AI and Privacy – 10 Questions to Ask

04 October 2024

UK data protection law applies to AI as it would any other technology.

Companies therefore need to ensure that they meet GDPR standards when processing personal data in the context of any AI used in their business. Although the over-arching legal framework is the same, the nature of AI technology means that businesses are presented with new privacy-related risks that need to be addressed. We set out in this section 10 key questions to ask yourself at the outset when developing or deploying AI solutions in your business. If you intend to use an off-the-shelf AI solution, see 'AI-as-a-Service – Key issues'.

1. What will your AI system do?

It is a basic question – but you need to understand the purpose of your AI system to understand why and how you will be processing personal data in the context of the system. A data protection impact assessment (DPIA) is crucial to assessing the privacy-related risks posed by your system and your proposed mitigations.

Where you are exploring the use of AI or engaging in pilot projects, the purpose of your AI system may not be apparent or may change over time. Similarly, your purposes for processing personal data may change over the lifecycle of the AI system, for example, training the system vs using the system to make decisions. You will need to review your DPIA periodically as your AI system evolves and consider the impact of such change on your compliance e.g. do you now need a different lawful basis to process the data? If you were relying on the legitimate interests lawful basis, do you need to perform a new legitimate interests assessment?

You should also consider whether the data processing carried out by your system is necessary and proportionate, or if there are potentially more privacy-preserving alternatives. Ultimately, your system should incorporate data protection by design and by default.

2. What are your roles and responsibilities?

Many companies are working collaboratively with tech providers to explore AI's potential for their business. This, and the inherent complexity of AI systems, mean that the usual analysis of whether parties are acting as controller or processor becomes a lot more challenging. You should be clear about who is responsible for making decisions about different aspects of the AI model, including if any are

made jointly. This will then drive the contractual framework you should put in place to document the rights and obligations of the parties e.g. data processing agreements and data sharing agreements.

3. What will be used as input data?

AI systems need vast amounts of data for training purposes. You will need to consider the data sets that will be input into your AI system and if you are compliant in respect of each of them. How much personal data and special category data exists in those data sets? Are you able to anonymise data before it is ingested into the system (thereby taking it outside the scope of the GDPR) or will this significantly impact the accuracy of the model? Are you able to use synthetic data sets to train the system instead as these present lower privacy risk? Bear in mind, however, that lower quality data will have an impact on the fairness of the AI system – discussed further below.

4. What will be the output data?

Consider the types of data that will be generated by your AI system, and your GDPR obligations in respect of each of them. The power of AI systems to analyse and find patterns across massive data sets also means that you may be processing personal data without expecting to do so. For example, where you use AI systems to make inferences about individuals or groups, and the inference relates to an identifiable person, this may be personal data or even special category data depending on the circumstances.

5. What are the data flows?

AI models require extensive processing power and are therefore typically hosted by the AI solution provider. Consider how personal data moves through your AI system. Is it sent from your systems to the provider, and if so is it commingled with other customers' data, or are you able to retain it in your own 'walled garden' instance of the AI model? Is it transferred out of the UK? If so you will need to ensure that a transfer mechanism under the GDPR has been put in place and a transfer risk assessment has been carried out. Note that the risks of any transfer are heightened because of the volumes of data processed by the AI system. Consider also how long you retain data and if this aligns with your data retention policies.

6. How do you ensure your AI system is fair?

Your AI systems (and any decisions made by them) must be fair and must not produce outcomes which are discriminatory or biased against individuals. Bias can occur at multiple points in the AI lifecycle and may not be apparent. For example, data sets that reflect historical biases or lack data for certain categories of data subjects may result in AI models being inadvertently trained to perpetuate bias. Therefore, you must plan to mitigate the risk of bias from the outset, for example, assessing the quality and neutrality of data inputs, engaging with a broad range of stakeholders to identify bias, and mapping out the potential effects of AI decisions on minority groups.

7. Will you be carrying out automated decision-making?

Article 22 of the GDPR restricts automated decision-making that produces legal or similarly significant effects for data subjects without any meaningful human involvement. If your AI system is likely to do this, you will need to assess the decision being made and how human involvement should be incorporated

into the process for it to have meaningful effect. Recent EU case law (that is influential on the UK regulator) also shows that "decision" may be interpreted broadly and can encompass even interim acts that play a determining role in the final decision. You should also ensure that any employees involved in the decision understand the importance of their review and that it is not merely a 'tick box' exercise.

8. How do you keep your AI system secure?

The sheer volumes of data used by any AI system exponentially increases the risk of a data breach. Any existing technical and organisational measures you implement to keep your systems secure must be updated to protect against novel security risks faced by AI systems (see 'Procuring AI – Commercial Considerations Checklist' for examples of these). You will need to adopt a more holistic approach to systems security as your AI solution will no doubt be integrated with various internal and third party systems. As industry standards around AI security are still being developed, ensure you remain current with the prevailing best practice from time to time.

9. What must you tell data subjects?

You will need to be transparent with data subjects as to how their personal data is processed by your AI systems. This may be difficult to do, as AI decision-making is sophisticated and frequently opaque. For this reason, you must deliberately design your AI system to be explainable and understandable by your data subjects and you should describe how the decisions made by your AI system are fair and avoid bias. Any consent sought from data subjects must be freely given, specific and informed, and is therefore predicated on the data subject being able to understand what they are consenting to.

10. How do you ensure data subject rights?

Your AI system must be designed to accommodate the data subject rights enshrined in the GDPR throughout the project lifecycle, for example, the rights to access, rectification, and erasure. This may be challenging depending on the volume of data processed, and if data sets are modified or commingled for training purposes. However, you would still need to take reasonable steps to comply with the data subject's request. Related to this, you will also need to consider how data subjects can withdraw consent and the impact of this on your AI system.

In addition to asking yourself these 10 questions, ensure you follow the Information Commissioner's guidance as and when published on developing, deploying and using AI, including the "AI and data protection risk toolkit" and guidance on explaining decision made with AI. The ICO has also just concluded a series of consultations on new guidance regarding generative AI which we will report on in due course.

Discover more insights on the AI guide