

# The Council of Europe Convention on AI: National Security Implications

Rosamund Powell

September 2024

*This is part two of a CETaS Explainer series on the **national security implications of international AI regulation initiatives**.*

*All views expressed in this CETaS Explainer are those of the author, and do not necessarily represent the views of The Alan Turing Institute or any other organisation.*

*This work is licensed under the terms of the Creative Commons Attribution License 4.0 which permits unrestricted use, provided the original author and source are credited. The license is available at: <https://creativecommons.org/licenses/by-nc-sa/4.0/legalcode>.*

---

## Introduction

The Council of Europe **Framework Convention on AI and Human Rights, Democracy and the Rule of Law** is expected to be the most widely adopted, legally binding treaty on AI to date, with **57 states** coming together to negotiate its contents, and **the UK, EU and US already signed on**.<sup>1</sup> This CETaS Explainer identifies key considerations for the national security community when preparing for the implementation of the Convention in the UK and internationally.

## What is the Council of Europe Convention on AI, Human Rights, Democracy and the Rule of Law?

The Council of Europe Convention on AI is an **international and legally binding** treaty. It aims to protect **human rights, democracy, and the rule of law** in light of risks posed by AI.<sup>2</sup> It was adopted on 17 May 2024 during the Council of Europe annual ministerial meeting and **opened for state signatures on 5 September 2024**.<sup>3</sup>

## What are the key provisions of the Convention?

- The provisions set out in the Convention are **high-level** and focused around **seven core principles**
  1. Human dignity and individual autonomy
  2. Transparency and oversight
  3. Accountability and reproducibility
  4. Equality and non-discrimination
  5. Privacy and personal data protection
  6. Reliability
  7. Safe innovation
- In addition to these principles, the Convention sets out several obligations for participating countries. For example, they must ensure those harmed by AI have a **right to redress** and must establish sufficient **human oversight** of AI-enabled decision-making.

## What is the scope of the Convention?

- On 5 September 2024, the Convention was signed by '**Andorra, Georgia, Iceland, Norway, the Republic of Moldova, San Marino, the United Kingdom, as well as Israel, the United States of America and the European Union**'. There is scope for more countries to sign in the future.<sup>4</sup>
- **Sectoral exclusions** are made for AI applications in **national security** and **defence** as well as for AI developed purely for **research and development**.
- The Convention **focuses on the use of AI systems by public authorities**. There is **more ambiguity in its coverage of private sector AI** as signatories are given the option to apply the Convention's principles to the private sector through 'other appropriate measures'.<sup>5</sup>
- The convention applies to **all stages of the AI lifecycle** from planning and design to operation, monitoring and retirement.<sup>6</sup>

## Where does the Convention sit in the broader regulatory landscape?

- The Convention is the **first treaty of its kind** given the number of countries brought together during negotiations (including not just European nations but some countries in Asia, Australasia and South America) and close focus on human rights. But it does still skew heavily towards European countries and the Global North and is not the only relevant AI regulation globally.
- Several other regulatory efforts will have significant impacts:
  - The **EU AI Act came** into force on 1 August 2024.<sup>7</sup>

- National governments have also made progress. Most notably, **the US Executive Order on AI** was issued in October 2023,<sup>8</sup> while **AI Regulation in China** emerged even earlier (including key regulations in 2021, 2022, and 2023).<sup>9</sup>
- Finally, certain local legislation will be significant, especially **state-level AI regulation in the US** which is progressing rapidly.<sup>10</sup>

Given the significant overlap in participating countries and timelines, the EU AI Act and Council of Europe Convention on AI are often grouped together. Nevertheless, these two initiatives diverge in several ways (Figure 1).<sup>11</sup>

Figure 1: Comparison of the Council of Europe Convention on AI and the EU AI Act

	Similarities	Differences
<b>Global reach</b> 	<ul style="list-style-type: none"> <li>• Most (or all) participating nations are in Europe and the Global North.</li> </ul>	<ul style="list-style-type: none"> <li>• More countries have signed the Council of Europe Convention on AI.</li> <li>• The EU AI Act has extraterritorial applications as it applies to companies across the world, so long as they operate within the EU.</li> </ul>
<b>Scope</b> 	<ul style="list-style-type: none"> <li>• Sectoral exclusions are set out for national security, defence and research and development.</li> </ul>	<ul style="list-style-type: none"> <li>• The private sector is covered more extensively in the EU AI Act, whereas the Convention gives signatories options for how it is applied to private sector AI.</li> <li>• The EU AI Act has specific provisions and exemptions for law enforcement and for open-source AI, the Council of Europe Convention does not.</li> </ul>
<b>Provisions</b> 	<ul style="list-style-type: none"> <li>• Both take a risk-based approach to AI regulation.</li> </ul>	<ul style="list-style-type: none"> <li>• The EU AI Act includes more comprehensive and specific obligations on AI developers.</li> <li>• The Convention gives more flexibility to ensure compatibility with domestic law.</li> </ul>

Source: Author's analysis.

## Timeline

Different countries will implement the Convention on different timelines, especially as some will simply need to implement or amend existing regulation, while others might need to introduce new, legally binding measures.<sup>12</sup> Nevertheless, several key milestones are detailed in the Convention (Figure 2).

Figure 2: Timeline for the implementation of the Council of Europe Convention on AI



Source: Council of Europe Convention on AI.

## The Principles

The Convention sets out a series of principles to be incorporated into the participating countries' domestic approaches to AI regulation (Figure 3).

Figure 3: The principles set out in the Council of Europe Convention on AI



Source: Council of Europe Convention on AI.

---

## The Obligations

In addition to these principles, the Convention establishes several obligations for all participating nations. These include the need for signatories to:

- Ensure that **existing domestic legal frameworks are up-to-date** and continue to protect human rights in a rapidly evolving AI landscape.<sup>13</sup>
- Protect **democratic processes** in the face of AI risks.<sup>14</sup>
- Make clear there exists a right to **launch a complaint** and to receive an **effective remedy** where human rights have been violated due to AI.<sup>15</sup>
- Preserve **human oversight** where AI-based decision-making substantially impacts human rights.<sup>16</sup>
- Establish **iterative risk management throughout the AI lifecycle**.<sup>17</sup>

These obligations are set out at a high level to allow compatibility with domestic law and policy.

---

## Exclusions

The Convention includes several sectoral exclusions (Figure 4). Broadly, these exclusions mirror those set out in the EU AI Act, but there are key differences.

The biggest difference is how the Convention deals with private sector AI. The Convention applies fully to private entities acting on the behalf of public authorities. However, for all other private sector AI, there is substantial leeway for signatory states. Specifically, participating countries can either enact the Convention as normal to the private sector *or* apply the same obligations through 'other appropriate measures'.<sup>18</sup> This more lenient approach to governing the private sector was discussed heavily during the final stages of negotiation, with the US reported to be pushing for private sector exclusions from the Convention.<sup>19</sup>

Figure 4: Exemptions from the Council of Europe Convention on AI by sector

Sector	Defence	National security	Law enforcement	Research and development	Private sector AI
In scope for the Convention?	X	X	✓	X	PARTIALLY INCLUDED

Source: Council of Europe Convention on AI.

## The National Security Exclusion

### How is the scope of the exclusion defined?

Any AI systems related to the protection of 'national security interests' are out of scope of the Convention, on the understanding that existing international human rights law still applies.<sup>20</sup>

### Debate surrounding the national security exclusion

The exclusion for national security has been debated extensively, both by those participating in treaty negotiations and by civil society groups.

1. During negotiations, the specific parameters of the national security exclusion evolved.<sup>21</sup> Several factors motivated those negotiating the Treaty. For example, the EU Commission wanted the Convention to align closely with the EU AI Act, which excludes national security.<sup>22</sup>
2. Outside of negotiations, civil society groups have objected to the exclusion for national security,<sup>23</sup> arguing that the sorts of AI applications which are prevalent in national security settings have significant human rights implications that must be addressed.<sup>24</sup>

---

## Why Does the Convention Still Matter for National Security?

Despite these exclusions, the Council of Europe Convention on AI will have national security implications in all participating countries. It will be crucial for the national security community to consider which of the Convention's provisions they might adopt voluntarily, while tracking the progress of this Convention for what it reveals about global trends in AI regulation.

## Principles for AI and Human Rights

Principles-based approaches to AI policy are nothing new, nor do the seven principles set out in the Convention show significant divergence from the many sets of principles for ethical AI which have emerged elsewhere (for example from the OECD, UNESCO, Global Partnership on AI and elsewhere).<sup>25</sup> Nevertheless, the seven principles in the Convention are likely to be particularly influential given they form the basis of a legally binding treaty.

Despite national security being excluded from the Convention, these principles are still applicable to national security agencies' work on AI. In some contexts, the value of AI ethics principles in national security is already recognised explicitly. For example, the US intelligence community has committed publicly to a set of principles for ethical AI,<sup>26</sup> while the UK's GCHQ has set out their own approach to AI ethics.<sup>27</sup> Numerous defence departments have also published their own AI principles.<sup>28</sup>

For those who are yet to publicly commit to a set of principles on AI and human rights, the seven principles set out in the Convention could provide a useful framework to guide AI practices in a national security context, and agencies may consider making voluntary public commitments to those principles, rather than developing new principles from scratch.

However, setting out a set of principles for AI and human rights is only a first stage in developing a practical approach to AI governance. In many national security settings, existing laws and policies are already in place to protect human rights and individual freedoms. Commitment to high-level principles is insufficient unless it is complemented

by additional internal mapping of how these principles are implemented in practice, and in a manner that is consistent with existing regulation and policy.

## **New best practice for fundamental rights impact assessment**

Given the high-level nature of the Convention, one of the most practically useful resources for the national security community won't be found in the Convention itself, but instead within accompanying efforts to develop a methodology for 'fundamental rights, democracy and rule of law impact assessments' for AI systems.<sup>29</sup>

While conducting human rights impact assessments won't be mandatory in national security contexts, methodologies which have been prepared in association with the Council of Europe provide detail on how to identify human rights risks associated with new AI projects.<sup>30</sup> These methods could be incorporated within national security approaches to AI assurance.<sup>31</sup>

## **Revealing fault lines for future AI regulation initiatives**

Beyond being a practical resource for the national security community, the Convention reveals how challenging it is to reach international consensus on AI regulation, especially on topics such as national security and private sector innovation.<sup>32</sup> The debates around the Convention provide useful insight into the most contentious issues likely to arise in any future regulatory discussions concerning AI and national security.

Despite bringing together just 57 nations during negotiations, with a heavy skew towards Global North countries, the degree of compromise needed to reach consensus on the final Convention resulted in significant criticism, especially around decisions made to soften the obligations of the private sector.<sup>33</sup>

The Convention is the first initiative to 'combine EU and US priorities' on AI regulation, so the need for compromise should not come as a surprise. Nevertheless, as one of the earliest international AI regulation initiatives, it offers important lessons, pointing to the sorts of compromises that may need to be made in the future.<sup>34</sup>



---

## Conclusion

While this Convention is unlikely to have such significant, practical implications for the national security community as the EU AI Act, its progress should still be monitored closely. The Convention's provisions are broad, leaving significant scope for how States will enact them. As this process begins, the UK national security community should clarify its own position on the principles set out in the Convention, and how they can best be applied in their domain. National security agencies should also closely follow how different countries enact this treaty for what it reveals about convergence and divergence in international approaches to AI governance.

---

## References

1. Council of Europe, "Committee on Artificial Intelligence," <https://www.coe.int/en/web/artificial-intelligence/cai>; Reuters, "US, Britain, EU to sign first international AI treaty," 5 September 2024, <https://www.reuters.com/technology/artificial-intelligence/us-britain-eu-sign-agreement-ai-standards-ft-reports-2024-09-05/>.
2. Council of Europe, "The Framework Convention on Artificial Intelligence," <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>.
3. Council of Europe, "Committee on Artificial Intelligence," <https://www.coe.int/en/web/artificial-intelligence/cai>.
4. Council of Europe, "Council of Europe opens first ever global treaty on AI for signature," Newsroom, 5 September 2024, <https://www.coe.int/en/web/portal/-/council-of-europe-opens-first-ever-global-treaty-on-ai-for-signature>.
5. Council of Europe, "The Framework Convention on Artificial Intelligence," <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>.
6. Council of Europe, "Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law," *Council of Europe Treaty Series*, no. 225 (2024), <https://rm.coe.int/1680afae67>.
7. Rosamund Powell, "The EU AI Act: National Security Implications," *CETaS Explainers* (August 2024), <https://cetas.turing.ac.uk/publications/eu-ai-act-national-security-implications>.
8. The White House, "Fact sheet: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence," 30 October 2023, <https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/>.
9. Matt Sheehan, *China's AI Regulations and How They Get Made* (Carnegie Endowment for International Peace: 2023), <https://carnegieendowment.org/research/2023/07/chinas-ai-regulations-and-how-they-get-made>.
10. Cobun Zweifel-Keegan, "US State AI Governance Legislation Tracker," IAPP Article, last updated 25 June 2024, <https://iapp.org/resources/article/us-state-ai-governance-legislation-tracker/>.
11. Jacques Ziller, "The Council of Europe Framework Convention on Artificial Intelligence vs. The EU Regulation: Two Quite Different Legal Instruments," *CERIDAP - ISSN 2723-9195* (June 2024), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4822757](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4822757).
12. Council of Europe, "Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law," <https://rm.coe.int/1680afae67>.

13. Council of Europe, “The Framework Convention on Artificial Intelligence,” Article 4, <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>.
14. Council of Europe, “The Framework Convention on Artificial Intelligence,” Article 5, <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>.
15. Council of Europe, “The Framework Convention on Artificial Intelligence,” Article 14, <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>.
16. Council of Europe, “The Framework Convention on Artificial Intelligence,” Article 15, <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>.
17. Council of Europe, “The Framework Convention on Artificial Intelligence,” Article 16, <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>.
18. Gibson Dunn, “Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law,” 3 June 2024, <https://www.gibsondunn.com/council-of-europe-framework-convention-on-artificial-intelligence-and-human-rights-democracy-and-rule-of-law/>; Council of Europe, “The Framework Convention on Artificial Intelligence,” Article 3, <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>.
19. Luca Bertuzzi, “EU prepares to push back on private sector carve-out from international AI treaty,” *Euractiv*, 10 January 2024, <https://www.euractiv.com/section/artificial-intelligence/news/eu-prepares-to-push-back-on-private-sector-carve-out-from-international-ai-treaty/>; Digital Watch Observatory, “EU challenges US-led bid to exclude private sector from potential international AI treaty,” 11 January 2024, <https://dig.watch/updates/eu-challenges-us-led-bid-to-exclude-private-sector-from-potential-international-ai-treaty>.
20. Council of Europe, “The Framework Convention on Artificial Intelligence,” Article 3, <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>.
21. See “Committee on Artificial Intelligence Draft Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law, Draft Explanatory Report: Item to be considered by the GR-J at its meeting on 29 April 2024” for earlier wording on the exclusion for national security from the Convention: [https://search.coe.int/cm/#{%22CoEIdentifier%22:\[%22090001680af0734%22\],%22sort%22:\[%22CoEValidationDate%20Descending%22\]}](https://search.coe.int/cm/#{%22CoEIdentifier%22:[%22090001680af0734%22],%22sort%22:[%22CoEValidationDate%20Descending%22]}).
22. Luca Bertuzzi, “EU prepares to push back on private sector carve-out from international AI treaty,” *Euractiv*, 10 January 2024, <https://www.euractiv.com/section/artificial-intelligence/news/eu-prepares-to-push-back-on-private-sector-carve-out-from-international-ai-treaty/>.
23. European Center for Not-for-Profit Law, “Open Letter to COE AI Convention Negotiators: Do Not Water Down Our Rights,” 25 January 2024, <https://ecn.org/news/open-letter-coe-ai-convention-negotiators-do-not-water-down-our-rights>.
24. Angela Muller, “The Council of Europe’s Convention on AI: No free ride for tech companies and security authorities!,” AlgorithmWatch Press Release, 5 March 2024, <https://algorithmwatch.org/en/council-of-europe-ai-convention/>.
25. OECD, “OECD AI Principles Overview,” <https://oecd.ai/en/ai-principles>; Global Partnership on AI, “About GPAI,” <https://www.gpai.ai/about/>; UNESCO, “Ethics of Artificial Intelligence: The Recommendation,” <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>.
26. INTEL.gov, “Principles of Artificial Intelligence Ethics for the Intelligence Committee,” <https://www.intelligence.gov/principles-of-artificial-intelligence-ethics-for-the-intelligence-community>.
27. GCHQ, “Pioneering a New National Security: The Ethics of AI,” 2021, <https://www.gchq.gov.uk/artificial-intelligence/index.html>.
28. HM Government, *Ambitious, safe, responsible: our approach to the delivery of AI-enabled capability in defence* (Ministry of Defence: 15 June 2022), <https://www.gov.uk/government/publications/ambitious-safe-responsible-our-approach-to-the-delivery-of-ai-enabled-capability-in-defence/ambitious-safe-responsible-our-approach-to-the-delivery-of-ai-enabled-capability-in-defence#using-ai-ethically>; Joint Artificial Intelligence Center (JAIC), “Ethical Principles for Artificial Intelligence,” [https://www.ai.mil/docs/Ethical\\_Principles\\_for\\_Artificial\\_Intelligence.pdf](https://www.ai.mil/docs/Ethical_Principles_for_Artificial_Intelligence.pdf).
29. Ad Hoc Committee on Artificial Intelligence Policy Development Group (CAHAI-PDG), *Human Rights, Democracy and Rule of Law Impact Assessment of AI systems* (Council of Europe: 2021), <https://rm.coe.int/cahai-pdg-2021-02-subworkinggroup1-ai-impact-assessment-v1-2769-4229-7/1680a1bd2d>.

30. David Leslie et al., *Human rights, democracy, and the rule of law assurance framework for AI systems: A proposal* (Zenodo: February 2022), <https://zenodo.org/records/5981676#.ZFuJeXbMK39>.

31. Rosamund Powell and Marion Oswald, "Assurance of third-party AI systems for UK national security," *CETaS Research Reports* (January 2024), <https://cetas.turing.ac.uk/publications/assurance-third-party-ai-systems-uk-national-security>.

32. Mahmoud Javadi, "What the Council of Europe's new treaty tells us about global AI governance," *The Loop*, 17 June 2024, <https://theloop.ecpr.eu/what-the-council-of-europes-new-treaty-tells-us-about-global-ai-governance/>; Christopher Lamont, "The Council of Europe's draft AI Treaty: balancing national security, innovation and human rights?," *Global Governance Institute Commentary*, 18 March 2024, <https://www.globalgovernance.eu/publications/the-council-of-europes-draft-ai-treaty-balancing-national-security-innovation-and-human-rights>.

33. European Center for Not-for-Profit Law, "Council of Europe approves AI Convention, but not many reasons to celebrate," *EDRi*, 10 July 2024, <https://edri.org/our-work/council-of-europe-approves-ai-convention-but-not-many-reasons-to-celebrate/>.

34. Mahmoud Javadi, "What the Council of Europe's new treaty tells us about global AI governance," *The Loop*, 17 June 2024, <https://theloop.ecpr.eu/what-the-council-of-europes-new-treaty-tells-us-about-global-ai-governance/>.