numerique and the Canadian DADM. The details of these regimes were set out in response to Q2, but we now draw attention to three precautionary requirements the DADM places on operators of automated

[60] J.B. Wiener and M.D. Rogers *'Comparing precaution in the United States and Europe'* Journal of Risk Research 5 (4), 317–349 (2002) available at: https://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=1985&context=faculty_scholarship#:~:text=Today%2C%20the%20conventional%20wisdom%20is,%3B%20Levy%20and%20Newell%2C%202000%3B

[61] See further https://eur-lex.europa.eu/EN/legal-content/summary/the-precautionary-principle.html

[62] J. Tomlinson and J. Maxwell, '*Experiments in Automating Immigration Systems'* (2022).

[63] *UK Artificial Intelligence Regulation Impact Assessment* (23rd March 2023).

[64] *UK Artificial Intelligence Regulation Impact Assessment* (23rd March 2023) at paragraph 172 and paragraphs 187-192.

decision making in federal institutions.[65]

(a) Firstly, the DADM mandates the completion AIAs and, importantly, to release the results. We have set out how the UK might learn from this approach where we have argued for the ATRS to be made compulsory and placed on a statutory footing.

(b) Secondly, the DADM requires users to inform people when they are being subjected to wholly or partially automated decision making and to explain how such a decision will be made. In so doing, the Canadian Government aims to increase public confidence in AI use by explaining it. The Government has associated itself with this aim in the White Paper and might learn from this example.

(c) Thirdly, federal institutions are obliged by the DADM to test data and information used by the AI system for accidental bias and monitor the outcomes to highlight any such bias. If the Government considers fairness a priority in public sector AI use, then it must undertake to do the same.

70. There are also lessons to be learned from the deficiencies of the Canadian regime. It has a limited scope and does not apply to provincial governments or other non-federal government institutions. The UK Government aims to promote "collaboration between government, regulators and business." If it is to be successful, it must also promote collaboration within the public sector and promote a precautionary approach to AI use.

71. France's Law for a Digital Republic (2016) is also more precautionary than that set out in the UK White Paper. It requires that public use of AI be transparent and affords individuals a right to have algorithmic decision making explained.[66] This entails all public bodies to publish a list of the AI tools and their rules, even where the AI system is used to partially reach the decision.

72. PLP recommends that the Government go further in safeguarding the people whose lives are affected by, or will be affected by, public sector use of AI from the uncertain harms that could result from such use. The White Paper is right to identify that AI should be regulated according to its function and not its form, however, due consideration has not been given to how public sector use might differ from private sector use.

Question 9

Do you agree that the functions outlined in Box 3.1 would benefit our AI regulation framework if delivered centrally?

Question 10

What, if anything, is missing from the central functions?

73. We will address Qs 9 and 10 together.

74. PLP acknowledges the importance of central oversight. Without such central oversight, the regulation

[65] PLP carried out a detailed analysis of this regime in our submission to the House of Commons Science and Technology Committee and highlighted it as a positive example of AI governance in the public sector, available at: https://publiclawproject.org.uk/content/uploads/2023/01/PLP-submission-Science-and-Technology-Committee-Call-for-evidenceas-submitted.pdf (at paragraphs 28 to 31).

[66] Law No. 2016-321 of 7 October 2016 for a Digital Republic https://www.legifrance.gouv.fr/loda/id/JORFTEXT000033202746/

of AI will remain piecemeal, with different approaches existing and developing in different sectors. As identified in the White Paper, "conflicting or uncoordinated requirements from regulators… may leave risks unmitigated, harming public trust".[67] However, PLP believes that centralised AI regulation (with new legislative requirements as set out in response to Qs 7 and 8) would be more effective than the proposed model of central functions and devolved responsibility to existing regulators, and therefore does not agree with the premise of this question.

75. We do not believe that the revised cross-sector principles and central functions will effectively address the risk posed by AI used in public decision-making and administration. The White Paper itself acknowledges the limitations to the "current patchwork of regulation", [68] and we do not envisage the central functions to adequately address these gaps. On the contrary, centralised AI regulation would maximise the potential for regulatory coordination, allow for cross-cutting AI risks to be addressed and minimise the likelihood of AI systems falling through the gaps and operating without regulatory oversight.

76. The UK AI impact assessment, in setting out the four policy options considered, recognises that the new central functions (option 2) places only a "a clear *incentive* to engage" on regulators, and that existing regulatory remits may be "inadequate" and legislative measures may be necessary to "address regulatory gaps".[69] Yet the decision has been made to "[d]elegate to existing regulators with a duty to regard the principles, supported by central AI regulatory functions". This approach will not allow the AI regulation to achieve its stated aim for three reasons.

77. First, AI is already being used in ADM systems in central and local government. These systems have been in operation at varying levels for many years and continue to be rolled-out at pace in high-stakes contexts such as immigration, welfare and policing.[70] The central functions are described as a mechanism to "coordinate, monitor and adapt the framework as a whole",[71] and eventually "inform improvements to the framework by government, including legislative measures to address regulatory gaps".[72] Whilst we appreciate the challenging nature of the task and echo the desire to ensure regulation is future proofed, this iterative regulatory regime will allow for AI-related risks and harms to persist. Given the severity of the risks posed by AI in public authority decision-making, and the harm caused within the AI-related scandals in the Netherlands (SyRI), USA (Michigan Integrated Automated System) and Australia (Robodebt), we must see more urgent action taken sooner to ensure they do not go unaddressed.

78. Second, specific concerns arise about the central functions being carried out by the Government when applied to public sector use of AI and ADM systems. Where such systems are used by government and other public authorities, the question of how the central functions would be carried out impartially must be addressed. As expressed throughout this response, PLP is concerned that the White Paper has completely failed to address the need for a separate public sector focus in the regulation of AI which, in this instance, results in a lack of clarity in how the central functions would operate.

79. Third, if a specific and centralised regulator is not introduced, it is possible that the ICO's remit, as the

[67] DSIT (n 1), para 27, page 15.
[68] DSIT (n 1), para 70, page 42.
[69] Department for Science, Innovation & Technology, UK Artificial Intelligence Regulation Impact Assessment (23 March 2023), para 99, page 29 (emphasis added).
[70] Public Law Project, Tracking Automated Government 'TAG' Register (9 February 2023) http://trackautomatedgovernment.org.uk/.
[71] DSIT (n 1), para 71, page 43.
[72] Department for Science, Innovation & Technology, UK Artificial Intelligence Regulation Impact Assessment (23 March 2023), para 100, page 30.

UK's independent regulator for Data Protection and Freedom of Information will be expanded to include oversight of AI in line with the revised cross-sector principles proposed in the White Paper. We believe that short of an independent centralised AI regulator, the ICO is, amongst existing regulators, best placed to carry out this function if adequately resourced both financially and in terms of legal powers. However, the structure of the proposed central functions would mean the Government is responsible for reviewing the effectiveness of the ICO's regulation of government use of AI. It is a significant omission that the Government has not considered the circular nature of government self-assessing the regulation of its use of AI, and the implications for impartiality of the central functions. Yet, the White Paper does acknowledge that there "may be more value in a more independent delivery of the central functions", but again this is signalled to be contemplated at a later date.[73]

Question 17

Do you agree that our approach strikes the right balance between supporting AI innovation; addressing known, prioritised risks; and future-proofing the AI regulation framework?

80. As set out in the introduction to our consultation response, PLP believes that the innovation focus of the proposed AI regulation means the use of AI by public authorities has been overlooked. The stated aim of the White Paper is to address the risk of AI in a proportionate manner to ensure the continued growth and development of AI in the UK,[74] and we support this. However, we are concerned that the pro-innovation aim materialises as an unnecessary 'light-touch' approach to the regulation of public authority use of AI and ADM. For that reason, we do not think the appropriate balance has been struck.

81. Instead, we believe the AI regulation White Paper is a missed opportunity to implement effective regulation of AI use in the public sector, which would allow for individuals to feel the intended benefits of introducing automation into public administration whilst sparing them the unintended effects that have been realised elsewhere.

Question 18

Do you agree that our approach strikes the right balance between supporting AI innovation; addressing known, prioritised risks; and future-proofing the AI regulation framework?

82. PLP does not agree.

83. First, we disagree that existing regulators are best placed to apply the principles. As set out in our introduction and in response to Qs 5 and 6, existing regulators do not have the necessary statutory backing and resourcing to adequately apply the principles and regulate AI within their current remit. As articulated by the Ada Lovelace Institute, "[w]ithout new legislation, regulators cannot be obliged to

[73] DSIT (N 1), para 75, page 53.
[74] *UK Artificial Intelligence Regulation Impact Assessment* (23rd March 2023), paragraph 61.

enact the principles set out in the White Paper".[75] Lord Clement Jones echoed this sentiment stating that "the suggested form of governance of AI is a set of principles and exhortations which various regulators – with no lead regulator – are being asked to interpret in a range of sectors under the expectation that they will somehow join the dots between them. They will have no new enforcement powers".[76]

84. In the joint statement we convened, 'Key principles for an alternative AI White Paper',[77] PLP and 29 other individuals and civil society organisations, called for a specialist regulator to enforce the regulatory regime and ensure people can seek redress when things go wrong. This principle arose as a point of consensus during the roundtables we hosted in 2022, with attendees citing the lack of statutory powers and financial resources currently available to regulatory bodies. Given the specificity and complexity of this domain, if effective AI regulation is to be achieved through non-binding principles, an independent expert regulator is required. This regulator needs to be adequately resourced and given the right tools to enforce the regulatory regime, including powers to proactively audit public ADM tools and their operation.

85. Second, we do not agree that government is best placed to provide oversight and deliver central functions. As set out in response to Q10, the structure of the proposed central functions would mean the Government are responsible for reviewing the effectiveness of ICO's regulation of government use of AI. The proposed AI regulation could be implemented more effectively if delivered by a specialist regulator, that has oversight of the entire AI landscape and is independent from the procurement and operation of AI systems. This would allow for the regulatory cohesion desired by the 130+ organisations and individuals that responded to the 2022 policy paper, whilst also guaranteeing an independent approach to evaluating how the framework is working, where it is effective and where it may need improving.

Question L1

What challenges might arise when regulators apply the principles across different AI applications and systems? How could we address these challenges through our proposed AI regulatory framework?

86. PLP works across areas of the public sector, and thus is aware that different parts of the sector give rise to different problems. We reiterate the concerns of the Ada Lovelace Institute that the proposed regulatory framework "inherits gaps in existing UK regulation, and therefore risks leaving prominent AI harms unaddressed".[78] The White Paper itself acknowledges that some regulators have warned they may "lack the statutory basis to consider the application of the principles".[79] There is currently no specific regulator for general-purpose AI systems and the White Paper contains no new provisions for

[75] Ada Lovelace Institute, 'Policy briefing – the UK government's White Paper on AI regulation' (March 2023) https://www.adalovelaceinstitute.org/wp-content/uploads/2023/03/Briefing-on-AI-White-Paper-March-2023-Ada-Lovelace-Institute.pdf.
[76] Lord Clement Jones, 'The long-awaited AI Governance White Paper falls far short of what is needed' (3 April 2023) https://www.politicshome.com/thehouse/article/longawaited-ai-governance-white-paper-falls-far-short-needed.
[77] Key principles for an alternative AI White Paper (June 2023) https://publiclawproject.org.uk/content/uploads/2023/06/AI-alternative-white-paper-in-template.pdf.
[78] Ada Lovelace Institute, 'Policy briefing – the UK government's White Paper on AI regulation' (March 2023) https://www.adalovelaceinstitute.org/wp-content/uploads/2023/03/Briefing-on-AI-White-Paper-March-2023-Ada-Lovelace-Institute.pdf.
[79] DSIT (n 1), page 36.

them.

87. The most effective way to address these challenges would be to create a specialist regulator, that has oversight of the entire AI landscape and is able to dedicate resource and expertise to principles across different AI applications and systems.

Question L2.1

Do you agree that the implementation of our principles through existing legal frameworks will fairly and effectively allocate legal responsibility for AI across the life cycle?

88. PLP does not agree.

89. The existing legal framework governing the use of AI provides a patchwork of vital provisions that require ADM systems to be transparent and not to discriminate and/or breach other individual rights. However, we consider the governance of AI, which is necessary to ensure that the legal framework works in practice, operates less effectively. Current governance of AI in the UK does not provide an adequate level of transparency, accountability, or protection against unlawfulness, unfairness, and discrimination.

90. We are concerned that the implementation of the principles through these existing frameworks will not go far enough to ensure their achievement. This concern is made increasingly salient by the Government's watering down of data protection through the DPDI Bill. This is a particular concern for public sector use of AI, as the ICO's main statutory backing is derived from the DPA 2018 and UK GDPR.

91. Although fragmented, provisions under FOIA 2000, DPA 2018, UK GDPR, EA 2010, HRA 1998 and ECHR do provide a patchwork of vital, albeit imperfect, safeguards. If AI is to be regulated through non-statutory principles, the existing safeguards found in legal frameworks must be fortified.

Question L2.2

How could it be improved, if at all?

92. The clearest way to fairly and effectively allocate legal responsibility for the use of AI in the public sector would be to create a clear and cohesive legal framework that places a statutory duty on public authorities developing and operating AI to comply with transparency measures. Such measures should include a statutory duty on the public body to inform the person subject to the decision that ADM has been used, and how it is being used; mandatory publication of the tool on the ATRS; and a statutory duty to publish a risk assessment (including the data protection, equality, human and child rights impacts) of the tool and measures of impact post-deployment.

93. To secure accountability for AI-related harms there must be adequate statutory avenues to challenge them, together with effective enforcement mechanisms and the possibility of sanctions. Applying Mashaw's 'six inquiries' of accountability as set out in response to Q3 (paragraph 33), would require public authorities to be transparent about their use of AI and ADM systems to adequately take responsibility for any harms that occur from their operation. To allow for this harm to be mitigated, public authorities must be under the legal obligation to reveal the reasons behind decisions to a

selected counterpart (this could be the community as a whole). This must be a statutory entitlement on the part of the counterpart to request that the reasons are revealed. Finally, the accountable public authority must be 'sanctionable', which may range from public criticism to information or enforcement notices (similar to the powers held by the ICO) and legal remedies for the affected individuals. The authority to scrutinise compliance with the statutory requirement for transparency and accountability, and to take action where public authorities do not act in accordance, would be best placed with an independent specialist AI regulator.

94. Ultimately, legal responsibility would be much more effectively allocated and secured through the introduction of specific obligations in statute, that require adherence from public authorities developing, deploying, and operating AI. The legal framework should seek to bring coherence and clarity by building upon and working with existing data protection safeguards and our public law framework.

Contact

Mia Leslie
Research Fellow
m.leslie@publiclawproject.org.uk

Joseph Summers
Research Assistant
j.summers@publiclawproject.org.uk

Isabelle Agerbak
Policy and Parliamentary Lead
i.agerbak@publiclawproject.org.uk

Public Law Project is an independent national legal charity.

We are researchers, lawyers, trainers, and public law policy experts.

For over 30 years we have represented and supported individuals and communities who are marginalised through poverty, discrimination, or disadvantage when they have been affected by unlawful state decision-making.

Our vision is a world where the state acts fairly and lawfully. Our mission is to improve public decision making, empower people to understand and apply the law, and increase access to justice.

We deliver our mission through casework, research, policy advocacy, communications, and training, working collaboratively with colleagues across legal and civil society.

Public Law Project contributes and responds to consultations, policy proposals, and legislation to ensure public law remedies, access to justice, and the rule of law are not undermined.

We provide evidence to inquiries, reviews, statutory bodies, and parliamentary committees and we publish research and guides to increase understanding of public law.

Public Law Project's research and publications are available at:

www.publiclawproject.org.uk/resources-search/