



marian/Moment via Getty Images

● INSIGHT

International: CoE Convention on AI - what to expect after its signature?

November 21, 2024



Summary

The Council of Europe's Framework Convention on Artificial Intelligence, signed by ten countries including the EU, UK, US, Canada, Australia, Japan, and Mexico, aims to set global legal standards for AI governance to protect human rights, democracy, and the rule of law. It introduces broad definitions, general obligations, and principles for signatories to ensure AI systems' alignment with human rights and democratic integrity, adopting a risk-based approach without prescribing specific penalties. The Convention will be legally binding only upon ratification, with varying implications for AI legislation in different countries, such as the EU's alignment with the Convention through the EU AI Act, the UK's forthcoming legislation, and the US's fragmented approach due to the lack of federal AI law.

On September 5, 2024, ten countries signed the Council of Europe's (CoE) Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (the Convention). The first signatories to the Convention include the EU, the UK, the US, Canada, Australia, Japan, and Mexico. The Convention was drafted by representatives from the 46 Member States of the CoE, the EU, and 11 non-Member States. The Convention aims to establish a minimum legal standard for the governance of artificial intelligence (AI) globally. It creates unifying language on the regulation of AI to protect human rights, democracy, and the rule of law, and represents the first international treaty on AI. While signing the Convention indicates a state's intention to be bound by the Convention, once the Convention comes into force, it will only be legally binding on states that have ratified it. Daniel P. Cooper, Stacy Young, and Sam Jungyun Choi, from Covington & Burling LLP, look at the content of the Convention and how it may shape the development of AI regulatory frameworks.

The substance of the Convention

Scope

The Convention covers 'the activities within the lifecycle of [AI] systems that have the potential to interfere with human rights, democracy and the rule of law' (Article 3(1) of the Convention). It defines an AI system as 'a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that may influence physical or virtual environments' (Article 2 of the Convention) - this definition is similar to that under the EU Artificial Intelligence Act (EU AI Act), which is based on the Organisation for Economic Co-operation and Development's (OECD) definition of AI. The Convention's [Explanatory Report](#) emphasizes that this definition is deliberately broad and is intended to 'future-proof' the Convention by 'ensur[ing] legal precision and certainty, while also remaining sufficiently abstract and flexible to stay valid despite future technological developments' (Explanatory Report, paras. 14 and 24).

Obligations and principles

The Convention introduces two general obligations on signatories. Specifically, they must adopt or maintain measures to ensure that AI systems: (i) function in a manner that is consistent with the protection of human rights and (ii) do not undermine the integrity of democratic processes and the rule of law.

The Convention also sets out seven principles that require signatories to adopt or maintain measures to ensure human dignity and individual autonomy, transparency and oversight, accountability and responsibility, equality and non-discrimination, privacy and data protection, reliability, and safe innovation. These themes are also central to the AI Act.

The Convention requires signatories to adopt or maintain measures to achieve these principles but does not impose granular requirements on how they should do so. By way of example, with respect to the principle of transparency and oversight, signatories must adopt or maintain measures with regard to the identification of content generated by AI systems. The Explanatory Report notes that such measures may include techniques such as labeling and watermarking (Explanatory Report, para. 59).

Risk-based approach

The Convention adopts a risk-based approach, requiring signatories to adopt measures that are 'graduated and differentiated as may be necessary in view of the severity and probability of the occurrence of adverse impacts on human rights,

democracy and the rule of law throughout the lifecycle of [AI] systems' (Article 1(2) of the Convention). Defining whether an AI system has a 'significant impact' or 'significant effect' on human rights is left to each signatory (Explanatory Report, para. 101).

Penalties

The Convention does not prescribe specific penalties for violations of its obligations or principles. Instead, as set out under the principle of accountability and responsibility, signatories have an obligation to adopt or maintain measures to 'ensure accountability and responsibility for adverse impacts on human rights, democracy and the rule of law resulting from activities within the lifecycle of [AI] systems' (Article 9 of the Convention). The Explanatory Report clarifies that this refers to the use of 'judicial and administrative measures, civil, criminal and other liability regimes' (Explanatory Report, para. 66), though the specifics of these regimes, including penalties, are left to each signatory's own legal system.

What's next

The Convention will enter into force on the first day of the month following the expiration of a period of three months after the date on which five signatories, including at least three CoE Member States, have ratified, accepted, or approved it (Article 30 of the Convention). Only once the Convention enters into force will ratifying signatories be bound by it.

The ratification process will differ depending on the country. For example, in the UK, the Government will need to lay the Convention before Parliament for 21 sitting days before it can be ratified. There may be greater obstacles to ratifying the Convention in countries that have federal or otherwise more complicated administrative structures, such as the US, and it is probable that some signatories will not be able to ratify the Convention as a result.

It remains to be seen how the Convention will shape the development of AI regulatory frameworks among its signatories once it comes into force. While it is unlikely to introduce significant changes in the EU, where the EU AI Act is already largely harmonized with the Convention, it has the potential to influence the

development of AI legislation particularly in countries where it is still in its nascent stages. By way of overview:

- **EU** - The Commission [notes](#) that the Convention will be implemented in the EU by means of the EU AI Act, which entered into force on August 1, 2024, and that it will prepare a proposal for a Council decision and seek the consent of the European Parliament for it to be formally recognized in the EU.
- **UK** - While there is currently no standalone AI legislation in the UK, the Government announced in the King's Speech in July 2024 that it would introduce new legislation to 'place requirements on those working to develop the most powerful AI models.' In addition, on signing the Convention, the UK Government released a [statement](#) affirming its commitment to ratifying the Convention, noting that '[o]nce the treaty is ratified and brought into effect in the UK, existing laws and measures will be enhanced,' and that it 'will work closely with regulators, the devolved administrations, and local authorities as the Convention is ratified to ensure it can appropriately implement its new requirements.' It therefore appears likely that the UK Government will ratify the Convention and take steps to implement it in upcoming AI legislation.
- **Japan** - While the Japanese Government has published various strategies and policies on AI, and various ministries have published guidance on the use of AI, no specific AI legislation has yet been passed. That said, there appears to be an appetite to develop a comprehensive AI legislative framework, and the Government's AI Strategy Council began discussing the development of AI regulations over the summer of 2024.
- **Australia** - In September 2024, the Australian Government published a [proposal](#) to introduce 'mandatory guardrails for AI in high-risk settings' and a set of [Voluntary AI Safety Standards](#).
- **Mexico** - In May 2024, the National Artificial Intelligence Agenda for Mexico 2024-2030 was presented in the Mexican Senate. This agenda sets out a multi-pronged strategy to regulate AI and ensure that it brings about societal benefits and includes recommendations in the areas of public policy, regulation, and governance.
- **US** - The federal government has [acknowledged](#) that the signing of the Convention is but an 'initial step' to becoming a party to the Convention. There is currently no federal AI law in the US, and proposals have taken a highly fragmented approach thus far - there are currently several proposed federal AI laws, though none have been able to successfully obtain bipartisan support. In the absence of consensus at the federal level, AI is currently being regulated

through a patchwork of state laws and proposed laws, with over 40 states introducing AI-related bills this year.

In all, as the first multilateral treaty on AI, the Convention is a landmark achievement. However, its influence in shaping and harmonizing AI legislation among its ratifying signatories - once it comes into force - will depend largely on how strongly such legislation is enforced domestically.

Daniel P. Cooper

Partner dcooper@cov.com

Stacy Young

Associate slyoung@cov.com

Sam Jungyun Choi

Associate jchoi@cov.com Covington & Burling LLP, Brussels

Erin Lynch, a trainee with Covington in London, assisted with the writing of this piece.

Topics:

Artificial Intelligence