



Strasbourg, 25 January 2019

T-PD(2019)01

**CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION
OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING
OF PERSONAL DATA**

(Convention 108)

GUIDELINES ON ARTIFICIAL INTELLIGENCE AND DATA PROTECTION

Directorate General of Human Rights and Rule of Law

Artificial Intelligence¹ (“AI”) based systems, software and devices (hereinafter referred to as AI applications) are providing new and valuable solutions to tackle needs and address challenges in a variety of fields, such as smart homes, smart cities, the industrial sector, healthcare and crime prevention. AI applications may represent a useful tool for decision making in particular for supporting evidence-based and inclusive policies. As may be the case with other technological innovations, these applications may have adverse consequences for individuals and society. In order to prevent this, the Parties to Convention 108 will ensure and enable that AI development and use respect the rights to privacy and data protection (article 8 of the European Convention on Human Rights), thereby enhancing human rights and fundamental freedoms.

These Guidelines provide a set of baseline measures that governments, AI developers, manufacturers, and service providers should follow to ensure that AI applications do not undermine the human dignity and the human rights and fundamental freedoms of every individual, in particular with regard to the right to data protection.²

Nothing in the present Guidelines shall be interpreted as precluding or limiting the provisions of the European Convention on Human Rights and of Convention 108. These Guidelines also take into account the new safeguards of the modernised Convention 108 (more commonly referred to as “Convention 108+”)³.

I. General guidance

1. The protection of human dignity and safeguarding of human rights and fundamental freedoms, in particular the right to the protection of personal data, are essential when developing and adopting AI applications that may have consequences on individuals and society. This is especially important when AI applications are used in decision-making processes.
2. AI development relying on the processing of personal data should be based on the principles of Convention 108+. The key elements of this approach are: lawfulness, fairness, purpose specification, proportionality of data processing, privacy-by-design and by default, responsibility and demonstration of compliance (accountability), transparency, data security and risk management.
3. An approach focused on avoiding and mitigating the potential risks of processing personal data is a necessary element of responsible innovation in the field of AI.
4. In line with the guidance on risk assessment provided in the Guidelines on Big Data adopted by the Committee of Convention 108 in 2017⁴, a wider view of the possible outcomes of data processing should be adopted. This view should consider not only human rights and fundamental freedoms but also the functioning of democracies and social and ethical values.
5. AI applications must at all times fully respect the rights of data subjects, in particular in light of article 9 of Convention 108+.
6. AI applications should allow meaningful control by data subjects over the data processing and related effects on individuals and on society.

¹ The following definition of AI is currently available on the Council of Europe’s website <https://www.coe.int/en/web/human-rights-rule-of-law/artificial-intelligence/glossary>: “A set of sciences, theories and techniques whose purpose is to reproduce by a machine the cognitive abilities of a human being. Current developments aim, for instance, to be able to entrust a machine with complex tasks previously delegated to a human.”

² These Guidelines follow and build on the Report on Artificial Intelligence (“Artificial Intelligence and Data Protection: Challenges and Possible Remedies”) available at : <https://rm.coe.int/artificial-intelligence-and-data-protection-challenges-and-possible-re/168091f8a6>

³ Amending Protocol CETS n°223 to the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data.

⁴ <https://rm.coe.int/t-pd-2017-1-bigdataguidelines-en/16806f06d0>

II. Guidance for developers, manufacturers and service providers

1. AI developers, manufacturers and service providers should adopt a values-oriented approach in the design of their products and services, consistent with Convention 108+, in particular with article 10.2, and other relevant instruments of the Council of Europe.
2. AI developers, manufacturers and service providers should assess the possible adverse consequences of AI applications on human rights and fundamental freedoms, and, considering these consequences, adopt a precautionary approach based on appropriate risk prevention and mitigation measures.
3. In all phases of the processing, including data collection, AI developers, manufacturers and service providers should adopt a human rights by-design approach and avoid any potential biases, including unintentional or hidden, and the risk of discrimination or other adverse impacts on the human rights and fundamental freedoms of data subjects.
4. AI developers should critically assess the quality, nature, origin and amount of personal data used, reducing unnecessary, redundant or marginal data during the development, and training phases and then monitoring the model's accuracy as it is fed with new data. The use of synthetic data⁵ may be considered as one possible solution to minimise the amount of personal data processed by AI applications.
5. The risk of adverse impacts on individuals and society due to de-contextualised data⁶ and de-contextualised algorithmic models⁷ should be adequately considered in developing and using AI applications.
6. AI developers, manufacturers and service providers are encouraged to set up and consult independent committees of experts from a range of fields, as well as engage with independent academic institutions, which can contribute to designing human rights-based and ethically and socially-oriented AI applications, and to detecting potential bias. Such committees may play an especially important role in areas where transparency and stakeholder engagement can be more difficult due to competing interests and rights, such as in the fields of predictive justice, crime prevention and detection.
7. Participatory forms of risk assessment, based on the active engagement of the individuals and groups potentially affected by AI applications, should be encouraged.
8. All products and services should be designed in a manner that ensures the right of individuals not to be subject to a decision significantly affecting them based solely on automated processing, without having their views taken into consideration.
9. In order to enhance users' trust, AI developers, manufacturers and service providers are encouraged to design their products and services in a manner that safeguards users' freedom of choice over the use of AI, by providing feasible alternatives to AI applications.
10. AI developers, manufacturers, and service providers should adopt forms of algorithm vigilance that promote the accountability of all relevant stakeholders throughout the entire life cycle of these applications, to ensure compliance with data protection and human rights law and principles.
11. Data subjects should be informed if they interact with an AI application and have a right to obtain information on the reasoning underlying AI data processing operations applied to them. This should include the consequences of such reasoning.
12. The right to object should be ensured in relation to processing based on technologies that influence the opinions and personal development of individuals.

⁵ Synthetic data are generated from a data model built on real data. They should be representative of the original real data. See the definition of synthetic data in OECD. 'Glossary of Statistical Terms'. 2007. http://ec.europa.eu/eurostat/ramon/coded_files/OECD_glossary_stat_terms.pdf ("An approach to confidentiality where instead of disseminating real data, synthetic data that have been generated from one or more population models are released").

⁶ This is the risk of ignoring contextual information characterising the specific situations in which the proposed AI-based solutions should be used.

⁷ This happens when AI models, originally designed for a specific application, are used in a different context or for different purposes.

III. Guidance for legislators and policy makers

1. Respect for the principle of accountability, the adoption of risk assessment procedures and the application of other suitable measures, such as codes of conduct and certification mechanisms, can enhance trust in AI products and services.
2. Without prejudice to confidentiality safeguarded by law, public procurement procedures should impose on AI developers, manufacturers, and service providers specific duties of transparency, prior assessment of the impact of data processing on human rights and fundamental freedoms, and vigilance on the potential adverse effects and consequences of AI applications (hereinafter referred to as algorithm vigilance⁸).
3. Supervisory authorities should be provided with sufficient resources to support and monitor the algorithm vigilance programmes of AI developers, manufacturers, and service providers.
4. Overreliance on the solutions provided by AI applications and fears of challenging decisions suggested by AI applications risk altering the autonomy of human intervention in decision-making processes. The role of human intervention in decision-making processes and the freedom of human decision makers not to rely on the result of the recommendations provided using AI should therefore be preserved.
5. AI developers, manufacturers, and service providers should consult supervisory authorities when AI applications have the potential to significantly impact the human rights and fundamental freedoms of data subjects.
6. Cooperation should be encouraged between data protection supervisory authorities and other bodies having competence related to AI, such as: consumer protection; competition; anti-discrimination; sector regulators and media regulatory authorities.
7. Appropriate mechanisms should be put in place to ensure the independence of the committees of experts mentioned in Section II.6.
8. Individuals, groups, and other stakeholders should be informed and actively involved in the debate on what role AI should play in shaping social dynamics, and in decision-making processes affecting them.
9. Policy makers should invest resources in digital literacy and education to increase data subjects' awareness and understanding of AI applications and their effects. They should also encourage professional training for AI developers to raise awareness and understanding of the potential effects of AI on individuals and society. They should support research in human rights-oriented AI.

⁸ On the notion of algorithmic vigilance, as adoption of accountability, awareness and risk management practices related to potential adverse effects and consequences throughout the entire life cycle of these applications see also 40th International Conference of Data Protection and Privacy Commissioners, Declaration on Ethics and Data Protection in Artificial Intelligence, guiding principle no. 2. See also the Report on Artificial Intelligence (footnote 2), Section II.4