THOMSON REUTERS
PRACTICAL LAW™

Transparent AI use in the public sector (UK)

by Alex Lawrence-Archer, AWO

Status: **Publication date (14 October 2024)** | Jurisdiction: **England, Wales**

This document is published by Practical Law and can be found at: **uk.practicallaw.tr.com/w-044-3714**
Request a free trial and demonstration at: **uk.practicallaw.tr.com/about/freetrial**

A note analysing a public sector body's statutory and common law duties to be transparent about its use of AI and government guidance on complying with these duties. The note discusses when transparency is required proactively and responsively, and the standard of explanation expected of a public body about its AI models and automated decisions.

Scope of this note

A public body may use AI to support its decision-making, the exercise of its statutory powers and functions and service delivery. This might, for example, involve using natural language processing to summarise or draft text, or assess trends and patterns and the monitoring of live data. A public body's AI use could result in decisions that have significant effects on individuals or groups, and may involve the processing of personal data.

This note sets out the extent to which the law requires a public body to be transparent about its use of AI, and how government guidance elaborates on those requirements. It explains when a public body must provide this information both proactively and responsively, what information is likely to be required and the standard it should meet, and the processes for disclosing it. The note also outlines the consequences of non-compliance.

The note is relevant to:

- Public bodies within the scope of the legal requirements.
- Private actors providing technology and services to public bodies within the scope of the legal requirements.
- Individuals affected by the decisions of public bodies within the scope of the legal requirements.

This note uses the term AI broadly to cover any data-driven automated analysis or decision-making (see Article, Automated decision-making in the public sector). The legal issues do not change where more complex technology is used, such as machine learning or generative AI (GenAI), although the use of that technology may make compliance more challenging.

For further information and resources on AI, see AI toolkit (UK).

Background on public sector AI use

The potential impact and related legal complexity of the use of AI in the public sector is often greater than in the private sector. This is because of the nature of public sector powers and decision-making, and their impact on public life and individuals' rights. The impact is intensified by the significant volume of information (particularly personal data) that is available to public authorities to power AI systems.

The importance of these issues was recognised in a white paper published by the Department for Science, Innovation and Technology (DSIT) in March 2023 (DSIT: AI white paper: a pro-innovation approach to AI regulation (DSIT white paper); see Legal update, Government publishes AI white paper: a pro-innovation approach to AI regulation). For further information on the DSIT white paper, see Practice note, AI: UK regulatory developments.

Successive governments have committed to harnessing the power of AI, including by using it to deliver public services, so its use is only likely to grow. Despite this, the nature and extent of AI use in the public sector is often not well understood.

Proactive transparency

There are relatively few legal requirements for extensive proactive transparency, such as by publishing information on a website or in a register, about public sector AI use.

Common law transparency and publication requirements

In certain cases, a public body may be subject to a common law duty of transparency, particularly in relation to the criteria it uses to determine individuals'

legal rights. For example, in *R (Ames) v Lord Chancellor and others [2018] EWHC 2250 (Admin)*, the High Court held that it was unlawful for a public body to withhold an algorithmic "calculator" used to determine legal aid fee offers.

Similarly, a public body may have a duty to publish its policies (which may also arise under a statutory duty). This is more likely where a previous policy has been published and practice has since departed from that policy (see *XY v Secretary of State for the Home Department [2024] EWHC 81 (Admin)* and Practice note, Public authority policy documents: drafting and challenging policy content). An example might be introducing an AI system to a published decision-making policy and process specification.

The precise content of what must be published varies significantly depending on the AI system in use.

Duty to give reasons

Administrative law recognises a duty to give reasons for some decisions of public bodies, specifically to explain the reasons for a decision to those affected by it. The duty may derive from either statute or common law but its application is far from universal. For further information on the duty to give reasons, see Practice note, Duty to give reasons.

However, where it applies, it seems clear that the duty to give reasons for a decision in which AI was used would extend to explaining that fact, including how the system had either reached a decision or arrived at a recommendation and how that recommendation had been considered by a human decision-maker.

There is no generally applicable standard of reasoning. Without a statutory specification, the precise content and detail required to explain the use of an AI system may vary substantially between cases, depending on the circumstances. However, the principle to be applied was set out in *South Bucks District Council and another v Porter [2004] UKHL 33*:

> "The reasons for a decision must be intelligible and they must be adequate. They must enable the reader to understand why the matter was decided as it was and what conclusions were reached on the 'principal important controversial issues', disclosing how any issue of law or fact was resolved." (*at paragraph 36*.)

The duty to give reasons is owed to the person whom the decision affects. While this does not create a requirement to publish information about AI use, a public body which has already published that information is likely to find it easier to comply with the duty to give reasons in individual cases.

UK GDPR: data processing information provided to data subjects

AI use by a public body generally involves the processing of personal data. This is more likely where there is any automated decision-making on an individual basis, since individuals are necessarily "singled out" by that decision-making. That processing engages the retained EU law version of the General Data Protection Regulation ((EU) 2016/679) (UK GDPR): the public authority is the data controller and the individuals whose data is processed are the data subjects.

For further information on the UK GDPR and its key concepts, see Practice note, UK GDPR: a guide for non-data protection lawyers. For information on the interaction between AI and data protection, see Practice note, AI and data protection (UK).

This note is based on one public body deciding the purposes and means of processing personal data (making it the data controller with obligations under the UK GDPR). However, in practice, multiple public bodies may be involved in the processing, in which case each body must consider whether it is a controller for that processing and, if so, comply with transparency requirements. (Compliance may be done jointly if two public bodies are jointly determining the purposes and means of processing.)

Transparent processing is one of the core principles of the UK GDPR. Article 5(1)(a) provides that "Personal data shall be processed lawfully, fairly and in a transparent manner in relation to the data subject". Articles 13 and 14 elaborate on this, requiring a controller to proactively provide information to data subjects subject to certain exemptions (such as where data subjects already have the information). At a minimum, this information must include:

- The identity of the controller.
- The purpose(s) of the processing and its legal basis.
- Any recipients of data subjects' data. This will be particularly relevant if data is being shared between public bodies as independent controllers (see Practice note, Overview of data sharing arrangements: UK GDPR and DPA 2018).
- The categories of data processed (where the data was not obtained from data subjects).

This minimum information set requires relatively little transparency about the details of any AI system, but it would likely need to include at least a general description of AI use to be intelligible to data subjects.

A controller must proactively provide **further** information to data subjects to the extent "necessary

to ensure fair and transparent processing" (meaningful information requirement). Most notably, this includes:

> "The existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject." (*Articles 13 and 14.*)

The meaningful information requirement is relevant to many but not all instances of public sector AI use (see Practice note, UK GDPR and DPA 2018: profiling and automated decision-making). It requires consideration of several factors, which are somewhat open to interpretation, including:

- Whether Article 22(1) and (4) are engaged.
- What constitutes "meaningful information".
- To what extent the information is required to ensure fair and transparent processing.

The Advocate General's opinion in Dun & Bradstreet Austria GmbH and Magistrat der Stadt Wien (Case C-203/22) EU:C:2024:745 emphasises that the meaningful information requirement requires the controller to provide concise, intelligible and contextualised information that enables the data subject to exercise their other rights. While this is not directly binding in the UK, the additional guidance is useful and emphasises the core principles of the requirement.

The principles-based nature of the meaningful information requirement must be considered in the context of the general principle of transparency under the UK GDPR, which is not limited to the matters set out in Articles 13 and 14. Therefore, even if the meaningful information requirement is not triggered, the general principle of transparency may require a proactive explanation of AI use to data subjects.

The UK GDPR therefore requires public bodies using AI to proactively provide some information to data subjects on their AI processing. The requirement is stricter where the meaningful information requirement is engaged. However, exactly what information is required to be provided to comply with the relatively flexible principles in the UK GDPR is unclear on the face of the legislation.

For more information on transparency under the UK GDPR and in relation to AI, see:

- Practice note, AI and data protection (UK).
- Checklist, Complying with the UK GDPR's transparency requirements.
- Complying with the UK GDPR's transparency requirements toolkit.

FOIA publication schemes

Any public authority subject to the Freedom of Information Act 2000 (FOIA) must adopt and maintain a publication scheme comprising specified information (section 19). The scheme must be approved by the Information Commissioner (IC). The IC's model scheme includes a requirement to publish information on how a public authority makes decisions. For further information, see Practice note, FOIA: model publication scheme.

The section 19 requirement arguably extends to publishing information on the use of AI systems as part of the scheme. However, the section 19 requirement currently has little impact in promoting transparency on the use of AI in the public sector, and even the IC has expressed scepticism about the role of publication schemes in relation to AI use (Committee on Standards in Public Life: Artificial Intelligence and Public Standards (February 2020), *at page 55*).

Internal or responsive transparency

A public body is legally required to provide information on its use of AI responsively as well as proactively.

To comply with these requirements, a public body must have a good internal understanding of its use of AI systems. This may be described as internal transparency on AI use. Arguably, a public body may benefit from proactively publishing this information, or at least some of it, as this is likely to make compliance with responsive transparency requirements easier.

Data protection legislation

Data protection impact assessments

A data protection impact assessment (DPIA) must be carried out where:

> "a type of processing, in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons ..." (*Article 35, UK GDPR*).

Public sector AI use often meets this description and is therefore likely to require a DPIA.

A DPIA must include "a systematic description of the envisaged processing operations", namely a detailed description of how any AI system used by the public body works (*Article 35(7)(a),* UK GDPR). It must be carried out and documented before any processing occurs. Conducting a DPIA requires a high degree of internal transparency about the AI system.

While the ICO encourages publication of DPIAs, this is not a legal requirement.

For more information on DPIAs, see:

- Practice note, AI and data protection (UK): Accountability and DPIAs.
- Practice note, Data protection impact assessments (DPIA) (UK).
- Standard document, Data protection impact assessment (DPIA) (UK).
- Article, AI and privacy compliance: getting data protection impact assessments right.

Data subject access requests

Any data subject whose personal data is processed by a public authority can request copies of their personal data and information about the processing (*Article 15,* UK GDPR). This information is similar but not limited to the information listed in Articles 13 and 14 of the UK GDPR (see UK GDPR: data processing information provided to data subjects). For example, where the meaningful information requirement is engaged, that information must be provided in response to a subject access request, if it has not already been provided to data subjects proactively.

Information sought through Article 15 is not limited to the logic of decision-making. It extends to the (specific) recipients of personal data and the presence of any international transfers (*Harrison v Cameron and other [2024] EWHC 1377 (KB)*, see Legal update, Rights of others subject access exemption applies when there is a significant risk of intimidation (High Court)). Complying with a subject access request necessitates in-depth internal transparency about how public sector AI systems work.

Data subjects also have a range of other rights including the rights to rectification and erasure.

For more information on data subject rights, see Practice notes:

- Data subject rights (UK).
- AI and data protection (UK): Data subject rights.

Accountability

The general principle of accountability in Article 5(2) of the UK GDPR also requires a public body acting as a data controller to "be able to demonstrate compliance" with the data protection principles, including lawfulness, fairness, accuracy and security. The breadth of these general principles is extensive, and it is therefore difficult to see how a public body using AI could comply with the accountability principle without a very high degree of internal transparency about the systems it uses. However, legal challenges and enforcement regarding the accountability principle as a freestanding issue are very rare. For more information on accountability, see Data protection accountability toolkit (UK).

Law enforcement and intelligence services processing

Law enforcement and intelligence services processing are not covered by the UK GDPR. However, Parts 3 and 4 of the Data Protection Act 2018 (DPA 2018) provide analogous regimes for law enforcement and intelligence services respectively, which include obligations to provide information to data subjects, both proactively and in response to a subject access request (see Practice note, Data Protection Act 2018: overview). The DPA 2018 also requires relevant bodies to notify data subjects of solely automated significant decisions taken about them. However, these are relatively rare since few decisions in this context are solely automated.

The DPA 2018 includes extensive exemptions from the obligation to provide information on law enforcement and intelligence services processing to avoid those purposes being prejudiced. Therefore, the content of information that must be provided (if any) varies greatly depending on the AI system and context. (As an example, the Metropolitan Police agreed in 2022 to greater compliance with rights under the DPA 2018 in relation to its "gangs violence matrix", in response to a claim (settled out of court) brought by Liberty. This shows that transparency can be successfully challenged and the standard of transparency improved.)

At a minimum, law enforcement and intelligence agencies must have a good internal understanding of their own AI systems so they can either comply with the requirements of the DPA 2018 or account to the relevant regulator or courts for any exemption to those requirements on which they rely.

For further information, see Practice note, Data Protection Act 2018: overview.

Equality law

A public body must have due regard to the need to:

- Eliminate discrimination, harassment, victimisation and any other conduct that is prohibited by or under the Equality Act 2010 (EqA 2010) (*section 149(1)(a)*).
- Advance equality of opportunity between persons who share a relevant protected characteristic and persons who do not share it (*section 149(1)(b)*).
- Foster good relations between persons who share a relevant protected characteristic and persons who do not share it (*section 149(1)(c)*).

This is known as the Public Sector Equality Duty (PSED), which is explored in more detail in Practice note, PSED: general public sector equality duty under section 149 of the Equality Act 2010.

A notable risk of using an AI system is that it may create bias towards or against certain groups, and that such bias may go unnoticed or be reinforced by a sense of scientific "objectivity". This form of bias may interact with the characteristics protected by the EqA 2010.

It is therefore difficult to see how any public body using AI can comply with the PSED without an in-depth understanding of its relevant systems, to identify whether these systems risk bias and consequently breach the PSED in respect of one or more protected characteristics.

This issue was demonstrated and explored in detail in *R (Bridges) v South Wales Police [2020] EWCA Civ 1058*. The Court of Appeal (CoA) held that the defendant police force had breached the PSED by failing to use information about the AI facial recognition system it deployed which would have enabled it to understand the risks of discrimination on the basis of sex and race (*at paragraph 199*, in particular). The CoA indicated that the PSED could require a public body to have access to the data used to train an AI system, or at least an independent assessment of that data. It was insufficient that the recommendation of the AI system required "confirmation" by an individual officer of the defendant.

A public authority may also mitigate the risk of breaching the EqA 2010 by regularly performing equality impact assessments (EIAs), which is considered good practice. However, there is no requirement to proactively publish an EIA.

Public law principles

The actions and policies of a public body can be challenged by way of judicial review if these are perceived to have breached any of the public law principles. Therefore, where a public body has used AI in the context of a decision or policy under challenge, it is likely to be required to account for the operation of that system in response.

Some grounds of judicial review are particularly relevant to the use of AI:

- Overreliance on an AI system's recommendation could breach the principle of non-delegation or could amount to an unlawful fettering of discretion (see Practice note, Delegation of statutory powers).
- It may be argued that the use of an AI system involves taking into account irrelevant considerations, or a failure to take account of relevant ones (see Practice note, Decision-making by public bodies: avoiding legal challenge).

Even if a public law claim did not challenge the use of an AI system per se, it could still require a public body to explain its use of AI where that use is relevant to the subject matter of the challenge.

In Bridges, an important ground of judicial review was that the use of the facial recognition system breached the European Convention on Human Rights (ECHR). Responding to the ground required the defendant police force to give a detailed account of its use of the AI system. In 2020, the Home Office's use of an AI visa streaming tool was challenged, similarly requiring the public body to give an account of the tool (which it later stopped using).

The need to be able to demonstrate compliance with administrative law is therefore an important driver of the need for a high degree of internal AI transparency in public bodies.

Procurement compliance

The government has encouraged contracting authorities purchasing AI technology to take the following steps:

- Draft invitation to tender (ITT) questions that elicit information from potential bidders about the relevant algorithms and models, including variable selection, AI techniques and the data that the supplier trained the algorithm on.
- Avoid "black box" algorithms and underline the need in ITTs for an "explainable approach to AI development".
- During the selection and evaluation phases, ask suppliers to demonstrate that they avoid unfair discriminatory outputs, and include processes to ensure accountability for algorithms' outputs.

(Department for Digital, Culture, Media and Sport and others: Guidelines for AI procurement (June 2020).)

The Public Contracts Regulations 2015 (SI 2015/102) do not require suppliers' responses to be proactively published (and this is also not intended under the Procurement Act 2023). However, a public authority's awareness of this supplier information significantly assists its internal transparency on AI use.

Limitations to AI disclosure requirements

Duties to provide detailed information about AI systems, whether proactively or responsively, are likely to conflict with other interests and legal requirements affecting public bodies as follows:

- That information may be the subject of commercial confidence where the AI system is provided to the public body by a private sector actor. This was a

feature in Bridges, where the police force could not obtain the training data used for its facial recognition system.

- Providing that information may undermine the integrity of the AI system, for example, by compromising its security. This dynamic is common to many areas where public bodies use AI tools in an "adversarial" context in which it would be disadvantageous for people to know how to manipulate or "game" an AI system.

How these limitations on transparency fit into the legal framework(s) which require transparency is a complex issue. Under the UK GDPR, the principle of transparency is general and flexible. Courts and regulators take into account a wide variety of factors such as the balance of power between the controller and data subject, the complexity of the processing, and the (potential) impact on rights and freedoms (see Checklist, Complying with the UK GDPR's transparency requirements). The meaningful information requirement is equally flexible: it need only be provided to the extent necessary to ensure fair and transparent processing (*Articles 13 and 14*). In the context of a data subject access request, transparency does not need to be provided if to do so would adversely affect others' rights.

Other requirements, such as the duty to give reasons and the PSED, vary as to their substantive content, depending on the circumstances. This may allow contrary considerations to be taken into account. In *Bridges*, the Court of Appeal stated that:

> "We acknowledge that what is required by the PSED is dependent on the context and does not require the impossible ..." (*at paragraph 181*).

The principles-based nature of the various transparency requirements means that courts and regulators are likely to strike a balance between the need for transparency and arguments against it. The government's flagship ATRS recognises the need for balance:

> "Some use cases for algorithmic tools include identifying potential risky applications for a service or highlighting possible fraud. In such circumstances, providing too much information about how an algorithmic tool works, or the specifics of the datasets it draws on, could compromise the operational effectiveness of the tool. For example, a malicious user might modify their behaviour to avoid triggering a fraud warning.
>
> In most cases, such issues can be managed by being careful about the level of detail provided in the algorithmic transparency record, especially around the technical design or data used. Wider information, for example on how the algorithmic tool is used in the overall decision-making process may still be safe to release and relevant."

(Central Digital and Data Office (CDDO) and DSIT: Algorithmic Transparency Recording Standard Hub (updated March 2024).)

Similarly in the EU, the Advocate-General's opinion in Dun & Bradstreet Austria envisages that where compliance with transparency obligations could infringe other rights such as trade secrets, all relevant information should be provided to a court or regulator, so that an independent arbiter can strike a proper balance between (for example) commercial confidentiality and data subject rights.

However, there are few if any examples of how this balance has been struck in practice. This means that a public body must make its own judgement about the level of transparency to provide on its AI use, and these conclusions may be challenged by citizens and civil society groups seeking a better understanding of AI's role in public administration.

Guidance on AI transparency in the public sector

Although (and perhaps because) the substantive content of proactive and responsive transparency requirements is not very prescriptive, a significant volume of guidance has been published on how public bodies can comply with their legal requirements on AI transparency. The guidance comprises:

- Regulatory guidance from the regulators of data protection and equality law.
- Other guidance and standards developed by government.

Regulatory guidance on AI transparency

ICO guidance on AI transparency

While the ICO has promulgated statutory codes of conduct in some areas, which have a special status under the DPA 2018, its guidance on public sector AI use is not binding on public bodies. However, it would likely be considered by the ICO in any regulatory investigation or enforcement action, and possibly by the courts when assessing compliance with transparency requirements under the UK GDPR or DPA 2018 (or even from other areas), but this is less certain. For further information on regulatory enforcement, see Practice note, UK GDPR, DPA 2018 and PECR: enforcement, sanctions and remedies.

The ICO has published a wide range of guidance on AI which is relevant to public bodies (see Practice note,

AI and data protection (UK): ICO guidance relevant to AI). Though its general guidance on AI and data protection refers to transparency, it broadly restates the requirements of the UK GDPR and DPA 2018.

The ICO's most relevant guidance is Explaining decisions made with AI (with the Alan Turing Institute) (ExplainAI guidance) as well as its consultation on GenAI (though this is not yet finalised). The ExplainAI guidance was developed to aid compliance with both the UK GDPR and the other legal requirements for transparency.

Equality and Human Rights Commission guidance on AI transparency

Guidance from the Equality and Human Rights Commission (EHRC) does not have special statutory status and the EHRC has significantly fewer powers of enforcement than the ICO (see Practice note, Equality and Human Rights Commission). Its guidance is less useful as a guide to enforcement, since questions of transparency deriving from equality law are more likely to be resolved through litigation than by the EHRC.

However, the EHRC's commentary on public sector AI transparency in Artificial intelligence in public services (September 2022) may help public bodies seeking to comply with transparency obligations for AI driven by the PSED. The EHRC's The Public Sector Equality Duty and data protection (September 2024) does not provide substantive guidance on what transparency requires in practice (see Legal update, Equality and Human Rights Commission publishes guidance on PSED and data protection).

Government guidance and standards on AI transparency

Other government guidance on public sector AI transparency is not binding on public bodies. It would not necessarily be considered by a regulator with an AI transparency remit such as the ICO, which has set out its own guidance, nor would it necessarily be followed by courts in determining the extent of AI transparency obligations. However, a public body is likely to be in a better position to comply where it follows government guidance, which may form the basis for future legally binding requirements.

General government guidance on AI transparency

Various government bodies have promulgated guidance partly addressing public sector AI transparency, including:

- DSIT white paper.
- Cabinet Office and CDDO: Generative AI Framework for HM Government (GenAI framework).
- CDDO: Data Ethics Framework.
- Alan Turing Institute: Understanding artificial intelligence ethics and safety (AI ethics guidance).
- Cabinet Office and others: Ethics, Transparency and Accountability Framework for Automated Decision-Making (AI transparency framework).

There is significant overlap between these and other sources of guidance.

ATRS

The ATRS is a voluntary means for public bodies to proactively publish information on their use of AI. It is designed to increase public trust in the use of AI, rather than to help public bodies comply with specific transparency obligations (see Practice note, Legal aspects of AI (UK): Algorithmic Transparency Reporting Standard Hub).

The ATRS has seen very limited use: only nine reports have been made against it from across the whole of government as at 3 October 2024. This compares unfavourably with a tracker developed by the Public Law Project using open-source information and FOIA requests. The government has stated that it is "mandatory" for all government departments to comply with the ATRS (though merely encouraged for other public bodies) (DSIT: A pro-innovation approach to AI regulation: government response (February 2024), *section 6*).

It is unclear how the ATRS will be made mandatory, how this will be enforced and how the ATRS relates to other government guidance on AI use. However, given the extent and variety of that guidance, a degree of standardisation in what the government expects of public bodies as best practice would likely be welcomed.

Practical effects of AI transparency guidance

Legal requirements for AI transparency are principles-based, context-dependent and flexible. This has led to a significant amount of guidance being published and arguably some difficulty ascertaining where the law ends and where guidance begins.

Although none of the guidance is strictly binding on public bodies, they have good reasons for following it because:

- The basis of a challenge to a public body on its use of AI may be the argument that following the guidance is what the law requires in order to comply with the general principle of transparency in a specific case. If a public body can demonstrate that it has already complied with government guidance, this may avoid legal challenges or at least negate specific grounds of claim.

- Guidance may form the basis for future compulsory requirements. This has already been observed for government departments and the ATRS, which could in time be made mandatory for other public bodies.
- A public body taking a maximalist approach to what it must publish about its AI use is likely to find it easier to comply when called on to provide responsive transparency, such as when challenged on its policies by way of judicial review.
- In the long term, public trust in the public sector's use of technology is likely to increase through greater transparency. This is explicitly recognised in the DSIT white paper and its renewed commitment to the ATRS.

Despite this, a public body must make nuanced judgements about how much of the guidance to follow in each case, particularly given that the guidance itself (such as the ExplainAI guidance) explicitly provides for varying levels of transparency depending on the context.

It is unrealistic (and likely unpracticable) for a public body to follow all relevant guidance all the time given it is sometimes contradictory. Further, much government guidance is relatively general or cross-refers to one or two of the most authoritative sources. For example:

- The DSIT white paper emphasises the importance of transparency but does not discuss substantive issues, referring instead to the ATRS (*at paragraphs 100-104*).
- The AI transparency framework covers transparency relatively briefly, cross-referring to the ExplainAI guidance.

For a public body using AI, it is likely to be more practical to focus on the more authoritative and comprehensive guidance and standards, namely the ExplainAI guidance and the ATRS, and consult other guidance only as necessary (see Practice note, AI: UK regulatory developments).

Given the uncertainty about the full extent of public bodies' transparency obligations, private technology providers are likely to be well-positioned where they can show that their products enable public authorities to comply with the guidance.

Disclosing the existence and use of AI systems

Where personal data is processed through the use of AI, the UK GDPR sets the minimum standard for informing data subjects about the key facts of that processing, such as the responsible controller and its purpose. This is often done by way of general privacy notices.

With the ATRS reporting standard purportedly mandatory for central government, the existence of AI systems should be disclosed under the ATRS by any central government department. Other public bodies have a choice about whether to use a privacy notice or the ATRS (or both), but there may be benefits to these bodies of following the ATRS in case its use and application is expanded in the future.

The relevant law and guidance means that, a public body must audit at a minimum:

- Whether and where they are using AI.
- Whether their AI use:
  - involves the processing of personal data;
  - is in the context of a duty to give reasons; and
  - represents departure from published policy, triggering the duty of conformity or publication.
- The nature and significance of any AI-assisted (or fully automated) decisions.

If the public body is a government department, it must disclose the existence of all AI systems under the ATRS. If it is not a government department, the public body may disclose its use of AI either through the ATRS or by otherwise providing it to data subjects or those affected by the relevant decision or policy (which may be done through a privacy notice in some cases) if its AI use either:

- Involves the processing of personal data.
- Is in the context of a duty to give reasons.
- Represents departure from published policy, triggering the duty of conformity or publication.

Compliance with the ATRS does not necessarily equate to compliance with the UK GDPR or public law principles. This is because merely publishing the information online or through the ATRS may not constitute informing the subject of the processing or the decision, particularly where there is a more direct data collection or decision-making relationship.

As best practice, a public body should also consider how it can be open about its use of AI through its annual FOIA publication scheme.

In most cases, a public body must go beyond merely disclosing the existence of AI systems and explain the models used and how it reaches its recommendations and decisions.

Effects of law and guidance on public bodies

Explaining AI models and decisions

A public body must determine whether, as well as disclosing its use of AI, it is legally required to publish

or otherwise provide an explanation of how its systems work. This obligation may arise from (for example) the UK GDPR general principle of transparency, the meaningful information requirement or the duty to give reasons (among others). Both the UK GDPR and relevant public law principles are relatively broad, so even if there is no clear legal duty to provide an explanation, a public body should consider doing so. This has the additional benefit of improving its internal and responsive transparency.

The standard of explanation required for AI systems may be similar, even where they are derived from different legal frameworks. Standards in different areas may converge, with administrative law drawing on precedent from data protection law and vice versa.

As a starting point, a public body should use a combination of the ATRS and the ExplainAI guidance to develop its explanation. In summary:

- The ATRS recognises that, generally, a public body must provide an explanation of its AI systems. Both the general and more detailed sections provide for explanations of various aspects of an AI system. However, the ATRS offers only limited guidance on how AI decision-making should be explained.
- The ExplainAI guidance provides more detail than the ATRS on the kinds of explanation which might be appropriate for certain AI systems and how to prioritise which information to provide in specific contexts. A range of examples is provided, and guidance is given on how to develop and implement AI systems in ways which enable explanations to be readily developed.

The most appropriate explanation types to include in an ATRS return (or other transparency document) depends on the source of the legal requirement. For example:

- A **fairness-based** explanation, which accounts for how potential bias has been identified and reduced, would be most appropriate in demonstrating compliance with the PSED.
- A **responsibility-based** explanation, accounting for how an AI system feeds into human decision-making, would be most appropriate in response to a challenge relating to delegation and fettering of discretion.
- An **outcome-based** explanation, accounting for the principal reasons for an individual recommendation, would be most appropriate in the case of the duty to give reasons.

To comply with the general principles of transparency in data protection and public law, a range of explanation types are likely to be required.

The amount of detail required varies according to the domain in which the AI system is used and its impact on individuals, which is explicitly recognised by the ExplainAI guidance. A public body should also consider the kinds of transparency that its stakeholders are likely to need. For example, considering concerns that have been raised by individuals regarding automation in the relevant domain can help them to be addressed in transparency information.

Where particularly complex factors are present, further guidance should be consulted:

- The Alan Turing Institute's Understanding artificial intelligence ethics and safety provides deeper analysis in some areas than the ExplainAI guidance and may therefore be useful where providing an explanation proves especially challenging.
- The Generative AI Framework for HM Government, where the use of generative AI specifically brings unique challenges in complying with transparency principles.

Although the ATRS and ExplainAI guidance allow for flexibility, following them implies that a public body has a high and exacting standard of transparency because:

- The public body may require access to training data. If this is unavailable, it may require independent evaluations of that data for issues such as accuracy and bias.
- Commercial confidentiality does not override the legal requirement to provide explanations. The public body should not procure AI systems where such concerns would prevent it from providing a meaningful explanation of how the systems operate.
- The public body must be able to understand the AI systems it uses and explain them in an intelligible way. It is no defence to say that the system is a "black box" or too complex to be explained. The public body must make reasonable enquiries to gather the information required to comply with the ATRS and ExplainAI guidance (and other relevant guidance).
- While some allowance may be made for withholding information to protect the integrity of AI systems and prevent fraud or manipulation of the systems, this should not result in the failure to provide any explanation of how the systems operate.

The stringency of the standards set by this guidance reflects that relatively little case law exists in this area.

Completing the ATRS, as supplemented by the ExplainAI guidance and other government guidance if needed, goes a long way towards a public body complying with most legal requirements for proactive transparency about public sector AI use. But a public body should consider the following points to ascertain if by abiding by that guidance it has met its legal requirements:

- Whether information which is required to be provided to individuals has been provided. Filing an ATRS return is not necessarily sufficient in all cases.
- Whether the information provided is intelligible to the individuals entitled to it. This is an explicit requirement of the UK GDPR and implicit in other legal requirements for transparency. This is a particular issue given the emphasis the ATRS places on detailed and somewhat formulaic technical information, which may not be intelligible to those affected by public sector AI use.

Enforcement and consequences of non-compliance

A public body's failure or refusal to provide transparency about its AI use to the required legal standard may result in legal challenge and regulatory action by the ICO.

A legal challenge could either be:

- A private law challenge under the UK GDPR. This would most likely be seeking a compliance order from the court that the information be provided to the data subject under section 167 of the DPA 2018. In theory, the court could award compensation for loss, including distress, caused by a transparency breach of the UK GDPR. But it is difficult to conceive of a kind of compensable loss that would be caused by a breach of the transparency principle alone, so this is unlikely in practice.
- A judicial review challenge. If successful, the Administrative Court could order the public body to provide the information (for example, to provide reasons or to publish an updated policy). Compensation would not be available unless the failure to provide transparency constituted or was closely linked to another cause of action, such as a breach of the ECHR under the Human Rights Act 1998 (see Practice note, Human Rights Act 1998: overview).

The ICO can in principle order a public body to take steps to comply with the UK GDPR, for example by improving the transparency information it provides about its AI use (see Practice note, UK GDPR, DPA 2018 and PECR: enforcement, sanctions and remedies).

The ICO has published Regulating AI: The ICO's strategic approach (April 2024), which indicates a strong focus on guidance, advice and support to help organisations to comply. In cases of non-compliance, the ICO could take formal enforcement action against a public body. However, the public body would likely be given extensive prior opportunity to remedy its transparency information without the need for formal enforcement.

In an extreme case, the ICO could also issue a fine to a public body for a breach of AI transparency requirements.