The Future of Generative AI Regulation

Author: Michael McCagh, Lawyer, KPMG Law

It is no secret that generative AI has become increasingly popular of late. The speed of developments in the field may be of great benefit to communications professionals, many of whom are already leveraging generative AI to create content such as text, audio, images or video, utilising chatbots to improve customer service, completing translations between languages at scale or performing client segmentation for marketing purposes. Generative AI presents the potential to save significant quantities of time and resources. Arguably, employees will be less burdened by administrative tasks, allowing them to spend more time to perform the tasks that rely upon their professional skills and judgement. However, the benefits of generative AI come with substantial risks. For example, generative AI may hallucinate to provide incorrect or misleading responses, its outputs may unintentionally replicate the personal data used to train it, it may be tampered with by malicious actors or it may produce materials that go against social values or assist the commission of a criminal offence.

To confront the risks and increase public trust in generative AI, a large number of countries have released guidelines or ethical frameworks on the use of AI and most are considering whether to introduce AI-specific laws. Whilst those regulations will facilitate public trust in organisations' systems, they will likely come at a cost to the business. Before an organisation properly weighs up whether to implement a generative AI system into its operations, it is important to understand and consider all of the regulatory costs and limitations, including those that will only become applicable in the future. Two important and contrasting examples are developing in the European Union and the United Kingdom. Australia is no doubt paying attention to these developments, in considering its next moves.

The EU AI Act

The European Union's Artificial Intelligence Act (**AI Act**) has not yet passed into law, so it is not possible to be completely certain of its final form. However, that has not prevented it from drawing attention from all over the globe, largely due to its particularly prescriptive nature and heavy penalties (up to 6% of global annual revenue[1]). Practitioners have also experienced the manner in which the General Data Protection Regulation (**GDPR**) became somewhat of a global standard and now wonder whether the AI Act might be of similar influence. In any event, the AI Act will have extraterritorial effect,[2] so all producers, intermediaries and users of artificial intelligence systems will need to comply with its obligations if they wish for their product to affect subjects located in the world's largest economic market. There is a good chance that the obligations imposed by the AI Act upon generative AI may become the standard by which most entities will find themselves needing to abide.

The AI Act famously allocates AI use cases according their risk profile into three categories: unacceptable-risk; high-risk and low-risk. Foundation models (being AI models trained on broad data such that they can be applied across a wide range of use cases), including generative AI, have their own regime outside of that classification, though the obligations imposed upon them for the most part resemble those that apply to high-risk systems.

Article 28(b) of the AI Act proposed by the EU Parliament sets out the following obligations that generative AI systems must comply with:

1. **Identify and mitigate risks**: a formal risk assessment will be required to document this process. We have seen organisations face regulatory scrutiny in the data protection space on the basis that risk assessments have allegedly inadequately identified the active risks.[3]
2. **Datasets with appropriate data governance**:[4] this is an important obligation that carries the higher penalty of up to 6% of global annual revenue of the infringing organisation. The obligation requires that the training datasets must be free of errors and not lead to biases or discrimination. The training data is more important than the algorithm itself in terms of preventing biases to certain segments of the population.
3. **Efficient in terms of energy usage**: generative AI systems are known to use significant quantities of energy, which will need to be limited to the extent possible.
4. **Technical documentation**: which describes how the AI works, how it was developed and details of a post-monitoring plan. The post-monitoring plan is designed to ensure ongoing compliance with the AI Act and detect biases and discrimination.
5. **Quality management system**: which is a plan for ensuring compliance with the obligations of the AI Act, including organisational and technical measures.
6. **Register with EU database**: which will need to identify the authorised representative within the EU and declare the purpose of the AI system, the status of the system and the electronic instructions for it use.
7. **Transparency requirements**: generative AI systems will need to inform humans that they are interacting with an AI system, in the same way that cookies notices are found on most websites.
8. **Train to not generate content in breach of EU law**: generative AI systems will need action to ensure they do not contravene of range of laws. This obligation reinforces those of other laws and ensures the producer of the generative AI system is also responsible for content ultimately produced by a user.
9. **Publish a summary of training data protected by copyright law**: depending upon the volume and type of the training data, this could create a burdensome obligation to identify all the copyrighted materials. There are, of course, numerous other intellectual property issues that merit significantly longer discussion.

1 Proposed Artificial Intelligence Act (EU), Art. 71.

2 Proposed Artificial Intelligence Act (EU), Art. 2.

3 See the regulatory action taken by the Information Commissioner's Office in the United Kingdom against Snap: https://ico.org.uk/about-the-ico/media-centre/news-and-blogs/2023/10/uk-information-commissioner-issues-preliminary-enforcement-notice-aganst-snap/

4 See also the equivalent obligation for high-risk systems the proposed Artificial Intelligence ACT (EU), Art. 10.

It is important to emphasise that all of these obligations fall upon the entity that is actually producing the AI, not the entity that is using it to produce content. Further, an organisation that utilises a foundation model for the purposes of adapting it for a generative AI purpose will be a provider and therefore be subject to the obligations. However, users must also be aware of the above obligations. Users are obliged to ensure that providers have correctly categorised the AI system, so are formally required to undertake a due diligence process when using a generative AI system produced by another entity.[5]

Users will also need to follow the instructions that providers have set out in the technical documentation. Further, responsibility will fall upon users to complete a Data Protection Impact Assessment (**DPIA**)[6] where personal data is to be used in the training data.[7] The user will also need to notify human subjects when they are interacting with a generative AI system.[8]

AI Regulation in the United Kingdom

In March 2023, the UK Government released its policy paper entitled "A Pro-innovation Approach to AI Regulation", which sets out a comparatively outcome-focussed and flexible regime. This paper describes the UK's plans as not involving the enactment of AI-specific legislation. Rather, existing regulators will be required to oversee the implementation of AI in their respective sectors to ensure that all AI systems adhere to the five key principles, namely: (1) safety, security and robustness; (2) transparency and explainability; (3) fairness; (4) accountability and governance; and (5) contestability and redress.

Regulators will have the benefit of expertise in their respective sectors, though this means that there will be differences in the way that sectors are monitored and enforced. To ensure some level of consistency, the central UK Government will release guidance, standards and tools that regulators will be expected to draw upon. Given that the UK Government policy paper expressly recognises the serious risks posed by generative AI, one might expect that generative AI will be the subject of centralised templates and guidance. This may come in the form of specific rules that describe in detail when a generative AI system is deemed to have satisfied the five aforementioned key principles.

Compliance with AI-specific Regulations

The digital era has meant that services are often likely to reach many jurisdictions in various parts of the world. It is for exactly that reason that most AI-specific regulations will have extraterritorial effect. Organisations may need to formulate a strategy to comply with a range of regimes and one way will be satisfy the most rigorous regime. Both the AI Act and the UK's AI policy will need to be considered in detail by multinational organisations and entities that wish to use their generative AI systems in those jurisdictions. The AI Act should be of particular focus since it is a key market, will influence the regulations implemented in other jurisdictions (including Australia) and to date appears to be the most prescriptive example of AI-specific regulation. Therefore, the obligations in these respective regulations must be properly considered when evaluating potential use cases for generative AI and when onboarding vendors that adopt generative AI. Early preparation is likely to prove far more efficient than correcting existing mechanisms once legislative measures fully materialise.

5 Proposed Artificial Intelligence Act (EU), Art. 29(6)(a).
6 See General Data Protection Regulation, Art. 35.
7 Proposed Artificial Intelligence Act (EU), Art. 29.
8 Proposed Artificial Intelligence ACT (EU), Art. 52.