

Council of Europe Convention on Artificial Intelligence

Thomas Wahl

News

On 5 September 2024, the first-ever, international, legally-binding instrument on Artificial Intelligence (AI) was [opened for signature](#) by the Council of Europe (CoE). The CoE's "[Framework Convention on Artificial Intelligence](#)" provides a common baseline to ensure that activities within the lifecycle of AI systems are fully consistent with human rights, democracy and the rule of law.

Each Party to the Convention is obliged to adopt or maintain appropriate legislative, administrative or other measures to give effect to the provisions set out in this Convention. These measures shall be 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. This may include specific or horizontal measures that apply irrespective of the type of technology used.

Definition

The Convention defines "artificial intelligence 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.

Territorial Scope

The Framework Convention on Artificial Intelligence is open to accession to the 46 Council of Europe Member States, the European Union, and states around the world that are not members of the Council of Europe. Involved in elaborating the Convention were for instance the non-member states Argentina, Australia, Canada, Costa Rica, the Holy See, Israel, Japan, Mexico, Peru, the United States of America, and Uruguay. Since the subject of the convention falls within the exclusive competence of the European Union, [only the European Union will become party](#) to the Convention.

Material scope

The Convention applies to the use of AI systems in the public sector – including companies acting on its behalf – and in the private sector. Excluded from the scope are:

- Artificial intelligence systems related to the protection of a Party's national security interests (but the Party will be obliged to ensure that AI activities respect international law and democratic institutions and processes);

AUTHOR

Thomas Wahl

Senior Researcher
Max Planck Institute for the
Study of Crime, Security and
Law

Published in
2024, Vol. 19(3) eucrim pp 194
– 196

ISSN: 1862-6947

<https://eucrim.eu>



- Research and development activities, except when the testing of AI systems or similar activities may have the potential to interfere with human rights, democracy or the rule of law;
- Matters relating to national defence.

Main obligations

The Framework Convention sets forth general obligations and common principles that each Party is obliged to implement in regard to AI systems. Parties must ensure, for instance, that the activities within the lifecycle of AI systems are consistent with obligations to protect human rights, as enshrined in applicable international law and in their domestic law. They must also adopt or maintain measures that protect the integrity, independence and effectiveness of democratic institutions and processes.

The Convention establishes transparency and oversight requirements tailored to specific contexts and risks, including identifying content generated by AI systems. Parties must adopt measures to identify, assess, prevent, and mitigate possible risks and assess the need for a moratorium, a ban or other appropriate measures concerning uses of AI systems where their risks may be incompatible with human rights standards.

Parties are also obliged to ensure accountability and responsibility for adverse impacts and that AI systems respect equality, including gender equality, the prohibition of discrimination, and privacy rights.

Other provisions of the Convention relate to topics such as public consultation and digital literacy/skills.

Remedies and safeguards

The Convention sets the parameters for accessible and effective remedies for violations of human rights resulting from the activities within the lifecycle of AI systems. It is considered important, for instance, that the relevant content in the information-related measures should be context-appropriate, sufficiently clear and meaningful, and critically, provide a person concerned with an effective ability to use the information in question to exercise their rights in the proceedings in respect of the relevant decisions affecting their human rights.

Procedural safeguards must include that persons interacting with AI systems are notified that they are interacting with such systems rather than with a human.

Risk assessments

Similar as the EU legislation (for the EU AI Act → [eucrim 4/2023, 316-317](#)), the CoE Convention follows a risk-based approach. The Convention introduces minimum requirements for risk assessments: Parties are obliged to identify, assess, prevent and mitigate *ex ante* and, as appropriate, iteratively throughout the lifecycle of the AI system the relevant risks and potential impacts to human rights, democracy and the rule of law by following and enabling the development of a methodology with concrete and objective criteria for such assessments. These obligations are key to enable the implementation of all relevant principles, including the principles of transparency and oversight as well as the principle of accountability and responsibility.

The Convention also provides that, in the risk and impact assessment process, attention should be paid both to the dynamic and changing character of activities within the lifecycle of AI systems and to the shifting conditions of the real-world environments in which systems are intended to be deployed. Requirements are introduced regarding not only the documentation of the relevant information during the risk management

processes, but also the application of sufficient preventive and mitigating measures in respect of the risks and impacts identified.

Follow-up

In order to ensure its effective implementation, the Convention establishes a follow-up mechanism in the form of a “Conference of the Parties”, composed of representatives of the Parties. In addition, each Party will be obliged to provide a report to the Conference of the Parties within the first two years after becoming a Party and then periodically thereafter with details of the activities undertaken to give effect to the use of AI systems in the public and private sector. Last but not least, Parties are required to adopt or maintain effective mechanisms to oversee compliance with the obligations in the Framework Convention. Oversight bodies must be functionally independent from the relevant actors within the executive and legislative branches.

Background

Negotiations on the Convention began back in September 2022 under the auspices of the [Committee on Artificial Intelligence \(CAI\)](#) established by the Council of Europe in Strasbourg. The negotiating process not only brought together government representatives from the Council of Europe member states and non-member states (see above), and from the European Commission (negotiating on behalf of the European Union), but also representatives from civil society, academia, industry, and other international organisations who participated as observers.

On 5 September 2024, the Convention was [signed by 10 states](#) (including Israel and the United States of America as non-member states of the Council of Europe) and by the European Commission on behalf of the European Union.

The Convention will enter into force after five states have given their consent to be bound by the Convention (e.g. after ratification); at least three out of these five states must be member states of the Council of Europe.

For the EU, the Convention means that it will be implemented by means of the EU AI Act which entered into force on 1 August 2024 and contains generally fully harmonised rules for the placing on the market, putting into service and use of AI systems in the EU (→ [eucrim 2/2024, 92-93](#)).

Eucrim will regularly update the accessions to the Framework Convention on Artificial Intelligence (CETS No. 225) on its [website documenting ratifications](#) of CoE Conventions.

About eucrim

eucrim is the leading journal which regularly informs about current developments in European criminal and “criministrative” law.

All news items are freely accessible at: <https://eucrim.eu/news/>

Stay informed by emailing to eucrim-subscribe@csl.mpg.de to receive alerts for new releases of issues.

The project is co-financed by the [Union Anti-Fraud Programme \(UAFP\)](#), managed by the [European Anti-Fraud Office \(OLAF\)](#).



Co-funded by
the European Union