



Strasbourg, 30 June 2022

CAI(2022)07

RESTRICTED

COMMITTEE ON ARTIFICIAL INTELLIGENCE (CAI)

ZERO DRAFT [FRAMEWORK] CONVENTION ON ARTIFICIAL INTELLIGENCE, HUMAN RIGHTS, DEMOCRACY AND THE RULE OF LAW

Proposal by the Chair

Preamble	3
Chapter I: General provisions	4
Article 1 – Purpose and object	4
Article 2 – Definitions	4
Article 3 – Principle of non-discrimination	5
Article 4 – Scope	5
Chapter II: Fundamental principles	5
Article 5 – Design, development and application of artificial intelligence systems	5
Article 6 – Additional requirements for design, development and application of artificial intelligence systems in the public sector	6
Article 7 – Procedural safeguards.....	6
Article 8 – Restrictions.....	7
Article 9 – Relationship with other legal instruments.....	7
Article 10 – Wider protection	7
Chapter III: Risk and impact assessment and related measures	8
Article 11 – Risk and impact assessment.....	8
Article 12 – Measures in respect of artificial intelligence systems posing significant levels of risk.....	8
Article 13 – Measures in respect of artificial intelligence systems posing unacceptable levels of risk	8
Article 14 – Prohibited artificial intelligence practices.....	9
Article 15 – Measures in respect of artificial intelligence systems and combinations of such systems posing significant levels of risk	9
Chapter IV: Follow-up mechanism and cooperation.....	10
Article 16 – Consultation of the Parties.....	10
Article 17 – International co-operation.....	11
Article 18 – National supervisory authorities	11
Chapter V: Final provisions.....	11
Article 19 – Effects of the Convention.....	11
Article 20 – Amendments.....	12
Article 21 – Revision of the Appendix	12
Article 22 – Dispute settlement.....	12
Article 23 – Signature and entry into force	12
Article 24 – Accession.....	13
Article 25 – Territorial application.....	13
Article 26 – Reservations	13
Article 27 – Denunciation	14
Article 28 – Notification.....	14
Appendix: Methodology for risk and impact assessment of artificial intelligence systems	15

Preamble

The member States of the Council of Europe and the other signatories hereto,

Considering that the aim of the Council of Europe is to achieve greater unity between its members, based in particular on respect for human rights and fundamental freedoms, democracy and the rule of law;

Recognising the value of fostering cooperation with the other Parties to this Convention;

Conscious of the accelerating developments in science and technology and the profound changes brought about by the application of artificial intelligence systems through the provision of new tools, which have the potential to promote human prosperity as well as individual and social well-being by enhancing progress and innovation;

Concerned by the risk that certain uses of some artificial intelligence systems also have the potential for unduly interfering with the exercise of human rights and fundamental freedoms, undermining democracy and violating the rule of law through such harmful practices and effects as, for instance, unrestricted mass surveillance, information distortion, possible unlawful discrimination, the general weakening of human agency, unlawful electoral interference and digital exclusion;

Conscious of the fact that human rights, democracy and the rule of law are inherently interwoven and convinced of the need to establish, as a matter of priority, a common legal framework establishing fundamental principles and rules governing design, development, and application of artificial intelligence systems which would effectively preserve the shared common values and at the same time be conducive to innovation;

Conscious also of the fact that in view of its framework character the Convention may be supplemented by further binding or non-binding instruments to address challenges relating to the application of artificial intelligence in specific sectors;

Welcoming efforts undertaken by other international and supranational Organisations and fora, which further advance international understanding and cooperation in this area;

Mindful of the need to ensure a proper balance between respect for human rights as enshrined in the 1950 Council of Europe Convention for the Protection of Human Rights and Fundamental Freedoms and its protocols, the 1966 United Nations International Covenant on Civil and Political Rights and other applicable international human rights treaties and various economic, security and other interests in the development and use of artificial intelligence;

Underlining that the present Convention is intended to complement those conventions in order to fill in any legal gaps in view of the specific challenges raised by design, development and application of artificial intelligence systems;

Mindful also of the right to privacy and the protection of personal data, as conferred, for example, by the 1981 Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data and amending protocols;

Have agreed as follows:

Chapter I: General provisions

Article 1 – Purpose and object

1. This Convention establishes certain fundamental principles and rules aimed at ensuring that design, development and application of artificial intelligence systems is fully consistent with respect for human rights, the functioning of democracy and the observance of rule of law.
2. Each Party shall take the necessary measures in its domestic legislation to give effect to the principles, rules and rights set out in this Convention.
3. In order to ensure effective implementation of its provisions by its Parties, this Convention establishes a follow-up mechanism.

Article 2 – Definitions

For the purposes of this Convention:

- a. “artificial intelligence system” means a machine-based system that is capable of informing or autonomously generating outputs such as predictions, recommendations or decisions using machine and/or human-based data and inputs and influencing the environment that it interacts with;
- b. “lifecycle” means all phases of existence of an artificial intelligence system between its design and decommissioning;
- c. “artificial intelligence provider” means any natural or legal person, public authority or other body that develops an artificial intelligence system or that has an artificial intelligence system developed with a view to putting it into service/commissioning it;
- d. “artificial intelligence user” means any natural or legal person, public authority or other body using an artificial intelligence system in their own name or under their authority;
- e. “artificial intelligence subject” means any natural or legal person whose human rights and fundamental freedoms, legal rights or interests are impacted by decisions made or substantially informed by the application of an artificial intelligence system.

Article 3 – Principle of non-discrimination

The implementation of the provisions of this Convention by the Parties shall be secured without discrimination on any ground such as sex, gender, sexual orientation, race, colour, language, age, religion, political or any other opinion, national or social origin, association with a national minority, property, birth, state of health, disability or other status, or based on a combination of one or more of these grounds.

Article 4 – Scope

1. Parties undertake to apply this Convention to design, development and application of artificial intelligence systems throughout their lifecycle, regardless of whether these activities are undertaken by public or private actors.
2. The present Convention shall not apply to design, development and application of artificial intelligence systems used for purposes related to national defense.

Chapter II: Fundamental principles

Article 5 – Design, development and application of artificial intelligence systems

Bearing in mind the need to safeguard and uphold human dignity, human rights and fundamental freedoms, preserve and foster robust and accountable democratic institutions and safeguard the rule of law as the institutional basis for assuring both democratic participation and the effective protection of human rights and fundamental freedoms, in their respective jurisdictions Parties shall:

1. ensure that any research, design, development and application of artificial intelligence systems or combinations of such systems is compatible with core values of democratic societies. In particular, Parties shall ensure that such research, design, development and application are not aimed at undermining or curtailing fundamental rights and freedoms, the functioning of democracy or the observance of rule of law;
2. ensure that any research, design, development and application of artificial intelligence systems or combinations of such systems is grounded in the principle of sustainability and solidarity and the need to protect the environment;
3. ensure that design, development and application of artificial intelligence systems respects the principle of equality, including gender equality and rights related to discriminated groups and people in vulnerable situations;
4. ensure that privacy of individuals is protected including through applicable personal data protection and data governance laws and standards and that data protection principles and rules are applied in respect of design, development and application of artificial intelligence systems and that appropriate guarantees and safeguards have been put in place for data subjects;
5. ensure a continuous chain of accountability, responsibility and legal liability for any unlawful harm in respect of design, development and application of artificial intelligence systems throughout their lifecycle and that appropriate redress mechanisms are available;

6. ensure that, where appropriate, adequate oversight mechanisms as well as transparency and auditability requirements tailored to the specific risks arising from the context in which the artificial intelligence system is designed, developed and applied are in place;
7. with a view to encouraging research and fostering innovation, ensure that a controlled regulatory environment for testing artificial intelligence systems is available for artificial intelligence providers and users under the supervision of their competent authorities;
8. ensure that adequate safety, security, data quality, data integrity, data security, cybersecurity and robustness requirements are in place regarding design, development and application of artificial intelligence systems;
9. strive to ensure that fundamental questions raised by the design, development and application of artificial intelligence systems are the subject of appropriate public discussion and multi-stakeholder consultation in the light, in particular, of relevant social, economic, ethical and legal implications;
10. encourage and promote digital literacy and digital skills for all segments of the population.

Article 6 – Additional requirements for design, development and application of artificial intelligence systems in the public sector

Bearing in mind the increased risks posed to human rights, democracy and the rule of law by the application of artificial intelligence systems or combinations of such systems in the public sector and the need to ensure proper accountability, responsibility and legal liability in relation to the application of such systems, Parties shall:

1. ensure that the deployment and application of artificial intelligence systems in the public sector do not undermine human rights and fundamental freedoms, legal rights and interests which may be guaranteed under the laws of any Party or under any other agreement to which it is a Party and respect the principle of the rule of law.
2. ensure that the deployment and application of artificial intelligence systems in the public sector have an appropriate legal basis and that a careful preliminary consideration of the necessity and proportionality of the use of such system is carried out in view of the context of the deployment.
3. ensure compliance with the standards set out in this article insofar as any private entity acting on their behalf is concerned.
4. ensure that all relevant artificial intelligence users receive appropriate training in operating any artificial intelligence systems employed as well as the respective limitations of such systems.

Article 7 – Procedural safeguards

1. Parties shall ensure that, where appropriate, the usage of an artificial intelligence system is duly recorded and communicated to the artificial intelligence subjects concerned. The

exercise of the right of access to the relevant records, including the grounds on which it may be exercised, shall be governed by domestic law;

2. Parties shall also ensure that in cases where artificial intelligence system substantially informs or takes decision(s) affecting human rights and fundamental freedoms, legal rights and interests, the artificial intelligence subject in question is informed about the application of an artificial intelligence system in the decision-making process and that there is a right to human review of such decisions.
3. Parties shall ensure that, where appropriate, relevant explanations and justifications are offered by the artificial intelligence provider and/or user in plain, understandable, and coherent language and are tailored to the context. Such communication shall contain sufficient information in order to provide the artificial intelligence subject in question with an effective possibility of challenging the decision(s) affecting the subject's human rights, legal rights and interests insofar as any use of artificial intelligence technology is concerned.
4. Parties shall ensure that any person has the right to know that one is interacting with an artificial intelligence system rather than with a human and, where appropriate, shall provide for the option of interacting with a human in addition to or instead of an artificial intelligence system.
5. The exercise of the rights set out in paragraphs 1 to 4, including the grounds on which they may be exercised, shall be governed by domestic law.

Article 8 – Restrictions

Restrictions on the exercise of the rights specified in Article 7 paragraphs 1 to 4 may be provided for by law where necessary and proportionate in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.

Article 9 – Relationship with other legal instruments

Nothing in the present Convention shall be construed as limiting or derogating from any of the human rights and fundamental freedoms as well as legal rights which may be guaranteed under the laws of any Party or under any other agreement to which it is a Party.

Article 10 – Wider protection

None of the provisions of this Convention shall be interpreted as limiting or otherwise affecting the possibility for a Party to grant a wider measure of protection than is stipulated in this Convention.

Chapter III: Risk and impact assessment and related measures

Article 11 – Risk and impact assessment

1. Without any prejudice to already existing relevant national and international legal instruments, and irrespective of whether the respective activities are undertaken by public or private actors, Parties undertake to adopt a methodology applicable to design, development and application of artificial intelligence systems, for identifying, analysing and evaluating risk and assessing impact of the application of artificial intelligence systems in relation to the enjoyment of human rights, the functioning of democracy and the observance of rule of law.
2. The methodology shall be based on the model set out in the Appendix to the present Convention. It shall set out clear, concrete and objective criteria for identifying such artificial intelligence systems or combinations of such systems that in view of their concrete applications pose significant levels of risk to the enjoyment of human rights, the functioning of democracy and the observance of the rule of law.
3. Each Party shall take appropriate measures, particularly in the field of training of national supervisory authorities, artificial intelligence providers and artificial intelligence users, with a view to ensuring that the relevant actors are capable of identifying, analysing and evaluating risk and assessing impact of the application of artificial intelligence systems in relation to the enjoyment of human rights, the functioning of democracy and the observance of rule of law, in accordance with the present methodology.

Article 12 – Measures in respect of artificial intelligence systems posing significant levels of risk

1. Parties shall, in accordance with their domestic law, provide for the imposition on artificial intelligence providers and users of an obligation to apply all necessary preventive and mitigating measures to an artificial intelligence system, deemed in accordance with the methodology referred to in Article 11, paragraph 2, to present significant levels of risk of interfering with the enjoyment of human rights, the functioning of democracy and the observance of the rule of law, aimed at mitigating that risk, as well as to document and keep the records of the respective process.
2. Each Party shall also ensure that the relevant requirements in respect of such artificial intelligence systems take into account the measures set out in Article 15 (a) to (e) and are proportionate to the nature of the risk they pose to human rights, democracy and the rule of law.

Article 13 – Measures in respect of artificial intelligence systems posing unacceptable levels of risk

1. Parties shall, in accordance with their domestic law, provide for the possibility of imposing a full or partial moratorium or ban in respect of artificial intelligence systems, which in accordance with the aforesaid methodology and in view of the measures set out in Article 15 (a) to (e) are deemed to present unacceptable levels of risk of interfering with the enjoyment of human rights, the functioning of democracy, and the observance of the rule of law. A moratorium or ban should, however, only be considered, where on an objective basis

an unacceptable risk to human rights, democracy or the rule of law has been identified and, after careful examination, there are no other measures available for mitigating that risk.

2. Parties shall, in accordance with their domestic law, establish appropriate and independent review procedures in order to enable the reversal of a moratorium or ban provided that relevant risks are sufficiently reduced or appropriate mitigation measures become available.

Article 14 – Prohibited artificial intelligence practices

The following artificial intelligence practices shall be banned:

1. the use of artificial intelligence systems by public authorities using biometrics to identify, categorise or infer emotions of individuals;
2. the use of artificial intelligence systems for social scoring to determine access to essential services leading either to (a) detrimental or unfavourable treatment of certain natural persons or whole groups in social contexts which are unrelated to the contexts in which the data was originally generated or collected, or (b) detrimental or unfavourable treatment of certain natural persons or whole groups that is unjustified or disproportionate to their social behaviour;
3. any other use of artificial intelligence systems by public authorities for such purposes as are not compatible with core values of democratic societies, do not pursue a legitimate aim and are not necessary in a democratic society.

Article 15 – Measures in respect of artificial intelligence systems and combinations of such systems posing significant levels of risk

Parties shall take such measures as are aimed at:

- a. minimizing and, to the extent possible, preventing any unlawful harm or damage to human rights, fundamental freedoms, legal rights and interests of natural or legal persons, democratic institutions and processes, the administration of justice, public health and the environment, which could result from the inappropriate application of artificial intelligence systems;
- b. preserving individual freedom, human dignity and autonomy and in particular the ability to reach informed decisions free from undue influence, manipulation or detrimental effects which adversely affect the right to freedom of expression and assembly, democratic participation and the exercise of other relevant human rights and fundamental freedoms resulting from the inappropriate application of artificial intelligence systems;
- c. ensuring that all interested parties, groups and individuals enjoy equal and fair access to public debate and inclusive democratic processes, taking, in particular, due account of the relevant implications of the technological developments in the area of artificial intelligence and the role of public and private entities that help shape the public debate in their respective jurisdictions;

- d. ensuring that design, development and application of artificial intelligence systems do not compromise the principle of equality before the law, including gender equality and rights related to discriminated groups and people in vulnerable situations in so far as they are used to inform or take decisions affecting human rights and fundamental freedoms, legal rights and interests of natural or legal persons;
- e. ensuring that the rights to privacy and to personal data protection are adequately respected during design, development and application of artificial intelligence systems throughout their lifecycle in both public and private sectors, notably through additional safeguards for special categories of data such as medical data.

Chapter IV: Follow-up mechanism and cooperation

Article 16 – Consultation of the Parties

1. Parties shall consult periodically with a view to:
 - a. making proposals to facilitate or improve the effective use and implementation of this Convention, including the identification of any problems and the effects of any declaration made under this Convention;
 - b. making proposals for the amendment of this Convention in accordance with Article 20;
 - c. formulating their opinion on any proposal for the amendment of this Convention which is referred to them in accordance with Article 20, paragraph 3;
 - d. examining and revising, as necessary and in accordance with the procedure prescribed in Article 21, the methodology for risk and impact assessment of artificial intelligence systems contained in the Appendix to this Convention;
 - e. expressing an opinion on any question concerning the interpretation and application of this Convention and facilitating the exchange of information on significant legal, policy or technological developments.
2. The Consultation of the Parties shall be convened by the Secretary General of the Council of Europe whenever (s)he finds it necessary and in any case when a majority of the Parties or the Committee of Ministers request its convocation.
3. The Consultation of the Parties shall adopt its own rules of procedure.
4. Parties shall be assisted by the Secretariat of the Council of Europe in carrying out their functions pursuant to this article.
5. A contracting Party which is not a member of the Council of Europe shall contribute to the financing of the Consultation of the Parties in a manner to be decided by the Committee of Ministers upon consultation of that Party.

Article 17 – International co-operation

1. Parties shall, as appropriate, co-operate to the fullest extent in the realisation of the purpose of this Convention.
2. Parties shall, on a regular basis, exchange information between them concerning design, development and application of artificial intelligence systems which they assess as posing significant levels of risk to the enjoyment of human rights, the functioning of democracy and the observance of rule of law.

Article 18 – National supervisory authorities

1. Parties shall establish or designate national supervisory authorities tasked, in particular, with overseeing and supervising compliance with the requirements of the risk and impact assessment of artificial intelligence systems in accordance with Article 11 and 12.
2. Parties shall, in accordance with their domestic law, provide for efficient procedures for the imposition of a moratorium or a ban on design, development and application of an artificial intelligence system in accordance with Articles 13 and 14.
3. Parties shall ensure the national supervisory authorities have sufficient resources and properly trained personnel to carry out their activities.
4. The national supervisory authorities shall be independent and impartial in the exercise of their duties.

Chapter V: Final provisions

Article 19 – Effects of the Convention

1. If two or more Parties have already concluded an agreement or treaty on the matters dealt with in this Convention or have otherwise established their relations on such matters, or should they in future do so, they shall also be entitled to apply that agreement or treaty or to regulate those relations accordingly. However, where Parties establish their relations in respect of the matters dealt with in the present Convention other than as regulated therein, they shall do so in a manner that is not inconsistent with the Convention's objective and principles.
2. Parties which are members of the European Union shall, in their mutual relations, apply Community and European Union rules in so far as there are Community or European Union rules governing the particular subject concerned and applicable to the specific case, without prejudice to the object and purpose of the present Convention and without prejudice to its full application with other Parties.
3. Nothing in this Convention shall affect other rights, obligations and responsibilities of a Party and individuals under international law.

Article 20 – Amendments

1. Amendments to this Convention may be proposed by any Party, the Committee of Ministers of the Council of Europe or the Consultation of the Parties.
2. Any proposal for amendment shall be communicated by the Secretary General of the Council of Europe to the Parties.
3. Moreover, any amendment proposed by a Party, or the Committee of Ministers, shall be communicated to the Consultation of the Parties, which shall submit to the Committee of Ministers its opinion on the proposed amendment.
4. The Committee of Ministers shall consider the proposed amendment and any opinion submitted by the Consultation of the Parties and may approve the amendment.
5. The text of any amendment approved by the Committee of Ministers in accordance with paragraph 4 shall be forwarded to the Parties for acceptance.
6. Any amendment approved in accordance with paragraph 4 shall come into force on the thirtieth day after all Parties have informed the Secretary General of their acceptance thereof.

Article 21 – Revision of the Appendix

1. The methodology for risk and impact assessment of artificial intelligence systems contained in the Appendix shall be regularly examined by the Consultation of the Parties and, as necessary, revised, as provided for in Article 16 , paragraph 1, d.
2. The Consultation of the Parties may adopt any revision to the methodology for risk and impact assessment of artificial intelligence systems contained in the Appendix by unanimity and shall communicate such revision to the Secretary General of the Council of Europe.
3. The revised methodology for risk and impact assessment of artificial intelligence systems contained in the Appendix shall come into force on the thirtieth day after it has been notified to the Secretary General.

Article 22 – Dispute settlement

In the event of a dispute between Parties as to the interpretation or application of this Convention which cannot be resolved by the Consultation of the Parties, as provided for in Article 16 , paragraph 1, e, they shall seek a settlement of the dispute through negotiation or any other peaceful means of their choice, including submission of the dispute to an arbitral tribunal whose decisions shall be binding upon the Parties to the dispute, or to the International Court of Justice, as agreed upon by the Parties concerned.

Article 23 – Signature and entry into force

1. This Convention shall be open for signature by the member States of the Council of Europe, by the European Union and by non-member States which have participated in its elaboration.

2. This Convention is subject to ratification, acceptance or approval. Instruments of ratification, acceptance or approval shall be deposited with the Secretary General of the Council of Europe.
3. This Convention shall enter into force on the first day of the month following the expiration of a period of three months after the date on which six Signatories, including at least four member States of the Council of Europe, have expressed their consent to be bound by the Convention in accordance with the provisions of paragraph 2.
4. In respect of any Signatory which subsequently expresses its consent to be bound by it, the Convention shall enter into force on the first day of the month following the expiration of a period of three months after the date of the expression of its consent to be bound by the Convention in accordance with the provisions of paragraph 2.

Article 24 – Accession

1. After the entry into force of this Convention, the Committee of Ministers of the Council of Europe, after consulting with and obtaining the unanimous consent of the Parties to the Convention, may invite any State which is not a member of the Council of Europe and which has not participated in its elaboration to accede to this Convention. The decision shall be taken by the majority provided for in Article 20.d of the Statute of the Council of Europe and by the unanimous vote of the representatives of the Parties entitled to sit on the Committee of Ministers.
2. In respect of any State acceding to the Convention under paragraph 1 above, the Convention shall enter into force on the first day of the month following the expiration of a period of three months after the date of deposit of the instrument of accession with the Secretary General of the Council of Europe.

Article 25 – Territorial application

1. Any State or the European Union may, at the time of signature or when depositing its instrument of ratification, acceptance, approval or accession, specify the territory or territories to which this Convention shall apply.
2. Any Party may, at a later date, by a declaration addressed to the Secretary General of the Council of Europe, extend the application of this Convention to any other territory specified in the declaration. In respect of such territory the Convention shall enter into force on the first day of the month following the expiration of a period of three months after the date of receipt of the declaration by the Secretary General.
3. Any declaration made under the two preceding paragraphs may, in respect of any territory specified in such declaration, be withdrawn by a notification addressed to the Secretary General of the Council of Europe. The withdrawal shall become effective on the first day of the month following the expiration of a period of three months after the date of receipt of such notification by the Secretary General.

Article 26 – Reservations

No reservation may be made in respect of any provision of this Convention.

Article 27 – Denunciation

1. Any Party may, at any time, denounce this Convention by means of a notification addressed to the Secretary General of the Council of Europe.
2. Such denunciation shall become effective on the first day of the month following the expiration of a period of three months after the date of receipt of the notification by the Secretary General.

Article 28 – Notification

The Secretary General of the Council of Europe shall notify the member States of the Council of Europe, the non-member States enjoying observer status with the Council of Europe, the European Union, the non-member States which have participated in the elaboration of this Convention, and any State having been invited to sign this Convention in accordance with the provisions of Article 23, or having been invited to accede to it in accordance with the provisions of Article 24, of:

- a. any signature;
- b. the deposit of any instrument of ratification, acceptance or approval;
- c. any date of entry into force of this Convention in accordance with Article 23, paras. 3 and 4, and Article 24, para. 2;
- d. any amendment adopted in accordance with Article 20 and the date on which such an amendment enters into force;
- e. any revision of the Appendix adopted in accordance with Article 21 and the date on which such revision enters into force;
- f. any denunciation made in pursuance of Article 27 ;
- g. any other act, notification or communication relating to this Convention.

In witness whereof the undersigned, being duly authorised thereto, have signed this Convention.

Done in [place], this ... day of [month] 202[4], in English and in French, both texts being equally authentic, in a single copy which shall be deposited in the archives of the Council of Europe. The Secretary General of the Council of Europe shall transmit certified copies to each member State of the Council of Europe, to the non-member States which enjoy observer status with the Council of Europe, to the European Union and to any State invited to sign or accede to this Convention.

Appendix: Methodology for risk and impact assessment of artificial intelligence systems

Under development