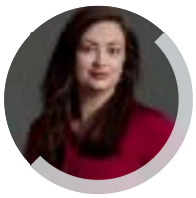
Developing a Responsible Governance Framework for AI

Authors: Rob Corbet,
Partner, Arthur Cox LLP and member of the Data Protection & Information Security Working Group.

Rosemarie Blake,
Senior Associate, Arthur Cox LLP.



Developing a Responsible Governance Framework for AI[1]

The proliferation of AI into the mainstream of business and consumer technologies and the emergence of AI-specific legislation has raised a new question at the boardroom table - what should we be doing today to ensure we can leverage the potential of AI technologies in a manner that minimises legal risk?

AI Laws are Coming

The EU is leading the way with the proposed AI Act, the text of which is likely to be finalised in the coming months and is expected to become binding across the EU within two years. The draft AI Act sets out 6 general principles that are, in general terms, reflective of the views of policy makers, legislators and standardisation bodies and which should form part of an organisation's AI risk assessment framework.

These principles focus on fundamental rights and can be summarised as follows:

Human agency and oversight	Technical robustness and safety	Privacy and data governance
Transparency (which includes traceability and explainability)	Diversity, non-discrimination and fairness	Social and environmental well-being

These specific rules will supplement existing laws in the areas of privacy, consumer protection, equality and product liability so while the GDPR is an important point of reference for AI technologies that involve the processing of personal data, an AI governance framework requires a broader lens.

Ethical Framework

While the legislation is still evolving, when it comes to policy and ethical issues, there is no shortage of relevant reference points. For example, the European Commission's Declaration on Digital Rights formed the basis for the approach to foundation model AI providers in the EU's draft AI Act and the EU has also published Ethics Guidelines for Trustworthy AI. The EU and US are currently discussing a joint code of conduct for AI firms, the OECD are also working on a code, and the World Economic Forum recently published a set of Guidelines for the Procurement of AI Solutions by the Private Sector. All of these developments, while not legally binding, provide instructive guidance to organisations looking to build and deploy a principles-based ethical governance framework in advance of the EU AI Act being adopted.[2]

Adopting A Risk Based Approach to AI

Against such a rapidly evolving landscape, organisations would be well advised to develop a risk-based assessment framework which involves assessing and managing all the potential risks associated with the use of AI. This should at least include a consideration of the following issues:

1) Beyond GPDR - Consider all AI uses cases: Don't limit risk assessments to those under the GDPR where there is already a robust set of rules governing automated decision making, data minimization, data protection impact assessments etc. Most of the AI Principles outlined above are unrelated to data protection laws so they require an assessment of new forms of risk which will, once the laws have commenced, attract their own new forms of corporate liability. Trying to undertake retrospective risk assessments after the AI Act becomes binding is unlikely to be a sufficient mitigation.

2) Data Protection Regulators are Watching: Having said that, it is already clear that the existing regulatory regime around the GDPR and ePrivacy Directive is a fertile environment for active compliance and enforcement. Even in cases where apparently anonymized data is used, a documented risk assessment may be important as a mitigating factor in the context of any potential regulatory fine or other sanction. Conversely a lack of consideration for any controls on the application of AI could act as an aggravating factor. Notably, Article 71 (6) of the draft AI Act is broadly similar to Article 83(2) of the GDPR in setting out factors to be taken into account when calculating administrative penalties.

Regulatory Action

As mentioned, privacy regulators have already been active in interrogating the approach to privacy and data protection. The most notable example was the Italian regulator's ("Garante") recent sanction of ChatGPT in Italy. Garante issued a ban to OpenAI on processing Italian data subjects' personal data through ChatGPT, which was subsequently lifted on OpenAI's implementation of urgent privacy measures. These measures focused on established GDPR rules on transparency, legal basis and data subject rights (including opt-out rights and rights of rectification and erasure).

Conclusion

Privacy regulators will have little patience for organisations who, as controllers or processors, do not adopt appropriate risk mitigations for AI processing that is subject to GDPR rules. However, it is also clear that an entirely new regulatory framework, with a similar sanctions regime as exists under GDPR, is coming. Accordingly, organisations who are embracing the opportunities presented by AI technologies today (and few are not) should proceed with a cautious enthusiasm.

At a minimum, they should adopt an AI Policy for their organisation that identifies the boundaries beyond which they are not currently willing to go (e.g. prohibiting the processing of confidential information or personal data within generative AI models) while documenting their corporate policy position on the key issues that are likely to form the bedrock of future regulatory investigations and litigation. In particular, anticipating obvious and less obvious risks around privacy, bias/discrimination or other unfair treatment and documenting the measures being adopted to mitigate those risks will be an invaluable legal asset as the legislative environment matures. However, documents alone will not suffice so this should be coupled with a governance framework that creates transparency and accountability around the use of AI by the organisation. Bedding this policy internally into the organisation and ensuring it tracks through to third parties in the procurement channels will set a solid foundation for a successful and lawful AI strategy.

REFERENCES
1. Rob Corbet, Partner and Rosemarie Blake, Senior Associate, Arthur Cox LLP.
2. See also:
The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems
The Montreal Declaration for Responsible AI
The AIGA AI Governance Framework
NIST Artificial Intelligence Risk Management Framework