How to use AI and personal data appropriately and lawfully

Increasingly, we are seeing an uptake in applications of artificial intelligence (AI). The use of AI has the potential to make a significant difference to society. However, to realise the benefits, there needs to be confidence that AI is being deployed appropriately and lawfully.

The use of AI transcends across different regulatory remits and therefore, regulators need to have a grounding in the lawful and appropriate use of personal data and AI systems, and its opportunities and risks. This grounding will help regulators when they need to assess the use of AI in the sector(s) they regulate. It will also help regulators who are using AI as part of their regulatory remit to do so appropriately and lawfully.

ico.
Information Commissioner's Office

How to improve how you handle AI and personal information

You can improve how you handle AI and personal information by following these tips:

1. Take a risk-based approach when developing and deploying AI

You should assess whether you need to use AI for the context you will deploy it in. AI is generally considered a high-risk technology and there may be a more privacy-preserving and effective alternative.

If you use AI, you need to assess the risks and implement appropriate technical and organisational measures to sufficiently mitigate them. You cannot realistically remove every risk, and data protection law does not require you to do so.

You should:

- carry out a data protection impact assessment (DPIA) to help identify and minimise the risk of non-compliance with data protection legislation that your use of AI poses, as well as preventing harms to individuals; and
- consult with different groups who may be affected by your use of AI to better understand the risks.

You must complete a DPIA when your processing is likely to result in high risk to people. When a DPIA is legally required, you must carry it out before you deploy an AI system and include appropriate technical and organisational measures to mitigate or manage the risks you identify. If you identify a risk that you cannot sufficiently mitigate, then you are legally required to consult with the ICO before any processing takes place.

Further reading: What do we need to consider when undertaking data protection impact assessments for AI?

2. Think carefully about how you can explain the decisions made by your AI system to individuals affected

It can be difficult to explain how an AI system arrived at a decision, especially when it uses machine learning. However, people still have a right to obtain a meaningful explanation.

You should:

- be clear, open and honest with people from the start about how and why you use their personal data;
- consider what explanation is needed in the context that you will deploy your AI system in;
- assess what people are likely to expect your explanation to look like;
- assess the potential impact of the decisions made by your AI system to understand how comprehensive your explanation needs to be; and
- consider how you will handle individual rights requests.

Further reading: What goes into an explanation?; How do we ensure individual rights in our AI systems? | ICO

3. Collect only the data you need to develop your AI system and no more

AI systems often require lots of data. This can seem at odds with data protection law which requires you to minimise the amount of personal data you use. However, you can still use AI.

You should:

- ensure that the personal data you use is accurate, adequate, relevant and limited. This will vary depending on the context you are using AI in; and
- consider which privacy-preserving techniques are appropriate for the context you are using AI to process personal data in. For example, in the training phase you could use perturbation (adding 'noise'), synthetic data or federated learning to minimise the personal data being processed.

Further reading: What data minimisation and privacy-preserving techniques are available for AI systems?

4. Address risks of bias and discrimination at an early stage

There are several ways that an AI system can be biased or can lead to discrimination, including imbalanced datasets and datasets reflecting past discrimination. Addressing these risks early can make a huge difference.

You should:

- assess whether the data you are gathering is accurate, representative, reliable, relevant, and up-to-date with the population or different sets of people that you will apply the AI system to; and
- map out the likely effects and consequences of the decisions made by the AI system for different groups and assess whether these are acceptable.

Further reading: How should we address risks of bias and discrimination?

5. Take time and dedicate resources to preparing the data appropriately

The quality of the output of an AI system is dictated by the quality of the input. Taking time and dedicating resources to preparing the data to train an AI system will result in better outputs.

You should:

- have clear criteria and lines of accountability about the labelling of data involving protected characteristics or special category data (or both);
- consult with members of protected groups to define the labelling criteria; and
- involve multiple human labellers to ensure consistency and to assist with unusual cases.

Appropriately labelled data can lead to fairer outcomes for people. You must ensure that the data you hold remains accurate, up-to-date, and relevant.

This can also be a way of demonstrating your accountability and transparency requirements.

6. Ensure that your AI system is secure

AI systems can exacerbate security risks or create new ones. There is no one-size-fits-all when it comes to security measures. However, you are legally required to implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk.

You could:

- carry out a security risk assessment, which includes an up-to-date inventory of all AI systems to allow you to have a baseline understanding of where potential incidents could occur;
- carry out model debugging (ie the processing of finding and fixing problems in your model), either by someone internal, or an external security auditor, on a regular basis; and
- proactively monitor the system and investigate any anomalies.

Further reading: How should we assess security in AI?

7. Ensure that any human review of decisions made by AI is meaningful

When using AI for decision-making, you should decide whether to use it to support a human decision-maker, or whether it will make solely automated decisions. People have the right not to be subject to solely automated decisions with legal or similarly significant effects. If you are making such decisions, people can request a human review of the decision made about them.

To ensure human review is meaningful, you should ensure human reviewers are:

- adequately trained to interpret and challenge outputs made by the AI system;
- senior enough and have the authority to override an automated decision.

- taking into account other additional factors that weren't included as part of the input data.

Further reading: Automated decision-making and profiling | ICO

8. Work with the external supplier to ensure your use of AI will be appropriate

Procuring an AI system from a third party does not absolve you of responsibility for complying with data protection law. In most cases, you are likely to be the controller as you will decide how to deploy the AI system. Therefore, you need to be able to demonstrate how the AI system complies with data protection legislation.

You should:

- choose an appropriate supplier by carrying out due diligence ahead of any procurement;
- collaborate with the external supplier to carry out an assessment prior to deployment (eg a DPIA);
- agree and document roles and responsibilities with the external supplier (eg who will answer individual rights requests or who will carry out security tests);
- ask to see documentation from the external supplier that demonstrates they took a data protection by design approach; and
- consider whether there will be any international transfers of personal data. If there are, ensure people's personal information rights are protected, or if one of a limited number of exceptions applies.

Further reading: Contracts and liabilities between controllers and processors

Artificial intelligence and personal information – frequently asked questions

We want to add to these questions in this document over time. Therefore, we encourage you to submit questions you would like our help with AI@ico.org.uk.

1. If we plan to use AI, do we have to carry out a data protection impact assessment (DPIA)?

In the vast majority of cases, the use of AI will involve a type of processing likely to result in a high risk to individuals' rights and freedoms. It will therefore trigger the legal requirement for you to undertake a DPIA. You need to make this assessment on a case-by-case basis.

You are required to undertake a DPIA if your use of AI involves:

- systematic and extensive evaluation of personal aspects based on automated processing, which produce legal or similarly significant effects for individuals;
- large-scale processing of special categories of personal data; or
- systematic monitoring of publicly-accessible areas on a large scale.

Beyond this, AI can also involve several processing operations that are themselves likely to result in a high risk. For example, use of new technologies or novel application of existing technologies, data matching, invisible processing, and tracking of location or behaviour. You are required to carry out a DPIA when these involve activities such as evaluation or scoring, systematic monitoring, or large-scale processing or both.

In any case, if you have a major project that involves the use of personal data it is also good practice to carry out a DPIA. If you decide that your use of AI does not require a DPIA, you should document the reasons why. DPIAs are an ideal opportunity for you to consider and demonstrate your accountability for the decisions you make in the design or procurement of AI systems.

Further reading – ICO guidance

When do we need to do a DPIA? | ICO

What are the accountability and governance implications of AI? | ICO

2. Do the outputs of an AI system have to comply with the accuracy principle under data protection law?

'Accuracy' has different meanings in the contexts of data protection and AI.

Accuracy in data protection is one of the fundamental principles, requiring you to ensure that personal data is accurate and, where necessary, kept up-to-date. You must take all reasonable steps to make sure the personal data you process is not "incorrect or misleading as to any matter of fact" and, where necessary, is corrected or deleted without undue delay.

Example

If a person moves house from London to Manchester, a record saying that they currently live in London will obviously be inaccurate. However, a record saying that they once lived in London remains accurate, even though they no longer live there.

Accuracy in AI (and, more generally, in statistical modelling) broadly refers to how often an AI system guesses the correct answer, measured against correctly labelled test data.

Data protection's **accuracy principle** applies to all personal data, whether it is information about a person used as an input to an AI system, or an output of the system. However, this does not mean that an AI system needs to be 100% **statistically accurate** to comply with the accuracy principle.

Example

An organisation develops an AI system to predict which region of the country people live in. They base this on data taken from copies of utility bills to inform research about where energy prices are increasing the most. The organisation takes all reasonable steps to make sure that the output of the AI system is not incorrect or misleading as to any matter of fact. They also recognise that the system will not be 100% statistically accurate as they find it fails to effectively handle copies of utility bills that are low quality (eg where words are blurry). Therefore, the organisation puts in place measures to ensure that any incorrect outputs of the AI system can be quickly remedied. In this instance, the organisation has satisfied the accuracy principle, despite the AI system not being 100% statistically accurate.

In many cases, the outputs of an AI system are not intended to be treated as factual information about the person. Instead, they are intended to represent a statistically informed guess as to something which may be true about the person now or in the future.

If you use an AI system to make inferences about people, you need to ensure that the system is sufficiently statistically accurate for your purposes. This does not mean that every inference has to be correct. But you do need to factor in the possibility of them being incorrect and the impact this may have on any decisions that you may take on the basis of them. Failure to do this could mean that your processing is not compliant with the fairness principle. It may also impact on your compliance with the data minimisation principle, as personal data, which includes inferences, must be adequate and relevant to your purpose.

Further reading – ICO guidance

What is the difference between 'accuracy' in data protection law and 'statistical accuracy' in AI?

3. What steps can we take to avoid bias and discrimination in our use of AI?

There are different reasons why an AI system might lead to discrimination. Steps to avoid bias and discrimination will depend on what the causes are.

One reason is imbalanced training data. For example, the training data for an AI system designed to predict whether someone will pay back a loan may include a greater proportion of male borrowers. This is because in the past fewer women applied for loans and therefore the bank doesn't have enough data about women. It may be possible to balance it out by adding or removing data about under or over-represented subsets of the population. Although, you should consider the impacts this might have on the subsets of the population who you are collecting more data from.

Another reason is that the training data may reflect past discrimination. For example, training data for an AI system designed to assist in recruitment decisions may reflect discrimination where men have historically been considered to be more suitable candidates for certain roles. You could either modify the data, change the learning process, or modify the model after training.

There could be bias in the way variables are measured, labelled or aggregated, or developers could demonstrate biased cultural assumptions. If this is due to human labellers, you should ensure there are clear criteria and lines of

accountability about the labelling of data involving protected characteristics or special category data. You could consult with members of protected groups or their representatives to define the labelling criteria. When labelling data, you should create criteria that are easy to understand; include descriptions for all possible labels; examples of every label; and cover cases which require special handling. You should involve multiple human labellers to ensure consistency.

Further reading – ICO guidance

How should we address risks of bias and discrimination?

4. How can we comply with the data minimisation principle when developing an AI system?

You are required to identify the minimum amount of personal data you need to fulfil your purposes, and to process only that personal data, and no more.

However, AI systems generally require large amounts of data. At first glance it may therefore be difficult to see how AI systems can comply with the data minimisation principle. Yet, if you are using AI as part of your processing, you are still required to do so.

Whilst it may appear challenging, in practice this may not be the case. The data minimisation principle does not mean either 'process no personal data' or 'if we process more, we're going to break the law'. The key is that you only process the personal data you need for your purpose.

How you go about determining what is "adequate, relevant and limited to what is necessary" is therefore going to be specific to your circumstances. (Our existing guidance on data minimisation details the steps you should take).

In the context of AI systems, what is "adequate, relevant and limited to what is necessary" is therefore also case specific. One practical step is to map out all the areas of the AI pipeline where you might use personal data. You should schedule time at each significant milestone to review whether you still need this data for your purpose.

Further reading – ICO guidance

What data minimisation and privacy-preserving techniques are available for AI systems?

Principle (c): Data minimisation | ICO

5. What counts as a solely automated decision with legal or similarly significant effects?

The UK GDPR restricts you from making solely automated decisions, including those based on profiling, that have a legal or similarly significant effect on individuals.

A legal effect is something that affects someone's legal rights. Similarly significant effects are more difficult to define but include, for example, automatic refusal of an online credit application, and e-recruiting practices without human intervention.

A decision is solely automated if there is no meaningful input by a human in the final decision being made about a person. You should be aware that a decision does not fall outside this scope just because a human has 'rubber-stamped' it. The degree and quality of human review and intervention before a final decision is made about a person are key factors. This determines whether you are using an AI system for automated decision-making or merely as decision-support.

In more detail – ICO guidance

We have published detailed guidance on automated decision-making and profiling.

6. Is using AI legal?

There is no part of data protection legislation that explicitly regulates the use of artificial intelligence, and no part that prohibits its use.

Data protection law adopts a risk-based approach. This means you need to carefully assess the risks that your use of AI poses to people's rights and freedoms. You then need to implement appropriate measures to sufficiently mitigate or manage those risks. If you cannot do so, you must consult with the ICO prior to starting your processing. In some cases, the ICO may decide that you cannot go ahead with your proposed use of AI because the risks to people are too high.

7. Do we need people's permission to analyse their data with AI?

You need a lawful basis to process personal data whether it involves the use of AI or not. For some kinds of data, such as special category data, you need a lawful basis and must also satisfy a further condition. There are multiple lawful bases available, and the most appropriate one will depend on your circumstances.

Consent may be an appropriate lawful basis in cases where you have a direct relationship with the people whose data you want to process. However, it may be difficult to collect valid consent for more complicated processing operations, such as those involved in AI. For example, the more things you want to do with the data, the more difficult it is to ensure that people understand everything you are planning to do. As a result, it is harder to ensure that their consent is freely given and informed. Relying on consent also means allowing people to withdraw consent, which can be difficult to manage and means that you then need to identify a new lawful basis if you continue the processing.

An additional consideration is the requirement to inform individuals about how you are collecting and using their personal data. This applies whether you collected the data directly from the individual or from a different source. The information you provide to people must be concise, transparent, intelligible, easily accessible, and use clear and plain language.

In very limited circumstances, you may not need to provide people with privacy information. For example, if it would involve disproportionate effort to provide it to people.

Further reading – ICO guidance

What do we need to do to ensure lawfulness, fairness, and transparency in AI systems? | ICO

Special category data | ICO

Exemptions | ICO

8. Do we need to publish an AI policy?

There is no part of data protection legislation that explicitly regulates the use of AI, and so you are not required you to publish an AI policy.

You must ensure that you inform individuals about how you are collecting and using their personal data, including when you are using AI to process their personal data. How you present this information depends on how and why you are processing personal data. You may choose to publish an AI policy, but the law does not mandate it.

9. Do we need to understand how a third-party supplied AI system works?

The accountability principle under data protection law means the controller is responsible for complying with the data protection principles and must be able to demonstrate this compliance. When procuring an AI system from a third party, it is important to agree who will be the controller, joint-controller, or processor. In some cases, you will take on the role of controller. This may be the case if you make decisions about what the target output of the model will be (ie what is being predicted or classified), or about what features will be used in the model.

As the controller, you need some level of understanding of how the AI system works to comply with data protection principles. For example, you need to understand how the AI system makes a decision because you have an obligation to explain this to people.

In some cases, you may be a joint controller with the third party supplying the AI system. In these cases, you may allocate different responsibilities related to data protection compliance. However, both parties remain responsible for complying with data protection legislation.

Further reading – ICO guidance

What are the accountability and governance implications of AI? | ICO

Controllers and processors | ICO

Principle (a): Lawfulness, fairness and transparency | ICO