Implementing AI — what you need to know

Stephanie Pritchett and Ben Wootton, Partners at Pritchetts Law LLP, consider how organisations can help to protect themselves and ensure GDPR compliance when implementing or trialing AI systems in their organisations

In being presented with the task of writing an article on implementing artificial intelligence ('AI'), it was tempting to ask the AI-powered chatbot, ChatGPT, to create a first version for us and see what it came up with. However, ChatGPT's modus operandi is to form content from huge datasets scraped from the internet. Add in the fact that these datasets may not be accurate — and may have their own GDPR and copyright issues — and you can see why we elected to write the article ourselves!

The same concerns were reflected in the Italian Supervisory Authority's decision to ban ChatGPT in March 2023, citing various data protection issues including lack of transparency and inappropriate legal bases for data processing. Although the ban was subsequently lifted following changes implemented by its developer, OpenAI, ChatGPT remains under investigation by several EU Supervisory Authorities ('SAs') and the European Data Protection Board's ChatGPT Task Force.

Technological development in general, and the field of AI in particular, is evolving fast. With its potential to revolutionise various aspects of our lives, many organisations will be considering how they can implement AI applications to keep up with competitors and drive efficiencies. Organisations may even find that they are using AI tools already without realising it. For example, Microsoft Viva Insights is an AI-powered tool that can deliver a report summarising an individual's activity (or lack of) across the working week. Another app called Microsoft 365 Copilot — due to be rolled out soon — will draft suggested email responses based on previous emails. It all sounds useful, but what do organisations need to consider to ensure GDPR compliance? This article sets out some of the key issues.

A changing landscape

In the EU, the use of AI will be regulated by the proposed AI Act, which in its current form adopts a risk-based approach to AI. Talks are now underway between EU countries in the Council on the final form of the law, with the aim to reach an agreement by the end of 2023.

The UK is also taking strides forward, with the government setting out its principles-based (rather than risk-based) proposals for future AI regulation in March 2023 in its policy paper, 'A pro-innovation approach to AI regulation', and the House of Commons' interim report on the governance of AI on 31st August 2023, which the government is due to respond to in the next two months.

There is clearly significant cross-over between AI and data protection legislation. For example, there are considerations around geographical scope, comprehensive fairness and transparency obligations, accountability and governance principles, security, an assessment of risk to individuals, risk management systems/data protection impact assessments ('DPIAs'), administrative sanctions and national oversight. This level of intersection is not surprising: data are the fuel that powers the AI engines.

10 questions to consider if implementing AI

Earlier this year, the UK Information Commissioner's Office ('ICO') updated its guidance on AI and data protection. The ICO has also provided various reports and opinions, including a blog post listing several questions that organisations developing or using generative AI should consider. In the following sections, we explore those questions and add a couple of our own.

1. What is our lawful basis for processing personal data?

To process personal data lawfully, an organisation must be able to identify (and inform data subjects about) the appropriate lawful basis for such processing.

When controller organisations implement AI, they must determine the

(Continued on page 4)

Stephanie leads various training sessions with PDP, including 'Conducting Data Protection Impact Assessments', 'Data Protection Essential Knowledge — Level 1' and 'Data Protection Essential Knowledge — Level 2'. See the website for further information.

(Continued from page 3)

purposes and means of processing any personal data, and establish, justify and document the legal bases for that processing. This will include understanding the source of any raw or training data used in the systems, and justifying legal bases for processing those data. For example, data collected for one purpose (such as documents for a specific project) might be used to suggest drafting for another, unrelated project.

Organisations will also require a comprehensive understanding of the algorithms applied to the data, the types of personal data that could be created, how the data could be used, and how that use might evolve over time. For example, if relying on consent, organisations would need to consider how to collect valid consent. Is it specific enough for a potentially evolving set of outcomes? Is it freely given? Can it be withdrawn?

Alternatively, if relying on legitimate interests, organisations will need to perform a legitimate interests assessment ('LIA'). It is important to remember that organisations must assess whether they have a legitimate interest in pursuing their purpose for processing personal data, not for the use of an AI tool. The use of the AI tool must be necessary to achieve the purpose.

2. Are we a controller, joint controller or a processor?

The ICO's summary view of this complex question is "if you are developing generative AI using personal data, you have obligations as the data controller. If you are using or adapting models developed by others, you may be a controller, joint controller or a processor."

Organisations should assess what role they and others will play, and how they interlink, so that they can establish appropriate compliance measures. This might be difficult to determine and will require a detailed understanding of how the AI application works, and how it will be used.

> ***"Due to the potentially high risks involved, we expect the regulators to take a stringent approach. They are likely to demand collaboration between developers and controller organisations that develop and implement AI, and impose harsh sanctions on rule-breakers."***

UK and EU SAs have issued guidance on how to understand the complexities of these relationships. We expect this to be expanded in future to give more examples in an AI context. For now, organisations will need to assess their understanding of these roles rigorously and document the outcome.

3. Have we prepared a DPIA?

To assess a controller's relationship with third parties, and the impact on individuals' privacy, organisations should consider preparing a DPIA. Some uses of AI may trigger a mandatory DPIA. For example, a DPIA must be performed where the processing is likely to result in a high risk to the rights and freedoms of individuals. This is very likely to be the case where AI is used, according to the ICO and several EU SAs. In its list of examples, the ICO refers to "artificial intelligence, machine learning and deep learning" as a triggering factor if combined with other factors. It also says it "considers it best practice to do a DPIA, whether or not the processing is likely to result in a high risk".

If the DPIA's outcome suggests that the AI processing would result in high risks to individuals that cannot be mitigated, organisations must consult with the ICO (and other relevant EU SAs) prior to performing any such processing.

The ICO states that organisations "must assess and mitigate any data protection risks via the DPIA process before [they] start processing personal data. [The] DPIA should be kept up to date as the processing and its impacts evolve". This will be a particularly challenging requirement to meet, given the fast-paced changes in AI. Due to the potentially high risks involved, we expect the regulators to take a stringent approach. They are likely to demand collaboration between developers and controller organisations that develop and implement AI, and impose harsh sanctions on rule-breakers.

The UK government's policy paper on AI emphasises the importance of building trust, stating that "public trust in AI will be undermined unless [its] risks, and wider concerns about the potential for bias and discrimination, are addressed." It adds that to "maintain the UK's position as a global AI leader, we need to ensure that the public continues to see how the benefits of AI can outweigh the risks." The ICO's AI guidance reminds organisations of the need to ensure accountability, governance, transparency, lawfulness, security and protection of individual rights. It also sets out compliance requirements to ensure accuracy (including statistical accuracy) and fairness (ensuring bias and discrimination risks are properly considered). The DPIA should help organisations to scope out each of these areas.

DPIAs will need to assess risk mitigation safeguards such as counterfactual analysis, human verification of AI decision-making, circuit-breakers and robust dataset testing. These should be used at the start and continue throughout, to ensure that systems are not discriminatory.

When assessing risks to individuals, most organisations will find the ICO's AI and data protection risk toolkit valuable.

4. How will we ensure transparency?

The GDPR requires that those processing personal data do so fairly and not in a way that is unduly detrimental, unexpected or misleading to

individuals. Apart from rare exceptions, it also requires controllers to be transparent with data subjects from the outset, explaining in full how their personal data will be used in the context of AI.

The GDPR has different rules about how to communicate privacy information to individuals, depending on whether controllers collect their information directly or via a third-party source. In the context of AI, controllers must communicate with individuals about how their personal data have been obtained and will be used to train the AI system, and how they will be obtained and used within the system itself.

Controllers must provide this privacy information clearly. Various methods are available, including privacy notices, dashboards, pop-up notices, icons, and mobile and smart device functionalities.

They must also update privacy notices appropriately and ensure transparency from the initial design and training stages of AI systems through to their implementation. Articles 13 and 14 of the GDPR set out the requirements, including explaining any automated decision-making capabilities in the AI system.

The GDPR requires information to be communicated in "a concise, transparent, intelligible and easily accessible form, using clear and plain language" so that affected individuals can understand it easily. This may be challenging when tasked with describing complicated technical details and analytical processes where the AI is often designed to process data independently. There are similar challenges when communicating how personal data are being processed for the purpose of training the AI system.

Despite these challenges, transparency is key to ensuring the successful implementation of AI systems. Ideally, individuals should understand what is happening, while accepting and supporting it. The ICO's guidance 'Explaining decisions made with AI' is a helpful starting point.

5. How will we mitigate security risks?

The GDPR requires personal data to be processed securely using technical and organisational measures that:

- are appropriate to the nature, scope, context and purpose of the processing and to the risks to individual rights and freedoms; and
- enable organisations to restore access and availability to personal data in a timely manner in the event of a physical or technical incident.

Along with taking account of technology available and implementation costs, security measures must also be appropriate to the specific AI setup and the risks that the AI processing poses. Controllers will need to perform and document assessments to consider those risks, for example through robust AI supplier due diligence and DPIAs. They must also consider additional risks associated with AI, such as access to sensitive data sets, personal data leakage, model inversion, membership inference, data poisoning and adversarial attacks. In addition, they will need strong breach-handling processes and tried-and-tested prevention and detection systems.

6. How will we limit unnecessary processing?

The GDPR data minimisation principle requires that controllers only collect data that are adequate to fulfil their stated purposes, and that the data should be relevant and limited to what is necessary for those purposes. The nature of AI systems necessitates a large amount of data. If that data includes personal data, compliance with the data minimisation principle will be paramount.

Controllers will need to assess, perhaps in detail through the DPIA, how to ensure that:

- they can use the personal data as planned. Were they collected appropriately? Are they being used for a compatible purpose? Have they been kept for longer than necessary for the original purpose?;
- any personal data processed or generated using the AI is 'adequate, relevant and limited' to what is necessary for the identified purposes;
- data protection by design and default, and data minimisation, are implemented from the outset as part of the initial AI design processes. Less privacy-intrusive processes should be considered; and
- specific scrutiny is applied so that data minimisation processes are built into AI systems developed by third-party suppliers.

Controllers should also consider whether they could use anonymised datasets to train the AI systems.

7. How will we comply with individual rights requests?

Controllers must be able to appropriately handle individual rights requests and to explain these rights in their privacy notices. A controller must consider how to comply with individual rights such as subject access requests, and the rights to rectification and erasure. For example, how would they delete personal data that have been consumed within the AI dataset? Controllers must also ensure that people have the right to object to solely automated decisions.

A detailed DPIA should help to establish how to comply with individual rights requests in the context of the specific AI system. Again, the privacy notices should set out how to exercise all of these rights.

8. Will we use generative AI to make solely automated decisions?

The GDPR defines some rules to protect individuals in situations where controllers conduct solely automated decisions that have legal or similarly

(Continued on page 6)

(Continued from page 5)

significant effects on individuals, for example, certain e-recruiting practices and data profiling. It's a complicated area, but UK and EU SAs have produced helpful guidance. Automated decisions can be an effective tool for many organisations, for example to interpret policies and make fair, consistent decisions. They can, however, only use automated decisions if one of the three exceptions set out in the GDPR applies. In summary, these are where the decision is: necessary for a contract; authorised by law; and based on an individual's explicit consent (ensuring that the GDPR's conditions for consent have been met).

Controllers can only process special category personal data if one of the three exceptions above applies and they have the individual's explicit consent, or the processing is necessary for reasons of substantial public interest.

Because this type of processing is deemed high-risk, controllers should definitely conduct DPIAs. They should also provide affected individuals with "meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing" (Article 13(2(f))) and inform them about their GDPR rights.

Implementing suitable safeguards, including meaningful human intervention, is a must. Controllers should be able both to explain how automated decisions were made and to verify the results. This might be challenging, given complex and sometimes opaque AI algorithms. The regulators do not seem to expect organisations to provide technical explanations about how the AI system works, but to explain simply:

- the data used in the decision-making process;
- the source of the information and its relevance;
- the key decision points that formed the basis for the AI decision; and
- whether any alternative decisions could have been made and, if so, why they were not.

These explanations are important to ensure that affected individuals understand the reasoning behind automated decisions made. Controllers must inform affected individuals that they have the right to object to a decision made using automated decisions and provide a process to enable them to do so.

They must also ensure that someone suitably qualified and authorised to change the decision carries out a human review of any AI-generated automated decisions. Therefore, controllers must ensure that the AI system has been developed to enable escalation, resolution and override of automated decisions.

9. Are staff aware of our due diligence processes for AI suppliers?

If procuring an AI system from a third party, controllers must comply with the GDPR's requirement that appropriate due diligence ('DD') is conducted on AI suppliers and systems. The DD should be reviewed regularly given the historically fast-paced development of these systems. DPIAs that suppliers have conducted on their AI systems may form a crucial part of the DD process, so they should form part of the sales strategy for AI developers and suppliers. If the controller's DD flags issues with the AI system, alternative solutions or providers should be considered.

Controllers should inform all workers of GDPR guidelines for appointing processors, including ensuring appropriate contracts are in place that deal with liability around data protection issues, and address any restricted international data transfers. They should also update guidelines and DD checklists to take account of AI risks and mitigations.

10. What are the accountability and governance implications of AI?

The ICO has provided guidance on this topic, setting out how controllers can 'prove' compliance with the GDPR accountability principle in any AI system that processes personal data. If controllers have addressed questions 1-9 above, this final step should be a breeze!

Stephanie Pritchett and Ben Wootton
Pritchetts Law LLP
stephanie.pritchett@pritchettslaw.com
ben.wootton@pritchettslaw.com