THE USE OF ARTIFICIAL INTELLIGENCE TOOLS BY GOVERNMENT

A CASE STUDY OF THE HOME OFFICE'S ASYLUM PRACTICE

JOINT LEGAL OPINION FOR THE OPEN RIGHTS GROUP

CONFIDENTIAL & DRAFT

Introduction

1. The Open Rights Group (**ORG**) has commissioned this legal Opinion to explain how civil society actors can evaluate the use of Artificial Intelligence (**AI**) tools and systems by the Government, whether through dialogue or litigation. It is particularly concerned to prevent the deployment of AI in ways that risk violating the rights of migrants, refugees and asylum-seekers. The Government's use of AI in the process of determining refugee status has become a focus of concern.

2. It is well recognised in the academic literature that governments, both in the UK and elsewhere, are looking to AI systems to relieve some of the pressure on administrative systems and with the aim of improving and increasing their speed, efficiency and cost-effectiveness. These are obviously good aims in themselves, but they are becoming administratively essential in the face of ever-increasing numbers of individuals seeking asylum and an already substantial backlog of asylum cases awaiting decisions. The literature recognises both advantages and risks.[1]

3. The ORG considers that technology should be used to enhance, rather than diminish, the rights of society's most vulnerable. The ORG recognises that the Government use of AI may offer benefits including reducing costs. However, a priority for the ORG is to ensure that the Government does not use AI in a way which endangers the rights and welfare of individuals. In its view, cost-cutting should never justify the dilution of basic legal protections.

4. To provide a specific focus for these concerns, this Opinion addresses the use of two AI tools: the Home Office's Asylum Case Summarisation (**ACS**) tool which was intended to go 'live' in January 2026 and the complementary Asylum Policy Search (**APS**) tool which has been in place since at least November 2025.[2]

[1] See by way of example Bailey, E., Given-Wilson, Z. and Memon, A., 2025. Perceived credibility of asylum claimants: the role of decision-maker affect and asylum seeker's emotions *Psychology, Crime & Law*, pp.1-19, and Judijanto, L., 2025. Asylum Processing Algorithms and Epistemic Violence: A Review of AI's Role in Refugee Status Determination *Journal of Information, Technology and Policy*, pp.1-11.

[2] This information was provided on 25 November 2025 when Alex Norris (Labour MP) answered a question in Parliament about the ACS and APS as recorded here.

5. The Home Office's rationale for using these generative AI tools is to '*aid efficiency*'[3] yet, as we develop in this Opinion, the publicly available information calls into question their accuracy.

6. It is understood that these are not the only ways in which the Home Office uses AI systems in respect of its management of asylum cases. While this Opinion does not seek to offer a comprehensive analysis of all such use by the Home Office in respect of asylum management, it is hoped that its analysis of the extent to which the ACS and APS tools comply with the recognised principles for the best legal use of AI systems can provide a good pointer to the way in which other systems can be assessed.[4]

7. This Opinion's analysis will start with the overall legal framework currently in existence for AI use. Two points about this framework must be noted: first, though the Government has not yet introduced legislation to regulate the use of AI, it has introduced various detailed policies concerning its use, which at points are prescriptive in nature; secondly, there are existing legal protections, which though not AI-specific, operate to protect the welfare and rights of people who are subject to decisions shaped or supported by AI technology; and indeed there has already been some high-profile litigation testing how effectively existing laws regulate the state's use of AI.[5]

[3] 'Research and analysis: Evaluation of AI trials in the asylum decision making process', 29 April 2025. Referred to later in this Opinion as **'the Home Office Report'**.

[4] The Public Law Project operates a register called 'Tracking Automated Government (TAG) Register' which reveals many uses of technology by the Home Office in relation to asylum: https://trackautomatedgovernment.shinyapps.io/register/. It must also be noted in this respect that the Minister of State for Border Security and Asylum recently announced that, subject to further testing, AI will be used to estimate the age of asylum-seekers: see Statement made to Parliament on 22 July 2025: https://questions-statements.parliament.uk/written-statements/detail/2025-07-22/hcws885. Part of the rationale for using AI was that it will be '*the most cost-effective option*'.

[5] For example, *Glukhin v Russia* (2024) 78 EHRR 6 and *R (on the application of Bridges) v The Chief Constable of South Wales Police and others* [2019] EWHC 2341, and *R (on the application of Bridges) v The Chief Constable of South Wales Police and others* [2020] EWCA Civ 1058.

8. A recent development, which in our view is underappreciated in the domestic discourse, is consensus among international legal instruments of 11 common core principles *directly concerned* with the ethical and legal use of AI systems. These may be summarised under the titles: 'Democracy', 'Fairness, equality & non-discrimination', 'Human dignity & autonomy', 'Respect for human rights', 'Privacy & data governance', 'Sustainability', 'Robustness & digital security', 'Safety & reliability', 'Transparency & explainability', 'Accountability & responsibility', and 'Contestability, oversight & redress'.[6] We discuss their implications in this paper, and we refer to them collectively as '**the AI Ethical Principles**'.

9. One key point is that, while not incorporated into domestic legislation, these AI Ethical Principles (save for 'Democracy' and 'Sustainability') have been recognised by the Government Digital Service in the recently published AI Playbook for the UK Government ('**the UK AI Playbook**') as playing an essential role in securing good governance concerning the deployment and use of AI systems.[7]

10. Feryal Clark MP, then Parliamentary Under-Secretary of State for AI and Digital Government in the Department for Science, Innovation and Technology (**DSIT**), made the Government's commitment to them in her introduction to the Playbook:

> *'…The AI Playbook will support the public sector in better understanding what AI can and cannot do, and how to mitigate the risks it brings. It will help ensure that AI technologies are deployed in responsible and beneficial ways, safeguarding the security, wellbeing, and trust of the public we serve…'.*

11. There can be no doubt that this Playbook has been published with the intention that it will provide practical guidance to Government generally, and more specifically a series of benchmarks against which Governmental use of AI can be judged. The key directive statements are summarised at Annex B. This Opinion discusses the Playbook's implications and how the AI Ethical Principles provide the basic template by which civil society should engage with Government's use of AI. To demonstrate

[6] Annex A contains a Table indicating where in the International Agreements these principles can be found.

[7] It was first published on 10 February 2025, and it has been updated. The latest version is available here.

what can be done, this Opinion will focus on the ACS and APS tools by way of a 'case study' to illustrate how such a critique can be deployed to protect society's most vulnerable.

CONFIDENTIAL & DRAFT

Executive summary

Part I: The need to regulate AI systems that affect individual rights

12. Changes in technology have frequently caused seismic shifts in society. AI is no different. However, there are unique features to AI systems which mean that they pose greater and different risks in comparison to other technological advances. The way AI systems work is often opaque ('the Blackbox problem'); they can exacerbate power imbalances; they often process vast amounts of personal data giving rise to privacy risks and concerns about data security; they can 'hallucinate' (i.e. completely fabricate information); and sometimes they provide inaccurate, abusive or discriminatory outputs. It is obvious that where the rights of individuals are concerned such risks must be taken very seriously, and that the more vulnerable the individual and impactful the decision, the more care must be taken.

Part II: The AI Ethical Principles

13. There has been an intense global discourse about how to regulate AI in a way that harnesses its promised benefits, whilst also managing its inherent risks. From this has emerged a set of common ethical principles about the use of AI contained in international instruments made by the United Nations Educational, Scientific and Cultural Organization (**UNESCO**),[8] the Organisation for Economic Co-operation and Development (**OECD**),[9] and the Council of Europe.[10]

[8] See UNESCO, 2022. *Recommendation on the ethics of artificial intelligence*. United Nations Educational, Scientific and Cultural Organization:

https://unesdoc.unesco.org/ark:/48223/pf0000381137/PDF/381137eng.pdf.multi

[9] See OECD, Recommendation of the Council on Artificial Intelligence https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

[10] See Council of Europe, Framework Convention on Artificial Intelligence

https://rm.coe.int/1680afae3c

14. The United Kingdom is a member of UNESCO, the OECD and the Council of Europe, and is a signatory to the Council of Europe's Framework Convention on Artificial Intelligence. Whilst each of these instruments uses slightly different language, the common principles that run through them relate to:

- 'Democracy'
- 'Fairness, equality and non-discrimination'
- 'Human dignity and autonomy'
- 'Respect for human rights'
- 'Privacy and data governance'
- 'Sustainability'
- 'Robustness and digital security'
- 'Safety and reliability'
- 'Transparency and explainability'
- 'Accountability and responsibility'
- 'Contestability, oversight and redress'.

15. Each of these address specific and well known problems inherent in AI systems. We refer to them in this Opinion as **'the AI Ethical Principles'**.

16. While the UK Government has not directly transposed these international instruments into domestic law it has stated its commitment to them in the AI Playbook for the UK Government (**'the UK AI Playbook'**) which distils the principles to be followed by the Government when it deploys AI.

17. Annex A contains a table which maps the relationship between the UK AI Playbook and international instruments concerning these AI Ethical Principles.

18. In tandem with its adoption of the AI Ethical Principles, the UK AI Playbook contains a detailed step-by-step procedure for decisions as to the deployment, use, risk assurance and review of AI systems. These steps are also called 'Principles' although they more closely resemble procedural steps. The following Principles (or procedural steps) in the UK AI Playbook are particularly relevant to this Opinion and the Home Office's use of AI:

- Principle 1: You know what AI is and what its limitations are;
- Principle 2: You use AI lawfully, ethically and responsibly;
- Principle 4: You have meaningful human control at the right stage;
- Principle 6: You use the right tools for the job;

- Principle 7: You are open and collaborative; and
- Principle 10: (a) You use these principles alongside your organisation's policies and (b) have the right assurance in place.[11]

19. The procedural steps contained in the Playbook are detailed and clearly expressed. They contain important prescriptive statements about what Government must do as summarised in Annex B. They link directly with the AI Ethical Principles. In short, the UK AI Playbook sets out the Government's own standards which departments and public sector organisations are expected to follow in relation to their use of AI, and civil society can use the UK AI Principles to engage with the Government. As discussed below, the UK AI Principles will also be relevant to the interpretation and application of domestic legislation and common law principles.

Part III: Domestic legislation, legal principles and their relationship to the UK AI Playbook

20. In parallel with the UK AI Playbook, there are domestic legislation and legal principles, which, whilst not specifically mentioning AI, must be considered whenever the Government uses AI because of their field of application. These include the Equality Act 2010, the Human Rights Act 1998, data protection rules and general public law principles. We refer to this as the '**Domestic legal framework'**. The UK AI Playbook will be relevant when considering this legal framework in so far as the Government is deploying AI.

21. The UK AI Playbook also recognises that this wider framework of laws and principles must be applied. In Principle 2 it instructs those acting on behalf of the Government to '*...use AI lawfully, ethically and responsibly*'.

22. There is thus a rich tapestry of controls that the law places on the Government's use of AI and that are available for civil society when engaging with the Government, whether through dialogue or litigation. If deployed effectively, these existing legal protections, along with the AI Ethical Principles, as contained in the UK AI Playbook, can regulate the Government's AI use to protect the welfare and rights of people who are subject to decisions shaped or supported by technology. We analyse how this might happen in a case study in Part IV.

[11] Principle 10 is not divided into (a) and (b) within the UK AI Playbook. We consider this a mistake. Principle 10 includes two distinct notions, AI assurance being hugely significant and a key theme of this Opinion.

Part IV: Case Study: The Home Office's use of AI during the refugee status determination process

23. The law requires those taking decisions as to whether an asylum-seeker has refugee status ('**Decision-Makers**') to look at all material facts to determine whether an asylum-seeker has a '*well-founded*' fear of persecution in their own country. This is a subjective test with objective elements. Decision-Makers must examine (1) the information provided by the applicant during interviews and (2) information concerning the situation in countries from which asylum-seekers have fled.

24. The two tools discussed in this Opinion relate to these two elements: **ACS** is a generative AI tool that summarises information provided by applicants for Decision-Makers; **APC** is a generative AI tool that searches country information for Decision-Makers. The important point is that both AI tools create *new text* for the Decision-Maker to consider rather than simply indexing or organising the existing source information ('**the Source Material**'). In this way, they funnel, filter and regurgitate important facts which are material to the Decision-Maker's legal obligations when determining refugee status. They may 'filter out' crucial information. The output of the APC and APS is not shared with the asylum-seeker. In fact, we understand that they are not even informed that AI is going to be used for their application.

25. We also note that, from the available information, it appears there is a significant risk of inaccuracy in the output generated by the ACS and APS; during the pilot for the ACS, there were inaccurate summaries 9% of the time and 5% of users of the APS were '*not confident in tool accuracy*'.

26. This is concerning, but additionally we note that there is no detailed information available as to how the level of accuracy was measured or evaluated in relation to the ACS; it appears that some form of objective quantitative data must have been measured since the 9% metric has been published. Whether this inaccuracy arose from the Large Language Model (**LLM**) 'hallucinating', or whether it was simply an error in summarisation, is entirely unknown. The assessment of accuracy for the APS appears to have been limited to an assessment of confidence rather than an objective measure of inaccuracy.

27. Our analysis of these two tools in Part IV, based on the publicly available information, leads us to conclude that (1) in significant respects the UK AI Playbook is not being followed or (2) there are serious concerns that this is so.

28. Our specific observations on compliance with the AI Playbook’s principles, which we see as procedural steps, are summarised in the table below.

Principles 1, 4 & 10b: ‘You know what AI is and what its limitations are’, ‘You have meaningful human control at the right stage’ & ‘You have the right assurance in place’		
1	There has been a failure to assess quantitatively the extent to which the APS produces inaccurate outputs.	Paras 98 - 99
2	Whilst there has been some form of quantitative assessment of the ACS to assess the accuracy of outputs; there is insufficient information about what accuracy means in this context, the extent of inaccuracy and what benchmarking the Government is using (i.e. what ‘good looks like’).	Paras 100 - 104
3	There is no ability to cross-reference the summarised output from the ACS (and we assume the APS) to the original Source Material making it difficult for Decision-Makers to assess/verify accuracy. This is an important missing procedural safeguard.	Para 105
4	There is no system in place to allow asylum-seekers to check the output of the ACS or APS for accuracy since they do not know it happened and have no opportunity to see the text generated. This is an important missing procedural safeguard.	Para 106

5	There is a risk that the summaries produced by the ACS and APS will become part of the decision-making process since they filter and funnel information which could be relied on by the Decision-Makers at the expense of the Source Material. There is no clear guidance, that we can see, which tells Decision-Makers that they must fully consider all original Source Material. This is an important missing procedural safeguard. It also means that we cannot be satisfied that there are procedures in place to ensure meaningful human control.	Paras 107 - 111 & 121
5	We are not satisfied that there are adequate technical measures in place and auditing processes to ensure that Decision-Makers consider all original Source Material. This is an important missing procedural safeguard. It also means that we cannot be satisfied that there are procedures in place to ensure meaningful human control.	Paras 107 - 111
6	We are not satisfied that there has been an adequate assessment of whether the quality of the decisions is negatively affected by the ACS or APS. Given that the Decision-Makers' task is to assess relevant/material facts, our view is that quality – in this context – should measure the extent to which the ACS and APS disregarded or correctly identified/summarised that information. In other words, the Government has not properly approached AI Assurance.	Para 112
7	It is unacceptable that the ACS and APS have been rolled out despite only limited bias testing having been undertaken. In other words, the Government has not properly approached AI Assurance.	Para 113

Principle 6: 'You use the right tools for the job'

8	Despite the risks associated with generative AI, there is no publicly available information that the Government considered whether an analogue or non-AI solution could achieve the aim of greater efficiency.	Paras 114 - 120

Principle 7: 'You are open and collaborative'		
9	To the ORG's knowledge, the Home Office has not engaged with civil society in relation to the ACS and APS.	Para 122
10	There is no published Data Protection Impact Assessment or Equality Impact Assessment. The 'prompt' used in the ACS tool is unknown (for no obvious good reason). Asylum-seekers are not told about the use of the ACS tool. Technical information such as the training data source has yet to be published. Neither the ACS nor APS is listed in the repository for the Algorithmic Transparency Recording Standard.	Paras 123 - 132
11	There is no information about the AI Assurance that has happened since the pilot scheme.	Para 133

29. As to the existing domestic legal framework, which is incorporated into Principle 2, we conclude that:

Article 3

1	Asylum-seekers will often be seeking refuge in the UK as a means of escaping torture or inhumane and degrading treatment such that their Article 3 rights are engaged. If the Home Office adopts a practice of using generative AI tools which create a risk that Decision-Makers will consider inaccurate information and/or overlook relevant/material fact facts when determining people's asylum claims, and those risks are not mitigated by appropriate safeguards, such an assessment would be unlikely to have the necessary rigour to comply with the UK's procedural obligations under Article 3.	Paras 136 to 140

	Public law	
2	The Home Office is under a heightened *Tameside* duty of enquiry as regards the accuracy and functionality of the ACS and APS. The Home Office will be at significant risk of breaching its *Tameside* duty if it fails to undertake adequate assessments in respect of and monitor the accuracy of the APS and the ACS, the extent to which they impact the quality of asylum decisions, the risk of bias and discrimination posed by the AI tools, and the alternatives to the AI tools to achieve the Home Office's aims.	Paras 144 to 146

3	The use of the ACS and the APS give rise to a significant risk of process irrationality. If a Decision-Maker relies upon the ACS and APS summaries at the expense of a full examination of the Source Material, and those summaries have filtered out relevant information regarding the country of origin or asylum-seeker's interview, there will be a significant risk that the Decision-Maker will have failed to take relevant considerations and evidence into account when determining the asylum claim in question.	Paras 147 to 149
4	Given the apparent inaccuracies in the summaries generated by the APS and ACS, there is a significant risk that decisions which are based on those summaries will be based upon and vitiated by material errors of fact.	Para 150
5	As a matter of procedural fairness, we consider that asylum--seekers have a right to be informed that AI is being used in the determination of their claims, how it is being used, and to be provided with the output of the AI-generated summaries. That asylum-seekers appear not to have been so informed is likely to be unlawful.	Para 151

	Data protection	
6	There will be a breach if the ACS produces inaccurate summaries of personal data, if there is no explanation to the asylum-seekers that the AI tool will be used and / or they are denied access to the output from the ACS to correct any errors.	Para 153

Public Sector Equality Duty (PSED)		
7	The Government has not published an Equality Impact Assessment, so we cannot be satisfied that the PSED has been met and that there are no broader equality issues.	Paras 154 to 156

Regulators		
8	The Independent Chief Inspector of Borders and Immigration (ICIBI) should examine the way in which the Home Office is using AI. The Government has made it plain that it expects regulators to implement AI ethical principles in their work. We are not aware of any scrutiny of the ACS or APS by regulators.	Paras 157 to 159

30. The analysis in Part IV illustrates both why civil society has a role in holding Government to account and the role it can play in doing so.

Part I: The need to regulate AI systems

31. It has become increasingly obvious that powerful AI systems will have far-reaching effects on society. While examples of the benefits from such systems abound, fears have also heightened about the kinds of adverse effects that could occur from such rapid technological change.

32. These fears are driven by the unique features of AI systems that mean they pose greater and different risks in comparison to other technological advances.

33. One key feature is that AI systems have a **high degree of opacity** because there is significant difficulty in understanding the internal working of a system and the basis for its outputs (the so called 'Blackbox problem'). This opacity arises primarily from the machine learning[12] process by which AI tools are generally created, which is so complex and sophisticated that the tool can be opaque even to its creator.[13] For the public, this opacity can be exacerbated by the desire of the system's developers and deployers to keep commercial information secret and/or their fear that transparency could enable users of AI (or those subject to AI) to 'game the system'.[14]

34. AI can exacerbate **power imbalances**. The marriage of vast data and computing power means that AI can be used to perform intrusive tasks that would be practically

[12] The International Standards Organisation (**ISO**) has stated: '*Establishing a clear machine learning definition can be challenging. Machine learning (ML) is a type of artificial intelligence that allows machines to learn from data without being explicitly programmed. It does this by optimizing model parameters (i.e. internal variables) through calculations, such that the model's behaviour reflects the data or experience. The learning algorithm then continuously updates the parameter values as learning progresses, enabling the ML model to learn and make predictions or decisions based on data science…*'

[13] '*They arrive inscrutable to us, the model's developers. This means that we don't understand how models do most of the things they do*'; 'Tracing the thoughts of a large language model', Anthropic, 27 March 2025.

[14] Burrell, J. 2016. How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, *3*(1). https://doi.org/10.1177/2053951715622512 (Original work published 2016)

impossible for humans alone.[15] A powerful example of this power imbalance is the deployment of live facial recognition technology by state bodies enabling them to identify individuals within crowds on a scale never before possible.[16]

35. It is obvious too that the vast amounts of personal data that can be processed by AI tools gives rise to **privacy risks**.[17] There are also concerns about **data security** when so much personal data, which is sometimes sensitive, is processed or stored within AI tools.[18]

36. The occurrence of **‘hallucinations’** where information is fabricated in a way that is often plausible and hence difficult to spot is another concerning feature of some AI. Generative AI[19], like Large Language Models (**LLMs**), is prone to ‘hallucinate’. LLMs are AI models that predict words in a way that creates believable human language.

[15] Annual report of the United Nations High Commissioner for Human Rights and reports of the Office of the High Commissioner and the Secretary-General,

‘The right to privacy in the digital age’, 2022; available here.

[16] *Glukhin v Russia* (2024) 78 EHRR 6, para 35 quotes the report of the United Nations High Commissioner for Human Rights of 24 June 2020 entitled ‘Impact of new technologies on the promotion and protection of human rights in the context of assemblies, including peaceful protests’ (UN Doc. A/HRC/44/24) which stated: ‘*Assemblies traditionally have allowed participants a certain level of protection against being singled out or identified. This protection was already considerably weakened by many States that routinely made audiovisual recordings of assembly participants. The rise of facial recognition technology has led to a paradigm shift in comparison with practices of audiovisual recordings, as it dramatically increases the capacity to identify all or many participants in an assembly in an automated fashion*.’ (para 34).

[17] *R (on the application of Bridges) v The Chief Constable of South Wales Police and others* [2020] EWCA Civ 1058. This is a Court of Appeal case which recognised that live facial recognition technology, when deployed by the police, engages the right to privacy. At para 23, the Court observed: ‘*Facial biometrics bear some similarity to fingerprints because both can be captured without the need for any form of intimate sampling and both concern a part of the body that is generally visible to the public. A significant difference, however, is that AFR technology enables facial biometrics to be procured without requiring the co-operation or knowledge of the subject or the use of force, and can be obtained on a mass scale*.’

[18] Information Commissioner’s Office, ‘*How should we assess security and data minimisation in AI?*’, available here.

ChatGPT is a ubiquitous example but there are also an increasing number of others, such as Copilot, Gemini, Llama, Claude and Grok. The Government understands LLMs can fabricate information, stating in its 'Guidance to civil servants on use of generative AI'[20]:

> A generative AI tool, such as an LLM, will answer your question by probabilistically choosing words from a series of options it classifies as plausible. These tools cannot understand context or bias. Always treat with caution the outputs these tools produce and challenge the outputs using your own judgement and knowledge.

37. AI can also lead to **inaccurate**, **discriminatory** or **abusive** outputs. The recent controversy over the use of AI to 'nudify' women is a powerful illustration of this capability.[21]

38. Cumulatively, these attributes of AI systems make it difficult to develop, deploy, and use them responsibly. Fears are naturally greatest where there is no adequate framework to protect society and fundamental human rights.

39. It should be plain that *if* AI systems cannot be completely trusted, the technology should not be used to its fullest extent. That is why it has been widely recognised in the United Kingdom, as indeed it has in the European Union and many other

[19] A type of AI system that can create a wide variety of data, such as images, videos, audio, text, and 3D models. The technical definition adopted by NIST, which is a US standards body, is: '*The class of AI models that emulate the structure and characteristics of input data in order to generate derived synthetic content*.'

[20] Guidance to civil servants on use of generative AI, 29 January 2024, available online here. This was superseded by Generative AI Framework for HMG, which in turn was withdrawn on the publication of the AI Playbook.

[21] Ofcom launched an investigation into Grok in January 2026. Thereafter, X implemented measures to prevent Grok from being used to create intimate images of people. Ofcom's investigation remains ongoing. The press release is available here; see also the update here and news of the ICO's investigation here.

countries, that there must be a mechanism of assurance to establish what is an appropriate level of trust and then to maintain it.

40. The UK Government has not introduced specific legislation to regulate AI systems. However, and significantly, the Government has adopted - on an international and domestic basis – a set of AI Ethical Principles which apply exclusively to AI tools. These principles are intended to manage the specific risks that arise from AI systems. We will next consider how the AI Ethical Principles apply to Government use of AI.

CONFIDENTIAL & DRAFT

Part II: The AI Ethical Principles

41. There has been a long international discussion concerning the regulation of AI. One major take away from this discussion is the development of a set of commonly agreed ethical principles, which are set out in several important international instruments. In this Opinion, we refer to these as the **AI Ethical Principles**. In the next paragraphs we shall start by providing an overview of these AI Ethical Principles, before introducing the international instruments and the AI UK Playbook that reflects them.

Overview of the AI Ethical Principles

42. The common threads to various international instruments which are intended to manage the risks that arise from AI systems can be summarised as follows:

Democracy	This principle asserts that AI should not be used to manipulate people's autonomy or events to undermine democracy or the rule of law. This includes preventing the use of AI to create or spread misinformation.
Fairness	This principle has several facets but at its heart is the idea that AI should not be used to wrongly disadvantage or undermine individuals or groups. AI can create significant power imbalances especially when used by the state or influential private organisations. Fairness dictates that AI should not be used oppressively or to exacerbate power imbalances.

Equality and non-discrimination	This principle recognises that AI should not treat people less favourably in comparable circumstances because of their protected characteristics, for example, their sex or race or status as a disabled person. It also dictates that protected groups should not be indirectly disadvantaged by the apparently neutral use of technology, unless it can be justified as being a proportionate means of achieving a legitimate aim.
Human dignity and autonomy	This principle notes that AI should not undermine the fundamentals of humanity. It should not be used to manipulate or control. It should not infantilise or demean.
Respect for human rights	This principle records that AI should not undermine basic human rights such as the right to privacy.
Privacy and data governance	This principle recognises that vast amounts of data, often personal data, can be processed when creating AI. Further, personal data can be used when AI tools are then used. The data should be used appropriately and proportionately and subject to stringent safeguards.
Sustainability, robustness and digital security	This principle concerns the safe and secure use of data. It engages with concerns that AI can have a negative impact on the environment, society and the economy. It requires that AI be used and deployed sensibly.
Safety and reliability	This principle demands that AI is used in a way which is accurate. Since AI can be used to make decisions or support decision-makers, the system needs to be robust.

Transparency and explainability	This principle requires that the entire life cycle of an AI model should be as observable and understandable as possible. This means that those subject to AI can understand what data is used, how it is used, what is done to that data, how the outputs are generated, and what steps are taken to ensure that the outputs are accurate and non-discriminatory.
Accountability and responsibility	This principle asserts that people and organisations that use AI must be responsible for its use. Simply because a technology is clever or difficult to understand does not mean that the humans that create it, buy it and deploy it are not liable for when it goes wrong.
Contestability, oversight and redress	This principle requires that people and organisations that use AI must have in place AI Assurance and related mechanisms. AI Assurance is a key concept and must be understood by civil society if it is to hold the Government to account. It is described in more detail at 96 below.

43. This Table summarise the descriptions found on these themes in the international instruments. We have provided a description of where each can be found in those instruments in Annex A. Under the next cross-heading we describe these instruments and the degree to which the United Kingdom has been involved in their agreement and its future intentions with respect to them.

The OECD Recommendation

44. The first international standard on AI was published in May 2019 when the OECD adopted the Recommendation of the Council on Artificial Intelligence (**‘the OECD Recommendation**’) as a legal instrument.[22] It was subsequently amended on 3 May

[22] It is available online here.

2024. The United Kingdom, and many other countries such as the United States, are '*adherent members*' to the Recommendation and the Government is an active member of the OECD working party on AI governance which supports the implementation of the OECD principles.[23]

45. The Recommendation contains five principles. They are 'Inclusive growth, sustainable development and well-being' (Article 1.1), 'Respect for the rule of law, human rights and democratic values, including fairness and privacy' (Article 1.2), 'Transparency and explainability' (Article 1.3), 'Robustness, security and safety' (Article 1.4) and 'Accountability' (Article 1.5). These specific principles appear in Section 1 entitled 'Principles for responsible stewardship of trustworthy AI'. This title reflects the importance of human-centric development of AI and its beneficial deployment.

46. Each of the specific principles are directed to '*actors*' i.e. those who play an active role in the AI system lifecycle, including organisations and individuals that deploy or operate AI. It recognises the capacity of AI systems to undermine fundamental rights and requires that such actors should avoid such outcomes by assessing the risks of misuse inherent in any proposed system and develop safeguards and oversight mechanisms.

The UNESCO Recommendation

47. On 24 November 2021, UNESCO adopted the Recommendation on the Ethics of Artificial Intelligence, which it described as '*the first-ever global standard on AI ethics*' ('**the UNESCO Recommendation**').[24] It is applicable to all 194 member states of UNESCO, and was adopted by the United Kingdom.[25] The principles are (1)

[23] On 22 May 2024, the Under-Secretary of State for Science, Innovation and Technology told Parliament that '*At the OECD, we are an active member of the working party on AI governance, which supports the implementation of the OECD's AI principles. It enables the exchange of experience from best practice to advance the responsible stewardship of AI*', HC Hansard col. 127, see here.

[24] It is available online here.

[25] On 22 May 2024, the Under-Secretary of State for Science, Innovation and Technology said in Parliament: '*Let me turn my attention to UNESCO. The UK was actively involved in the development of its recommendation on the ethics of AI, and UK organisations such as the Alan Turing Institute have supported the development of implementation tools. As we have heard, we, along with all 192 other UNESCO member states, adopted the recommendations in November 2021, demonstrating our*

‘Proportionality and do no harm’, (2) ‘Safety and security’, (3) ‘Fairness and non-discrimination’, (4) ‘Sustainability’, (5) ‘Rights to privacy and data protection’, (6) ‘Human oversight and determination’, (7) ‘Transparency and explainability’, (8) ‘Responsibility and accountability’, (9) ‘Awareness and literacy’ and (10) ‘Multi-stakeholder and adaptive governance and collaboration’.

48. Of particular relevance to this Opinion are the comments within the UNESCO Recommendation concerning ‘transparency’:

> People should be fully informed when a decision is informed by or is made on the basis of AI algorithms, including when it affects their safety or human rights, and in those circumstances should have the opportunity to request explanatory information from the relevant AI actor or public sector institutions. In addition, individuals should be able to access the reasons for a decision affecting their rights and freedoms, and have the option of making submissions to a designated staff member …

49. This text highlights an important theme of the AI Ethical Principles that people subject to decision-making using AI should know about it and be able to make appropriate representations.

The Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law

50. The UK has long been a member of the Council of Europe (which is distinct from the European Union), which has also developed a Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (‘**the AI Convention**’). The AI Convention was made on 5 September 2024[26] when it was opened for signature and immediately signed by the UK[27] along with other states, although it has yet to be ratified by the UK.

commitment to developing a globally compatible system of responsible and ethical AI governance’, HC Hansard col. 123, see here.

[26] The current version to which this Opinion refers is available online here.

51. The preamble acknowledges AI's risk to human rights. Article 4 mandates states to ensure AI activities align with human rights obligations. Article 5 is 'Integrity of democratic processes and respect for the rule of law', Articles 6 to 12 outline principles for AI implementation, such as 'Human dignity and individual autonomy', 'Transparency and oversight', 'Accountability and responsibility', 'Equality and non-discrimination', 'Privacy and personal data protection', and 'Reliability'. Article 15 relates to 'Procedural safeguards' and provides that 'where an artificial intelligence system significantly impacts upon the enjoyment of human rights', states must ensure that 'effective procedural guarantees, safeguards and rights' are available to persons so affected. Article 16 relates to AI Assurance and requires states to 'adopt or maintain measures for the identification, assessment, prevention and mitigation of risks posed by artificial intelligence systems by considering actual and potential impacts to human rights, democracy and the rule of law'.

The UK AI Playbook

52. On 10 February 2025, the Government Digital Service published the UK AI Playbook, and this has been subsequently updated. The UK AI Playbook '*offers guidance on using AI safely, effectively and securely for civil servants and people working in government organisations*'.[28] The UK AI Playbook was produced after extensive internal consultation with the major United Kingdom Government departments[29]

[27] The UK Government press release is here.

[28] Its official title is 'Guidance: AI Playbook for the UK Government'; it has been subsequently updated, see the Update Log on p.6. The current version to which this Opinion refers is available online and is available here.

[29] Crown Commercial Service; Cabinet Office including the Number 10 Data Science and i.AI teams; Department for Business and Trade; Department for Education; Department for Environment, Food and Rural Affairs; Department for Energy Security and Net Zero Department of Health and Social Care; Department for Levelling Up, Housing and Communities; Department for Science, Innovation and Technology including the Government Office for Science and the Responsible Technology Adoption Unit; Department for Transport; Department for Work and Pensions; Foreign, Commonwealth and Development Office; Government Legal Department; HM Land Registry; HM Revenue and Customs; HM Treasury; Home Office; Ministry of Defence; and Ministry of Justice.

including the Home Office, arm's length bodies, devolved administrations and public sector bodies,[30] industry and academic institutes.

53. The UK AI Playbook is intended to help Government departments and public sector organisations harness the power of a wider range of AI technologies safely, effectively, and responsibly.[31]

54. The Playbook adopts a directive style instructing Government departments and public sector organisations on the procedural *steps* and standards that should be applied with deploying and using AI systems:

Principle 1	You know what AI is and what its limitations are
Principle 2	You use AI lawfully, ethically and responsibly
Principle 3	You know how to use AI securely
Principle 4	You have meaningful human control at the right stage
Principle 5	You understand how to manage the AI life cycle
Principle 6	You use the right tools for the job
Principle 7	You are open and collaborative
Principle 8	You work with commercial colleagues from the start
Principle 9	You have the skills and expertise needed to implement and use AI
Principle 10	(a) You use these principles alongside your organisation's policies and (b) have the right assurance in place[32]

[30] Driver and Vehicle Licensing Agency; Government Communications Headquarters; Government Internal Audit Agency; HM Courts and Tribunals Service; Information Commissioner's Office; Met Office; National Health Service; Office for National Statistics and the Scottish Government.

[31] See 'Foreword'.

[32] Principle 10 is not divided into (a) and (b) within the UK AI Playbook. We consider this a mistake. Principle 10 includes two distinct notions, AI assurance being hugely significant and a key theme of this Opinion.

55. The Scottish Government has also developed an online Scottish AI Playbook first published in March 2022 and subsequently amended.[33] The UK AI Playbook is more extensive and develops a wider range of principles hence our focus on it here.

56. The most relevant principles to this Opinion are the following:

Principle 1	You know what AI is and what its limitations are
Principle 2	You use AI lawfully, ethically and responsibly
Principle 4	You have meaningful human control at the right stage
Principle 6	You use the right tools for the job
Principle 7	You are open and collaborative
Principle 10 (b)	You have the right assurance in place

57. Within these principles there are many prescriptive rules dictating how the Government should use AI. These are set out in Annex B to make this Opinion readable. They should be carefully consulted by civil society as they are an important means by which civil society can engage with and hold the Government to account regarding its use of AI.

58. Principle 2, in so far as it requires the Government to 'use AI lawfully', merely reiterates the obvious requirement on the Government to act in accordance with the law. In Part III Domestic legislation, legal principles and their relationship to the UK AI Playbook, we further consider the ways in which the AI Ethical Principles in the UK AI Playbook may inform the interpretation and application of various legal principles including – importantly – the application of public law principles.

[33] Available here.

Part III: Domestic legislation, legal principles and their relationship to the UK AI Playbook

59. There is domestic legislation which will govern AI, for example, data protection laws, the Equality Act 2010, the Human Rights Act 1998 along with sector specific legislation such as the Employment Rights Act 1996.

60. This Opinion cannot comprehensively outline this legal framework. However, we are clear that civil society should be broadly aware of this framework, which we outline below with the aim that civil society can quickly identify potential 'touch points', when the Government deploys AI in the public sphere, where greater legal analysis is required.

61. **Data protection rights** are engaged whenever personal data is processed by an AI tool, with enhanced rights where the personal data is 'special category data', such as data concerning race or nationality.[34] Subject to various exemptions, key rights in the UK GDPR which will apply in the AI field are as follow:

Article 5(1)(a)	The personal data must be processed in a lawful, fair and transparent manner
Article 5(1)(d)	The personal data must be accurate
Articles 6 and 9	There must be a lawful basis for processing the data
Articles 13 and 14	Certain information must be provided where personal data is collected from the data subject or otherwise
Article 15	There is a right to access the personal data
Article 16	There is a right to rectification of inaccurate personal data
Article 17	There is a right to be forgotten
Article 21	There is a qualified right to object to data processing

[34] See UK GDPR, Data (Use and Access) Act 2025 and the Data Protection Act 2018. Please note that there are bespoke rules for certain contexts like immigration, taxation, activities by the police etc. This is simply a broad summary of the core provisions in non-specialist contexts.

Articles 22A to 22D	There are qualified prohibitions and controls where automated decision-making[35] happens and this results in a significant decision being taken in relation to a data subject.
Article 35	Data Protection Impact Assessments (DPIA) are mandatory where data processing poses 'a high risk to the rights and freedoms of natural persons'. As a matter of best practice, these are usually published but this is not mandatory.[36]

62. Civil society should also be aware that the Government has published a summary of the data protection rules in the context of AI on its knowledge hub, which is said to complement and update the UK AI Playbook. It is accessible at https://ai.gov.uk/knowledge-hub/how-to/legal/.

63. The use of AI by the Government must also, of course, comply with the **human rights** contained in the European Convention on Human Rights and incorporated into domestic law through section 1 of, and Schedule 1 to, the Human Rights Act 1998. The Articles that are most likely to be relevant when AI is deployed include the following:

Article 3	No one shall be subjected to torture or to inhuman or degrading treatment or punishment[37]
Article 6	Right to a fair trial

[35] This is defined to mean where a decision is taken 'solely' based on automated processing i.e. there is no meaningful human in the loop.

[36] The Information Commissioner's Office (**ICO**) advises: *'Although publishing a DPIA is not a requirement of UK GDPR, you should actively consider the benefits of publication. As well as demonstrating compliance, publication can help engender trust and confidence. We would therefore recommend that you publish your DPIAs, where possible, removing sensitive details if necessary.'*: see here.

[37] This is explored later when we consider the Home Office's use of AI in the context of examining whether an asylum-seeker is a refugee.

Article 8	Right to respect for private and family life[38]
Article 9	Freedom of thought, conscience and religion
Article 10	Freedom of expression
Article 11	Freedom of assembly and association
Article 14	Prohibition of discrimination

64. Beyond domestic legislation, as expressly recognised by the AI Playbook, the Government's use of AI in the exercise of public functions must conform with established **public law** principles, which exist at common law.[39] These principles are longstanding yet are subject to continuing development.[40] Relevantly, these principles were conceived on the premise of there being a rational and human decision-maker responsible for the exercise of public functions. It remains to be seen how these principles will be interpreted and applied in the context of AI.

[38] For example, it is well-established that the state use of Live Facial Recognition Technology, which is a form of AI, engages with Article 8 (right to private life) and as such there must be accessible, sufficiently certain rules governing its use and adequate safeguarding: *Glukhin v Russia* (2024) 78 EHRR 6 and *R (Bridges) v The Chief Constable of South Wales Police and others* [2019] EWHC 2341.

[39] Public Law is briefly discussed in the AI Playbook. At page 62 it says: '*Public law principles explain how public bodies should act rationally, fairly, lawfully and in compatibility with human rights. These are guidelines for public bodies on how to act within the law. Many of these public law principles overlap with the ethical principles set out in this guidance. As a result, your lawyers will likely be able to guide you on how to apply the ethical principles based on their knowledge of public law, the court cases that have occurred and the detail of the judgments. For example, public law involves a principle of procedural fairness. This is not so much about the decision that is eventually reached but about how a decision is arrived at. The transparency and explainability of the AI tool may well be key in being able to demonstrate that the procedure was fair. Similarly, an inability to determine how AI tools have arrived at their decisions or outputs may introduce risk into the decision-making process. Public law also considers rationality. Rationality may be relevant in testing the choice of an AI system, considering the features used in a system, and considering the outcomes of the system and the metrics used to test those outcomes. If you're considering using AI in decision making, public law can also guide you. For example, it can help you determine whether a particular decision should be delegated to a decision maker, rather than letting an AI tool make an automated decision. When operating in a regulated environment, such as a procurement process, automated decision making or assessments could be subject to legal challenge if procedural fairness, lack of bias and rationality cannot be evidenced.*'

[40] *A v British Broadcasting Corporation* [2015] AC 588, [56].

65. While it is beyond the scope of this Opinion to provide a comprehensive account of public law principles, the following are likely to be particularly relevant:

   a. Public authorities must not act in a manner that is unreasonable. There are two limbs to public law reasonableness. A decision will be unreasonable if: (i) it is '*outside the range of reasonable decisions open to the decision-maker*'; or (ii) the '*process by which the decision was reached*' was unreasonable, which includes the duty to take into account relevant considerations and ignore irrelevant considerations.[41] The latter is often referred to as 'process rationality'.

   b. Process rationality is closely related to the *Tameside* duty; i.e. decision-makers are under a duty to take sufficient and reasonable steps of inquiry to acquaint themselves with material and information relevant to their decisions.[42] The making of reasonable and necessary inquiries is an essential condition of reasonableness.[43]

   c. Public authorities must act in a manner that is procedurally fair, which means *inter alia* not being biased and – in some circumstances – ensuring affected persons have a right to be heard, are adequately informed, and are provided with adequate reasons for a decision.[44]

[41] *R (Law Society) v Lord Chancellor* [2019] 1 WLR 1649, [98]; restated in *R (Law Society) v Lord Chancellor* [2024] EWHC 155 (Admin), [226]-[228].

[42] *Secretary of State for Education and Science v Tameside Metropolitan Borough Council* [1977] AC 1014, 1065B. The relevant principles are summarised in in *R (Balajigari) v Secretary of State for the Home Department* [2019] EWCA Civ 673, [70].

[43] *Law Society No.2*, [201]-[202], [235].

[44] *R (Osborn) v Parole Board* [2014] AC 1115, [67]-[68]; *R v Secretary of State for the Home Department, ex p Doody* [1994] 1 AC 531, 560D-G; *R (Institute of Dental Surgery) v UFC* [1994] 1 WLR 242.

d. A public authority will err in law where it makes a decision based on a material error of fact giving rise to unfairness.[45]

e. Public authorities must not abdicate, fetter or (improperly) delegate their functions and/or discretion.[46] As to the former, it is a ‘rule of domestic administrative law that a statutory power of decision-making must be exercised by the person on whom the power has been conferred’ and with their ‘independent judgment’.[47]

66. Underlying those grounds of review are higher-level principles of good administration. One is transparency, which is closely related to procedural fairness. As per *R (Justice for Health Ltd) v Secretary of State for Health* [2016] EWHC 2338 (Admin), per Green J at para 141:

CONFIDENTIAL & DRAFT

[45] *E v Secretary of State for the Home Department* [2004] QB 1044, [66].

[46] *British Oxygen v Minister of Technology* [1971] AC 610; *Noon v Matthews* [2014] EWHC 4330 (Admin). The well-established exception to the non-delegation rule is where it is permissible according to the *Carltona* doctrine: *Carltona v Commissioners of Works* [1943] 2 All ER 560.

[47] *Shahid v Scottish Ministers* [2016] AC 629, [68] and [72].

> … The principle of transparency has evolved out of Strasbourg jurisprudence, but it is now well established as a common law principle. It is said to amount to a component of the "rule of law" and the principle of "legal certainty". In *Nadarajah v Secretary of State for the Home Department* [2005] EWCA Civ 363 at [68] Lord Justice Laws stated that it was a *"requirement of good administration"* (to which the courts would give effect) that *"public bodies ought to deal straightforwardly and consistently with the public".* The principle serves a number of important purposes … Clear notice of a policy or decision is also required so that the individual knows the criteria that are being applied and is able to both make meaningful representations to the decision maker before the decision is taken and subsequently to challenge an adverse decision (for instance by showing that the reasons include irrelevant matters). Where the principle applies it might require the publication of the policy that a decision maker is exercising; it might require that the policy be spelled out in greater detail so that the limits of a discretion may be demarcated; it might require the decision-maker to be more specific as to when he/she will or will not act.

67. What reasonableness, fairness and other public law principles require is context-sensitive, as is the intensity with which the courts will review the lawfulness of a decision.[48] The demands of reasonable and fairness, and the intensity of the court's review, will invariably rest on among other things the nature of the function in question, the statutory and policy context, the interests of the affected individuals, and the gravity of the impact of the decision upon them. In our view, that analysis is likely to be informed by the UK AI Playbook and the prescriptive requirements within it. There is a very broad range of things that departments and arm's length bodies must or must not do. Annex B illustrates in a non-exhaustive way the range of such specific directions in the AI Playbook.

68. We consider that these public law principles significantly constrain the Government's use of AI in its decision-making in the face of the risks identified in Part I: The need to regulate AI systems. In our view, public authorities would be well-advised to carefully consider the following issues before deciding to use AI within their decision-making process:

[48] *R (Howard League for Penal Reform) v Lord Chancellor* [2017] 4 WLR 92, [39]; *R (KP) v Secretary of State for the Home Department* [2025] EWHC 370 (Admin), [58]-[63], [76].

a. The extent to which decision-makers need to make inquiries as to the accuracy/functionality of AI tools they rely upon. Also, the extent to which they may rely upon AI to discharge the *Tameside* duty;

b. The extent to which decision-makers who rely upon AI within their decision-making process can demonstrate that they have taken into account relevant considerations and excluded irrelevant ones (having regard to the Blackbox problem). This will necessarily include the question whether the decision-maker has taken into account the accuracy and appropriateness of the relevant AI tool in question before relying upon it;

c. Determining whether a decision-maker has based a decision on a material error of fact arising from AI hallucinations or inaccuracies;

d. Whether, as a matter of procedural fairness, affected individuals have a right to know that AI is being used in reaching decisions that affect them, how it is being used, the algorithms used, and the prompts used by decision-makers; and

e. The compatibility of reliance on AI and automated decision-making with the non-abdication/delegation/fettering principles.[49]

69. We anticipate that each of those issues will likely be the subject of judicial review proceedings in due course.

70. Finally, **equality law** prohibits the use of AI by public authorities in a way which is discriminatory in relation to protected characteristics such as sex, race, sexual orientation, disability, gender reassignment, religion or belief, maternity and pregnancy. Key rights[50] in the Equality Act 2010 are as follows:

[49] Certain of these issues are explored in Lord Sales' Keynote Lecture for the Government Legal Department's Annual Conference on 5 November 2025 ('AI and Public Law: Automated Decision-Making in Government'), as well as by Azeem Suterwalla KC and Will Perry in their chapter 'Public law and procurement law', in The Law of Artificial Intelligence (2nd ed., Sweet and Maxwell 2024), ed. Dr Matthew Lavy and Matt Hervey.

[50] These rights do sometimes differ depending on the protected characteristic being considered.

Section 13	Prohibits less favourable treatment because of a protected characteristic.
Section 15	Prohibits unfavourable treatment because of something arising from disability unless there is justification. In our view, whenever the justification defence is engaged under the Equality Act 2010, and the UK AI Playbook is also applicable, the principles contained within it including the prescriptive requirements set out at Annex B will be highly relevant. A failure to comply with the principles and prescriptive requirements may well lead to the justification defence failing.
Section 17	Prohibits unfavourable treatment because of pregnancy and maternity.
Sections 19/19A	Prohibits neutral practices which put groups at a disadvantage unless there is justification.
Sections 20 – 21	Duty to make reasonable adjustments for persons with a disability.
Section 26	Prohibition on harassment.
Section 149	A public authority must, in the exercise of its functions, have due regard to the need to eliminate discrimination etc. This is referred to as the Public Sector Equality Duty (**PSED**).[51]

[51] The Equality and Human Rights Commission has published guidance on equality and AI for public sector bodies which contains detailed guidance on the PSED. It is available here.

Part IV: Case study: The Home Office's use of AI during refugee status determination

71. So far, this Opinion has focused on the AI Ethical Principles and the parallel domestic legal framework. To illustrate how this framework should apply, in Part IV we will consider the Home Office's recent use of AI in the asylum process when determining refugee status.

The asylum process in the UK

72. To appreciate the impact of AI in the asylum process, it is necessary to first have an overview of the system.

73. People who wish to stay in the UK as a refugee must apply for asylum. The definition of a '*refugee*' is dictated by the 1951 Convention Relating to the Status of Refugees (**the Refugee Convention**). In very general terms, the asylum-seeker must be unable to live safely in any part of their own country because they fear persecution and this fear must be '*well-founded*'.[52] This persecution must be because of the asylum-seeker's race, religion, nationality, political opinion or anything else that puts them at risk because of the social, cultural, religious or political situation in their country, for example their sexual orientation (sometimes referred to as 'a Convention reason').

74. Closely related to this is the principle of non-refoulement in Article 33(1) of the Refugee Convention, according to which the UK shall not expel or return a refugee to a territory '*where his life or freedom would be threatened on account of his race, religion, nationality, membership of a particular social group or political opinion*'. Closely related to this, again, is the prohibition against removal under Article 3 of the European Convention on Human Rights, according to which the UK must not remove persons to other States where there are substantial grounds for believing they would be at

[52] 1951 Convention Relating to the Status of Refugees, Article 1A(2); Nationality and Boarders Act 2022, ss. 30 – 38. There are very detailed immigration provisions in the UK but for the purposes of this Opinion, the summary provided here is adequate.

real risk of torture or inhuman or degrading treatment (discussed further at para 136 and following).[53]

75. The above principles have been given effect in our domestic law by several statutes of Parliament and the Immigration Rules.[54] By section 6 of the Human Rights Act 1998, it is unlawful for the Home Office to act incompatibly with an ECHR right. Further, s.82(1) of the Nationality, Immigration and Asylum Act 2002, read with s.84(1), confers a right of appeal against the refusal of both protection claims (including claims that removal would breach the Refugee Convention) and human rights claims (claims that removal would be unlawful under the Human Rights Act 1998).[55]

76. The Home Office is responsible for considering asylum claims.[56] Once asylum has been claimed, there are four broad stages to the process:

(i) **Step 1: Screening or Registration Interview**: After providing biometric information, the applicant will be interviewed and asked to produce relevant documentation.

[53] *R (AAA (Syria)) v Secretary of State for the Home Department* [2023] 1 WLR 4433, [19]-[23].

[54] Section 2 of the Asylum and Immigration Appeals Act 1993 provides: '*Nothing in the immigration rules (within the meaning of the 1971 Act) shall lay down any practice which would be contrary to the [Refugee] Convention.*' Immigration Rules, paragraph 328 dictates that all asylum applications have to be decided in accordance with the Refugee Convention as well as the Immigration Rules.

[55] For a summary, see: *AAA*, [27]-[29].

[56] This Opinion provides a high-level description of the process for adults only; there are bespoke rules for children. Please note that there is also a Streamlined Asylum Process for *'manifestly well-founded claims*', which includes certain nationals. This Opinion does not consider this 'fast track' process.

(ii) **Step 2: Substantive Interview**: This is the point at which the applicant is meant to outline in detail their claim for refugee status. This is emphasised by the Immigration Rules and Government guidance.

The interview is conducted in private and may happen face to face or by video conference. Applicants are advised to keep the day free for their interview.[57] It should be conducted in the applicant's first language or a language which they are able to understand. It is usually audio-recorded and then typed up. The ORG understands that the Home Office is gradually introducing technology-driven transcription for substantive interviews (called 'digital interviewing capability')[58] although precisely how this works, and the role of AI, is not clear[59]. The Government has said that transcripts '*can extend to 50 pages or more*'.[60] The transcript and audio are retained and made available to the applicant or their legal representative.

(iii) **Step 3: Decision**: A decision-maker acting on behalf of the Home Office will determine the asylum application ('**the Decision-Maker**'). They will not

[57] 'Guidance: Information booklet about your asylum application', updated 28 October 2025.

[58] 'Guidance: Asylum interviews', updated 16 October 2025.

[59] AI transcriptions tools give rise to various legal considerations. Researchers in the United States have noted that voice recognition technology can be more inaccurate for certain racial groups. A summary of the paper is available here: https://fairspeech.stanford.edu/. In the context of asylum interviews, the relevance of this research is obvious. To the ORG's knowledge, there is no publicly available information concerning the accuracy of the 'digital interviewing capability' deployed by the Home Office and whether accuracy differs for certain racial or national groups. In the context of asylum interviews, the relevance of this research is obvious. To the ORG's knowledge, there is no publicly available information concerning the accuracy of the 'digital interviewing capability' deployed by the Home Office and whether accuracy differs for certain racial or national groups. See also the paper from the Ada Lovelace Institute published on 11 February 2026 called, 'Report: Scribe and prejudice? Exploring the use of AI transcription tools in social care' which outlines the various risks of AI transcription, available here.

[60] 'Research and analysis: Evaluation of AI trials in the asylum decision making process', 29 April 2025.

necessarily be the person who conducted the substantive interview. The decision is provided in writing. It could be to grant refugee status or humanitarian protection or limited permission to stay. Alternatively, the application could be rejected.

Decision-Makers are not judges. However, they must have knowledge with respect to the standards applicable in the field of asylum and refugee law.[61]

(iv) **Step 4: Appeal**: A decision that an asylum-seeker is not a refugee can ordinarily be appealed to the First-tier Tribunal (there is also a subsequent right of appeal). This is a judge-led process.[62]

77. At Stage 3, to determine whether an applicant has refugee status, the Decision-Maker must consider whether they are unable to live safely in any part of their own country because they fear persecution and this is '*well-founded*'. In other words, the test is *subjective* in that the applicant must fear persecution but also has an *objective* element in that the fear must be '*well-founded*'.

78. The objective element is explained by the UNHCR[63] in the 'Handbook on Procedures and Criteria for Determining Refugee Status and Guidelines on International Protection' (**the Handbook**)[64] as follows:

[61] Immigration Rules, paragraph 339HA.

[62] Section 82(1) Nationality, Immigration and Asylum Act 2002. Please note that there are currently proposals in place to reform the appeal process through the introduction of an independent appeals body.

[63] The Office of the UN High Commissioner for Refugees. It was established in the 1950's to provide international protection under the auspices of the United Nations to refugees.

[64] The Handbook is available on the UNHCR website. The Handbook is not incorporated into UK law, but the Court of Appeal has commented that it is '*particularly helpful as a guide to what is the international understanding of the Convention obligations, as worked out in practice …*': see *R v Secretary of State for the Home Department ex parte Robinson* [1998] QB 929, page 938B/C.

> 42. As regards the objective element, it is necessary to evaluate the statements made by the applicant. The competent authorities that are called upon to determine refugee status are not required to pass judgement on conditions in the applicant's country of origin. The applicant's statements cannot, however, be considered in the abstract, and must be viewed in the context of the relevant background situation. A knowledge of conditions in the applicant's country of origin – while not a primary objective – is an important element in assessing the applicant's credibility …
>
> 43. … The laws of the country of origin, and particularly the manner in which they are applied, will be relevant. The situation of each person must, however, be assessed on its own merits …

79. Unsurprisingly, the Handbook is clear that to assess objectively whether a fear of persecution is '*well-founded*', there must be consideration of the information provided by the applicant and the situation in the asylum-seeker's country.

80. The Refugee Convention does not dictate what specific processes countries should use to determine refugee status. The Handbook does, however, stress that decision-making must proceed on an informed basis with a careful consideration of the evidence.

81. The domestic picture in the UK proceeds on the same basis as the Handbook.

82. Determining refugee status under the Refugee Convention is '*particularly fact sensitive*' as stated by the House of Lords in *R (on the application of Sivakumar) v Secretary of State for the Home Department* [2001] EWCA Civ 1196 (para 7). The House of Lords also endorsed the Court of Appeal in *Karanakaran v Secretary of State for the Home Department* [2000] 3 ALL ER 448 where Sedley LJ emphasised the need to take into account all material presented by an applicant and country information:

> … The question whether an applicant for asylum is within the protection of the 1951 Convention is not a head-to-head litigation issue. Testing a claim ordinarily involves no choice between two conflicting accounts but an evaluation of the intrinsic and extrinsic credibility, and ultimately the significance of the applicant's case … Such decision-makers, on classic principles of public law, are required to take everything material into account. Their sources of information will frequently go well beyond the testimony of the applicant and include in-country reports, expert testimony and - sometimes - specialist knowledge of their own (which must of course be

> disclosed). No probabilistic cut-off operates here: everything capable of having a bearing has to be given the weight, great or little, due to it. What the decision-makers ultimately make of the material is a matter for their own conscientious judgment, so long as the procedure by which they approach and entertain it is lawful and fair and provided their decision logically addresses the Convention issues….

83. As to the right methodology when examining whether an asylum-seeker's fear of persecution is well-founded, the Nationality and Borders Act 2022 dictates that a Decision-Maker must adopt a two-stage approach.

84. The first stage is to determine whether the applicant has – on the balance of probabilities – a characteristic which would cause them to fear persecution for reasons of race, religion, nationality, membership of a particular social group or political opinion (s.32(2)(a)) and, that the applicant does fear persecution in their country as a result (s.32(2)(b)).

85. The second stage is for the Decision-Maker to determine whether there is a reasonable likelihood that if the applicant were returned to their country they would be persecuted because of the characteristic and would not be protected (s.32(4); s.34).

86. The Immigration Rules make it mandatory to consider all relevant/material facts, which include facts about the applicant's characteristics which are said to cause them fear, and information relating to the asylum-seeker's country of origin when determining refugee status.[65] To this end, the Home Office publishes information and notes on various countries concerning the risk of persecution there ('**Country Policy Information Notes**'; '**CPIN**').

87. It is against this backdrop – the need for Decision-Makers to consider all relevant/material information to determine refugee status including information provided by the applicant and the CPIN – that we move to examine the Home Office's AI tools.

[65] The Immigration Rules have a status '*akin to that of law*' as per Sedley LJ in *Secretary of State for the Home Department v Pankina* [2010] EWCA Civ 719 (para 17).

An introduction to the ACS and APS

88. The ORG only knows of limited publicly available information about the ACS and APS. The main source is a report published on 29 April 2025 called '*Research and analysis: Evaluation of AI trials in the asylum decision making process*' ('**the Home Office Report'**)[66] along with information on the Government's AI Knowledge Hub ('**the Knowledge Hub'**)[67].

89. From this, it is understood that the **ACS** summarises asylum interview transcripts using generative AI[68], specifically a Large Language Model[69] i.e. it generates new text which purports to be a summary of the original text. The new text does not sign-post or highlight the source of the original information.[70]

90. The **APS** is also generative AI.[71] It is described as being a '*chat based interface*' which examines CPIN and related information to generate new text.[72] There is no publicly available information about what '*chat based*' means but it is likely that there is a

[66] Available online here.

[67] Available online here and here.

[68] Knowledge Hub.

[69] Home Office Report, section 1.1.

[70] Home Office Report, section 4.1.

[71] Knowledge Hub.

[72] Home Office Report, section 1.2.

traditional 'prompt' system by which the user enters free text[73] to formulate a request, for example: '*tell me about whether there is a risk of violence against gay men in country X in year Y*'.

The Home Office's objective when using the ACS and APS

91. At the end of September 2025, there were 62,171 cases, relating to 80,841 people, awaiting an initial decision on an asylum application.[74] The Home Office analysis reveals that despite a record number of asylum claims, the number of cases awaiting an initial decision fell 36% between September 2024 and September 2025. The stated aim of the ACS and APS is to further '*speed up*' the process of considering asylum applicants and '*improve efficiency*'.[75] What precisely this means requires unpicking.

92. It is no surprise that a fulsome review by the Decision-Makers of the transcript generated in the Substantive Interview and of CPINs is time consuming. The Home Office must perceive – on some level – that this is 'too long' since the aim of the ACS and APS is to reduce time. However, decision-making about refugee status is – by its nature – going to take time. It is not some form of arid procedural decision-making or a tick-box exercise: it is a process involved in determining whether those seeking asylum are genuinely at risk of harm and need protection. So, it cannot be legitimate to apply a new tool in this process with the aim simply to reduce the time taken to collect and assess information in absolute terms; it will always be necessary to have regard to the need to maintain or improve the quality of the decision-making. The Home Office's overarching aim – bearing in mind the legal framework addressed in para 77 above – ought therefore to be, nothing less, than to ensure that Decision-Makers *consider all relevant/material facts as efficiently as possible*. If it were to

[73] There is no suggestion in the publicly available information that prompts are pre-written in any way.

[74] 'Accredited official statistics: How many cases are in the UK asylum system?', 27 November 2025, available online.

[75] Home Office Report, 'Executive summary'.

maintain an aim such as this, which marries effectiveness (which here means the accuracy of summaries in terms of identifying material/relevant information) with efficiency, in our view it would almost certainly be a legitimate aim for the Government.[76]

93. Articulating the aim fully is not merely semantics: it is critical to understanding whether the tools comply with the AI Ethical Principles, for example, whether the AI is '*right for the job*', whether the risks (and benefits) are being adequately identified and balanced, and whether appropriate AI Assurance is being undertaken.[77]

Evaluating the ACS and APS against the UK AI Playbook

94. Having identified the aim which underpins the Home Office's AI tools, we now look to assess them by reference to the UK AI Playbook's principles which reflect the AI Ethical Principles.

Principles 1 and 10b: 'You know what AI is and what its limitations are' and 'You have the right assurance in place'

95. Principle 1 requires the Government to identify all potential risks associated with the proposed AI tools.[78] Principle 10b sits alongside Principle 1 and requires steps to be taken to assess and then mitigate identified risks. This principle requires the

[76] For example, the Court of Appeal in *R (The Motherhood Plan & Kerry Chamberlain) v Her Majesty's Treasury and the HMRC* [2021] EWCA Civ 1703 endorsed the aims of effectiveness, delivery, ease of verifying objective data, fraud reduction, the need to avoid perversity and reduce costs when the Government used technology to assess the profits lost by self-employed traders as a result of the pandemic (para 129). However, it should be borne in mind that every AI tool will need to be considered on its own merits; there are no generalisations which can be made here.

[77] The need to identify a legitimate aim and then balance whether that aim is achieved in a proportionate way, having regard to any negative impacts on individuals or groups of individuals, will be crucial in various legal contexts. For example, where a tool is, on its face, indirectly discriminatory contrary to the Equality Act 2010, there will need to be an objective justification for it to be lawful. A similar balancing exercise is used where there is interference with certain human rights.

[78] UK AI Playbook, page 10.

Government to *'understand, monitor and mitigate the risks that using AI tools can bring … You should have clearly documented review and escalation processes in place, and have an AI review board or programme-level board'*.[79]

96. AI Assurance is explained in 'Guidance: Introduction to AI assurance' published by the Department for Science, Innovation and Technology (**DSIT**)[80] in this way:

> The term 'assurance' originally derived from accountancy but has since been adapted to cover areas including cyber security and quality management. Assurance is the process of measuring, evaluating and communicating something about a system or process, documentation, a product or an organisation. In the case of AI, assurance measures, evaluates and communicates the trustworthiness of AI systems.

97. AI Assurance engages with many of the AI Ethical Principles such as safety, robustness, transparency and explainability. The practical tools by which AI Assurance takes place are often referred to as 'assurance mechanisms'.[81]

Inaccuracy caused by the use of generative AI; lack of transparency over the assurance process including benchmarking

98. We start by observing that there are a whole series of accuracy risks when AI is used to summarise information. The ORG has identified, in particular:

    a. Hallucinations: This will happen when the AI tool simply fabricates information.

[79] UK AI Playbook, pages 14-15.

[80] Published on 12 February 2024 and available online here.

[81] DSIT's 'Guidance: Introduction to AI assurance', Paragraph 3.2.

b. Output inaccuracy: This will happen where the AI tool inaccurately summarises text, which might be through omitting information or mispresenting it. This can happen in several ways, for example:

   i. Position bias: Research has shown that LLMs can focus too much on information at the beginning or end of text and accordingly struggle with information contained 'in the middle'.[82]

   ii. Understanding saliency (i.e. the relevance of information): Humans and AI tools can assess relevancy in different ways. Research suggests that the way in which 'prompts' are crafted will be relevant to how effectively saliency is judged.[83]

   iii. Processing ambiguity: Narratives will often involve ambiguity. Human listeners can often decipher the correct use of a word from the context. Research has identified that LLMs struggle to resolve ambiguity.[84]

   iv. Difficulty understanding subtext: Within any narrative there will be explicit parts of the 'story' and subtext, namely, what is meant through implication or the emotion conveyed. Research has shown than when LLMs summarise narratives they struggle with subtext.[85] This is likely to

[82] Wan, D., Vig, J., Bansal, M., & Joty, S. (2024). On Positional Bias of Faithfulness for Long-form Summarization: https://arxiv.org/abs/2410.23609

[83] Xu, L., Karim, M. A., Dingliwal, S., & Elangovan, A. (2024). Salient Information Prompting to Steer Content in Prompt-based Abstractive Summarization: https://arxiv.org/abs/2410.02741

[84] Sutriawan Sutriawan, Supriadi Rustad, Guruh Fajar Shidik, Pujiono Pujiono, Muljono Muljono (2024). Review of ambiguity problem in text summarization using hybrid ACA and SLR. Intelligent Systems with Applications, Volume 22. https://www.sciencedirect.com/science/article/pii/S266730532400036X

[85] Subbiah, M., Zhang, S., Chilton, L. B., & McKeown, K. (2024). Reading Subtext: Evaluating Large Language Models on Short Story Summarization with Writers. Transactions of the Association for Computational Linguistics, 12, 1290–1310: see https://direct.mit.edu/tacl/article/doi/10.1162/tacl_a_00702/124837/Reading-Subtext-Evaluating-Large-Language-Models

be a real issue when summarising interviews provided by asylum-seekers as no doubt much crucial information is to be derived from subtext rather than explicit information.

c. Cultural biases: LLMs have been observed to be less accurate in relation to cultures outside of those described in English language texts.[86] It is thought that this form of discrimination arises because during the machine learning process, the AI tool is trained on training data predominantly derived from English language texts.[87] As such, the tool does not 'learn' about situations which are not represented in that data, which is problematic as language and culture are closely tied[88]. It is obvious that cultural bias will be a real risk when considering information provided directly by asylum-seekers.

99. Based on the available information, one of the most significant risks is inaccuracy within the output generated by the ACS and APS. The Knowledge Hub reveals that during the pilot for the ACS there were inaccurate summaries 9% of the time and 5% of users of the APS were 'not confident in tool accuracy'. What type of 'inaccuracy' is being displayed here (see para 98 above) is entirely unclear. There is certainly no sense from the publicly available information that the assessors had any regard to the many and nuanced ways in which 'inaccuracy' could arise.

100. There is no information about how the degree of accuracy was measured or evaluated in relation to the ACS, although it appears that some form of objective quantitative data was measured since the 9% metric has been published. The extent to which this inaccuracy arose from the LLM 'hallucinating' or whether it was simply an error in summarisation (and what type of error) is entirely unknown.

[86] Li, C., Chen, M., Wang, J., Sitaram, S., & Xie, X. (2024). 'CultureLLM: Incorporating cultural differences into large language models': https://proceedings.neurips.cc/paper_files/paper/2024/file/9a16935bf54c4af233e25d998b7f4a2c-Paper-Conference.pdf.

[87] Multilingual cultural data is often expensive to procure – see fn. 82.

[88] Hershcovich, D., Frank, S., Lent, H., de Lhoneux, M., Abdou, M., Brandl, S., & Søgaard, A. et al. (2022). 'Challenges and Strategies in Cross-Cultural NLP': https://arxiv.org/abs/2203.10020

101. The assessment of accuracy for the APS appears to have been limited to an assessment of confidence rather than an objective measure of inaccuracy. Given that DSIT's 'Guidance: Introduction to AI assurance' highlights the importance of measuring objective quantitative data,[89] we consider this to fall short of the Government's own standards and expectations.

102. At present there are no benchmarks which are universally accepted (domestically or internationally) in relation to *levels of accuracy*,[90] nor is there consensus over how accurate an AI tool needs to be in any given context for it to be deemed 'safe'. However, the 9% measure of 'inaccuracy' (whatever that means in this context) identified by the Home Office for the ACS is worryingly high given the importance of determining refugee status and the significance of getting it wrong (i.e. being returned to a country where there is a risk of persecution).

103. It is concerning that there is no publicly available information about what benchmark is – in the Home Office's view – adequate or why. This is despite the Government's 'i.AI incubator for Artificial Intelligence'[91] rightly highlighting the need to define a benchmark *before* using a tool[92]:

CONFIDENTIAL & DRAFT

[89] Section 4.1.

[90] There is a major workstream currently being undertaken in Europe to create benchmarks as part of an effort to introduce standardisation i.e. European harmonised standards which will create legal certainty and reduce compliance costs. More information about this project is available here. There is also a UK-based AI Standards Hub (available here) which is supported by the Government but this is intended as a repository rather than intending to contain obligatory standards.

[91] The Incubator for Artificial Intelligence is based out of the Department for Science, Innovation and Technology (**DSIT**). It is a unit of expert engineers (applied AI, evaluation, platform), cross-cutting functions (design, product, delivery, ops) and civil service strategists.

[92] 'i.AI's approach to evaluation', 20 June 2025 available here.

> **Challenge 2: What *does* good look like?**
>
> Determining what our baselines are before the tool is introduced requires us to consider what we should compare our tools to: The work of an expert? An average worker? What is the level of ambition for the tool, and what is realistic?
>
> When estimating quality, if a human benchmark is the gold standard, i.e. is the 'ground truth' or 100% performance, it is impossible for a new tool to be considered 'better'. Conversely, when we consider efficiency, AI often vastly speeds up the process compared to a human, which may be a larger difference for senior vs junior staff comparisons but those differences are not likely to be meaningful (how much do we care if the product is 100x or 150x faster than a human?).
>
> Quality and efficiency are not generally well-measured in the public sector. In i.AI we are keen to set the bar high on how we are measuring our impact in order to support public sector AI adoption by highlighting what realistic impacts can be expected by using our tools.

104. Without this basic information, civil society cannot fully engage in a discourse with the Government about its approach towards AI Assurance.

Inadequate procedural safeguards to manage the risk of inaccuracy

105. Just as importantly, there are no appropriate procedural steps to guard against inaccuracy infecting the ultimate decision. In other words, there are insufficient assurance mechanisms. In particular, the ability to identify inaccuracies is limited by the design of the system which does not cross-reference back to the original Source Material (thereby inhibiting Decision-Makers from verifying the outcome of the AI tools). This was identified as a limitation of the ACS by users during the pilot ('*Some interviewees, however, highlighted minimal time-saving due to the summary not providing source references*'[93]). It is assumed that the APS is designed in the same way.

106. A further missing procedural safeguard is that the asylum-seeker is not provided with the output of the ACS and APS before a decision is made. If this happened, the applicant or their advisors could identify any errors. It would also allow third parties

[93] Home Office Report, para 4.1.

to independently understand how accurate the tools are in terms of identifying and summarising material and relevant facts.

Risk that inaccurate information becomes part of the decision-making process and inadequate procedural safeguards to manage this risk

107. One point emphasised by the Home Office is that these AI tools '*do not, and cannot replace any part of the decision-making process*'.[94] We have concerns about the boldness of this statement. While as a matter of law we agree with 'cannot', there is serious doubt about the 'do not'. The stated purpose of these tools is to filter and funnel information to save time. Plainly there is a risk that these tools could become part of the decision-making if the Source Material is discarded, and the Decision-Maker relies unduly (or exclusively) on the funnelled information. This could happen as a matter of policy or simply by default as time-stressed Decision-Makers work through a case load.

108. Whilst the Home Office asserts that Decision-Makers can only use the tools to aid decision-making, we have no information about the technical measures used to prevent them from *only* looking at the output of the tools. There is a suggestion that the tools might compel Decision-Makers to access CPIN or related information before making a decision, but this does not necessarily mean that they *compel* full consideration of the original Source Material.[95] Indeed, there is something of a tension between saying on the one hand – these tools should save time because they summarise information through generative AI and maintaining, on the other hand, that all material is to be considered fully especially where there is no cross-referencing back to the Source Material.

109. Put simply, if a Decision-Maker must carefully consider the entirety of the Source Material to confirm the accuracy of the AI-generated summaries in each case, it is difficult to conceive how any efficiencies would be realised. The corollary is this: we suspect that Decision-Makers are *not* compelled to fully consider the Source Material

[94] Home Office Report, 'Executive summary'.

[95] '*In line with the 'human in the loop' principle, APS has been designed so that decision-makers cannot use the tool by itself to make a decision and need to access the full CPIN or COIR*'. Home Office Report, section 1.2. There are similar comments in section 1.1 concerning the ACS.

and *will* rely upon the AI-generated summaries to a material extent, otherwise there would be little purpose in relying upon the APS or ACS.

110. This problem is further compounded by automation bias, which is a phenomenon by which humans can over trust the outputs from computer-based systems in certain contexts.[96] In the present context, it is easy to perceive of situations in which a Decision-Maker might assume that the summary provided by the AI tools is accurate and has identified all material/relevant facts without properly checking.

111. It is essential that AI Assurance mechanisms manage both the risk of inaccuracy and the possibility that Decision-Makers could disregard the original Source Material. To manage those risks there should be clear guidance that Decision-Makers cannot rely on the output of the ACS or APS alone. There should be effective technical measures in place to compel Decision-Makers to view the Source Material, understand where the Source Material is located (in so far as it summarised by the ACS and APS) along with an audit trail recording each step of the Decision-Making process so as to validate and verify that Source Material has not been discarded.

No adequate assessment of the impact on the quality of decision-making

112. A related point is that there does not appear to be an adequate assessment of whether the quality of the decisions is adversely affected by the ACS or APS tools. According to the Home Office, there was no impact – positive or negative – on the quality of decision-making.[97] This conclusion was reached using ‘Calibre’, which is an assurance tool. The metric here appears to have been simply whether a similar proportion of decisions reached the same outcome in the test group versus the control group.[98] We consider this inadequate. Given that the Decision-Makers’ task is to assess relevant/material facts, our view is that quality - in this context - should

[96] Romeo, G. and Conti, D., 2025. Exploring automation bias in human–AI collaboration: a review and implications for explainable AI. AI & SOCIETY, pp.1-20, see https://doi.org/10.1007/s00146-025-02422-7

[97] Home Office Report, ‘Executive summary’.

[98] Home Office Report, footnote 1.

measure the extent to which the ACS and APS disregarded or correctly identified / summarised that information. To our knowledge, this has not been assessed. There are very real costs associated with making bad decisions. There is the cost to the asylum-seeker but also the public purse in increased asylum appeals.[99]

No adequate assessment of the extent to which the tools discriminate

113. Discrimination is often a risk when AI is used. The Home Office says that it assessed whether nationality and / or the asylum reason impacted the output of the APS system by examining whether there was a different impact on certain groups when the tools were used versus when they were not used. It concluded that there was no impact.[100] A similar conclusion is reached in relation to the ACS although the Home Office notes that '*… pilot numbers were too small to fully test for variation in outcome for all characteristics, and this would require ongoing monitoring*'.[101] We consider it unacceptable that these tools would be rolled out when it is known that the bias testing was insufficient.

Principle 6: 'You use the right tools for the job'

114. This principle is concerned with choosing '*the most appropriate technology to meet your needs … you should be open to the conclusion that, sometimes, AI is not the best solution for your problem: it may be more easily solved with more established technologies*'.[102] It is inextricably intertwined with Principles 1 and 10b; ensuring that the AI tool is '*right for the job*' engages with the aim, the risks and mitigation measures. It requires the Government to ask whether there are other more appropriate ways of achieving the aim including analogue measures.

[99] On 18 February 2026, The Times reported that each asylum appeal costs around £4,000, see here.

[100] Home Office Report, para 4.2.

[101] Home Office Report, para 4.1.

[102] UK AI Playbook, pages 12-13.

115. Applying that approach here, we are concerned that there is an element of 'using a sledgehammer to open a nut' in relation to these tools. We have no difficulty accepting that reading a lot of information can be cognitively challenging and tools which improve the readability of interview transcript notes and CPINs must be important.[103] We note that the Home Office Report believes that 23 minutes per case was saved with the ACS tool and 37 minutes per case with the APS.[104] Reducing cognitive load would – one imagines – also improve decision-making as it would increase the prospect of identifying correctly all relevant/material facts.

116. However, we have seen no evidence that the Home Office has considered whether there are alternatives to AI – which do not carry the risks we have identified above – yet also achieve the aim of permitting Decision-Makers to identify all relevant/material facts in an efficient (time-saving) manner.

117. For example, the CPINs are published by the Home Office; if they are dense or difficult to digest, surely the solution is to re-write them or introduce more signposting/subheadings rather than use AI to regurgitate (perhaps inaccurately) their content? Why is AI-powered searching via the APS any more useful than traditional search tools like the 'find' function which is embedded in most software and has no risk of AI-style hallucinations? In other words, we query whether an AI tool like the APS is the right solution to voluminous CPINs and related information.[105]

118. We have similar concerns about the ACS. If the intention is to remove cognitive load due to lengthy transcripts, why is using AI to generate a summary useful (especially where the summary has no source references)? A tool that indexed or labelled the transcript might be invaluable (for example, '*Paragraphs 10, 34 and 56 contain the references to the applicant being tortured due to his sexuality*') but the tool

[103] These concerns are highlighted in the Home Office Report, paras 1.1 and 1.2.

[104] Home Office Report, Executive summary.

[105] This is an observation that was made in the pilot: '*... only 42% felt it gave them the right amount of information. This is echoed by some interviewees who noted that the summary output did not reference where to access the information in the transcript. This was a conscious decision of the tool designers for the pilot to reduce the technical complexity but could be revisited in future*'.

does more than this – it filters, funnels and regurgitates information into an unsourced summary.[106] The utility of this additional functionality is not easy to understand bearing in mind the need for the Decision-Makers to identify all relevant/material facts.[107]

119. Equally, if the Government needs to reduce the backlog of undecided asylum cases (see para 90 above), an analogue solution might be to simply recruit more Decision-Makers or more skilled ones or improve the training of existing workers. Similarly, there might be a more appropriate use of AI, for example only using it to assess those cases which are likely to be straightforward and uncontroversial i.e. the cases where the claimant has a clear right to refugee status.[108] Principle 6 requires the Government to consider a wide range of solutions rather than merely default to a technological solution.

120. There may be good reasons to opt for technology (AI or otherwise) but at present there is no publicly available information which demonstrates that this analysis has been taken. It is critical that civil society has such information to test whether the Government has complied with the procedure required – under Principle 6 – to consider whether the proposed AI is '*right for the job*'. Afterall, if AI is not the right tool

[106] This is an observation that was made in the pilot: '*A small minority (5%) of survey respondents were not confident in the tool's accuracy, and some stated they did not see the benefit of using the tool over searching in the CPINs directly*'. See Home Office Report, para 5.2.

[107] We should say, as an aside, that we can see that in some judicial and quasi-judicial contexts, the ability to summarise voluminous information might be very useful. For example, when case managing litigation where there has been extensive correspondence, most of which may be of marginal significance or only historic interest, an ability to summarise matters would be invaluable for a time-pressed judge so they could focus their pre-reading on the matters of real contention. However, that scenario is very far removed from the way in which ACS is being used here. Moreover, in that scenario there are likely to be procedural safeguards in place to ensure that the judge does not overlook any information of significance, for example, a hearing at which the parties or their lawyers can direct the judge's attention to important correspondence or correct any misapprehensions.

[108] Canada uses AI in this way: see https://www.canada.ca/en/immigration-refugees-citizenship/corporate/transparency/digital-transparency-advanced-data-analytics/processing-applications.html

and it leads to higher rates of appeal against Decision-Makers' decisions on asylum applications, it will increase the cost to the public purse rather than decrease it. Indeed, we note that at the time the Home Office Report was completed, the Home Office could not say that the tools offered value for money.[109]

Principle 4: 'You have meaningful human control at the right stage'

121. This principle is concerned with ensuring that there is always '*a human in the loop*'.[110] This does not mean that there is a human who 'rubber stamps' decisions. It means that a human should always have meaningful input and oversight. This principle does not prevent a *degree* of automation, but it does require that there be scope for meaningful human intervention within the overall process. Our concerns, outlined at paras 107 to 112, that the ACS and APS could be relied on exclusively (or too heavily) by the human Decision-Makers are equally applicable under Principle 4. Meaningful technical and procedural safeguards are required to ensure that the human Decision-Maker remains fully in the driving seat.

Principle 7: 'You are open and collaborative'

122. This principle encourages the Government to engage with civil society, stating: '*Where possible, you should engage with the wider civil society including groups, communities, and non-governmental, academic and public representative organisations that have an interest in your project. Collaborating with people both inside and outside government will help you ensure we use AI to deliver tangible benefits to individuals and society as a whole. Make sure you have a clear plan for engaging and communicating*

[109] Section 6.

[110] UK AI Playbook, pages 11-12. The Explanatory Memorandum published with the Council of Europe's Framework Convention commenting on Article 15 of the Convention, made this point as follows: '*Where an artificial intelligence system substantially informs or takes decisions impacting on human rights, effective procedural guarantees should, for instance, include human oversight, including ex ante or ex post review of the decision by humans. Where appropriate, such human oversight measures should guarantee that the artificial intelligence system is subject to built-in operational constraints that cannot be overridden by the system itself and is responsive to the human operator, and that the natural persons to whom human oversight has been assigned have the necessary competence, training and authority to carry out that role.*'

with these stakeholders at the start of your work.'[111] To the ORG's knowledge, there has been no compliance with this part of Principle 7 at all.

123. The principle also addresses transparency more generally stating that there should be openness '*with the public about where and how algorithms and AI systems are being used in official duties*'.[112] We have serious concerns about the level of transparency provided by the Home Office in relation to the ACS and APS.

124. Beyond the Home Office Report, the AI Knowledge Hub and some brief answers to Parliamentary Questions[113], there is a notable lack of public information from the Government. Answers to Freedom of Information Requests that have been published by 'What Do They Know', including a request from the ORG, underline further the lack of transparency.[114]

125. Drawing this information together along with our analysis so far, we have six serious concerns about the level of transparency provided by the Home Office.

126. **First**, a Data Protection Impact Assessment (**DPIA**) and Equality Impact Assessment (**EIA**) were undertaken when the tools were piloted but the Government declined to publish them in July 2025 when requested on the basis that the tools were at an early stage.[115] However, to the ORG's knowledge, they have still not been published despite the tools apparently being in use. These are crucial documents in terms of ensuring transparency.

[111] UK AI Playbook, page 13.

[112] UK AI Playbook, page 13.

[113] On 25 November 2025, Alex Norris (Labour MP) answered a question in Parliament about the ACS and APS. Available online here. It repeated some information contained in the Home Office Report.

[114] Available online: FOI2025/07437, FOI2025/13832 and FOI2025/15287.

[115] FOI2025/07437

127. **Secondly**, the Government also declined to publish details of the ‘prompt’ given to the ACS tool to tell it to summarise interviews due to concerns about cyber security and abuse of the system.[116] In our view, this concern is hard to understand. We cannot immediately see how disclosing the ‘prompt’ would lead to these risks. In contrast, disclosing the ‘prompt’ within the ACS would allow civil society to critique and offer constructive feedback on the appropriateness of the tool. This is an important part of accountability which is one of the AI Ethical Principles.

128. **Thirdly**, asylum-seekers are **not** told about the use of the ACS tool in relation to their applications according to one Freedom of Information Request[117]:

> Interviewees were not informed of the use of the ACS tool throughout the pilot. We will not be informing asylum seeker (sic) when the tool rolls out as this is an additional aide and does not replace or negate the requirement for the full interview transcript to be reviewed by the decision maker.

129. There can be no sensible justification for this lack of transparency. A key plan of trustworthy and ethical AI is that people know when it is being used especially in the hands of the state. The Government knows this since its ‘Guidance to civil servants on use of generative AI’[118] states:

> **How answers from generative AI can mislead**. These tools can produce credible looking output. They can also offer different responses to the same question if it is posed more than once, and they may derive their answers from sources you would not trust in other contexts. Therefore, be aware of the potential for misinformation from these systems. Always apply the high standards of rigour you would to anything you produce, and reference where you have sourced

[116] FOI2025/07437

[117] FOI2025/07437

[118] This was published on 29 January 2024 and is available online here.

> output from one of these tools. (Emphasis added).

130. **Fourthly**, the ORG's request asked for, amongst other matters, model cards, technical specifications, and internal guidance describing the ACS and APS tools including a description of their functions and training data sources, any Standard Operating Procedures and user guidance provided to Decision-Makers. It was made on 7 October 2025 (re-issued on 14 October 2025) and has recently been answered on 6 February 2026 following a complaint to the Information Commissioner's Office and an internal review request. It contains little substantive information. Disappointingly, it explains that no information will be provided on the training data sources on the basis that '*Training data is proprietary information of the organisations who created them*'. The failure to provide this information is in tension with the Government's recently published detailed guidance on how to ensure appropriate training sets are used: 'Guidance: Guidelines and best practices for making government datasets ready for AI' (19 January 2026).[119] The Guidance explains what steps need to be taken to ensure that data sets are appropriate for use by the Government. If civil society cannot understand even basic information about data sets, then this type of guidance will have limited impact.

131. A full answer to the OPG request would have addressed some of the concerns raised in this Opinion, for example, the extent to which Decision-Makers are informed about the risks of inaccurate information and to ensure that the tools do not replace consideration of the source information. It is disappointing that the information has not been provided.

132. **Fifthly**, at the time of writing, neither the ACS nor the APS is listed in the repository for the Algorithmic Transparency Recording Standard.[120] The purpose of the Algorithmic Transparency Recording Standard Hub is to help public sector organisations provide clear information about how and why they are using

[119] Available here.

[120] The only Home Office tool is CARS(V) which is a semi-automated workflow routing solution which identifies the likely complexity of visitor visa applications based on applicants' declared attributes.

algorithmic tools.[121] There are various exemptions.[122] There is no explanation as to why the ACS or APS does not feature in the repository. A Freedom of Information Request, which asked whether the ACS or APS would be added to the Algorithmic Transparency Recording Standard (and if not, why not) was refused on 5 December 2025 on the basis that, along with other queries, it was too onerous.[123] It is disappointing that this information has not been provided.

133. **Finally**, there is a lack of transparency concerning the AI Assurance process and related mechanisms as identified at para 100 and following above. It is also important to note that the Home Office Report recommends that '*any limitations of the tools identified in the evaluation should be addressed before a full rollout*' (para 9.2). There is no information that we can identify as to whether this has happened. It also recommends '*continuous monitoring and data capture in the early stages to ensure accuracy, quality and the use of the tools remains the same, and to ensure impact does not differ among case characteristics*' (para 9.2). Again, there is no information explaining how or whether this information will be made publicly available. It says that there should be '*full evaluation after deployment … to capture the tools outside the pilot environment*' (para 9.2). Again, there is no information explaining how or whether this information will be made publicly available. These points compound our concerns about the lack of transparency.

Summary of conclusions concerning the application of the AI Ethical Principles to the ACS and APS

134. Having identified the aim which underpins the Home Office's AI tools, we assessed the ACS and APS against the UK AI Playbook's procedural principles and its adoption of the AI Ethical Principles. In summary, our conclusion is that – based on the information currently publicly available – the UK AI Playbook is either not being followed, or we have serious concerns that there is non-compliance due to the absence of publicly available information.

[121] See online commentary here.

[122] They are detailed here.

[123] FOI2025/15287

Principles 1, 4 & 10b: ‘You know what AI is and what its limitations are’, ‘You have meaningful human control at the right stage’ & ‘You have the right assurance in place’		
1	There has been a failure to assess quantitatively the extent to which the APS produces inaccurate outputs.	Paras 98 - 99
2	Whilst there has been some form of quantitative assessment of the ACS to assess the accuracy of outputs, there is insufficient information about what accuracy means in this context, the extent of inaccuracy and what benchmarking the Government is using (i.e. what ‘good looks like’).	Paras 100 - 104
3	There is no ability to cross-reference the summarised output from the ACS (and we assume the APS) to the original Source Material, making it difficult for Decision-Makers to assess/verify accuracy. This is an important missing procedural safeguard.	Para 105
4	There is no system in place to allow asylum-seekers to check the output of the ACS or APS for accuracy since they do not know it happened and have no opportunity to see the text generated. This is an important missing procedural safeguard.	Para 106

5	There is a risk that the summarises produced by the ACS and APS will become part of the decision-making process since they filter and funnel information which could be relied on by the Decision-Makers at the expense of the Source Material. There is no clear guidance, that we can see, which tells Decision-Makers that they must fully consider all original Source Material. This is an important missing procedural safeguard. It also means that we cannot be satisfied that there are procedures in place to ensure meaningful human control.	Paras 107 - 111 & 121
5	We are not satisfied that there are adequate technical measures in place and auditing processes to ensure that Decision-Makers consider all original Source Material. This is an important missing procedural safeguard. It also means that we cannot be satisfied that there are procedures in place to ensure meaningful human control.	Paras 107 - 111
6	We are not satisfied that there has been an adequate assessment of whether the quality of the decisions is negatively affected by the ACS or APS. Given that the Decision-Makers' task is to assess relevant/material facts, our view is that quality – in this context – should measure the extent to which the ACS and APS disregarded or correctly identified/summarised that information. In other words, the Government has not properly approached AI Assurance.	Para 112
7	It is unacceptable that the ACS and APS have been rolled out despite only limited bias testing having been undertaken. In other words, the Government has not properly approached AI Assurance.	Para 113

Principle 6: 'You use the right tools for the job'

8	Despite the risks associated with generative AI, there is no publicly available information that the Government considered whether an analogue or non-AI solution could achieve the aim of greater efficiency.	Paras 114 - 120

Principle 7: ‘You are open and collaborative’		
9	To the ORG’s knowledge, the Home Office has not engaged with civil society in relation to the ACS and APS.	Para 122
10	There is no published Data Protection Impact Assessment. There is no published Equality Impact Assessment. The ‘prompt’ used in the ACS tool is unknown (for no obvious good reason). Asylum-seekers are not told about the use of the ACS tool. Technical information such as the training data source has yet to be published. Neither the ACS nor APS is listed in the repository for the Algorithmic Transparency Recording Standard.	Paras 123 - 132
11	There is no information about the AI Assurance that has happened since the pilot scheme.	Para 133

Compliance with the existing legal framework

135. We now move to consider the extent to which the ACS and APS comply with the existing legal framework that we have outlined in Part III: Domestic legislation, legal principles and the UK AI Playbook. This assessment will also analyse compliance with the AI Playbook’s procedural Principle 2, which requires the Government to act lawfully.

Article 3 European Convention on Human Rights

136. The prohibition on torture or inhuman or degrading treatment or punishment is contained in Article 3. It is an absolute right. In other words, the state cannot justify interfering with Article 3 rights.

137. Asylum-seekers may well be seeking refuge in the UK as a means of escaping torture or ill-treatment and accordingly their Article 3 rights will be engaged. The basic principle is that Contracting Parties to the Convention must not deport or expel persons from their territories where substantial grounds have been shown for believing that the person, if removed, would face a real risk of being subjected to treatment contrary to Article 3 in the receiving country.[124] It follows from that substantive obligation that there must be an independent and rigorous assessment of whether a person faces a real risk of ill-treatment if removed to the receiving country.[125] That assessment 'inevitably involves an examination by the competent national authorities… of the conditions in the receiving country against the standards of Article 3'.[126] Further, public authorities must not adopt a policy which exposes persons to a significant risk of treatment prohibited by Article 3 ECHR.[127]

138. Plainly, if the Home Office were relying upon AI to determine asylum claims, that could not be compatible with the Article 3 procedural obligations of independent and rigorous assessment. On the other end of the spectrum, we do not see the use of AI within the determination process as necessarily incompatible with Article 3 insofar as it is subject to appropriate safeguards and human review, and the ultimate determination remains that of the human Decision-Maker.

[124] *Soering v the United Kingdom* (1989), (paras 90-91); *Ilias and Ahmed v Hungary* (2019), (para 126); *AAA*, [23].

[125] *Ilias,* para 127; *Chahal v. United Kingdom* (1996) paras 79 and 96; *M.S.S. v. Belgium and Greece* (2011), para 293.

[126] *Ilias*, para 127 (see also paras 139-141). See also: *Hirsi Jamaa and Others v Italy*, paras 118-119.

[127] *R (Munjaz) v Mersey Care NHS Trust* [2006] 2 AC 148 at [29], [80]; *R (A) v SSHD* [2021] 1 WLR 3931 at [79].

139. Turning to the APS and ACS, our opinion is this: If the Home Office adopts a practice of using generative AI tools which create a risk that Decision-Makers will consider inaccurate information and/or overlook relevant/material facts when determining people's asylum claims, and those risks are not mitigated by appropriate safeguards (including a robust appeals process), such an assessment would be unlikely to have the necessary rigour to comply with the UK's procedural obligations under Article 3. Adopting such a practice would expose asylum-seekers to a significant risk of treatment contrary to Article 3.

140. We have already explained above, by reference to the AI Ethical Principles contained in the UK AI Playbook, that we consider there are various missing procedural safeguards which are required to mitigate the risk of inaccurate outputs from the ACS and APS infecting the decision-making process. A further way that civil society could engage with the Government over this topic is by framing concerns about inadequate safeguarding by reference to Article 3 and insisting that clear procedural steps are taken to ensure an independent and rigorous assessment of asylum claims.

Public law

141. The UK AI Playbook and the prescriptive rules contained within it will likely be highly relevant to the application of common law requirements of reasonableness and procedural fairness in the AI context.[128]

142. In the particular context of the APS and ACS, the most relevant public law principles are those of process rationality and procedural fairness.[129] As asylum

[128] There is an argument that the Home Office is under a direct public law duty to follow the UK AI Playbook unless there are good reasons for departure (per *Mandalia v Secretary of State for the Home Department* [2015] 1 WLR 4546, [39]). However, at present, we consider it unlikely that the policy adherence duty arises in respect of the UK AI Playbook as a whole. Our view in that respect may change if future iterations of the UK AI Playbook contain more prescriptive requirements geared towards particular decision-making functions. Certainly, if the Home Office issued a policy in respect of its use of AI in the field asylum determination, the policy adherence duty would arise: applying the principles in *R (Riverside Park Ltd) v Secretary of State for Levelling Up, Housing and Communities* [2023] EWHC 2937 (Admin), [76]-[80].

[129] We have not focused on the non-abdicating/fettering/delegating of duties in this context. Those public law constraints will be most relevant to situations in which public authorities adopt

determination concerns the most fundamental of rights, the Home Office's decisions and use of AI will be subject to 'the most rigorous examination, to ensure that it is in no way flawed'.[130] In this respect, the court will expect the Decision-Maker 'to show by their reasoning that every factor which might tell in favour of an applicant has been properly taken into account'.[131] Similar considerations increase the demands of procedural fairness.

143. Against that background, we make four observations.

144. **First**, in our view, the Home Office is under a heightened *Tameside* duty of inquiry as regards the accuracy and functionality of the ACS and APS before they are adopted within its asylum determination process. The Home Office will be at significant risk of breaching its *Tameside* duty if it fails to undertake the following:

a. A quantitative assessment of the extent to which the APS and ACS produce inaccurate outputs;

b. An adequate assessment of whether the quality of asylum decisions is negatively affected by the ACS or APS;

c. An assessment of the risks of bias and discrimination within the AI tools (the latter is equally required under the PSED);

automated decision-making with limited human review. On the basis of our current understanding of how the ACS and APS are used, the Decision-Maker does not appear to be abdicating or delegating their decision-making function to an AI tool or foregoing their independent judgment, but rather AI tools are being used as a decision-making aid. Neither does the Decision-Maker appear to be fettering their discretion in using the ACS or APS. However, the extent to which those principles are engaged may change depending on how the ACS and APS are being used in practice.

[130] *Bugdaycay v Secretary of State for the Home Department* [1987] AC 514, 531. See also: *KP*, [58]-[63], and [76]-[79].

[131] *R (YH (Iraq) v Secretary of State for the Home Department* [2010] 4 All ER 448, [24]. See also: *R (Alnoor) v Secretary of State for the Home Department* [2025] EWHC 992 (Admin), [28].

d. An assessment of the effectiveness of alternatives to the AI tools in achieving the Home Office's aim of greater efficiency; and

145. After the initial rollout, continuous monitoring and data collection in respect of the above matters.[132]

146. Our view as to the necessity of such enquiries and the unreasonableness of failing to take them is supported by Principles 1, 4, 6 and 10b of the AI Ethical Principles and our analysis of them. As highlighted at paras 98 to 121 above, it appears that the Home Office has not undertaken adequate enquiries in respect of the ACS and APS.

147. **Second**, the use of the ACS and APS give rise to a significant risk of process irrationality. The critical starting point is that the human Decision-Maker is dutybound to take all statutory mandatory and obviously relevant material into account when determining an individual's asylum claim, and to exclude irrelevant considerations. Relevant material in this context plainly includes the CPINs and the asylum-seeker's testimony.[133] In our view, it is not sufficient that an AI tool has taken a consideration into account if it has not been properly and rationally considered by the Decision-Maker.

148. If a Decision-Maker uses the ACS and APS summaries as an aid in their decision-making but nonetheless properly considers the Source Material in its entirety, no issue would arise. However, if a Decision-Maker relies upon the ACS and APS summaries at the expense of a full examination of the Source Material, and those summaries have filtered out relevant information regarding the country of origin or asylum-seeker's interview, there will be a significant risk that the Decision-Maker will have failed to take relevant considerations and evidence into account when

[132] Compared to failures of assessment in: *Law Society No.2*, [208]-[210].

[133] See para 83 above in which *Karanakaran* is quoted: '… *Testing a claim ordinarily involves no choice between two conflicting accounts but an evaluation of the intrinsic and extrinsic credibility, and ultimately the significance of the applicant's case … Such decision-makers, on classic principles of public law, are required to take everything material into account*.' This is reinforced by Paragraphs 339I, 339J and 339JA of the Immigration Rules, as well as the following guidance: 'Guidance: Information booklet about your asylum application', Updated 28 October 2025; 'Guidance: Conducting asylum interviews: casework guidance: Immigration staff guidance on conducting asylum interviews', Updated 16 October 2025; 'Guidance: Assessing credibility and refugee status in asylum claim lodged on or after 28 June 2022', Updated 28 September 2023.

determining the asylum claim in question. As discussed at paras 105, 111 and 121 above, we do not presently consider there to be adequate procedural safeguards in place to guard against that risk and ensure Decision-Makers consider all Source Material. In short, the breach or potential breaches of Principles 1, 4 and 10b in the UK AI Playbook reinforce our analysis here.

149. Further, upon conducting the enquiries outlined at para 144 above, the Home Office would be dutybound to rationally consider those matters (for example, the accuracy of the ACS and APS) when deciding whether to adopt the AI tools within its asylum determination process.[134]

150. **Third**, in a context where the ACS summaries are inaccurate 9% of the time and 5% of users are not confident in the accuracy of the APS, and the summaries do not cross-refer to the Source Material, there is a significant risk that decisions which are based on the summaries generated by both AI tools will be based upon and vitiated by material errors of fact. This risk is heightened in circumstances where there are no apparent safeguards requiring Decision-Makers to examine fully the Source Material and where asylum-seekers are not provided with copies of the ACS summary with a view to correcting errors.

151. **Fourth**, given the gravity of the consequences for asylum-seekers if their claims are determined on the basis of inaccurate information and the nature of the interests at stake, we consider that – as a matter of procedural fairness – asylum-seekers have a common law right to be informed that AI is being used in the determination of their claims, how it is being used, and to be provided with the output of the AI-generated summaries. That conclusion applies with greatest force to the ACS given that it summarises sensitive information that the asylum-seeker has provided, and which the asylum-seeker is well-placed to correct.[135] We are fortified in that view by the common law principle of transparency and the transparency principles that run through the UNESCO Recommendation and UK AI Playbook. In our view, the fact that asylum-seekers appear not to be so informed is likely to be unlawful.

[134] Compared to: *Law Society No.2*, [234].

[135] Compare to the failures of disclosure in: *Roberts v Secretary of State for Work and Pensions* [2025] EWHC 51 (Admin), [51]-[57]; *R (AK) v Secretary of State for the Home Department* [2025] EWHC 1651 (Admin), [82]-[84]; *Eisai Ltd v National Institute for Health and Clinical Excellence* [2008] EWCA Civ 438, [36], [49]-[50] and [66]; *R (Ames) v Lord Chancellor* [2018] Lloyd's Rep FC 545, [75].

Data protection considerations

152. It will be recalled that the ACS summarises information provided by the applicants. It follows that 'personal data' will be processed by the AI tool when it is used to summarise an individual's interview. It is also likely that 'special category' personal data will be processed.[136] This is data that reveals, amongst other matters, racial or ethnic origin, political opinions, religious or philosophical beliefs and data concerning a person's sex life or sexual orientation.[137] This means that the UK's data protection regime is engaged. When 'personal data' is processed, various obligations are engaged including:

Article 5(1)(a)	The personal data must be processed in a transparent manner
Article 5(1)(d)	The personal data must be accurate
Article 13(1)(c)	There must be an explanation of the purposes of any data process which is intended
Article 15	A right to access the personal data
Article 16	A right to rectification of inaccurate personal data

153. In our view, there will be a breach of these articles if the ACS produces inaccurate summaries of personal data, if there is no explanation to the asylum-seekers that the AI tool will be used in relation to their application and / or if asylum-seekers are denied the ability to see the output from the ACS with a view to correcting any errors. Whilst we note that Schedule 2 to the Data Protection Act 2018 contains an exemption for immigration in relation to most of these rights[138], it only arises where

[136] Article 4 (1), UK GDPR.

[137] Article 9, UK GDPR.

[138] An earlier version was amended after a judicial review was initiated by the ORG and the3million: see *R (on the application of the Open Rights Group) v The Secretary of State for the Home Department and another* [2021] EWCA Civ 800.

exercise of those rights would be likely to prejudice the maintenance of effective immigration control or the investigation or detection of activities that undermine the maintenance of effective immigration control. We cannot presently foresee how that exemption could arise. This is another way in which civil society could frame its challenge to the ACS, namely by arguing that the current system breaches data protection rules.

Public Sector Equality Duty

154. In circumstances where the Government has declined to publish the Equality Impact Assessment, we cannot be satisfied that the PSED has been met.[139] It is also useful to recall the comments of the Court of Appeal in *R (Bridges) v The Chief Constable of South Wales Police and others* [2020] 1 WLR 5037 about the importance of the PSED in a case which involved the use of AI by the police, which stated at para 176:

> We accept (as is common ground) that the PSED is a duty of process and not outcome. That does not, however, diminish its importance. Public law is often concerned with the process by which a decision is taken and not with the substance of that decision. This is for at least two reasons. First, good processes are more likely to lead to better informed, and therefore better, decisions. Secondly, whatever the outcome, good processes help to make public authorities accountable to the public. We would add, in the particular context of the PSED, that the duty helps to reassure members of the public, whatever their race or sex, that their interests have been properly taken into account before policies are formulated or brought into effect. (Emphasis added)

[139] For examples where an inadequate EIA has resulted in a breach of the PSED, see: *R (TG) v Secretary of State for the Home Department* [2025] EWHC 596 (Admin), [318]-[321].

155. The Decision-Maker must have given 'adequate thought' and 'be informed' about the potential impact of the use of the APS and ACS on the objectives in the PSED and to what protected groups should be considered.[140] We further consider that the PSED 'involve[d] a duty of inquiry' to ensure the Home Office is sufficiently informed of the potential discriminatory impact of incorporating the APS and ACS into its decision-making, and that such risk continues to be monitored.[141]

156. In the absence of a published Equality Impact Assessment, we also cannot be satisfied that there are no broader equality implications to the use of the ACS or APS, for example, that it is more accurate for certain groups, nor that due regard has been given to such impacts.[142]

Dialogue with regulators

157. Lastly, it should also be noted that regulators that operate in areas affected by AI are subject to principles identified by the Government in its 'Policy paper: A pro-innovation approach to AI regulation', updated 3 August 2023 ('**the White Paper**'). The principles are listed at section 3.2.3 as:

1	Safety, security and robustness
2	Appropriate transparency and explainability
3	Fairness
4	Accountability and governance

[140] *R (Ward and Gullu) v Hillingdon London Borough Council* [2019] PTSR 1738, [71]-[72]. There must be 'proper appreciation of the potential impact of the decision on equality objectives': *R (Hurley and Moore) v Secretary of State for Business, Innovation and Skills* [2012] EWHC 201 (Admin), [77].

[141] *Ward and Gullu* at [71]-[72]. See also: *Hurley and Moore* at [89]; *R (DXK) v Secretary of State for the Home Department* [2024] EWHC 579 (Admin), [154]-[155].

[142] Please note that there are limited rights under the Equality Act 2010 in the context of immigration by virtue of Part 4 in Schedule 3 to the Equality Act 2010 which disapplies many provisions.

5	Contestability and redress

158. These principles very much mirror those contained in the UK AI Playbook. The White Paper explains that ‘Existing regulators will be expected to implement the framework underpinned by five values-focused cross-sectoral principles’ (para 48).

159. The Independent Chief Inspector of Borders and Immigration (**ICIBI**) independently monitors and reports on the efficiency and effectiveness of the UK’s border, immigration and citizenship system, making recommendations directly to the Home Secretary. To the ORG’s knowledge, the ICIBI is not presently examining the way in which the Home Office uses AI to determine refugee status. Civil society could engage with the ICIBI, highlighting the risks identified in this Opinion and the principles contained in the White Paper, to encourage an assessment of the appropriateness of the APS and ACS tools.

CONFIDENTIAL & DRAFT

Summary of conclusions concerning the application of the domestic legal framework to the ACS and APS

160. We assessed the ACS and APS against the domestic legal framework which runs in parallel to, and has touch points with, the UK AI Playbook. Ultimately, we concluded that – on the basis of the information currently publicly available – there are various legal risks if the tools remain in their current form:

Article 3		
1	Asylum-seekers may well be seeking refuge in the UK as a means of escaping torture or ill-treatment such that their Article 3 rights are engaged. If the Home Office were relying upon AI to *determine* asylum claims, that would not be compatible with Article 3. At the other end of the spectrum, we do not see the use of AI within the determination process as necessarily incompatible with Article 3 insofar as it is subject to appropriate safeguards and human review such that the ultimate determination remains that of the human Decision-Maker.	Paras 136 to 140

Public law		
2	The Home Office is under a heightened *Tameside* duty of enquiry as regards the accuracy and functionality of the ACS and APS. To satisfy this there must be a quantitative assessment of the extent to which the APS and the ACS produced inaccurate outputs; an adequate assessment of whether the quality of asylum decisions is negatively affected by the ACS or the APS; an assessment of the risk of bias and discrimination; an assessment of alternatives to the AI tools to achieve the Home Office's aims; and continuous monitoring and data collection in respect of these matters.	Paras 144 to 146

3	The use of the ACS and the APS give rise to a significant risk of process irrationality under public law. This is because the human Decision-Maker is duty bound to take into account all statutory mandatory and relevant material. The AI tools may prevent this happening. There should also be a consideration of whether the AI tools should be used altogether and it is not clear if this has happened.	Paras 147 to 149
4	The use of the AI tools gives rise to a significant risk that decisions are based on summaries which are inaccurate, which would breach public law principles.	Para 150
5	There is a lack of procedural safeguards, which breaches the requirement of procedural fairness within public law.	Para 151

<table>
<tr><td colspan="3">Data protection</td></tr>
<tr><td>6</td><td>There will be a breach if the ACS produces inaccurate summaries of personal data, if there is no explanation to the asylum-seekers that the AI tool will be used and / or they are denied access to the output from the ACS to correct any errors.</td><td>Para 153</td></tr>
</table>

<table>
<tr><td colspan="3">Public Sector Equality Duty (PSED)</td></tr>
<tr><td>7</td><td>The Government has not published an Equality Impact Assessment so we cannot be satisfied that the PSED has been met and that there are no broader equality issues.</td><td>Paras 154 to 156</td></tr>
</table>

Regulators		
8	The Independent Chief Inspector of Borders and Immigration (ICIBI) should examine the way in which the Home Office is using AI. The Government has made it plain that it expects regulators to implement AI ethical principles in their work. We are not aware of any scrutiny of the ACS or APS by regulators.	Paras 157 to 159

ROBIN ALLEN KC
CLOISTERS CHAMBERS

DEE MASTERS
CLOISTERS CHAMBERS

JOSHUA JACKSON
DOUGHTY STREET CHAMBERS

22 February 2026

Annex A: Sources of the AI Ethical Principles

		UK AI Playbook	White Paper	OECD	UNESCO	Council of Europe
1	Democracy	-	-	Article 1.2	-	Article 5
2	Fairness, equality & non-discrimination	Principle 2	Principle 3	Article 1.2	3	Articles 10 & 17, 18, 22
3	Human dignity & autonomy	Principle 2	-	Article 1.1	6	Article 7
4	Respect for human rights	Principle 2	-	Article 1.2	1	Articles 4 & 21
5	Privacy & data governance	Principle 2	-	Article 1.2	5	Article 11
6	Sustainability	-	-	Article 1.1	4	-
7	Robustness & digital security	Principles 2, 3, 5, 9 & 10	Principle 1	Article 1.4	8	Article 12
8	Safety & reliability	Principles 2, 5, 6, 9 & 10	Principle 1	Article 1.4	8 & 9	Articles 12 & 13
9	Transparency & explainability	Principles 4 & 7	Principle 2	Article 1.3	7 & 10	Article 8 & 19
10	Accountability & responsibility	Principles 1, 5, 6, 9 & 10	Principle 4	Article 1.5	7, 8 & 9	Articles 9 & 20
11	Contestability, oversight & redress	Principles 1, 5, 6, 9 & 10	Principle 5	Article 1.5	8 & 9	Articles 8, 14, 15, 16 & 26

Annex B: Key directive statements in the UK AI Playbook

The following passages from the UK AI Playbook all contain mandatory requirements saying that a person or department engaged with an AI system 'must' do something. They are therefore among the most important obligations that departments and arm's length bodies have. There are other passages in the Playbook which are similarly directive so this list should not be seen as exhaustive; its main purpose is illustrative to exemplify the very broad range of mandatory obligations as to the way that Government works with AI.

	Foreword
1	'As technology evolves, so too will our approach, ensuring we remain at the forefront of responsible innovation - always guided by the principle that technology must serve people.'

	Principle 2
2	'Your use of AI tools must be lawful and responsible.'
3	'You must ensure that AI systems generate a positive impact on stakeholders and civil society at large while minimising potential harms as much as possible. When defining and deploying AI systems, you must understand people's needs and priorities by conducting user research and engaging with the public as appropriate, including civil society groups, underrepresented individuals, those most likely to experience harm, NGOs, academia and industry.'

	Principle 3
4	'When building and deploying AI services, you must make sure that they are secure to use and resilient to cyber attacks, as laid out in the Government Cyber Security Strategy. Your service must comply with the Secure by Design principles, which were developed by the Central Digital and Data Office (CDDO), and the government's Cyber Security Standard. Different types of AI are susceptible to different security risks. Some threats – such as data poisoning, perturbation attacks, prompt injections and hallucinations – are specific to AI. However, AI systems can also amplify generic risks such as phishing and cyber-attacks. You must understand the risks associated with your use of AI and of adversaries potentially using AI against you.'

	Principle 4
5	'You need to monitor the AI's behaviour and have plans in place to prevent any harmful effects on users. This includes ensuring that humans validate any high-risk decisions influenced by AI and that you have strategies for meaningful intervention.'

6	‘You should fully test the product before deployment, and have robust assurance and regular checks of the live tool in place. Since AI models can sometimes produce unwanted or inaccurate results, incorporating feedback from users is crucial. You should have systems in place that allow users to report issues and prompt a human review.’

	Principle 5
7	‘You should understand how to monitor and mitigate for potential drift, bias, and, in the case of generative AI, hallucinations. You should also have a robust testing and monitoring process in place to catch these problems.’
8	‘If you develop a service, you must use the government Service Standard.’

	Principle 6
9	‘… you should also be open to the conclusion that, sometimes, AI is not the best solution for your problem: it may be more easily solved with more established technologies.’

	Principle 7
10	‘You should be open with the public about where and how algorithms and AI systems are being used in official duties. If you’re a central government department or an arm’s length body within scope, you’re required to use the Algorithmic Transparency Recording Standard (ATRS). This means you must document information about any algorithmic tools you use in decision-making processes and make it clearly accessible to the public. The ATRS is not a requirement for all arm’s length bodies and other public sector institutions yet, but we still encourage you to use it. You should also clearly identify any automated response to the public. For example, a response generated via a chatbot interface should include something like ‘this response has been written by an automated AI chatbot’.’

	Limitations of AI
11	‘… accuracy: it’s difficult to produce an AI system which provides 100% accurate outputs under all conditions. You must be clear about what objective measures you’re assessing the AI outputs against and any factors that impact them’

	Building the team
12	‘Your minimum viable AI team must be able to: • identify user needs and accessibility requirements • manage and report to stakeholders and other teams, collaborating with different field experts • design, build, test and iterate AI products, using agile methodologies • ensure the responsible development of lawful, ethical, secure and safe-by-design AI services • be able to collect, process, store and manage data ethically, safely and securely • test with real users and measure the performance of the service • support the live running of the service, iterate and retire it’

	Ethics
13	‘... to promote fairness, you may need to collect demographic data to accurately assess the impact of a tool on different groups, which would have a detrimental impact on privacy. You must consider early on whether trade-offs are appropriate, if the benefits outweigh the risks, and that you’re avoiding any unacceptable risks.’

	Safety, security and robustness
14	‘You must build safe, secure and robust AI solutions. This means that your AI systems must be resilient, sustainable and function reliably, even in unforeseen situations or against adversarial attacks.’

	Transparency and explainability
15	‘You must consider transparency issues before deploying an AI system.’
16	‘All central government departments and certain arm’s length bodies that are in scope of the Algorithmic Transparency Recording Standard (ATRS) must use it to ensure transparency around the algorithmic tools used in decision-making processes by public bodies.’

	Fairness, bias and discrimination
17	‘You must ensure fairness in the development and use of AI solutions to comply with legal and human rights requirements, including consumer and competition law, public and common law, and rules protecting vulnerable people.’

	Accountability and responsibility
18	‘Ensuring accountability and responsibility in the context of AI means that individuals and organisations can be held responsible for the effects that the AI systems they develop, deploy or use, have on people and society. You must think about this at the start of your project to: • encourage mindful creation and usage of AI systems • ensure that people who design and deploy AI systems can be held accountable for their outputs and impacts’

	Answerability
19	‘Identifying the specific actors involved in AI systems is vital to answerability. This includes model developers, application developers, policymakers, regulators, system operators and end-users. In each case, you must define their roles and responsibilities, and align these with legal and ethical standards.’

	Liability
20	‘You must … put the necessary oversight and human-in-the-loop processes in place to validate output in situations with high impact or risk. Where these risks are too high, you must reconsider if AI should be used at all. Refer to the Identifying use cases for AI section for more on this.’

	Contestability and redress
21	‘Contestability and redress are important because they help identify and correct ethical issues in AI systems after deployment. You must design appropriate mechanisms before deployment, and continue to maintain them throughout the full life cycle of your AI system.’
22	‘ … public awareness: to enable users to contest and seek redress about your AI system, you must ensure that they are aware of the presence of the AI system and the function that it plays in the services that they’re interacting with. This includes making users aware of mechanisms for contestability and redress clearly and in a timely fashion’
23	‘… mechanisms for appeal: you must establish and promote clear and accessible mechanisms for people to challenge the decisions made by AI systems, and ask wider questions concerning the training, deployment and impacts of AI systems employed by the UK government’
24	‘… change processes: you must ensure that mechanisms are in place to investigate any areas highlighted by users, and make changes to or decommission AI systems if unacceptable risks or harms are identified’

	Societal wellbeing and public good
25	‘You must ensure that AI systems generate a net positive impact on stakeholders and society at large, while minimising potential harms as much as possible. If the potential negative consequences are too high, you must consider terminating the project. … public benefit: you must ensure that the AI solutions you develop and/or use represent good value for money and benefit the public. This aligns with the UK government’s ambitions to use AI to help solve societal and global challenges – as long as the AI solution is safe, lawful and compatible with other ethical principles’

	Human rights
26	‘Public authorities must act in a way that is compatible with human rights. It’s possible that AI systems (especially those involving the use of personal data) may in some way affect at least one of an individual’s rights, as set out in the European Convention on Human Rights (ECHR). Examples of the rights most likely to be impacted are Article 8 (right to a private and family life) and Article 10 (freedom of expression).’

	Data protection and privacy
27	‘Be aware that organisations developing and deploying AI systems must consider the principles of data protection outlined in the UK General Data Protection Regulation (UK GDPR) and the Data Protection Act 2018, and minimise the risk of privacy intrusion from the outset.’

	Lawfulness and purpose limitation
28	‘If your [data protection impact] assessment indicates that there’s a high risk to the data protection rights of individuals, and that you’re unable to sufficiently reduce these risks despite mitigating actions, you must consult the ICO before you can start processing personal data.’

	Fairness
29	‘You must make sure that AI systems do not process personal data in ways that are unduly detrimental, unexpected or misleading to the individuals concerned. If AI systems infer data about people, you need to ensure that the system is accurate and is not discriminatory. You need to uphold the ‘right to be informed’ for individuals whose personal data is used at any stage of the development and deployment of AI systems. This is part of fulfilling the transparency and fairness principles.’

30	‘Biometric data is also considered special category data when processed for the purposes of identification. You must ensure that the technologies used to capture and process this data are overt, accurate, proportionate, fair and deploy a narrow ‘zone of recognition’. For example, if someone walks past a camera and their image does not meet the threshold for a potential match, their data needs to be promptly deleted.’

	Human oversight
31	‘Services using AI that affect a person’s legal status or their legal rights must only use AI to support decisions that must be made by a human decision maker. AI systems need to introduce deliberation processes into all stages of the life cycle so that the abilities of humans and machines are combined to reach the best results when performing tasks. However, the human input needs to be ‘meaningful’. Several factors determine how much human involvement there should be in AI systems, such as the complexity of the output, its potential impact, and the amount of specialist human knowledge (for example, legal and medical) required.’

	Security
32	‘Cyber security is a primary concern for all government services, as laid out in the Government Cyber Security Strategy. When building and deploying new services, including AI systems, the government has a responsibility to make sure these are secure to use and also resilient to cyber attacks. To meet this requirement, your service must comply with the government’s Secure by Design principles before it can be deployed.’

	Public AI applications and web services
33	‘A simple way to implement an AI solution is to use publicly available commercial applications – such as Google Gemini or ChatGPT in the case of generative AI. While you might think that these public tools are more secure, you should consider that you cannot easily control the data input to the models: you must rely on educating users on what data they can and cannot enter into these services. You also have no control over the outputs from these models, and you’re subject to their commercial licence agreements and privacy statements. For example, OpenAI will use the prompt data you enter directly into the ChatGPT website to improve their models, although individual users can opt out. When using public AI applications, you must not enter official information unless it has been published or is cleared for publication.’

	Embedded AI applications
34	‘In addition to embedded applications, there are also many AI-powered plugins or extensions to other software. For example, Visual Studio Code has a large ecosystem of community-built extensions, many of which offer AI functionality. You must take extreme caution before installing any unverified extensions as these can pose a security risk.’