AI & DATA PRIVACY: BALANCING TENSIONS

MARCH 2023

DIGITAL
Part of the Horizon Scanning series

Artificial Intelligence already impacts our daily lives – we interact with it every time we use our social media feeds or banking apps – and its use is only set to increase as new tools like ChatGPT catch the world's attention, and we discover new ways for machines to learn and assist us. Legislators and businesses alike therefore need to manage the new risks and opportunities AI development presents, including those relating to the processing of personal data.

This article looks at some of the particular data privacy concerns raised by AI and how the UK's data regulator (the ICO) and the UK government are responding to this evolving challenge. (See end box "*What is AI?*")

Does AI pose particular privacy concerns?

When advising on AI and its privacy risk profile, it is important to understand how and when personal data is used.

While not all AI systems use personal data, a significant number do. Personal data can be processed both when training an AI algorithm and when deploying the AI. AI can even determine whether information falls within the definition of personal data, as the ability of AI to recognise patterns in data, or link data sets, can potentially enable data that would not normally be considered personal data to become "identifiable".

The challenge for organisations using AI, and which are within the scope of the GDPR (UK or EU), is that a number of the typical characteristics of AI seem, at least at first glance, to be at odds with the main principles of data protection law.

We often think of AI involving the processing of large quantities of data, sometimes for new purposes, to produce outcomes where it can be unclear why or how that decision was reached. But how does this sit with the main GDPR principles (see ***"GDPR Principles"*** box) and other UK GDPR rules, for example around the rights of data subjects and automated decision making?

How, for example, can you satisfy the data minimisation principle if you are using a system that allows the AI (machine learning) to conclude what information is necessary from large data sets? And how can you be transparent if you do not know why or how a decision (relating to a loan, or job, application for example) was reached?

AI systems may also exacerbate known risks, making them more difficult to manage. Examples of this include:

- Security: AI can increase the potential for loss or misuse of large amounts of personal data (and these large data sets are often required to train AI systems), new AI related code and infrastructure can introduce software vulnerabilities and security practices and expectations may vary significantly between different developers as common practices around security have not yet been developed. See also our blog on the NCSC's security concerns around LLMs like ChatGPT.

- Bias: Data must be processed fairly and lawfully but if AI systems learn from data which is unbalanced or discriminatory, they may produce outputs which have discriminatory effects on individuals based on their gender, race, age, health, religion, disability, sexual orientation or other characteristics. This could be in breach of both the Equality Act 2010 and the UK GDPRs requirements around fairness.

AI can also make it harder to spot where risks exist. The way AI systems are sometimes developed and deployed can lead to personal data being processed in unusual ways. This can make it harder to understand when individual rights apply to the data, and for the individuals (data subjects) to exercise those rights.

In addition, particular issues arise if AI is used to make automated decisions about individuals which have a legal

GDPR Principles

The UK GDPR contains seven principles relating to the processing of personal data that must be followed (Article 5). These are:

- Lawfulness, fairness and transparency.
- Purpose limitation.
- Data minimisation.
- Accuracy.
- Storage limitation.
- Integrity and confidentiality (security).
- Accountability.

effect on those individuals (for example, in relation to whether the individual's job application is successful). The UK GDPR contains specific provisions[1] which state that people have the right not to be subject to those solely automated decisions. They can request a human reviewer. However, those rules can be hard to apply in practice – for example, what does 'solely automated' cover, and how can organisations offer a human reviewer where the AI system has been introduced because the organisation does not have a sufficient workforce to undertake the task (for example, where thousands of CVs are being screened in a recruitment process)?

The ICO's approach

The ICO has been focussed on the risks around AI for some time now. It has categorised AI as both a strategic and regulatory priority in previous years, and its latest ICO25 strategy document confirms that AI risks will continue to be a focus until 2025. That said, AI should no longer be viewed as a "new risk" and organisations should bear this in mind when assessing the risk profile of adopting an AI solution.[2]

When it first launched its main AI guidance in 2020, the ICO was keen to stress[3] that the underlying questions for even the most complex AI project are much the same as with any new project. Is data being used fairly, lawfully and transparently? Do people understand how their data is being used? Is their data being kept secure?

The ICO does recognise that AI presents particular challenges when answering these questions, with some areas (like data minimisation) appearing particularly problematic at first glance. However, its guidance aims to provide organisations with methodologies and techniques to help mitigate and manage those risks in many cases.

A key ICO message is the importance of considering data privacy at an early stage ("data protection by design and default"). It says that AI increases the importance of embedding this approach into an organisation's culture and processes and that carrying out a data protection impact assessment, which will be a legal requirement in many cases[4], should assist with this.

Retrofitting compliance is more costly and time consuming than building compliance in from the start and often leads to poorer outcomes both in terms of data protection compliance and the product itself. In some cases, it simply does not work. The ICO's enforcement action against Clearview AI is arguably an example of this. As well as fining the facial recognition company £7.5m, it also issued an enforcement notice against the company ordering Clearview to stop obtaining and using the personal data of UK residents and to delete the data of UK residents from its systems. See our blog for more information.

Areas of focus

The ICO's web page "Our work on Artificial Intelligence" confirms that AI is still a priority area for the ICO due to the potential to pose a high risk to individuals and their rights and freedoms, and that its current areas of focus are:

- fairness in AI – in March 2023 the ICO updated its main Guidance on AI and Data Protection following requests from UK industry to clarify requirements for fairness (discussed below);
- dark patterns;
- AI-as-a-service;
- AI and recommender systems;
- biometric data and biometric technologies; and
- privacy and confidentiality in explainable AI.

What guidance is available to help me?

The ICO has produced a number of resources, including its main guidance on AI and data protection and detailed guidance on explaining decisions made with AI. Its web page, "Our work on Artificial Intelligence" lists its main resources, and key examples are set out in the timeline over-page.

> *"[AI] is a priority area for the ICO due to the potential to pose a high risk to individuals and their rights and freedoms…. And as a regulator we will continue to respond to the demand for more work in this space. (ICO Web page "Our work on Artificial Intelligence)."*

While the guidance produced by the ICO includes lots of useful information, it has tended to be long, sometimes overlapping, and has not been drafted in a consistent style. Organisations may therefore find it hard to know where to start, or how the various bits of guidance fit together. However, in November last year it published tips on how to better use AI (discussed below). This guidance is relatively short and high level, making it a more accessible starting point for organisations. The ICO's AI risk mitigation toolkit, finalised last year, also aims to provide more practical support for organisations. Both new resources also link through to relevant sections of the more detailed guidance. In addition to its guidance and toolkits, the ICO's Regulatory Sandbox is a useful resource for AI companies.

[1] Article 22 UK GDPR

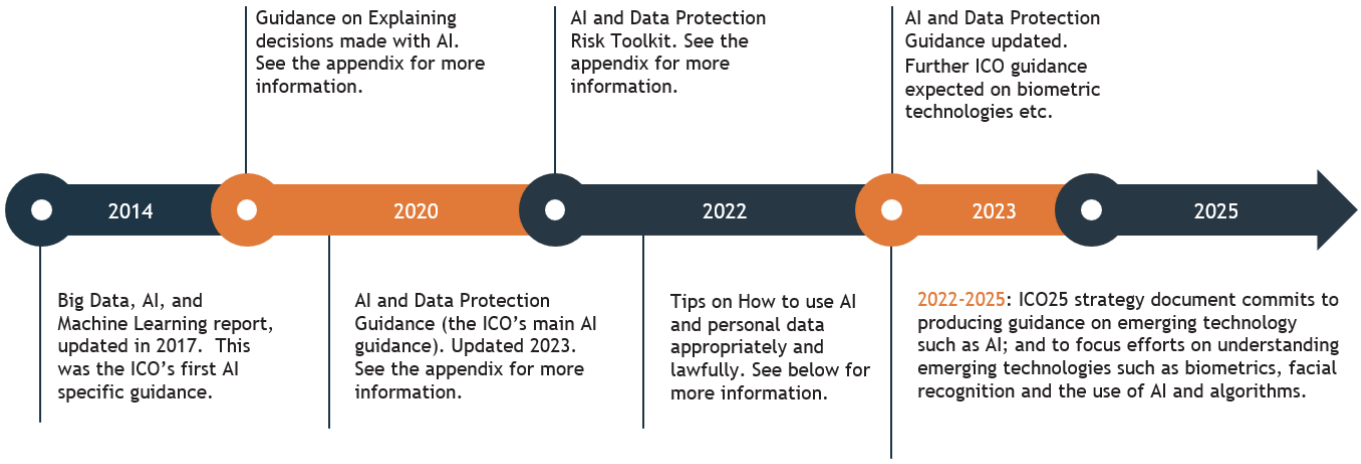
[2] Stephen Almond, the ICO's Director of Technology and Innovation speaking at a conference November 2022

[3] Guidance on AI and Data Protection: Information Commissioner's foreword in original version,

[4] Article 35(3)(a) UK GDPR

ICO AI guidance

AI Timeline



ICO top tips to improve the way you handle AI

In November 2022, the ICO published guidance on "How to use AI and personal data appropriately and lawfully." It condenses some of the key points from the "Guidance on AI and data protection" and "Explaining decisions made with AI" into eight easy to understand tips which organisations can use to improve the way they handle AI. These tips are:

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

   You first need to assess whether you need to use AI, as it is generally considered a high-risk technology. If you do, assess the risks and put in place measures to mitigate them. This includes carrying out a data protection impact assessment (DPIA) and consulting affected groups. Note: if you identify a risk that you cannot mitigate, you are legally required to consult the ICO before any processing takes place.

2. **Think about how you can explain the decisions made by your AI system?**

   While this can be difficult, particularly where machine learning or black-box AI is used, you must still provide a meaningful explanation to those individuals. You also need to think about what people may expect your explanation to look like and how you will handle individual rights requests. The ICO's Explainability guidance may assist with this process.

3. ***Only collect the data you need to develop your AI system and no more***

   AI systems often need lots of data, which can seem at odds with the GDPR's data minimisation principle. However, you can still use AI - you just need to ensure that data is accurate, adequate, relevant and limited. While the latter two points in particular may seem hard to satisfy in practice, an FAQ section at the back of the guidance discusses this further, suggesting, for example, mapping out the areas of the AI pipeline where you may use personal data and scheduling in a time at each significant milestone to review whether you still need the data for that purpose. It also points to the data minimisation section of its main AI guidance and advises organisations to consider if there are any privacy enhancing technologies that can help.

4. ***Address risks of bias early on***

   You need to assess whether the data you are gathering is accurate, relevant and representative of the population that you will apply the AI system to. You should also map out the likely effects and consequences of the decisions made by the AI system for different groups and assess whether these are acceptable.

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

   This will result in better outcomes, as the quality of the output is dependent on the quality of the data inputted. Having clear criteria and lines of accountability about the labelling of data involving protected characteristics, special category data (or both) can help. Appropriately labelled data can lead to fairer outcomes.

6. ***Ensure that your AI system is secure***

   AI systems can exacerbate security risks or create new ones. However, you must still implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk. To help manage this, you could carry out a security risk assessment, including maintaining an inventory of all

AI systems you use, and/or carry out model debugging (i.e., finding and fixing the problems in your model).

7. ***Ensure that any human review of decisions is meaningful***

   People have the right not to be subject to solely automated decisions. They can request a human review of the decision made about them and that human review must be meaningful. Human reviewers must therefore be trained to interpret and challenge the AI outputs, and be senior enough to override them and to consider additional factors.

8. ***If you are buying in your AI, work with your AI supplier to ensure your use is appropriate***
   You are still responsible for your AI use, even if you procure the AI system from a third party. You should therefore do your due diligence to ensure you chose an appropriate supplier (e.g., check they took a privacy by design approach), collaborate with the supplier to carry out an assessment (e.g., your DPIA), agree and document your roles and responsibilities (e.g., who will answer individual rights requests) and consider whether there will be any international transfers.

Changes to the law?

The UK Government is currently using its post Brexit regulatory freedom to reform data protection legislation.

In September 2021 it published the Data: New Direction consultation, which suggested a number of AI related changes to the regime. For example, the consultation looked at whether organisations should be able to use the legitimate interest ground to process personal data without applying the usual balancing tests where the processing relates to bias monitoring, detection and correction in AI systems. It also considered whether to remove or amend the rules around automated decision making mentioned above (see our blog on consultation for more information).

However, many of these proposals were dropped in the Government's response to the consultation and by the time the draft Data Protection and Digital Information Bill was put before parliament in July 2022 few proposals remained. For example, the bill broadened the circumstances where automated decisions could be taken (e.g., limiting the restrictions on this to when special category data is used) but the proposed changes to the legitimate interest rules around bias monitoring were dropped.

The Data Protection and Digital Information Bill was paused in September 2022 due to political changes in the UK and a second version of the bill was published in March 2023 (see our blog). Other than some clarifications around what is meant by 'meaningful human involvement' (including a right for the Secretary of State to bring in specific regulations on this) there have not been many significant changes to the proposals around the automated processing rules which are relevant to AI.

More broadly, the UK is also looking to introduce new AI rules which could impact the ICO. Current plans, set out in a recent white paper on a "Pro Innovation Approach to AI Regulating" suggest that the UK will introduce five cross-cutting principles which all regulators must apply. See our blog on the white paper for more information.

See the Appendix below for more information on the main pieces of ICO AI guidance and the box on "What is AI?".

This article was written by Rob Sumroy and Natalie Donovan. Rob is a partner, and Natalie is a PSL Counsel, in Slaughter and May's Emerging Tech Team. This article is part of our wider Regulating AI series. For more information on the other content available in the series, see *Regulating AI | Slaughter and May - Slaughter and May Insights*

CONTACT

ROB SUMROY
PARTNER, GLOBAL CO-HEAD OF DATA PRIVACY HUB
T: +44(0) 20 7090 4032
E: rob.sumroy@slaughterandmay.com

NATALIE DONOVAN
PSL COUNSEL
T: +44(0) 20 7090 4058
E: Natalie.donovan@Slaughterandmay.com

Appendix

We set out below more detail on the main pieces of AI guidance produced by the ICO:

Guidance	Commentary
Explaining Decisions Made with AI guidance, 2020 (the "**Explainability Guidance**"). This guidance was jointly produced by the ICO and the UK's national institute for data science and AI, the Alan Turing Institute.	**What is it for?** The Explainability Guidance provides best practice tips for organisations, focused on increasing transparency and trust in AI systems by providing advice on how to explain AI use to affected individuals. **Who in my organisation should read it?** The guidance is in three parts, and each part has a slightly different audience. Part one is a general overview intended for all stakeholders, while part two looks at how the guidance can be applied in practice and is aimed more at technical teams. Part three looks at what explaining AI means for an organisation and is aimed at senior executives. However, all three parts will be of interest to compliance teams, data protection officers and risk advisors. **What does it cover?** Part 1 covers the basics. It explains some of the terminology, GDPR provisions and risks associated with AI explainability. It also lists a set of AI principles which should be applied when explaining AI and a number of different ways in which AI decisions can be explained (explanation types). Part 2 is quite practical. It shows you how to: (i) select the appropriate explanation for your sector and use case; (ii) choose an appropriately explainable model (as part of this it looks at some of the issues which arise with black-box models); and (ii) use certain tools to extract explanations from less interpretable models. It also sets out six tasks that organisations can undertake to help them have a systematic approach to developing AI models with explainability in mind and selecting, extracting and delivering explanations regarding AI decisions. As mentioned, this is aimed at more technical teams (although it has received some criticism for not being technical enough) and is quite long. DPOs and compliance teams should, however, find it useful. Part 3 focuses on what explaining AI means for your organisation. It looks at the various roles, policies, procedures and documentation that senior management can put in place to ensure an organisation is set up to provide explanations to individuals. While its target audience (senior management) may find it quite detailed, it will again be useful for DPOs and compliance teams. **More information:** Please see our client publication - Explaining AI: The importance of transparency and explainability – for a more detailed examination of the guidance. We have also produced this shorter blog. Note: The AI guidance (below) also includes information on transparency and explainability. The ICO recommends that the two should be read together.
Guidance on AI and Data Protection, 2020 – updated 2023 (the "**AI Guidance**")	**What is it for?** The AI Guidance forms part of the ICO's AI auditing framework and is designed to provide practical support for organisations auditing their own AI use. The framework comprises this detailed guidance, auditing tools and procedures that the ICO will use in its own audits and investigations and the AI toolkit (see below). It is not a statutory code which means there is no penalty for failing to adopt good practice recommendations if you find another way to comply with the law. It has been designed to complement existing ICO resources, such as the Explainability guidance mentioned above. The ICO has confirmed this guidance is likely to be updated again as the technology and legislation in this space evolves, and that it will support the Government in implementation of its AI white paper. **Who in my organisation should read it?**

Guidance	Commentary
	The guidance has two audiences: firstly, those with a compliance focus (DPOs, GCs, risk managers, senior management and the ICO's own auditors); and secondly, technology specialists. It emphasises the need for senior management, as well as researchers and data scientists, to consider data protection. **What does it cover?** The ICO updated this AI Guidance in March 2023 following requests from UK industry to clarify requirements for fairness in AI. The changes are also designed to meet key commitments from the ICO25 strategy around protecting people from vulnerable groups. The revised structure (which contains a number of new chapters) is based on the foundational principles of data protection and looks at the following topics: 1. What are the accountability and governance implications of AI. 2. How do we ensure transparency in AI? 3. How do we ensure lawfulness in AI? 4. What do we need to know about accuracy and statistical accuracy? 5. How do we ensure fairness in AI? 6. What about fairness, bias and discrimination? 7. What is the impact of Article 22 on the UK GDPR on fairness? 8. How should we assess security and data minimisation in AI? 9. How do we ensure individual rights in our AI system. There is also an Annex which looks at fairness in the AI lifecycle and a glossary of terms. The key takeaway is for organisations to consider privacy from the initial design phase as trying to retrofit compliance often leads to non-compliant products. **More information:** The ICO explains the changes it made in March 2023 at the start of its AI Guidance.
AI and Data Protection Risk Toolkit, May 2022 v.1.0	**What is it for?** The toolkit is designed to be a practical tool to help organisations consider relevant risks when developing AI systems and what practical steps they can take to mitigate them. It is intended to complement DPIAs (which are often required when AI is used), not replace them, and forms part of the ICO's main AI Guidance. **Who in my organisation should use it?** The ICO have stated that the toolkit is designed for use by your organisation's risk and governance teams, model development teams and your senior leadership (who need to understand the risks if they are signing off on the deployment of AI solutions). **What does it cover?** The toolkit focusses on risks to individual's rights and freedoms: • The toolkit works through an AI lifecycle. The stages of the lifecycle are: (i) business requirements and design; (ii) data acquisition and preparation; (iii) training and testing; and (iv) deployment and monitoring. • Within each stage of the lifecycle, it identifies risk areas (accountability, purpose limitation, fairness etc.). These risks are taken from existing ICO AI guidance. It also includes sections on meaningful human review and individual rights. • It connects these risks to areas of the GDPR to help organisations map the risks onto their legal obligations (i.e., the areas they need to be accountable for and demonstrate their compliance with). • For each risk area identified, the toolkit identifies a number of practical steps which can help mitigate the risk. These are classified into steps you *must* take (i.e., it is a legal requirement), steps you *should* take (i.e., it's best practice); and steps you *could* take (i.e., it is optional good practice to follow these steps).

Guidance	Commentary
	• The toolkit also links to further ICO guidance and lets organisations record the action/steps they will take to mitigate risks, which can also help organisations demonstrate compliance. **More information** The toolkit is an excel spreadsheet. It is called v1.0 as both an alpha and beta version were released as part of its development process. The ICO's launch event for the toolkit is available to watch on the ICO's website. The launch event sets out plans to develop and publish case studies showing how to use the toolkit in practice.
How to use AI and personal data appropriately and lawfully, November 2022	Details on this latest guidance piece of ICO guidance are set out in the article above.

What is AI?

The concept of artificial intelligence (AI) has existed since the 1950s but rapidly increasing computational power, and reducing costs for processing and storing data, mean that it is now a practical reality.

AI-based systems can be purely software-based, acting in the virtual world, for example, voice assistants, image analysis software, search engines, and speech and face recognition systems. Alternatively, AI can be embedded in hardware devices - for example, advanced robots, autonomous cars, drones or internet of things applications. It is used today in a variety of sectors. The field of AI is generally subdivided into two categories, general and narrow AI. General AI is AI that has such broad applicability that it could successfully perform any tasks or solve any problem requiring human intelligence. By contrast, narrow AI refers to algorithms that are designed to solve a particular problem, such as playing a game.

AI can be achieved using a number of different technologies, from machine learning to natural language processing. Machine learning is a subset of AI and is a set of techniques and tools that allow computers to "think" by creating self-learning mathematical algorithms based on accumulated data. Machine learning is often used in image recognition, speech-to-text and credit risk classifying AI.

Does the ICO define AI?

In its guidance on Explaining Decisions made with AI, the ICO describes AI as "an umbrella term for a range of technologies and approaches that often attempt to mimic human thought to solve complex tasks." However, in its main Guidance on AI and Data Protection, it confirms that data protection law does not use the term "AI" and so none of the data protection related legal obligations depend on exactly how it is defined. It also confirms that its main guidance focuses on the data protection challenges that machine learning based AI creates. The ICO does, however, acknowledging that other types of AI may give rise to other data protection challenges.

It will also be interesting to see, going forward, whether the UK Government adopts a definition of AI which will be relevant to all regulators. In its recent white paper on a "Pro Innovation Approach to AI Regulating" it suggests that AI should be defined by reference to two characteristics - adaptivity and autonomy – as they generate the need for a bespoke regulatory response. See our blog on the white paper for more information.

London
T +44 (0)20 7600 1200
F +44 (0)20 7090 5000

Brussels
T +32 (0)2 737 94 00
F +32 (0)2 737 94 01

Hong Kong
T +852 2521 0551
F +852 2845 2125

Beijing
T +86 10 5965 0600
F +86 10 5965 0650