Council of Europe Framework Convention on AI, Human Rights, Democracy and the Rule of Law: Cheat Sheet

Created by: **Bianca-Ioana Marcu**, Policy Manager for Global Privacy

Overview: The Council of Europe's Framework Convention on AI, adopted on **17 May 2024**, is the world's first binding treaty on AI, human rights, democracy and the rule of law. It establishes a risk-based approach to regulating AI, as well as a number of general principles and obligations related to activities within the entire lifecycle of AI systems. The Framework Convention on AI will be open for States' signature as of 5 September 2024, also to countries not members of the Council of Europe.

Scope	Key Terms
Public sector application: Each Party to the Framework Convention on AI will have to apply its principles and obligations within the lifecycle of AI systems undertaken by public authorities, or private actors acting on their behalf (Art. 3(1)(a)). **Private sector application:** Two requirements have to be met: 1) the country in which the private actor is established or in which they develop or deploy AI products and services is a State Party to the treaty, an 2) they are designing, developing or deploying AI systems on behalf of that State Parties' public authorities (Art. 3). **Not within the scope:** The Convention does not apply to: activities within the lifecycle of AI systems related to the protection of a State's national security interests (Art. 3(2)); research and development activities regarding AI systems not yet made available for public use (Art. 3(3)); and matters of national defense. In such instances, international law and human rights obligations must still be respected.	**Framework Convention:** This type of treaty serves as an umbrella document which lays down principles and objectives, while leaving room for stricter and more prescriptive standards to domestic legislation. **Harmonized definition of an AI system:** The Framework Convention adopts the OECD definition of an AI system, aligning with the EU AI Act, President Biden's Executive Order on AI, and the Colorado AI Act. **Lifecycle of an AI system:** Includes the stages of design, development, use and decommissioning. **Risk-based approach:** Measures to be implemented in the lifecycle of AI systems 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" (Art. 1).

General Obligations and Common Principles
Activities within the lifecycle of AI systems must be consistent with obligations to protect human rights, respect the rule of law, and not be used to undermine democratic processes. Seven common principles are elaborated by the Framework Convention on AI: 1. Respect for **human dignity** and **individual autonomy** (Art. 7);

FPF Global

2. Maintain measures to ensure that adequate **transparency and oversight** requirements tailored to specific contexts and risks are in place (Art. 8);
3. Adopt or maintain measures to ensure **accountability and responsibility** for adverse impacts on human rights, democracy and the rule of law (Art. 9);
4. Ensure that activities within the lifecycle of artificial intelligence systems respect **equality**, including gender equality, and the **prohibition of discrimination** as provided under applicable international or domestic legislation (Art. 10);
5. Adopt or maintain measures to ensure that the **privacy of individuals and their personal data** are protected, including through international laws, standards and frameworks (Art. 11);
6. Take measures to promote the **reliability** of artificial intelligence systems and **trust** in their outputs, which could include requirements related to adequate **quality** and **security** (Art. 12);
7. Establish **controlled environments** for developing, experimenting and testing AI systems under the supervision of competent authorities (Art. 13).

The Risk-Based Approach in Practice

In order to assess and mitigate **actual and potential harms** posed by AI systems to human rights, democracy and the rule of law, the Framework Convention on AI introduces a set of broad requirements:

- Take due account of the **context and intended use** of an AI system (Art. 16(2)(a));
- Take due account of the severity and probability of potential impacts (Art. 16(2)(b));
- Consider, where appropriate, the **perspective of all relevant stakeholders**, in particular persons whose rights may be impacted (Art. 16(2)(c));
- Apply the risk-management requirements **iteratively** and throughout the lifecycle of AI systems (Art. 16(2)(d));
- Include **monitoring** for risks and adverse impacts (Art. 16(2)(e));
- Include **documentation** of risks, actual and potential impacts, and on the risk management approach (Art. 16(2)(f));
- Require **testing** of artificial intelligence systems before making them available for first use and when they are significantly modified (Art. 16(2)(g)).

International Cooperation

The Framework Convention on AI establishes two methods for furthering international cooperation on AI governance:
1. The establishment of a "**Conference of the Parties**", to be composed of representatives of the Parties to the treaty (Art. 23);
2. Assisting States that are non-Members of the treaty to act consistently with its obligations, with a view to supporting them in becoming Parties to it (Art. 25).

If you have any questions, please contact Bianca-Ioana Marcu at bmarcu@fpf.org.

FPF Global