

CDDH-IA(2024)07

06/11/2024

STEERING COMMITTEE FOR HUMAN RIGHTS

COMITÉ DIRECTEUR POUR LES DROITS HUMAINS

(CDDH)

**DRAFTING GROUP ON HUMAN RIGHTS AND ARTIFICIAL
INTELLIGENCE**

*GROUPE DE RÉDACTION SUR LES DROITS HUMAINS ET
L'INTELLIGENCE ARTIFICIELLE*

(CDDH-IA)

**Compilation of written contributions and presentations received from
experts of the exchange of views of the 1st meeting of CDDH-IA**

**Compilation de contributions écrites et présentations reçues des experts
de l'échange de vues de la 1^{ère} réunion du CDDH-IA**

TABLE OF CONTENTS / TABLE DES MATIÈRES

Technical aspects of AI related to human rights By Marko GROBELNIK.....	3
AI, the European Convention on Human Rights and third generation rights By Alberto QUINTAVALLA.....	16
AI and economic, social and cultural rights By Miriam KULLMANN	28
AI, human rights, democracy and the rule of law: A primer prepared for the Council of Europe – Connecting Principles, Rights and Obligations By David LESLIE	37
The Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law By Thomas SCHNEIDER.....	43
Implications of generative artificial intelligence for freedom of expression By Andrin EICHIN	52
Council of Europe instruments and policies concerning the right to privacy and to personal data protection By Peter KIMPIAN	58
European Ethical Charter on the use of artificial intelligence in judicial systems and their environment By Jan SPOENLE.....	74
The Draft Methodology for the Risk and Impact Assessment of Artificial Intelligence Systems from the point of view of Human Rights, Democracy and the Rule of Law (HUDERIA) By CAI Secretariat	77

Technical aspects of AI related to human rights | Marko GROBELNIK



Anatomy of the AI Definition

Marko Grobelnik

(Marko.Grobelnik@ijs.si)

Artificial Intelligence Department, Jozef Stefan Institute

UNESCO IRCAI



Name of the game:
Definition of AI



Informal definition of non-AI

- AI is exactly the **opposite** from what is happening in the video...
- ...instead of living beings mimicking machines, AI is intended to make machines imitating living beings.



AI Definitions from the literature

- *"The exciting new effort to make computers think...[as] machines with minds, in the full and literal sense."* (Haugeland 1985)
- *"[The automation of] activities that we associate with human thinking such as decision-making, problem-solving, learning."* (Bellman 1978)
- *"The art of creating machines that perform functions that require intelligence when performed by people."* (Kurzweil 1990)
- *"The study of how to make computers do things which, at the moment, people are better."* (Rich and Knight 1991)
- *"The study of the computations that make it possible to perceive, reason, and act."* (Winston 1992)
- *"Making machines intelligent; intelligence is that quality that enables an entity to function appropriately and with foresight in its environment."* (Nils Nilsson)

OECD AI Definition (OECD 2019)

(adopted also by G20 and EC)

“An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments.

It does so by utilising machine and/or human-based inputs to:

- i) perceive and/or analyse real and/or virtual environments;*
- ii) abstract such perceptions/analyses into models manually or automatically; and*
- iii) use model interpretations to formulate options for outcomes.*

AI systems are designed to operate with varying levels of autonomy.”

New OECD AI System definition (Oct 16th 2023)

(adopted by EU AI Act, G7, US NIST, Council of Europe)

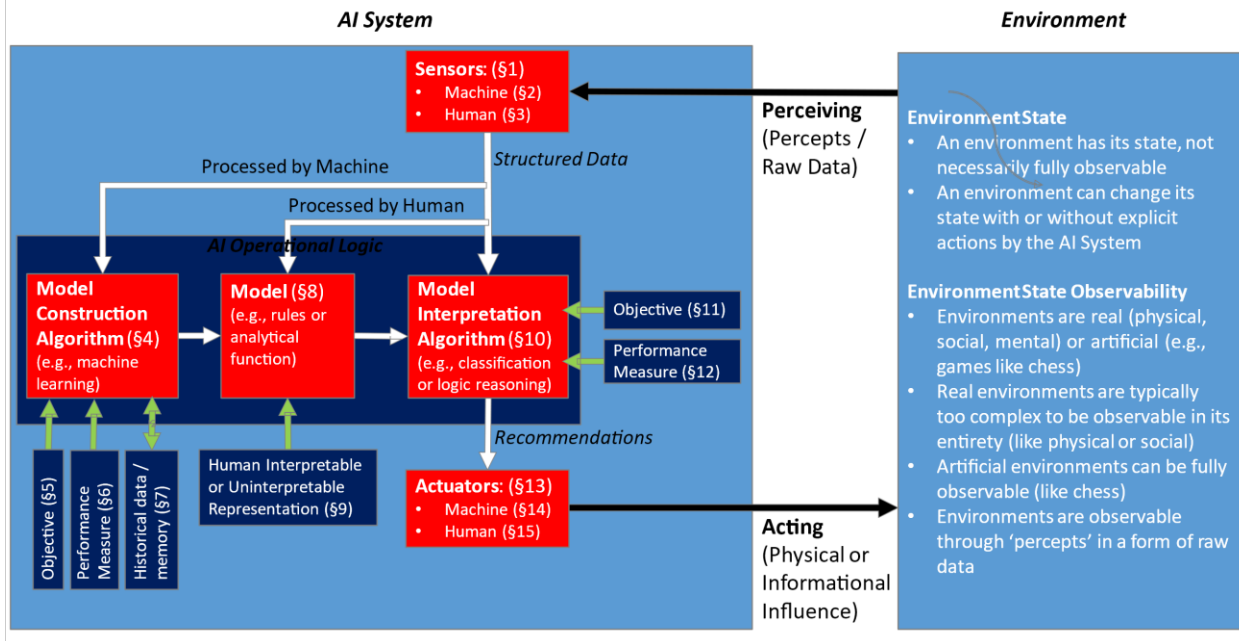
Proposed clean text:

An AI system is 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 [can] influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.

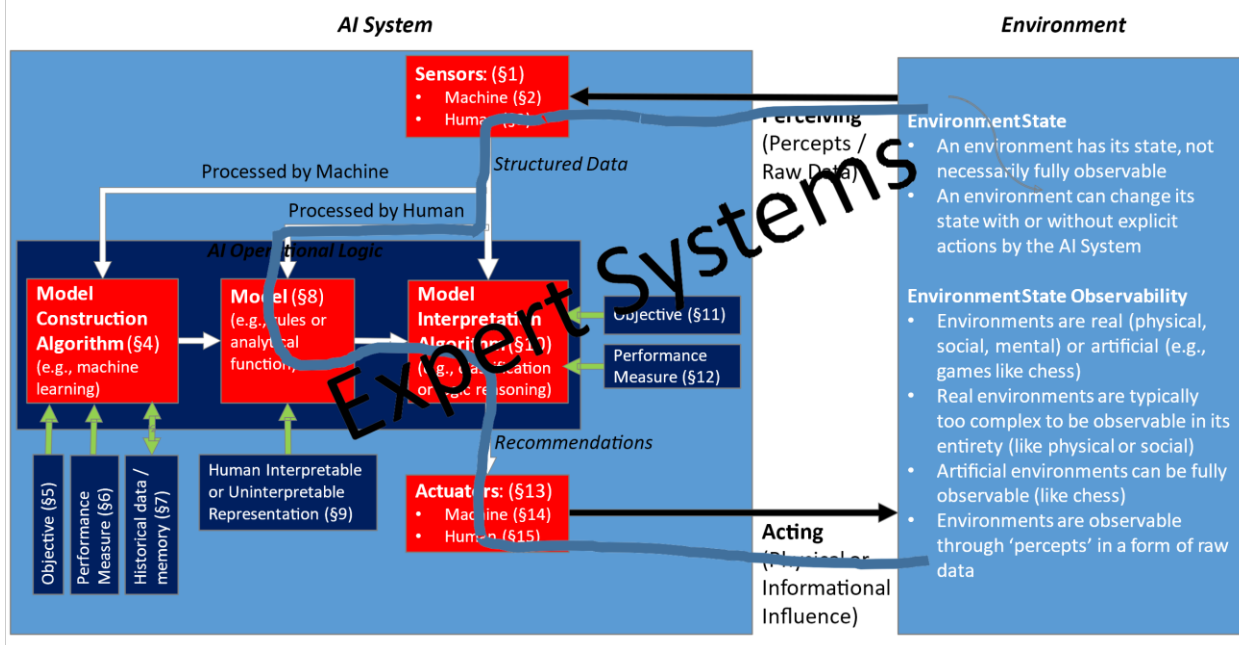
Proposed updates in blue: OECD AI System Definition from June 2019

*An AI system is a machine-based system that ~~can, for a given set of human-defined~~ **explicit or implicit** objectives, **infers, from the input it receives, how to generate outputs such as** ~~makes~~ **predictions, content, recommendations, or decisions that [can] influence** **physical** ~~real~~ or virtual environments. **Different AI systems** ~~are designed to operate with varying~~ **in their levels of autonomy and adaptiveness after deployment.***

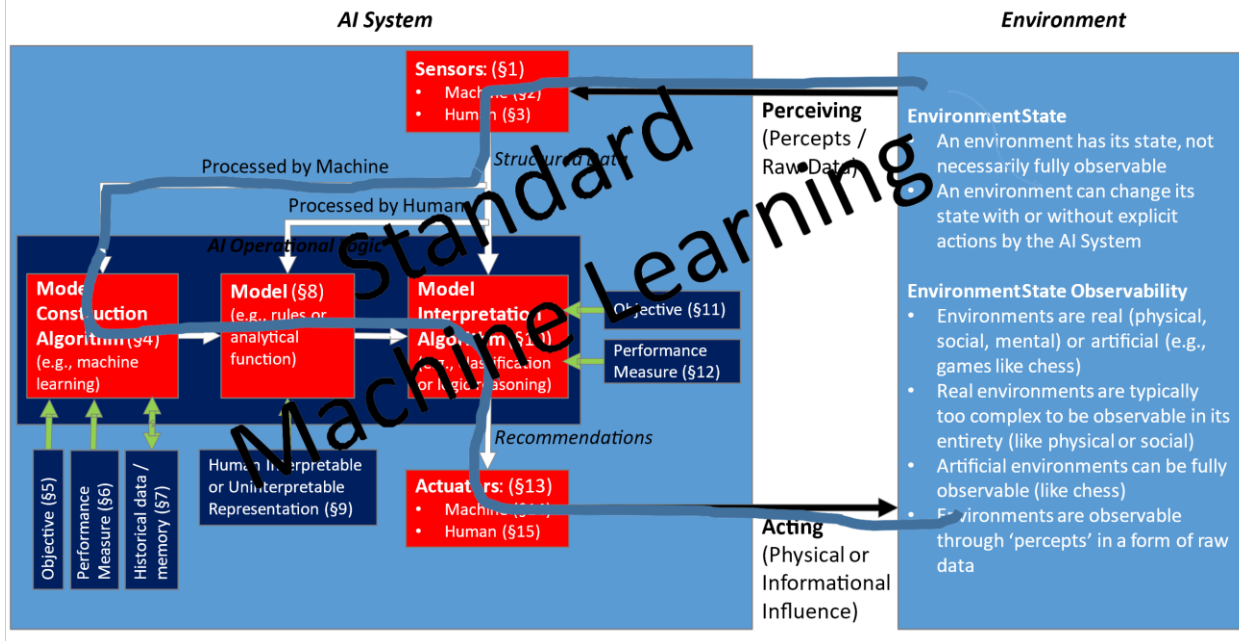
Anatomy of the AI System definition (as defined by OECD)



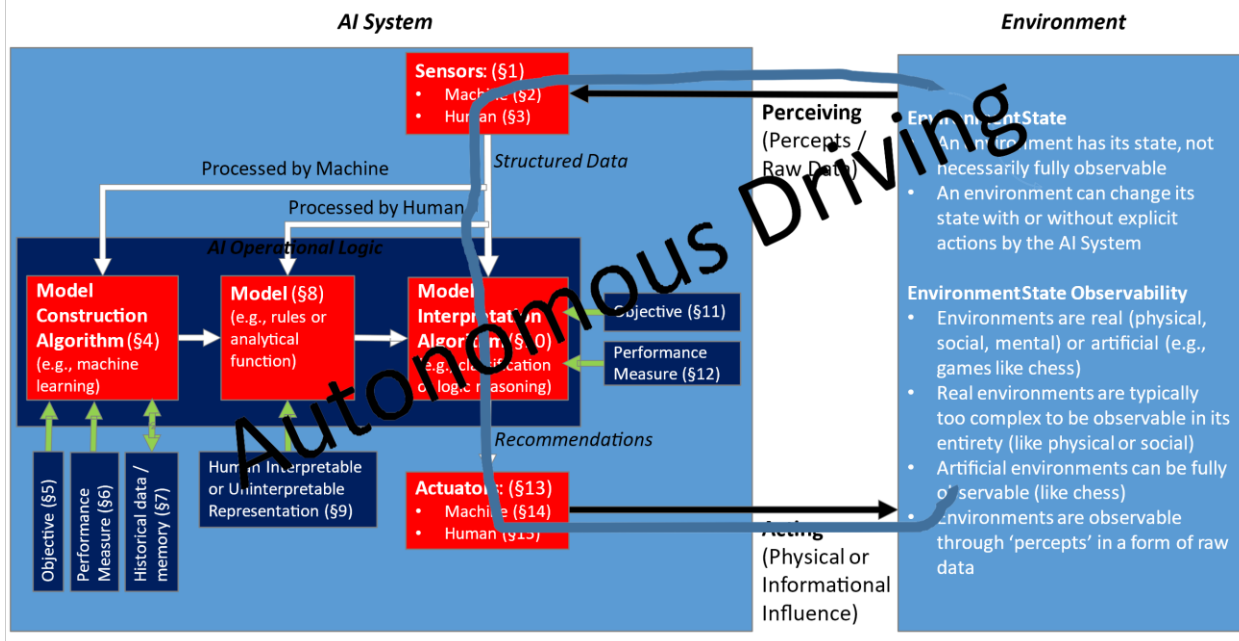
AI System as defined by OECD



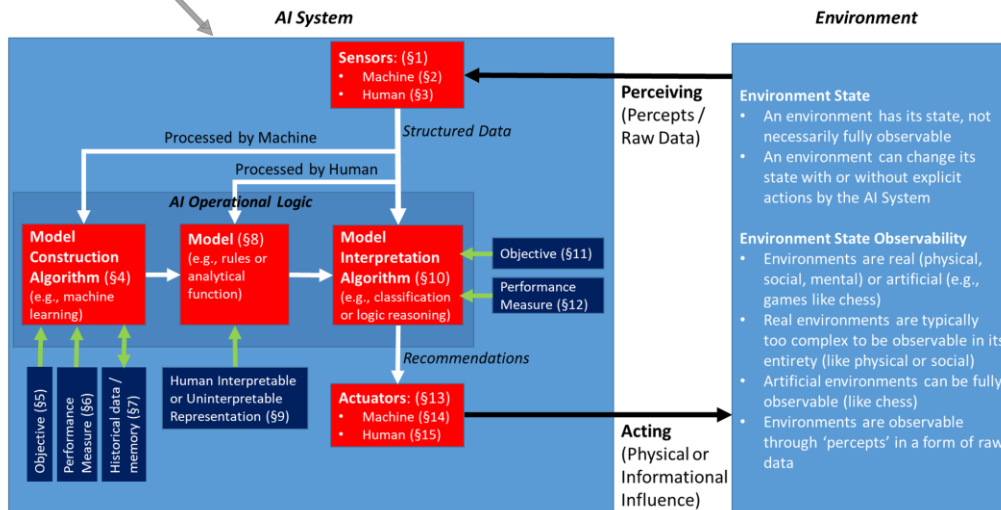
AI System as defined by OECD



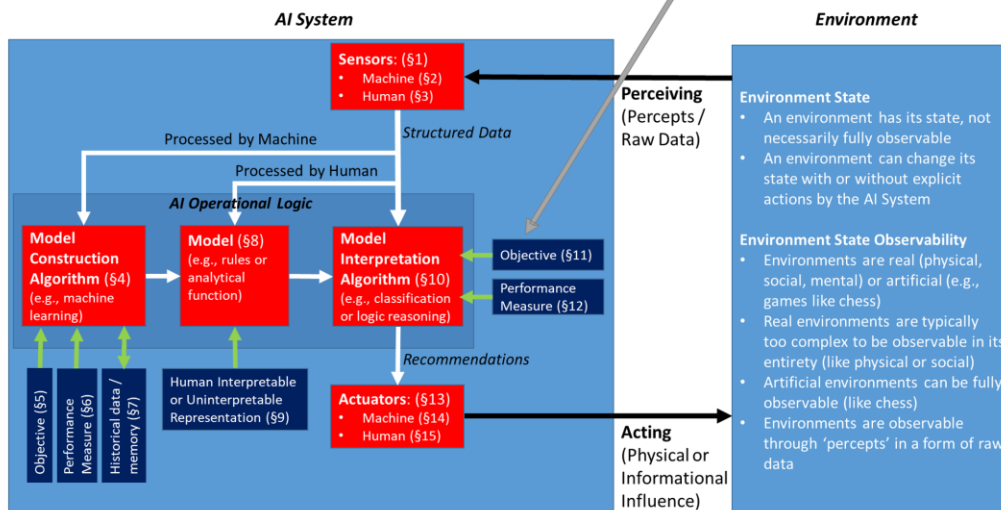
AI System as defined by OECD



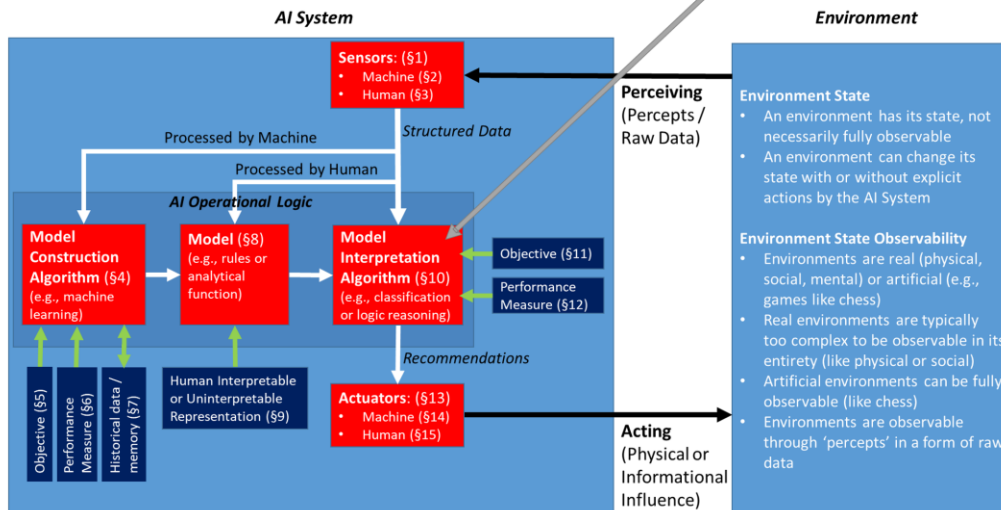
An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that [can] influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.



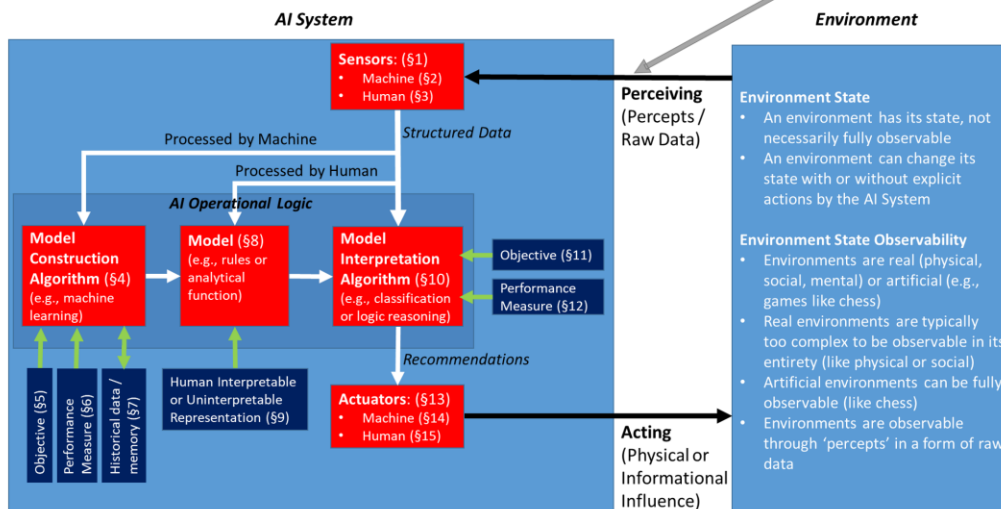
An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that [can] influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.



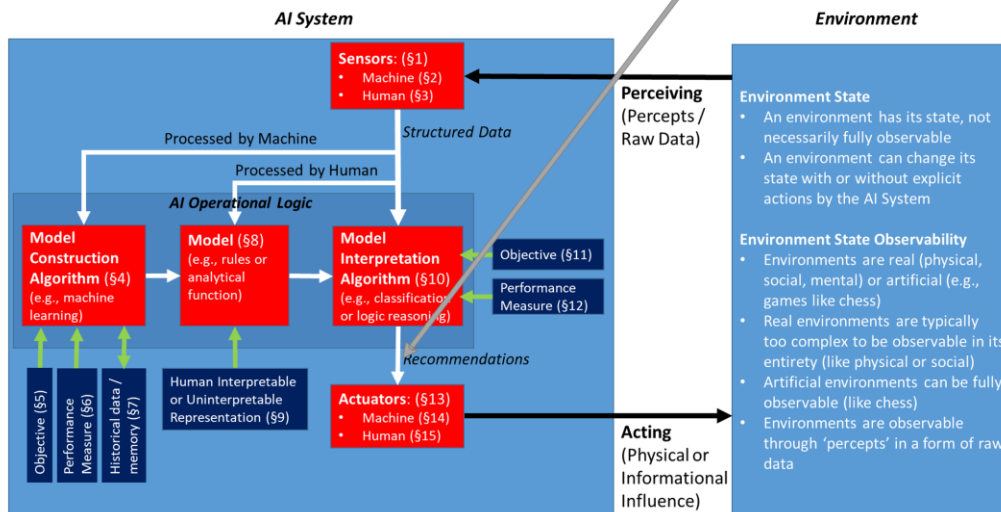
An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that [can] influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.



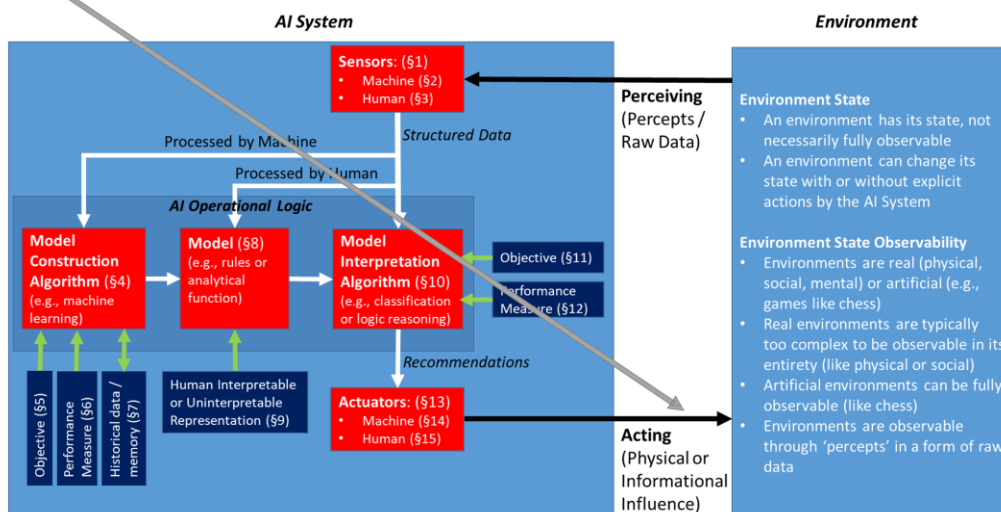
An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that [can] influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.



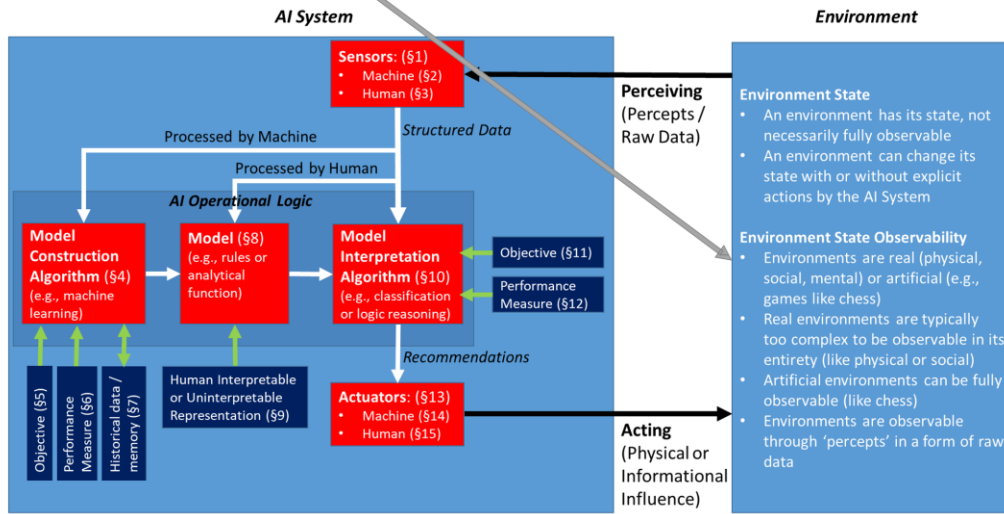
An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that [can] influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.



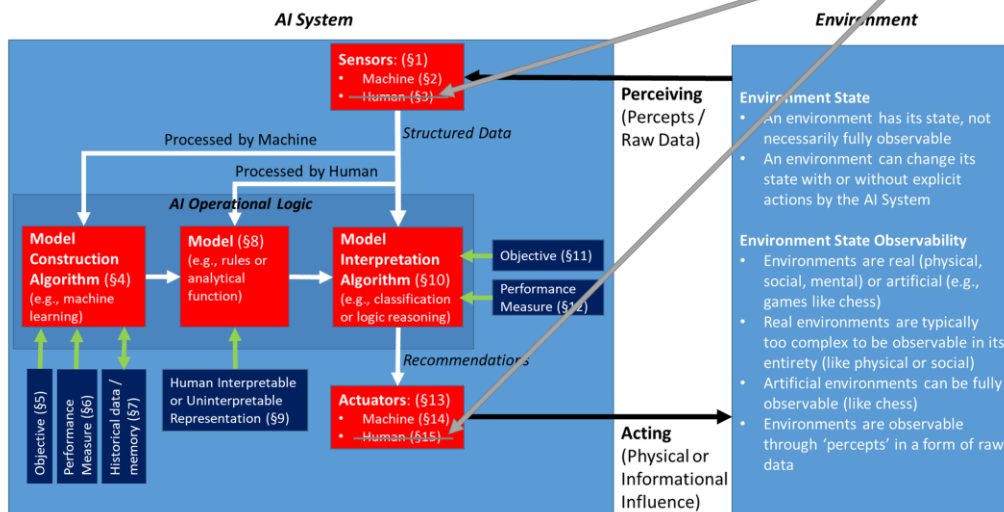
An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that [can] influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.



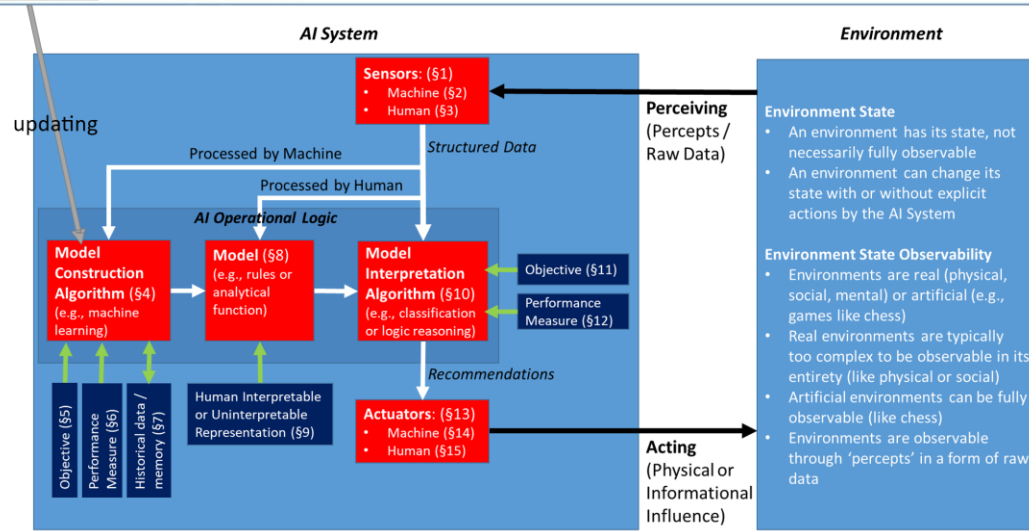
An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that [can] influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.



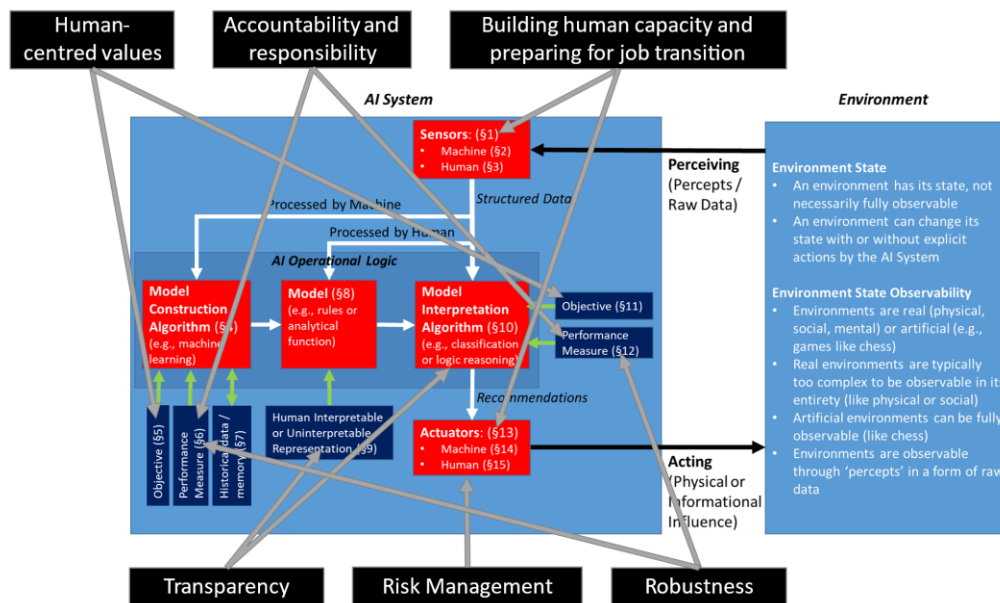
An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that [can] influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.



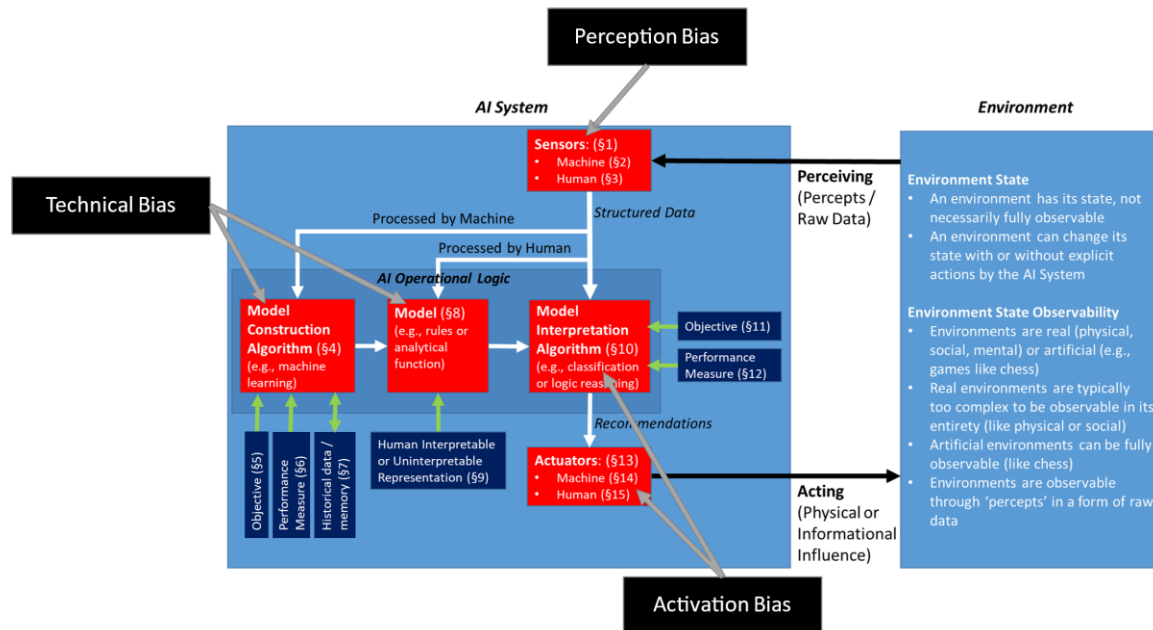
An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that [can] influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.



AI System and relation to higher level principles



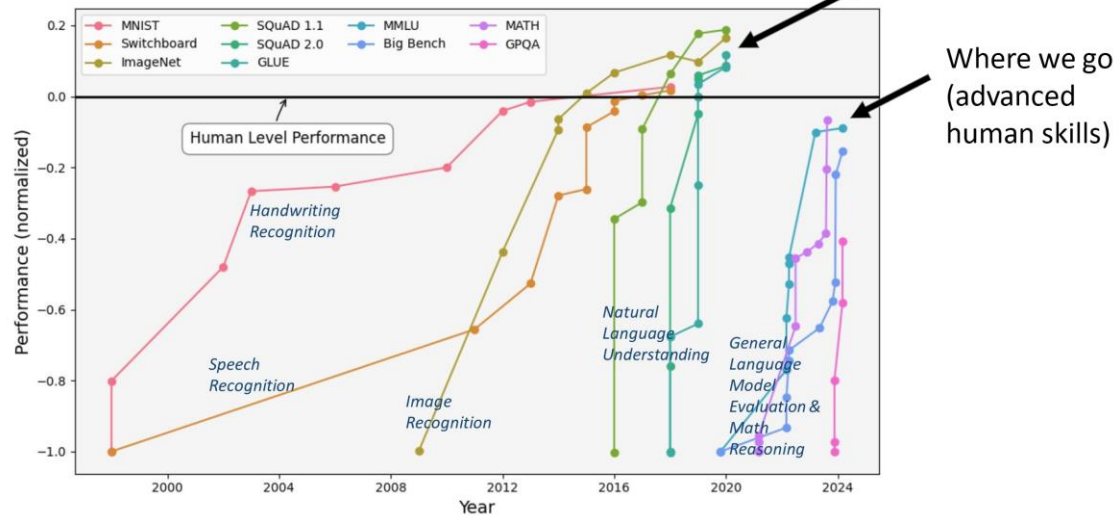
AI System and sources of various types of biases



Technical triggers of
AI & human rights

AI Systems vs. Human Capabilities

(Evolution of AI systems related to human skills)



<https://www.gov.uk/government/publications/international-scientific-report-on-the-safety-of-advanced-ai>

(Some of) the basic properties of AI systems which could endanger human rights

- Managing large scale of complexity
 - ...using the scale of data in the size of all human digital content
 - Humans cannot manage complexity beyond certain scale
- Black-box representation
 - ...suitable for machine, but not for human
 - Humans don't have feedback into machine (by explanation)
- Speed of inferencing
 - Surpassing humans in reaction time
- Autonomous Decision-Making
 - ...due to misalignment of human vs. machine value systems
- Unclear accountability
 - ...the chain of stakeholders (from humans to machine) in the process is long
- Robustness
 - ...AI systems are not perfect (especially GenAI) and is hard to guarantee stable results



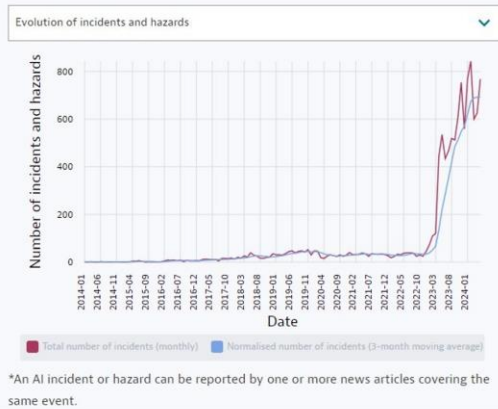
(Un)Known-(Un)Knowns – Model Representation vs. Phenomena Discovery

		Phenomena Discovery	
		Phenomena <i>Known to Humans</i> (what people already know, but want to model and understand)	Phenomena <i>Unknown to Humans</i> (what people typically don't know yet)
Model Representation	Human Interpretable (provided by a human to a machine)	Traditional Statistics, Traditional AI, Logic Reasoning	Advanced Statistical Methods, Unsupervised AI (e.g. anomaly detection)
	Human Uninterpretable (created by a machine to optimize the solution)	Modern AI (after 2010), Deep Neural Networks, Transformers, Reinforcement Learning	AI to come, e.g., AI with “multihop” reasoning, Online Reinforcement Learning

...this would allow to reach yet undiscovered concepts and relations and reach insights far from what humanity knows today

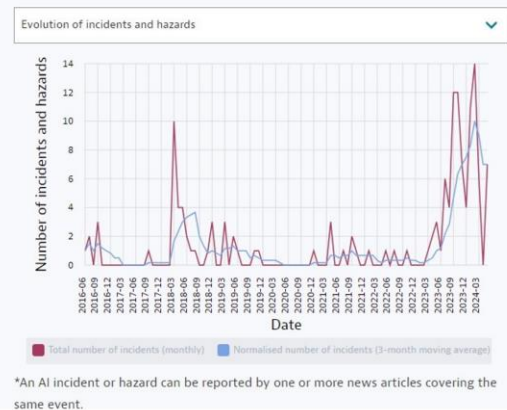
Brief statistics of AI incidents

Summary visualisations of AI incidents & hazards



Overall number of incidents increased after 2022

Summary visualisations of AI incidents & hazards



Number of deaths and physical harm stayed mostly stable

AI, the European Convention on Human Rights and third generation rights | Alberto QUINTAVALLA

The intersection between artificial intelligence, human rights and environment in the context of the European Convention on Human Rights

*Alberto Quintavalla**

The aim of this contribution is to provide an overview of the vast area at the intersection between human rights and artificial intelligence. Not all corners of the complex web of relationships that exist between human rights and artificial intelligence are given equal attention in research. Moreover, the latest technological developments have led public opinion and academic literature to focus on certain specific human rights, thereby overlooking the various human rights that may come into play when it comes to AI development and deployment.

The specific human rights that take the lion's share of this discussion are, on the one hand, the right to privacy and data protection and, on the other hand, non-discrimination. This focus is readily explainable. AI applications are data-driven, requiring an ever-increasing amount of data and personal information. At the same time, these AI applications allow for a more detailed analysis than ever before, as data from different sources can be combined, including statistical and publicly available data, easily giving a detailed picture of behavior of individuals or groups with the click of a button. It is thus very likely that violations of the right to privacy and data protection will arise from the collection and processing of personal data.

This is even more the case since personal data collection has become an essential activity of several governments for using AI technology to perform their tasks. A clear example is personal data collection for law enforcement authorities. The collection of personal data in that realm can serve several purposes, from prediction and preemption for terrorism prevention to datafication of evidence gathering and state surveillance. Although this may raise several concerns for the violation of the right to privacy and data protection, one should note that the existing regulation is solid, ranging from international law (such as the Council of Europe Convention 108+) to domestic law, including standards and frameworks.

Hopefully, this set of laws can provide effective protection mechanisms, as exemplified by existing case law such as the recent one from the European Court of Human Rights on facial recognition technology.¹ This does not however mean that society is safe from violations of the human right to privacy and data protection. For instance, data protection law does not apply to predictive AI models since they do not relate to identifiable persons such those models used in mortgage lending where different risk factors are attributed based on (e.g.) postal codes.² This latter example also highlights the related issue of discrimination in the use of AI technology.

* Assistant Professor and Co-Director of the Erasmus Centre of Law and Digitalization, Erasmus School of Law (Erasmus University Rotterdam).

¹ *Glukhin v. Russia*, App. No. 11519/20 (ECtHR 4 July 2023).

² Zuiderveen Borgesius, F. (2018). Discrimination, artificial intelligence, and algorithmic decision-making. *Council of Europe, Directorate General of Democracy*, 42, pp. 44-45.

Protection from discrimination is the other human right that takes the lion's share of the discussion. AI applications, as humans, are prone to discrimination and biases. There are various avenues where biases find their way into the AI life cycle.³ For instance, the bias can be included in the dataset used to train the AI application from scratch, for example with facial recognition systems, or it can be the human developers to incorporate them. An example of the latter is what occurred in the Netherlands with the so-called '*Toeslagenaffaire*'. There, nationality was considered a risk factor in the AI model and non-Dutch nationals received higher risk-scores when it came to detecting fraudulent or incorrect child benefits applications. The same goes for the so-called predictive policing AI, where research has historically shown that digital police systems exacerbate bias, based on incomplete data or the over-targeting of certain neighborhoods.⁴ The legislation on the protection from discrimination is well-developed. Nonetheless, one should acknowledge that ensuring protection from discrimination is a tall order, intensified by the complexities of the technical systems. For instance, an issue highlighted in the literature is when an AI system invents new classes that do not correlate with protected characteristics to differentiate between people.⁵ While this differentiation could be unfair, most non-discrimination laws would not protect those same characteristics.⁶

Hence, the role played by privacy and non-discrimination when discussing the interaction between AI and human rights is undeniably important because it shows that AI can negatively impact human rights. However, as this contribution tries to clarify, an overly piecemeal approach may fail to account for the more pervasive impact of AI technology on human rights and, in turn, have negative repercussions.

In fact, a more systematic view on the interaction between human rights and AI can have a series of advantages. First, and at a conceptual level, we should devote attention to all human rights pursuant to the principles of indivisibility, interdependency and interrelatedness.⁷ While the practical implementation of human rights may have a different speed⁸ – especially in light of the progressive realization of the economic, social and cultural rights, all human rights are highly intertwined besides being of equal relevance. It is thus unlikely that we could fully isolate a given human right without considering possible negative effects on other human rights due to their indivisibility, interdependency and interrelatedness.

Second, a more systematic view is necessary because in certain instances – and as this contribution explores later – there may be both positive and negative consequences on different

³ Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *Calif. L. Rev.*, 104, 671.

⁴ Alikhademi, K., Drobina, E., Prioleau, D., Richardson, B., Purves, D., & Gilbert, J. E. (2022). A review of predictive policing from the perspective of fairness. *Artificial Intelligence and Law*, 1.

⁵ Zuiderveen Borgesius, F. (2018). Discrimination, artificial intelligence, and algorithmic decision-making. *Council of Europe, Directorate General of Democracy*, 42, p. 36.

⁶ Ibid.

⁷ See, e.g., United Nations. Vienna Declaration and Programme of Action, World Conference on Human Rights. (1993) UN Doc. A/ CONF.157/ 24.

⁸ Quintavalla, A., & Heine, K. (2019). Priorities and human rights. *The International Journal of Human Rights*, 23(4), 679-697.

human rights stemming from the use of AI. In other words, the use of an AI application may lead to the enhancement of a human right while at the same time harming another one. This situation would thus call for the balancing of different human rights.

Third, and relatedly, a more systematic view could make it easier to identify some common risks and challenges among a wide array of human rights.⁹ In this way, mapping these risks and highlighting these challenges would serve as a compass for policymakers when it comes to developing effective safeguards to prevent and address possible human rights violations.

There have already been attempts to provide a more systematic view in scientific literature.¹⁰ By analyzing the impact of AI on human rights, it is possible to identify a series of general findings that, as previously indicated, can bring a series of advantages for a framework that is respectful and protective of all human rights, regardless of which generation they belong to.

First, and perhaps still trivial, AI does not limit itself to negatively impacting a given human right as is often the case for the right to privacy and protection from discrimination. It may well be that the impact of AI can be both positive and negative. Thereby, we should consider AI as having a two-way relationship with (a series of) human rights.

Take the right to fair trial as an example.¹¹ AI can improve the access to legal resources, identify relevant case law and – thereby – facilitate the work of lawyers and judges. For instance, AI can predict outcomes of a trial by statistically modelling previous decisions made by the judiciary.¹² In short, AI can make justice more accessible and, possibly, reduce the time invested in legal disputes. All these abilities can enhance the right to fair trial. However, AI applications can also lead to wrong and harmful outcomes. For instance, the literature points at several potentially discriminatory uses of AI in trial proceedings.¹³ In this way, AI would violate the right to fair trial and, relatedly, protection from discrimination.

This latter point is important because it also raises the issue that an AI application can infringe on more than a single human right at once – in the case mentioned, right to fair trial and non-discrimination. Another commonly raised example is the use of AI applications by authoritarian

⁹ Prifti K., Quintavalla, A., & Temperman, J., 'Artificial Intelligence and Human Rights: Understanding and Governing Common Risks and Benefits' in Quintavalla, A., & Temperman, J. (2023). *Artificial intelligence and human rights*, Oxford University Press, 441.

¹⁰ Quintavalla, A., & Temperman, J. (2023). Artificial intelligence and human rights. *Oxford University Press*; Rodrigues, R. (2020). Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. *Journal of Responsible Technology*, 4, 100005.

¹¹ Universal Declaration of Human Rights (adopted 10 December 1948) UNGA Res 217 A(III) (UDHR) art 10; International Covenant on Civil and Political Rights (adopted 16 December 1966, entered into force 23 March 1976) 999 UNTS 171 (ICCPR), art 14; Convention for the Protection of Human Rights and Fundamental Freedoms (European Convention on Human Rights, as amended) (ECHR), art 6.

¹² See, e.g., Colenette, J., Atkinson, K., & Bench-Capon, T. (2023). Explainable AI tools for legal reasoning about cases: A study on the European Court of Human Rights. *Artificial Intelligence*, 317, 35.

¹³ See, e.g., O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy, *Penguin*, pp. 20–25. For a more optimistic account, see Morin-Martel, A. (2024). Machine learning in bail decisions and judges' trustworthiness. *AI & Society*, 39(4), 2033.

regimes for surveillance and censorship purposes.¹⁴ There, the right to privacy, the freedom of assembly and association, and the freedom of expression can all be violated.¹⁵

Accordingly, AI and human rights have a two-way relationship. As indicated in the Framework Convention of the Council of Europe, AI technology offers unprecedented opportunities to protect and promote human rights but at the same time “certain activities within the life cycle of AI systems may undermine human dignity, individual autonomy and human rights”.¹⁶

The current description of the impact of AI may still seem rather straightforward. An AI application can either negatively or positively impact one or more human rights. However, one should note that there are cases in which AI causes a friction between two different human rights by enhancing one at the expenses of the other. This is the second general finding: AI and human rights can have a double-edged relationship. To put it shortly, there can exist a web of conflicting or mutually reinforcing relationships.

An example arises from the so-called “internet of bodies”, which is a network of devices such as smart watches, sleep trackers or digital pills worn by a person and connected to the Internet.¹⁷ With the internet of bodies, there is a continuous collection of data from the users that serves to enhance their health by providing constant feedback of information to the users and, potentially, third parties.

It is in this context that there may be a tradeoff between the right to privacy, on the one hand, and the right to health, on the other hand. Indeed, while individuals should always be entitled to revoke consent for the collection of data, the revocation of consent within the context of the internet of bodies may lead to the risk of deteriorating their health – thereby infringing their right to health at individual level.¹⁸ This is not the only trade-off between human rights that exist in this example. Another one can pertain to the right to health, on the one hand, and – what is named in certain jurisdictions – the right to a healthy environment, on the other hand. While the internet of bodies would enhance the right to health by collecting health data both at individual and collective levels, it could simultaneously lead to environmental degradation due to the high ecological footprint associated with AI technology.¹⁹

The fact that there exists a web of complex relationships between AI and human rights has relevant repercussions from a regulatory perspective. It forces policymakers and judges to consider how to deal with all the different human rights trade-offs arising from AI development and deployment. Due to this consideration, one may assume that the proportionality principle will

¹⁴ Koniakou, V. (2023). From the “rush to ethics” to the “race for governance” in Artificial Intelligence. *Information Systems Frontiers*, 25(1), 71.

¹⁵ *S. and Marper v. the United Kingdom*, App. Nos. 30562/04 & 30566/04 (ECtHR 4 December 2008).

¹⁶ Preamble, Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, Council of Europe Treaty Series - No. 225.

¹⁷ Matwyshyn, A. M. (2019). The internet of bodies. *Wm. & Mary L. Rev.*, 61, 77.

¹⁸ Vayena, E., Haeusermann, T., Adjekum, A., & Blasimme, A. (2018). Digital health: meeting the ethical and policy challenges. *Swiss medical weekly*, 148, w14571.

¹⁹ See *infra*.

play an ever-greater role in human rights law. Despite controversial,²⁰ the proportionality principle can offer a structured procedural avenue on how to handle trade-offs between human rights.²¹ Another advantage of the proportionality principle is that, by playing an active role in human rights litigation, it would allow for an actual balancing of interests in more practical settings.

There is a third general finding, as well: the relationship between AI and human rights is bilateral. This finding means that human right norms can contribute to addressing the impact of AI. One should only think about autonomous weapons and their impact on the right to life. It is for this reason that there is a clear call to ban autonomous weapons in warfare. A similar issue arises from the use of facial recognition technology. Accordingly, regulation can have a formative role in the development and deployment of AI technology. This latter consideration is, however, not exempted of challenges.

The current structure of human rights law makes it difficult to address the impact of AI. Although both AI and human rights started developing in the late 1940s, human right norms predate AI technology. It follows that certain notable gaps exist. The international human rights framework concerning AI is indeed far from being fully evolved, particularly the role and responsibilities of private businesses.

Private businesses played a central role in the AI sector being the key actors in the development stage and, sometimes, in the deployment stage. For instance, governments rely on non-state actors to undertake some governmental activities,²² including the criminal justice sector.²³ Moreover, non-state actors perform governance activities themselves to address governance gaps, such as in the development of smart cities.²⁴ It is in this context that human rights law should develop ways to deal with the role of private businesses.

The indirect horizontal effect of human rights and the existing UN human rights guidelines have thus far been used to mitigate this issue. As to the indirect horizontal effect of human rights, states have a positive obligation to protect human rights from harmful interference by non-state actors.²⁵ Hence, states should adopt more effective measures to protect individuals from harm by non-state actors. As to the existing regulatory framework, soft law instruments on corporate

²⁰ Tsakyrakis, S. (2009). Proportionality: An assault on human rights? *International Journal of Constitutional Law*, 7(3), 468.

²¹ Rivers, J. (2006). Proportionality and variable intensity of review. *The Cambridge Law Journal*, 65(1), 174, p. 176.

²² Office of the United Nations High Commissioner for Human Rights (OHCHR), 'B-Tech: Bridging Governance Gaps in the Age of Technology—Key Characteristics of the State Duty to Protect' (May 2021).

²³ Završnik, A. (2020). Criminal Justice, Artificial Intelligence Systems, and Human Rights. *ERA Forum*, 20, 67.

²⁴ Lane, L. (2022). Clarifying human rights standards through artificial intelligence initiatives. *International & Comparative Law Quarterly*, 71(4), 915, p. 917.

²⁵ Lane, L. (2018). The horizontal effect of international human rights law in practice: a comparative analysis of the general comments and jurisprudence of selected United Nations human rights treaty monitoring bodies. *European Journal of Comparative Law and Governance*, 5(1), 5.

responsibility to respect human rights have been adopted.²⁶ However, both strategies started only recently to emerge, and their actual impact has still to be assessed.

A last point should be discussed when it comes to looking at the interaction between AI and human rights. Thus far, the discussion has been on what type of interaction occurs between AI and *existing* human rights. There are however other emerging human rights issues that come to play when it comes to analyzing the impact of AI. One example that was mentioned previously is the issue of environmental protection or the ‘right to a healthy environment’.

In this context, there are several ways in which AI interacts with environmental protection. Unsurprisingly, these ways can be both positive – AI applications that help rangers indicate the most targeted areas in national parks by poachers and illegal loggers – and negative – misuse of the very same AI applications by illegal loggers and poachers themselves to avoid the said areas.²⁷ Nonetheless, the most noteworthy impact of AI applications arises from the high energy and water use required by the AI sector.²⁸

However, the issue is more difficult than it seems at first glance. To start with, it is very difficult to quantify the actual environmental impact of AI. For instance, there is disagreement on what factors to incorporate when it comes to quantifying the carbon footprints.²⁹ Secondly, there is currently no right to a healthy environment as such in the Council of Europe. That implies that individuals should ground claims on the infringement of other formally recognized human rights such as the right to privacy – as already happened in existing case law where individuals litigated cases of environmental deterioration.³⁰ However, in these very same cases, it is evident that environmental integrity is only instrumental to the determine an infringement on an individual’s human right(s).³¹

This approach makes it very hard for individuals to initiate litigation on environmental harm due to AI technology. The reason is that it would be difficult to give reasonable and convincing evidence of the occurrence of individual harm and, accordingly, fulfilling the so-called victim requirement as an admissibility criterion. This is even more so with the previously mentioned quantification problem. The issue of proving individual harm is also common in other cases of technology and mass surveillance. There, however, the European Court of Human Rights had

²⁶ Human Rights Council/UNGA, ‘Protect, Respect and Remedy: A Framework for Business and Human Rights’ (7 April 2008) UN Doc A/HRC/8/5; OHCHR, ‘Guiding Principles on Business and Human Rights: Implementing the United Nations’ “Protect, Respect and Remedy” Framework’ (21 March 2011) UN Doc A/HRC/17/31.

²⁷ Kwok, R. (2019). AI empowers conservation biology. *Nature*, 567(7746), 133.

²⁸ Quintavalla, A. (2023). Artificial Intelligence and the Right to a Healthy Environment. in Quintavalla, A., & Temperman, J. (2023). *Artificial intelligence and human rights*, Oxford University Press, 425.

²⁹ Dhar, P. (2020). The carbon impact of artificial intelligence. *Nature Machine Intelligence*, 2, 423.

³⁰ See, e.g., *Fadeyeva v Russia*, App. No. 55723/00 (ECtHR, 9 June 2005); *Kyrtatos v Greece*, App. No. 41666/98 (ECtHR, 22 May 2003).

³¹ Francioni, F. (2010). International human rights in an environmental horizon. *European Journal of International Law*, 21(1), 41.

been willing to relax this criterion.³² The *actio popularis* before the European Court of Human Rights does not seem a promising avenue for litigation, either.³³

There is a last challenge that should be considered in the context of the right to a healthy environment vis-à-vis AI technology. This is that human rights law tends to impose soft law instruments on states but not necessarily on corporate actors – something that affects the effectiveness of the existing regulatory framework to safeguard human rights from the impact of AI technology in general. These soft law instruments, moreover, are of general applicability and would require an adaptation to the specific sector of AI to prove more effective.

From the foregoing, one may conclude that AI and human rights is not just a two-way relationship; but it is also double-edged and bidirectional. This means that AI and human rights is a web of multilateral coexisting relationships, and it is society who decides what type of impact AI technology makes with the existing regulatory instruments. Moreover, there are other emerging human rights issues, such as environmental protection, that will need to be taken into consideration.

All these considerations call for developing a framework that would embrace the widespread impact of AI technology across different human rights. Such a framework should not only be broad in scope but also aim at its applicability at a wider geographical scale. The reason would indeed be to avoid large discrepancies in the regulatory approach of different countries, which could lead to an uneven impact of AI on human rights. Simultaneously, such a framework should take into consideration geographical diversity: some degree of heterogeneity is not just an inherent feature of international law but is also welcome. This need for contextualization also urges us to pay extra attention to the effectiveness of *ex post* human rights protection mechanisms to plug the gaps left by *ex ante* regulation.³⁴

³² *Case of Klass and others v. Germany*, App. No. 5029/71 (ECtHR 6 September 1978); *Case of Roman Zakharov v. Russia*, App. No. 47143/06 (ECtHR 4 December 2015).

³³ *Aksu v. Turkey*, App. Nos. 4149/04 and 41029/04 (ECtHR 15 March 2012); *Burden v. the United Kingdom*, App. No. 13378/05 (ECtHR 29 April 2008).

³⁴ Murray, D. (2020). Using Human Rights Law to Inform States' Decisions to Deploy AI. *AJIL Unbound*, 114, 158.

AI, the European Convention on Human Rights and third generation rights

Dr Alberto Quintavalla

Exchange of view – CDDH-IA Drafting Group (Council of Europe), 25 September 2024



The interaction between artificial intelligence and human rights

- Understudied subject matter in its comprehensiveness
- Focus on specific human rights– i.e., right to privacy and protection from discrimination
- AI applications are data-driven and biases are omnipresent
- **Right to privacy and non-discrimination** are instrumental for a “just” functioning of AI applications

The Erasmus logo, featuring the word "Erasmus" in a white, cursive script font on a red background.

The interaction between artificial intelligence and human rights



Is the impact more complex?

Erasmus

The interaction between artificial intelligence and human rights

Advantages of a more systematic view:

- Indivisibility principle
- Possible need for balancing different interests
- Common findings
- More informed decisions for policymakers

Erasmus

The interaction between artificial intelligence and human rights



AI has a two-way relationship with (a series of) human rights

e.g., right to a fair trial

e.g., right to a fair trial and protection from discrimination

Erasmus

The interaction between artificial intelligence and human rights



AI has a double-edged relationship with human rights

e.g., Internet of Bodies

e.g., right to privacy and right to health

e.g., right to health and “environmental rights”

proportionality principle

Erasmus

The interaction between artificial intelligence and human rights



AI has a bilateral relationship with human rights

e.g., lethal autonomous weapon systems and right to life

the “issue” of private businesses

Ezafun

The interaction between artificial intelligence and human rights

- Interaction between AI and **existing** human rights law
- Environmental protection as an emerging human rights issue
- AI brings both benefits and harms to the environment
- “Elephant in the room”: high energy and water use

Ezafun

The interaction between artificial intelligence and human rights

Difficult to address it from a human rights law perspective:

- Quantification problem
- No explicitly recognized right to a healthy environment

“environmental integrity is not seen as a value per se for the community affected or society as a whole, but only as a criterion to measure the negative impact on a given individual’s life, property, private and family life”

Kyrtatos v Greece App no 41666/98 (ECtHR, 22 May 2003)
- Role of transnational corporations

Ezafun

The interaction between artificial intelligence and human rights



AI and human rights is a web of multilateral co-existing relationship

Challenge of integrating the broad scope with contextual factors

Ezafun

AI and economic, social and cultural rights | Miriam KULLMANN

Steering Committee for Human Rights (CDDH) – Drafting Group on Human Rights and Artificial Intelligence

Exchange of views with experts, 25 September 2024, Strasbourg, Agora G.06

The use of automated decision-making systems and their impact on social and economic rights under the European Social Charter

Miriam Kullmann*

I. INTRODUCTION

Public authorities, such as government agencies, ministries or municipalities, as well as private entities, such as employers or credit institutions, increasingly make use of automated decision-making systems. For instance, public authorities may apply such systems for automating decisions on whether or not an individual is entitled to and will be granted an unemployment benefit or any other social benefit as well as access to social housing. Private entities, on the other hand, may use automated decision-making systems to sift through job applications, decide on job promotions or an improvement on employment granting, for instance, bonus payments, but equally job performance criteria may result in dismissals. Credit institutions use automated decision-making systems to assess the creditworthiness of individuals, deciding on whether or not someone is eligible for a mortgage and under what conditions. While such automated decision-making systems may offer benefits to those using them, such as efficiency and cost-reduction but also dealing with potential shortages on the labour market, these systems clearly have an impact on and thus bear serious risks for social and economic rights as laid down in the European Social Charter.³⁵

This contribution uses the term automated decision-making systems, referring to systems that execute decision models or paths based on algorithms.³⁶ An algorithm, as defined, for instance, by Solon Barocas and Andrew Selbst, 'is a formally specified sequence of logical operations that provides step-by-step instructions for computers to act on data and thus automate decisions'.³⁷ Notably, it involves the explicit decision of those working for public authorities and private entities, sometimes subject to public procurement processes or the consent of workers' representatives, to make use of automated decision-making and, in a next step, what kind of system is going to be used. What is important to highlight is that technology is not autonomous and that human

* Professor of labour law and social security law at Utrecht University and member of the European Committee of Social Rights.

³⁵ The European Social Charter is an international treaty that guarantees social and economic rights, namely a broad range of everyday human rights related to employment, housing, health, education, social protection and welfare. At the European level, the European Social Charter is the counterpart to the European Convention on Human Rights. See <www.coe.int/en/web/european-social-charter/home>.

³⁶ Algorithm Watch, *Automated Decision-Making Systems and Discrimination: Understanding causes, recognizing cases, supporting those affected* (June 2022) 5.

³⁷ Solon Barocas and Andrew D Selbst, 'Big Data's Disparate Impact' (2016) 104 *California Law Review* 671, 674. The term algorithm can also be used to refer to the mathematical definition, whereas the term computer systems can refer to automated decision processes reflecting decision policies implemented by (pieces of) software. See Joshua A Kroll et al., 'Accountable Algorithms' (2017) 165 *University of Pennsylvania Law Review* 633, 640.

beings have an influential role to play in designing the systems public authorities and private entities wish to employ.³⁸

Using automated decision-making systems was, initially, hoped to replace ‘the unreliable and often biased gut instinct’ or ‘anecdotal observation’.³⁹ Yet it cannot be ignored anymore that automating decision-making, even if at some point in this process a human is involved, can violate economic and social rights as for automated decision-making processes personal data is being processed. Relevant personal data are, e.g., a person’s age, gender, religion or belief, political views, trade union membership, place of living and/or work, family and/or marital status, the income situation and whether or not the person is in need of social security or social benefits, or an individual’s health condition. Systems that use (some of) the aforementioned data to decide on jobs, employment conditions, housing and health, may thus not only result in violating data protection rights and thus privacy rights, but it may equally result in discriminating certain (groups of) individuals and putting them in a disadvantaged position. Generally, automated decision-making systems can only be as good as the data it works with, meaning that where the data model that is being analysed is biased and/or the data that is used for the decision-making process is biased,⁴⁰ the systems may ‘result in disproportionately adverse outcomes concentrated within historically disadvantaged groups in ways that look a lot like discrimination’.⁴¹ Underlining the fact that ‘[a]lgorithms are not immune from the fundamental problem of discrimination’, Frank Pasquale stresses that ‘negative and baseless assumptions congeal into prejudice’.⁴²

Against this background, the central question that will be addressed in this contribution is how the use of automated decision-making systems do impact social and economic rights as laid down in the European Social Charter.⁴³ In doing so, this contribution will start by providing examples showing where automated decision-making has shown to have severe implications on social and economic rights of individuals (section II.). This is followed by an overview of the implications of the use automated decision-making systems for social and economic rights under the European Social Charter (section III.).

II. ACTUAL RISKS OF USING AUTOMATED DECISION-MAKING SYSTEMS IN AN ECONOMIC AND SOCIAL CONTEXT

In the following, a few examples will be provided to exemplify the actual impact of automated decision-making systems, highlighting how such systems automate decisions either on their own

³⁸ Felix Stalder, ‘Algorithmen, die wir brauchen’ (2017) NetzpPolitik.org, <<https://netzpolitik.org/2017/algorithmen-die-wir-brauchen/>>.

³⁹ Miriam A. Cherry, ‘People Analytics and Invisible Labour’ (2016) 61 *Saint Louis University Law Journal* 1, 1.

⁴⁰ Barocas and Selbst, ‘Big Data’s Disparate Impact’ 680ff. See also Kroll *et al.*, ‘Accountable Algorithms’ 680ff.

⁴¹ Barocas and Selbst, ‘Big Data’s Disparate Impact’ 673.

⁴² Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Cambridge, Massachusetts, Harvard University Press, 2015) 38.

⁴³ European Committee for Social Cohesion (CCS), Report on the Impact of Digitalisation and IT Developments on Social Rights and Social Cohesion (Study completed on 8 February 2022), CCS(2022)4, Strasbourg, 5 August 2022, <<https://rm.coe.int/ccs-2022-4-draft-report-digitalisation-en/1680a91c5d>>.

See also, more generally, Declaration by the Committee of Ministers on the risks of computer-assisted or artificial-intelligence-enabled decision making in the field of the social safety net, Decl(17/03/2021)2 (adopted 17 March 2021), <www.coe.int/en/web/portal/-/committee-of-ministers-computer-assisted-or-ai-enabled-social-services-decision-making-by-public-authorities-must-respect-human-rights> and Follow-up Recommendation to “Unboxing AI” (2019) by the Council of Europe Commissioner for Human Rights, Human Rights By Design: Future-Proofing Human Rights Protection in the Era of AI, May 2023, <<https://rm.coe.int/follow-up-recommendation-on-the-2019-report-human-rights-by-design-fut/1680ab2279>>.

or by making recommendations for humans to consider when responsible for taking the decision. It should be noted, however, that these examples do not offer an exhaustive picture of the risks involved when using AI.

Access to Employment and Promotion

The increasing digitalisation of the world of work is evident in the area of (access to) employment. This begins with the search for potentially suitable workers. Traditional print adverts with job vacancies in daily newspapers or specialist magazines have been largely replaced by online job portals, whether it concerns jobs in the public or private sector. Almost all (kind of) employing entities utilise the possibilities of digital job advertisements. Such online job advertisements not only allow a more detailed job description at lower costs, based on previous user behaviour, the job advertisements displayed on career websites, or their order and positioning can be specifically addressed to suitable applicants and adapted as specifically as possible to potential applicants.⁴⁴ A practice by Facebook resulted in several court cases in the USA due to discrimination. This is equally the case with the algorithm that had been used by Amazon, which discriminated female job applicants.⁴⁵

One step further would be using automated decision-making systems to match the ideal candidates with a particular job. Some systems are designed to quickly and efficiently find the most suitable employees for the respective positions. Automated decision-making systems are used to sort written applications, filter out information from applicants' CVs and prepare them for further processing to make statements about the accuracy of applicants and job offers as well as the comparability of applicants. Furthermore, so-called applicant management systems (AMS) offer the possibility of electronic applications via a digital user interface and allow the organisation and processing of applications received. And CV parsing systems allows CVs to be read automatically and transferred to pre-structured data fields with the help of semantic analyses. Simple parsing technologies only record data such as address and contact details, while more complex applications can also recognise and classify skills and competencies, ordering the information in a way that the hiring organisation deems relevant.

Moreover, there is software that tracks and measures the satisfaction and productivity of employees as well as through which the networking between individual employees and the team composition within companies can be improved. Information collected through such software may be used as input to make decisions on promotions, pay raises or, in the worst case, even dismissals.⁴⁶

Access to social protection

Also access to social protection is increasingly decided upon by automated decision-making systems. A widely known example from the Netherlands (the so-called *kinderopvangtoeslagaffaire*, childcare benefits scandal) has been the case where the Dutch tax

⁴⁴ This section has been largely based on Elisabeth Greif and Miriam Kullmann, 'Algorithmenbasiertes Personalrecruiting: Antidiskriminierungs- und datenschutzrechtliche Aspekte' (2021) ZAS 61, 62.

⁴⁵ See, e.g., Jeffrey Dastin, 'Insight - Amazon scraps secret AI recruiting tool that showed bias against women' *Reuters* (11 October 2018) <www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG>.

⁴⁶ Where employment is deemed fundamental to be able to make a living, supporting oneself and/or (close) family members, the potential of automated decision-making systems to, based on a certain combination of factors, exclude (groups of) individuals from having access to a job waries attention. The risk of discrimination in online job adverts lies primarily in the linguistic and/or graphic design of the adverts or in the exclusion of persons (groups) from access to individual adverts. In contrast, the risk of discrimination when using automated decision-making systems is closely linked to the functionalities on which these systems are based.

authority used an algorithm to spot suspected benefits fraud. The tax authority used a self-learning algorithm which created so-called risk profiles with the aim of identifying childcare benefits fraud. Yet, that algorithm used a mere suspicion of fraud based on the system's risk indicators, e.g. having a dual nationality or a low income, resulting in penalising families, pushing some of them into poverty due to the high level of debts these families had to the tax agency. While some had to sell their homes, a few others committed suicide, and more than thousand children were taken into foster care.⁴⁷ This fraud scandal resulted in the Dutch government resigning.⁴⁸ Also in Denmark, the perception of widespread welfare fraud has led to increasing the number of databases the Danish Public Benefits Administration can access, containing information on the people's homes, cars, relationships, employers, travel, and citizenship.⁴⁹ In both countries the accuracy of the automated decision-making systems have been criticised; accordingly, only people benefiting from social protection apparently do – deliberately or not – commit fraud. In an example from outside Europe, the US, Google's AI is being used for deciding, based on its own analysis, whether a person's unemployment claim should be approved, denied, or modified. While a human referee will review each recommendation made, signing and issuing a declaration may have huge implications for persons.⁵⁰

Access to housing and health care

Decisions on access to housing and health are also increasingly automatised. For instance, a bank may use algorithms to decide whether or not a person will get a mortgage while also social services may use automated decision-making systems to decide on who will be offered social housing.⁵¹ Similarly, decisions on who receives health care and what type of health care are also increasingly automatised. Besides access, but linked to that, health care itself is influenced by the use of AI. Where such AI is trained with biased data, there is a danger that the studies conducted may generalise individuals and their health conditions, neglecting to take into account that humans and their physical and mental characteristics may differ, resulting in potentially harming individuals.

III. THE IMPLICATIONS OF THE USE OF AUTOMATED DECISION-MAKING SYSTEMS FOR EUROPEAN SOCIAL CHARTER RIGHTS

So far, the (bodies within the) Council of Europe so far most elaborately has addressed civil and political rights as laid down in the European Convention on Human Rights (ECHR), paying only minor attention to the social and economic rights of the European Social Charter.⁵² Following the CDDH Discussion Paper, relevant European Social Charter provisions are Art 2, 3, 5, 11, 12, 13

⁴⁷ Melissa Heikkilä, 'Dutch scandal serves as a warning for Europe over risks of using algorithms', *Politico* 29 March 2020, <www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms>.

⁴⁸ BBC, 'Dutch Rutte government resigns over child welfare fraud scandal', 15 January 2021, <www.bbc.com/news/world-europe-55674146>.

⁴⁹ Gabriel Geiger, 'How Denmark's Welfare State Became a Surveillance Nightmare', *Wired* 7 March 2023, <www.wired.com/story/algorithms-welfare-state-politics>.

⁵⁰ Gizmodo, 'Google's AI Will Help Decide Whether Unemployed Workers Get Benefits', 10 September 2024, available at: <<https://gizmodo.com/googles-ai-will-help-decide-whether-unemployed-workers-get-benefits-2000496215>>.

⁵¹ Zoe Kleinman, Philippa Wain en Ashleigh Swan, 'Using AI for loans and mortgages is big risk, warns EU boss', BBC 14 June 2023, <www.bbc.com/news/technology-65881389>.

⁵² See CAHAI, *Feasibility Study*, Strasbourg, 17 December 2020, CAHAI(2020)23, 10, referring to Art 2, 3, 5, 11, 12, 13 and 20 European Social Charter.

and 20 European Social Charter.⁵³ Given the wide-ranging implications (the use of) AI may have on the life of individuals and their social and economic rights, there is reason to look at the European Social Charter rights more in-depth, thereby fundamentally supplementing the crucial task of the Steering Committee's Drafting Group for a Handbook on human rights and artificial intelligence.

Apart from the importance of social and economic rights in general, there are three additional reasons that underline the relevance of a focus that goes beyond civil and political rights. First, as research has shown, the 'theory' of the three generations of human rights has been misinterpreted. UNECSO's legal advisor Karel Vasák actually was not hierarchically classifying different human rights, rather what he meant was to distinguish the different 'categories' of rights that developed in three different stages within the UN. The first stage came with the adoption of the UN's Universal Declaration of Human Rights (UDHR) 1948, followed by the second stage which came with the adoption of the International Covenant on Civil and Political Rights and the International Covenant on Economic, Social and Cultural Rights in 1966. And, finally, the third stage which involved the emergence of rights on economic development, peace and a healthy environment.⁵⁴ Second, some rights that have been born as civil rights, such as the right to privacy (Art 8 ECHR), have invaded the field of social rights, mostly through employment relations in which these civil rights have been applied and interpreted.⁵⁵ And, third, some social rights, like the right to education, which may be found in Art 17 European Social Charter, has been seen as 'a necessary prerequisite of civil freedom'.⁵⁶ The same can be argued for the rights to access to gainful employment, social benefits, and housing.

The abovementioned examples explicate not only in what socio-economic areas of society automated decision-making systems can be and are being used. All examples mentioned do involve the risk of decision-making resulting in a discriminatory outcome through the use of certain (personal) data being used. While technically possible to use (personal) data to provide individuals with access to rights or deny them their rights, excluding people on the basis of one or more protected characteristics, such as age, disability, or gender is likely to violate Art E European Social Charter in conjunction with any other provision of the European Social Charter. The examples therefore show the dangers that their use may and do entail, often with serious and far-reaching consequences for those whose situation was decided by an algorithm. In many cases, automated decisions will provide a recommendation to be assessed by a human. Yet the question is whether this human has enough knowledge and skills to detect any potential flaws in the decision-making or whether this human is free to make an assessment on their own. Laws and other regulations may determine the scope within which decision-making has to take place, leaving almost no room for an individual human to make their own assessment.

In what follows, a brief overview will be provided on how automated decision-making can impact rights laid down in the European Social Charter. The list is not exhaustive and exemplary only for the purposes of this intervention.

⁵³ Steering Committee for Human Rights, Drafting Group on Human Rights and Artificial Intelligence (CDDH-AI), Discussion Paper, 6 September 2024, CDDH-IA(2024)03, 10. Art 2 is linked to Art 3, Art 5 is linked to Art 19(4b), Art 13 (also for States Parties that have not accepted Art 23) and 20.

⁵⁴ Steven L.B. Jensen and Charles Walton, 'Not 'Second-Generation Rights': Rethinking the History of Social Rights' in Steven L.B. Jensen and Charles Walton (eds), *Social Rights and the Politics of Obligation in History* (Cambridge, Cambridge University Press, 2022) 2.

⁵⁵ Maria Giovannone, 'The Right to Work' in Carole Nivard (ed), *The European Social Charter: A Commentary (Volume 2)* (Leiden, Boston, Brill Nijhoff, 2023) 85.

⁵⁶ Jensen and Walton, 'Not 'Second-Generation Rights': Rethinking the History of Social Rights'.

Employment

A central right of the European Social Charter can be found in Art 1 European Social Charter on the right to work. Art 1§2 European Social Charter refers to the right of the worker to earn his living in an occupation freely entered upon, which, in itself, also includes the right not to be discriminated. That European Social Charter provision is inherently linked to Art 20 European Social Charter (the right to equal opportunities and equal treatment in matters of employment and occupation without discrimination on grounds of sex), Art 8§2 European Social Charter (the right of employed women to the protection of maternity), Art 15§2 European Social Charter (the right of persons with disabilities to employment), Art 23 European Social Charter (the right of elderly persons to social protection), and Art 27 European Social Charter (the right of workers with family responsibilities to equal opportunities and equal treatment).⁵⁷

Prohibited grounds of direct and indirect discrimination are, inter alia, sex, race, ethnic origin, disability, religion, age, sexual orientation and political opinion. Discrimination is prohibited in connection with recruitment or with employment conditions in general (in particular, remuneration, training, promotion, transfer and dismissal or other detrimental action).⁵⁸ Discrimination based on the grounds of sex is laid down in Art 20 European Social Charter, covering: access to employment, protection against dismissal and occupational reintegration; vocational guidance, training, retraining and rehabilitation; terms of employment and working conditions, including remuneration; and career development, including promotion. Art 20 European Social Charter is a *lex specialis* to Art 1§2 European Social Charter.

As to foreign nationals' access to employment, a work permit may be required, but it is, in general, prohibited to ban nationals of States Parties from occupying jobs for reasons other than those set out in Art G European Social Charter.⁵⁹ This does most likely not apply where it concerns the jobs inherently connected with the protection of public interest or national security and involve the exercise of public authority. Also, when nationals of any States Party wish to engage in any gainful employment as worker or self-employed on the territory of another State Party, they should have access to that on an equal footing with nationals of the latter (Art 18 European Social Charter).

Where working time is considered a criterion that an automated decision-making system will take into account for deciding on, e.g., a promotion, that system would need to respect the limits set by the ECSR on extremely long working hours (16 hours within a period of 24 hours or where work for more than 60 hours per week is being performed⁶⁰). Besides these limitations that are set to protect the worker's health and safety, rules on overtime work and annual leave must respect the right of the worker not to be discriminated (Art E European Social Charter⁶¹). This equally applies to the right to fair remuneration in Art 4 European Social Charter where, e.g., the employer uses AI to monitor the productivity of the workers' job performance or where the employing entity uses a mobile application, think of platform businesses like Uber or Bolt, that manages the level of remuneration. More specifically, Art 4§3 European Social Charter ensures

⁵⁷ Digest of Case Law of the European Committee of Social Rights, June 2022, 47.

⁵⁸ Digest of Case Law of the European Committee of Social Rights, June 2022, 49.

⁵⁹ Art G European Social Charter provides that any restrictions or limitations to the European Social Charter rights are only allowed when prescribed by law and necessary in a democratic society for the protection of the rights and freedoms of others or for the protection of public interest, national security, public health, or morals.

⁶⁰ Digest of Case Law of the European Committee of Social Rights, June 2022, 57-58.

⁶¹ Art E European Social Charter provides that the enjoyment of the rights set forth in this Charter shall be secured without discrimination on any ground such as race, colour, sex, language, religion, political or other opinion, national extraction or social origin, health, association with a national minority, birth or other status.

that men and women are paid equally where their work is considered to be of equal value.⁶² Where employing entities using automated decision-making systems to terminate employment relationships, States Parties must ensure that a reasonable notice period is taken into consideration (Art 4§4 European Social Charter).⁶³ In case of termination of employment, Art 24 European Social Charter guarantees that the termination is subject to a valid reason on the sanction of being entitled to adequate compensation or other appropriate relief. No valid reason for termination of employment is the worker's temporary absence from work due to illness or injury, meaning that an automated decision-making system cannot make use of that factor when effectuating a dismissal.⁶⁴ Automated deductions from remuneration are, under the European Social Charter, only allowed where national laws, regulations, collective agreements, or arbitration awards do explicitly determine so (Art 4§5 European Social Charter).

To make sure that workers can perform their jobs in a safe and healthy environment, Art 3 European Social Charter determines that working conditions for all workers – employed and self-employed, all workplaces and all sectors must be safe and healthy. This provision will be relevant in situations where workers' work is supplemented or supported by technology that use, e.g., in manufacturing or transport. In addition to the States Parties, employers' and workers' organisations do have a role to play in defining and developing preventive policies.⁶⁵ This also includes the right to digitally disconnect, meaning that workers may not be penalised, if their working hours are being monitored and tracked, by an automated decision because they refuse to work outside their normal working hours or while on holiday or on other forms of leave.⁶⁶ Also here, the principle of non-discrimination applies (Art E European Social Charter).

While Art 5 European Social Charter guarantees the right for workers (in both public and private sector⁶⁷) and employers to freedom of association for the protection of their economic and social interests, being a member of a trade union may be information an employer considers to be relevant when defining the factors an automated decision-making system should be considering. In case the system's decision puts the worker in a disadvantaged position, it may have a deterrent effect on the worker on whose situation the decision has been made, but it may also deter other workers to become or stay member of a trade union, endangering the right enshrined in Art 5 European Social Charter. This, similarly, may be the case where trade unions and their member workers go on strike, in which case Art 6§4 European Social Charter might be violated. Concerning the right to bargain collectively under Art 6 European Social Charter, there is a role for workers' and employers' representatives to also collectively bargain on the use of automated decision-making systems in sectors and/or undertakings. The workers' and employers' representatives' right is closely linked to earlier provisions that were mentioned (Art 2, 3 and 4 European Social Charter) and relates also to Art 21 European Social Charter (information and consultation), Art 22 European Social Charter (participation in the determination and improvement of the working conditions and working environment), Art 24 European Social Charter (protection in cases of termination of employment), and Art 26 European Social Charter (dignity at work).

⁶² Digest of Case Law of the European Committee of Social Rights, June 2022, 75.

⁶³ Digest of Case Law of the European Committee of Social Rights, June 2022, 78-79.

⁶⁴ Appendix Art 24, §3 i) to European Social Charter; also Digest of Case Law of the European Committee of Social Rights, June 2022, 182.

⁶⁵ Digest of Case Law of the European Committee of Social Rights, June 2022, 71.

⁶⁶ ECSR, Statement on COVID-19 and social rights adopted on 24 March 2021, <<https://rm.coe.int/statement-of-the-ecsr-on-covid-19-and-social-rights/1680a230ca>>.

⁶⁷ Digest of Case Law of the European Committee of Social Rights, June 2022, 82.

Health and social security

Apart from health and safety at work, the European Social Charter contains more general provisions guaranteeing protection relating to an individual's health. For those engaged in gainful employment – including self-employment – and their dependants, Art 12§1 European Social Charter ensures the right to social security.⁶⁸ A number of risks are covered: medical care, sickness benefit, unemployment benefit, old age benefit, employment injury benefit, family benefit, and maternity benefit.⁶⁹ The European Social Charter requires that a social security system guarantees an effective right to social security with respect to the benefits provided under each branch.⁷⁰ Access to social security is subject to the right not to be discriminated based on Art 20 European Social Charter as well as Art E European Social Charter. Equal treatment with regard to social security under Art 20 European Social Charter implies the absence of any direct and indirect discrimination on grounds of sex, particularly as far as the scope of schemes, conditions of access to schemes, the calculation of benefits and the length of entitlement to benefits are concerned.⁷¹ It follows that public authorities making use of automated decision-making systems must make sure that these systems are non-discriminatory.⁷² The latter is equally relevant when granting social and medical assistance using automated decision-making systems (Art 13§1 European Social Charter) to persons without adequate resources.⁷³ The only valid criterion for being eligible to assistance is need. States Parties may define conditions for granting specific types of assistance, making eligibility dependent on a minimum age limit.⁷⁴ Benefiting from social welfare services also requires having (non-discriminatory) access to it (Art 14 and E European Social Charter).⁷⁵

Housing

Everyone has a right to adequate housing (Art 31 European Social Charter). Access to adequate housing should be promoted in particular to the different groups of vulnerable persons, such as low-income persons, unemployed persons, single parent households, young persons, persons with disabilities including those with mental health problems.⁷⁶ Yet actual access means that, where public authorities or private parties use automated decision-making systems to offer housing and/or provide mortgages, access should be granted on the principle of non-discrimination (Art E European Social Charter). States Parties must also take measures to prevent and reduce homelessness. Here, they are to ensure that automated decision-making systems do not discriminate individuals, meaning that they have to make sure that all those individuals in danger of homeless will have access to measures adopted and where access is granted that the principle of non-discrimination is being guaranteed (Art E European Social Charter).

⁶⁸ Digest of Case Law of the European Committee of Social Rights, June 2022, 119 ff.

⁶⁹ Digest of Case Law of the European Committee of Social Rights, June 2022, 120.

⁷⁰ Digest of Case Law of the European Committee of Social Rights, June 2022, 120.

⁷¹ Digest of Case Law of the European Committee of Social Rights, June 2022, 175.

⁷² Usually social security systems/benefits are established by law, but the Charter applies equally if the benefits are established by way of collective agreement.

⁷³ Art 13 European Social Charter applies only to persons without resources, unlike Art 12 European Social Charter which applies to persons engaged in gainful employment.

⁷⁴ Digest of Case Law of the European Committee of Social Rights, June 2022, 125.

⁷⁵ When individuals may need to go to the hospital, it is very likely that the information on the reasons of their stay as well as the duration thereof are not only registered, but also transferred to insurers and other authorities, such as those responsible for migration. The (automated) transfer of data may withhold migrants or unlawfully present persons going to hospitals for fear that their personal data/status are shared by the hospitals with the migration authorities, creating a real risks for their access to health care.

⁷⁶ Digest of Case Law of the European Committee of Social Rights, June 2022, 201.

Employment of children

States Parties must ensure that, for instance, employers, like Uber or Bolt, using mobile applications respect the minimum age of admission to employment through proper identify verifications, which is 15 years, subject to exceptions for children employed in prescribed light work without harm to their health, morals or education (Art 7§1 European Social Charter). This is closely linked to the right to education (Art 17§2 European Social Charter), following which Art 7§3 European Social Charter determines that persons subject to compulsory education may not be employed if work would deprive them of the full benefit of their education. The latter may also be inhibited if the working hours of persons under 18 years of age are not limited with a view to their needs of their development (Art 7§4 European Social Charter). When working, young persons should be paid a fair wage or other appropriate allowances (Art 7§5 European Social Charter).

AI, human rights, democracy and the rule of law: A primer prepared for the Council of Europe – Connecting Principles, Rights and Obligations | David LESLIE

A Tale of Two Axial Ages

Amidst the current industrial revolution of frontier AI technologies and the swell of generative applications across veritably every domain of lived experience, some have begun to speak of Axial age of artificial intelligence. The story runs deep.

In many ways, the swift metamorphosis of the computer age into our present era of omnipresent internet communications, networked devices, and mobile connectivity has ushered in a societal sea change. We increasingly live in a dynamic and integrated digital world where connected machines containing countless sensors and sites of behavioural measurement intermingle with ever-present algorithmic systems and cloud computing platforms. Amidst such an incipient computational ubiquity, artificial intelligence (AI) technologies are progressively more being directed at an unbounded problem space—an increasingly limitless field of potential applications. Indeed, as AI systems are ever more coming to play a constituent role in the critical digital infrastructure of connected society writ large, the breadth of their impacts is expanding in kind. In this cyber-physical condition, algorithmically personalised services reach into the private lives of targeted decision subjects and have an active curatorial hand in real life processes of identity formation concurrently as scaled computational methods of attention-mining, relevance-ranking, popularity-sorting, and trend-predicting steer the digital public square, influencing and organising the way citizens can interact and engage in the political, cultural, and moral life of the community. Meanwhile, as AI models continue to swiftly grow in speed, scale, and complexity with the explosion of ‘big data’ and the exponential enlargement of computing power, scientists are

gaining new capacities to alter the very constitution of nature. Such unprecedented technological affordances are enabling them, at the microscopic level, to more effectively engineer the molecular building-blocks of life and to re-program the genetic code of all earthbound organisms. Likewise, they are facilitating scientists’ abilities, at the macroscopic level, to manipulate the climate at a planetary scale through the re-engineering of the composition the earth system itself.

While boosters and pundits alike have embroidered this expansionist trend—calling AI “the new electricity”, “the single most influential human innovation in history”, or simply humanity’s “final invention”—one thing remains clear: The pervasiveness and innovation- spawning character of AI makes it a general-purpose technology which is transforming who we are, how we live, and who we aspire to be. AI is, however, not simply a general purpose technology par excellence. Taken together, the potential penetration of AI into all sectors of society and the protracted aggregate impacts of its knock-on effects in other innovation domains mean that it is, more essentially, an axial technology which is placing humanity at a kind of turning point. That is, as these technologies are becoming gradually more pervasive across all domains of human activity, as the scope of their impacts increasingly stretch from the depths and shallows of personal intimacy, individual wellbeing, and social connection to the fate of biospheric flourishing itself, humanity is, for better or worse, being drawn closer and closer to a kind of critical threshold beyond which its social reality—its world of shared experience—could fundamentally and perhaps irretrievably change. AI is axial in the sense that it is places the direction of this change in lap of

the moral agency of the present. It places us humans—and the way our values, motivations, and choices steer the gathering energies of technological advancement in AI—at the axis of choice which will determine the trajectory of this change, whether it veers toward the advancement of human wellbeing, the public good, and the protection of human rights, democracy and the rule of law or toward possibilities for the potentially irreparable harm to fundamental rights and freedoms and for the humanly prompted generation of global-catastrophic risks.

In the short time I have today, I want to take a step back from this somewhat frothy narrative about the current Axial Age of Artificial Intelligence and tell the original story of what has for almost a century been understood as the Axial Period of human history—the origin story, if you will, of human rights, democracy, and the rule of law. For, as I will show, by historically reconstructing the axial underpinnings of human rights and modern democratic ways of life, we can access essential critical and analytical leverage points that enable a better understanding of the full range of risks now posed by AI technologies to the enjoyment of human rights, the functioning of democracy, and the integrity of the rule of law.

So let me begin telling this second, earlier tale, without further ado. Thinkers like Karl Jaspers, Hans Joas, and Juergen Habermas have justifiably situated the true origins of the modern tradition of fundamental rights and freedoms in the global dawning of the “Axial Age” (which occurred roughly around the middle of the first millennium BCE, three to four thousand years ago). As Jaspers noted, such an era of world-historical pivot would remarkably crystallise in the convergence of the revolutionary thinking of Zoroaster in Iran, of Buddha in India, of the Hebrew prophets in Israel, of Xenophanes, Plato, and the Stoics in Greece, and of Confucius in China. Across the teachings of these thinkers, there arose a novel and unprecedented appreciation of the need for societies to confront the limitations of their own finite and historically-bounded perspectives. These communities were now seen to be subject to universal ethical standards that transcended the ephemeral quality of local traditions, superstitions, and myths.

In the face of a wholly transcendent God or a higher moral order or an encompassing cosmological purpose, mortal humans were suddenly—and each equally—jolted from their ritualistically-instantiated collective self-centres. They were increasingly compelled to critically interrogate the ways they had previously conceived of their shared histories, their societal roles, their received doctrines, and their reciprocally self-attuning mores and social norms. This radical disruption of localised religious and cultural reference points spelled the beginning of the end for the predominance of mythically-based archaic and ancient civilisations.

While initially, this new Axial sense of critical self-awareness and this new orientation to context-transcending moral standards would be held in check by the institutional assimilation and retrenchment of the very revolutionary social movements that had instigated the Axial moment (think here, for instance, of the early institutionalisation of Christianity), the journey toward modern structures of selfhood and social order had irrevocably begun. Once the genie of interminable doubt, critical self-awareness, and continual self-questioning had been let out of the bottle of Axial civilizations, the global historical processes that we have long since come to understand as modernisation would be set in motion.

Gradually, over the next millennium-and-a-half, as technologies of transportation, navigation, conquest, and commerce advanced, the exposure of the Continental communities to alien cultures, histories and traditions intensified. By the time the late Middle Ages transitioned into the Renaissance, the theocratic and feudal ways of life that prevailed across Europe were subjected

to another wave of Axial self-questioning and critique—a resurgence of what Jaspers called the “struggle of rationally-clarified experience against myth.”

As the story first told by classical social theorists from Alexis de Tocqueville and Emil Durkheim to Max Weber goes, the commencement of Western modernity was occasioned by the advent of a set of transformative processes in the 16th and 17th centuries that functioned to overhaul the hierarchically organised societal forms that characterised pre-modern ways of life. Growing scepticism toward the pervasive religious and teleological order of things gave rise to novel pressures to cope with the hardships of contingent reality without recourse to the sway of theocratic law, divine commandment or cosmic purpose (Taylor, 1989).

Over time, these strengthening pressures to manage the difficulties and uncertainties of harsh existence by human lights alone yielded corresponding demands that individuals depend solely upon themselves and each other to collaboratively shape the meanings, values, and purposes by which they could come to both share experience and direct their common affairs. Whereas the attitudes, commitments and interaction-coordinating norms that typified pre-modern self-understandings were able to draw upon the presumptive authority of a holistically interconnected and religiously anchored framework of meanings to apprehend reality, to delimit social relations and to steer practical possibilities, modernising processes shattered such a schema by resituating the legitimate sources of objective knowledge and social order ever more exclusively in humanity's collective capacity to rationally redress the manifold challenges of existence through the exercise of open-ended and consensus-based problem solving (Habermas, 1992).

Such an open-ended challenge to collaboratively shape a world of shared meanings, goals, and values through unrestrained communication would hence give rise to the burdens of a new way of interacting whereby the reciprocal exchange of rationally-accessible beliefs became a condition of possibility for the intersubjectivity of understanding itself. It gave rise, that is, to the unprecedented modes of rational justification, interpretation, and interpersonal accountability that would eventually manifest in the universal and inclusive character of egalitarian reciprocity which structures the post-conventional moral perspective. Such a perspective lies behind the basic rights and freedoms instantiated in modern democratic legal orders and human rights regimes.

What is more, this challenge to cooperative world-making would manifest normatively as a kind of ethical pressure to individuate—to become fully autonomous and to assume the full station of one's individual authority—through social processes of everyday discursive engagement. That is, as participants in the budding age of experience, the modern age, were increasingly stripped of communicative recourse to theologically fixed systems of beliefs and social statuses, they were progressively more obliged to bear responsibility for the symbolic reproduction of their shared human world. They were compelled to exercise their individual authority in forming and sharing the beliefs that fostered these social reproductive processes simultaneously as they were obliged to take responsibility in justifying and defending those beliefs to others. The emergence of the equal and inviolable moral status of individuals (on which so much of the normative self-understanding of contemporary social, political, and legal institutions is predicated) first arose, in fact, out of such a formative pressure placed on dialogically implicated interactors, who were individually called upon in their everyday communications to count as offering and demanding reasons and were held to be accountable for those reasons in turn.

It is in this interdependent relationship of individuation, intersubjective communication and collaborative social reproduction that we find the origins of the interconnected character of human rights, democracy, and the rule of law. The legitimacy of modern democratic institutions is rooted

in the notion that each and every citizen is equally entitled to participate in the shared life of the community and in the steering of the collective decisions that impact them. Yet, for citizens to exercise this right to participate in the conduct of public affairs, they must first possess many other interrelated civil, political, social, cultural, and economic rights. They must enjoy freedom of thought, association, assembly, and expression. They must be granted equal respect before the law and protection from any forms of discrimination that would encumber their full and equitable participation in community life. They must have access to the material means of participation through the provision of proper education, adequate living and working standards, health, safety, and social security. They must be able to access effective judicial remedies if any of their basic rights are harmed. It is in this latter respect that the rule of law provides the institutional basis for safeguarding both democratic participation and the protection of fundamental rights and freedoms. An independent and impartial judiciary, which ensures citizens due judicial processes and fair and equal treatment under the law, acts as a guarantor of recourse whenever fundamental rights or freedoms could be breached.

With this historical reconstruction of the axial age underpinnings of human rights and modern democratic ways of life in mind, I want to move now to providing a more comprehensive picture of the full scope of hazards now posed by AI and data-intensive technologies to the enjoyment of human rights, the functioning of democracy, and the integrity of the rule of law. We are now seeing multiple warning signs that the societal supports and institutional anchors of human rights, democracy and the rule of law will wane rather than reign.

At the level of individual rights, human dignity, and identity formation, we see these warning signs in the way that the broad-scale proliferation of individual-targeting algorithmic curation is already impoverishing autonomy and reflective agency across a range of human activities. For instance, in the domain of e-commerce, strengthening regimes of consumer surveillance have fuelled the use of “large-scale behavioural technologies” that have enabled incessant practices of hyper-personalised psychographic profiling, consumer curation, and behavioural nudging. Such technologies have tended to exploit the emotional vulnerabilities and psychological weaknesses of targeted people instrumentalising them as monetisable sites of behavioural surplus and treating them as manipulable objects of prediction rather than as reflective subjects worthy of decision-making autonomy and moral regard.

Analogous postures have spurred state actors and other public bodies to subject their increasingly datafied citizenries to algorithmic nudging techniques that aim to obtain aggregated patterns of desired behaviour which accord with government generated models and predictions. Some scholars have characterised such an administrative ambit as promoting the paternalistic displacement of individual agency and the degradation of the conditions needed for the successful exercise of human judgment, moral reasoning, and practical rationality. In like manner, the nearly ubiquitous scramble to capture behavioural shares of user engagement across online search, entertainment, and social media platforms has led to parallel feedback loops of digital surveillance, algorithmic manipulation, and behavioural. The proliferation of the so-called “attention market” business model has prompted digital platforms to measure commercial success in terms of the non-consensual seizure and monopolisation of focused mental activity. This has fostered the noxious attachment of targeted consumer populations to a growing ecosystem of “distraction technologies” and compulsion-forming social networking sites and reputational platforms, consequently engendering, on some accounts, widespread forms of surveillant anxiety (Crawford, 2014), cognitive impairment (Wu 2019), mental health issues, and diminished adolescent self-esteem and quality of life.

The ubiquitous expansion of generative AI technologies now threatens to intensify these effects. Conversation agents like ChatGPT, Pi, and the growing number of so-called AI companions that are trained to carry out human-like dialogue can deceptively lead people to believe they are interacting with “sentient” or “intelligent” agents, exposing them to heightened risks of manipulation. The combination of algorithmic targeting, curation, and intrusive hyper-personalisation with this increased capacity to manipulate promises more effective behaviour steering at the cost of agency loss, impingements on psychological integrity, and dignity violations.

Over and above these threats to basic individual dignity, human autonomy, and identity-formation, the proliferation of data-driven behavioural steering at the collective level has also generated risks to democratic institutions and the communicative processes of socialisation and social integration on which they depend. In current digital information and communication environments, for example, social media and search engine platforms have mobilised opaque computational methods of relevance-ranking, popularity-sorting, and trend-predicting to produce calculated digital publics devoid of any sort of active participatory social or political choice. Rather than stewarding informational plurality and the deliberatively achieved political will of interacting citizens, this vast meshwork of connected digital services shapes these computationally fashioned publics in accordance with the drive to commodify monitored behaviour and to capture user attention. And, as this manufacture of digital publics is ever more pressed into the service of profit-seeking by downstream algorithmic mechanisms of hyper-personalised profiling, engagement-driven filtering, and covert behavioural manipulation, democratic agency and participation-centred social cohesion will be increasingly supplanted by insidious forms of social sorting and digital atomisation. Combined with complementary dynamics of wealth polarisation and rising inequality, such an attenuation of social capital, discursive interaction, and interpersonal solidarity is already underwriting the crisis of social and political polarisation, the widespread kindling of societal distrust, and the animus toward rational debate and consensus-based science that have come to typify contemporary post-truth contexts.

Indeed, as these and similar kinds of computation-based social sorting and management infrastructures continue to multiply, they promise to jeopardise more and more of the formative modes of open interpersonal communication that have enabled the development of crucial relations of mutual trust and responsibility among interacting individuals in modern democratic societies. This is beginning to manifest in the widespread deployment of algorithmic labour and productivity management technologies, where manager-worker and worker-worker relations of reciprocal accountability and interpersonal recognition are being displaced by depersonalising mechanisms of automated assessment, continuous digital surveillance and computation-based behavioural incentivisation, discipline, and control. The convergence of the unrelenting sensor-based tracking and monitoring of workers’ movements, affects, word choices, facial expressions, and other biometric cues, with algorithmic models that purport to detect and correct defective moods, emotions, and levels of psychological engagement and wellbeing may not simply violate a worker’s sense of bodily, emotional, and mental integrity by rendering their inner life legible and available for managerial intervention as well as productivity optimisation.

These forms of ubiquitous personnel tracking and labour management can also have so-called “panoptic effects”, causing people to alter their behaviour on suspicion it is being constantly observed or analysed and deterring the sorts of open worker-to-worker interactions that enable the development of reciprocal trust, social solidarity, and interpersonal connection. This labour management example merely signals a broader constellation of ethical hazards that are raised by the parallel use of sensor- and location-based surveillance, psychometric and physiognomic profiling and computation-driven technologies of behavioural governance in areas like education,

job recruitment, criminal justice and border control. The heedless deployment of these kinds of algorithmic systems could have transformative effects on democratic agency, social cohesion, and interpersonal intimacy, preventing people from exercising their freedoms of expression, assembly, and association and violating their right to participate fully and openly in the moral, cultural, and political life of the community.

Here again Generative AI promises to be a force multiplier of risks and potential harms. The irresponsible or malicious development and use of GenAI technologies could lead to the scaled production of disinformation, propaganda and false-but-true-sounding information, potentially flooding the digital public square with misleading and non-factual content. This could undermine social trust in the information ecosystem and tear at the fabric of reliable public communication on which modern democratic ways of life and social responsibility rest. Combined with the societally polarising dynamics algorithmically-steered social sorting, the integrity of a digital public sphere polluted with synthetic content, deep fakes, disinformation, and other AI-altered text, photos, and videos could be mortally imperilled.

To close here, let me offer a simple observation: What these instances of ethical peril at individual, social, and societal levels tell us is that the fragile but indispensable heritage of human rights, democracy and the rule of law is itself now under subtle but relentless assault—and we must center and draw upon its history to recognize the full picture of what, if we fail to intervene effectively, might be lost. In the end, our current Axial Age of AI has not condemned us to cast fundamental rights and democratic ways of life into the dustbin of civilisational history. It has, more importantly placed pivotal questions about what shape the technologically transfigured society of tomorrow will take at the fingertips our shared democratic agency, social creativity, and collective intelligence.

The Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law | Thomas SCHNEIDER



CDDH-IA

Strasbourg

25 September 2024

The Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law: how to protect democracy and human rights and foster innovation at the same time?

Ambassador Thomas SCHNEIDER

Chair of the Committee on Artificial Intelligence (April 2022 – September 2024)

1



COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

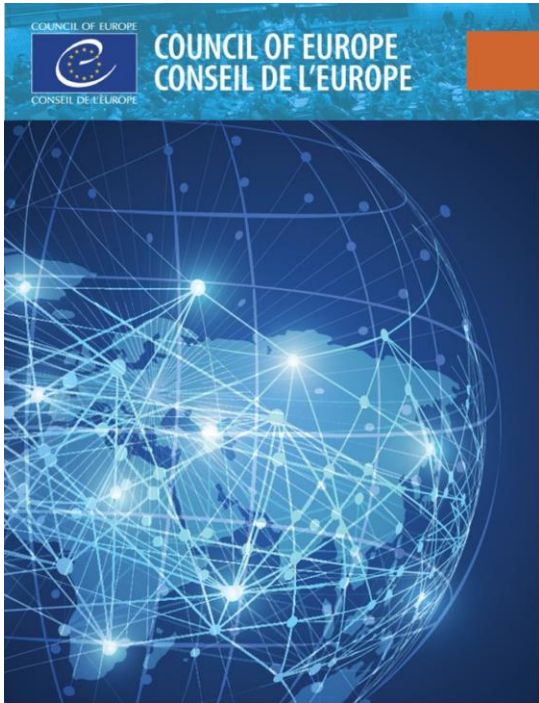
AI: Opportunities, risks & impact



AI is a(nother) disruptive technology

- Industrial revolution of the 19th century:
 - > engine-driven machines replace repetitive physical labour:
 - Machines to process raw material into goods or food
 - Machines that move beings or goods from one place to another
- Digital revolution of the 21st century:
 - > AI-driven machines replace repetitive cognitive labour:
 - Process or generate content
 - Analyze data and prepare or take decisions

2



**COUNCIL OF EUROPE
CONSEIL DE L'EUROPE**

Regulating AI: Yes, but how?

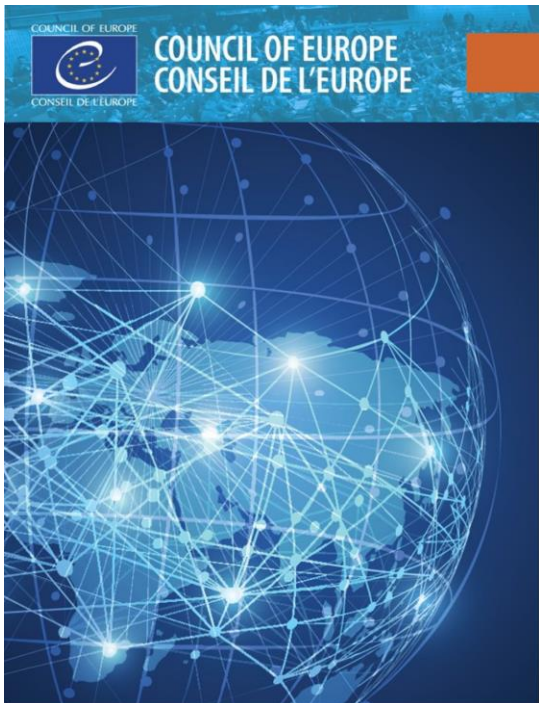
Regulating AI: Issues to consider

- What is really new with AI? Gaps?
What is the *delta* that is not covered in existing regulation?

Dealing with AI: What are precedents/analogies to learn from?

- How we did and do “regulate” engines:
 - There is no one single “engine law” !
 - 1000s of technical, legal and sociocultural norms & institutional settings for engines / machines / people / infrastructure, etc. (with different levels of international harmonization)
- Same for AI: Not one single law or standard, but mix of norms, (with different degrees of international harmonization)

3



**COUNCIL OF EUROPE
CONSEIL DE L'EUROPE**

Regulating AI: Yes, but how?

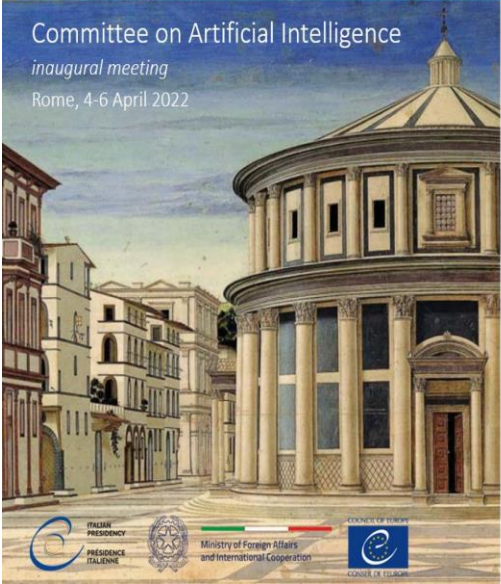
Regulating AI: Issues to consider II

- Despite many parallels:
 - AI-systems are not engines, data is not oil:
 - Data & AI are not physical, but dematerialized
 - > can be copied, moved and used anywhere
 - > evolve in nature and functioning
- We need new agile & differentiated approaches for governing & regulating new technologies!

4



Committee on Artificial Intelligence
inaugural meeting
Rome, 4-6 April 2022



ITALIAN PRESIDENCY
PRESIDENZA ITALIANE

Ministry of Foreign Affairs
and International Cooperation

COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

CoE: from CAHAI to CAI

Ad-Hoc Committee on Artificial Intelligence (2019-21)

- feasibility study** for and **potential elements** of a legal framework for the development, design and application of AI systems, based on CoE's standards on HR, Dem & RoL

Committee on Artificial Intelligence (2022-24)

- elaborate an appropriate legal framework (convention) on the development, design and application of AI
 - based on CoE standards on HR, Dem & RoL
 - conducive to innovation
 - global instrument, not just CoE member states
 - complementary to existing and future standards
 - involvement of all stakeholders
- Develop a methodology for a HUDERIA

5



Committee on Artificial Intelligence



CAI: Purpose and process

- Purpose:
 - nor create new substantive rights, nor weaken /undermine existing rights and protections in any country.
 - But: **make sure that the existing protection levels** of HR, Dem & RL would **also apply to** the development and use of **AI**
- importance of treaty becoming a **global instrument**: CoE plus ARG, AUS, CAN, CRC, ISR, JAP, MEX, PER, URU, USA & HS
- **Inclusive process**: other CoE bodies & IGOs (OECD, OSCE, UNESCO, etc.) & >70 reps from civil society, business, technical and academic community
 - > active participation, comments, text proposals to the draft until the very last day of negotiations

6



COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

Committee on Artificial Intelligence



Challenges:

- European vs. Global approach?
- Scope of application:
 - Application to Public vs. Private actors?
 - National security?
- Principle of non-discrimination: what grounds?
- Logic and nature of risk and impact assessment?
- Remedies: what requirements for AI actors?
- What regulation is conducive to innovation?
- Principle of protection of health and environment?

7



COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

CoE Framework Convention on AI



General points

- formulates **fundamental principles and rules**
 - safeguard human rights, democracy and the rule of law
 - conducive to progress and technological innovations.
- **complementary** to the already existing international human rights, democracy and rule of law standards and aims at filling-in any **legal gaps** that may have formed as a result of rapid technological advances
- remain **future-proof**: operating at high level, does not regulate technology and is essentially **technology neutral**
- **Graduated and differentiated approach** based on severity and probability of adverse impacts

8



COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

Outline of the CoE Framework Convention



Scope of the Convention

- The CoE Framework Convention applies to both **public and private sectors**: States commit to address risks by all actors – while having great **leeway** regarding by which means they intend to achieve the goals of the convention
- Limited exemptions from the scope are foreseen as regards **national security** and **research and development**.
- In line with the Statute of the CoE, matters relating to **national defence** do not fall within the scope of the Convention.

9



COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

Outline of the CoE Framework Convention

Definition of an AI system

- Article 2 of the CoE Framework Convention contains a **definition of an AI system**.
- In order to prevent the proliferation of many different definitions at global level, and with a view to **facilitating international co-operation** and **harmonising the relevant terminology**, the Drafters decided to use for the purpose of the Framework Convention the OECD's definition (as revised and adopted on 8 November 2023):

“artificial intelligence system” means 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. Different artificial intelligence systems vary in their levels of autonomy and adaptiveness after deployment

10



COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

CoE Framework Convention on AI



Fundamental Principles related to AI

- Human dignity and individual autonomy
- Transparency and oversight
- Accountability and responsibility
- Equality and non-discrimination
- Privacy and personal data protection
- Reliability
- Safe innovation

11



COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

Outline of the CoE Framework Convention

Remedies, procedural rights and safeguards

- 1) The obligation **to document** the relevant **information regarding AI system** and the relevant **information regarding its usage** and to make it **available to affected persons**,
- 2) Such information should be **sufficient to contest the decision(s)** made or substantially based on the use of the AI system and **the use of the system** itself,
- 3) **Effective possibility** to lodge a complaint to competent authorities.





Remedies, procedural rights and safeguards

- 4) The obligation to **provide effective procedural guarantees, safeguards and rights** to affected persons in connection with the application of an artificial intelligence system where an artificial intelligence system significantly impacts upon the enjoyment of human rights and fundamental freedoms,
- 5) The **provision of notice** that one is interacting with an artificial intelligence system.



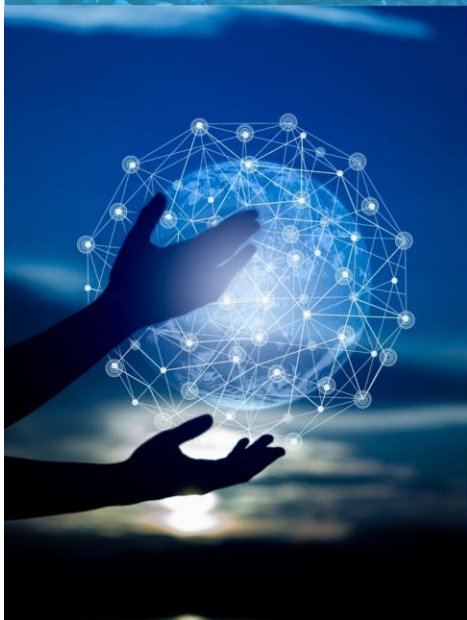
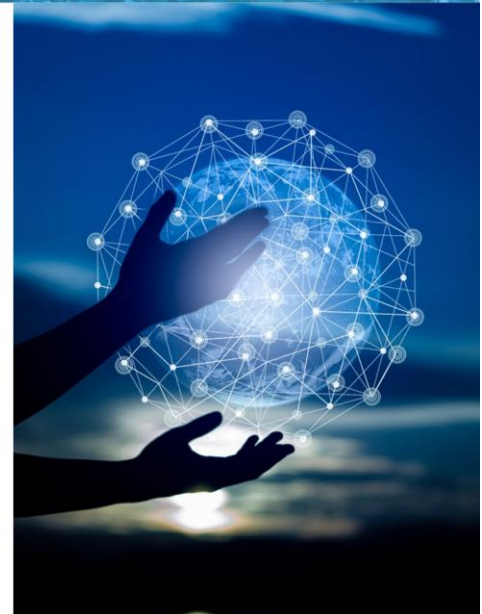
Risk and Impact Management Requirements

- Critical and crucial element of the governance of AI systems
- Obligation for relevant AI actors to carry out **Risk and Impact Assessments** in respect of actual and potential impacts on Human Rights, Democracy and the Rule of Law
- Relevant **assessments** should be carried **iteratively**



Risk and Impact Management Requirements

- Obligation to apply **sufficient preventive and mitigation measures** resulting from the application of the above **assessments**.
- Possibility for the authorities to **introduce bans or moratoria on certain application of AI systems** (so-called “red lines”).
- The CAI is working on a **Human Rights, Democracy and the Rule of Law Risk and Impact Assessment (HUDERIA)**, a stand-alone and non-binding instrument supporting the Framework Convention with **detailed guidance on risk and impact management issues**.

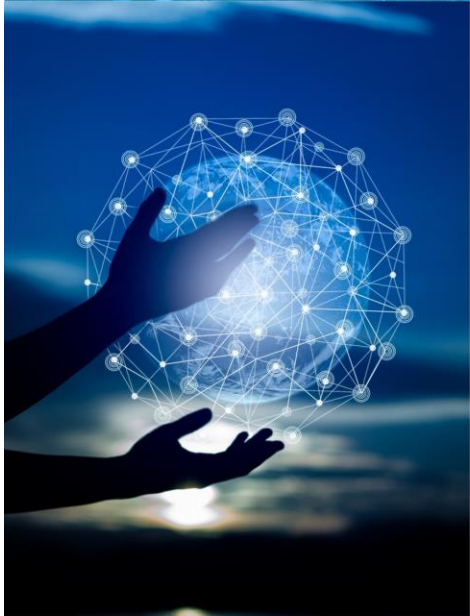


- 4 April 2022: 1st CAI meeting in Rome
- 30 June 2022: distribution of first «zero draft»
- 1 Jan 2023: beginning of negotiations in the CAI
- 14 March 2024: framework convention adopted by the CAI
- 17 May 2024: FC adopted by the CoE Committee of Ministers
- 5 September 2024: Signing Ceremony in Vilnius:
FC opened for signature. 36 countries signed on the first day
(AD, GE, IS, NO, MO, SM, UK, IL, US & 27 EU MS)
- Entry into force: 3 months after ratification of 5 countries
(of which at least 3 are CoE member states)



COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

Outlook and food for thought...



- CoE framework convention is one important step, but not the last...
- We will need to continue to develop technical, legal and socio - cultural norms – on local, regional and global levels.
- We need to develop international dialogue and cooperation, with a view to develop a globally shared vision about how to use AI for good (and not for bad)
- We need to work on interoperability of norms
- We need to develop new and more agile forms of governance and policy making fit for the technologies of the 21st century
(we are currently working with the governance framework built during the 1st industrial revolution in the 19th century)



COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

Information Society Department

Thank you for
your attention !



further resources:

www.coe.int/freedomofexpression

www.coe.int/internetgovernance

www.coe.int/dataprotection

www.coe.int/cybercrime

www.coe.int/AI

Facebook Page

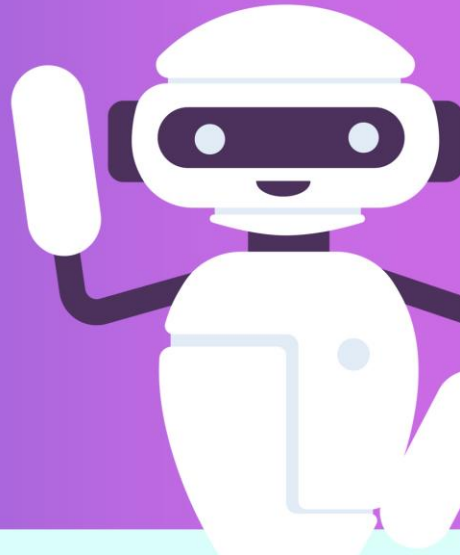
CoE Digital Governance

Implications of generative artificial intelligence for freedom of expression | Andrin EICHIN

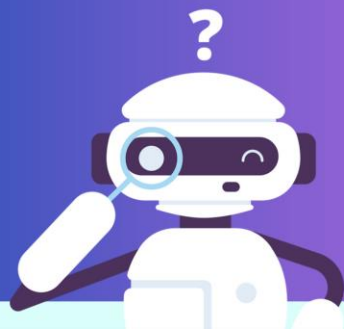
Committee of Experts

on the implications of generative AI on freedom of speech (MSI-AI)

Andrin Eichin
Chair of MSI-AI



Mandate



Deliverable requested by CDMSI

Draft Guidelines on the implications of generative artificial intelligence for freedom of expression

FoE in a new world?

To what extent is GenAI impacting FoE differently? Does this alter the necessary conditions to strengthen FoE in a society?

Narrow focus on freedom of expression

Primary focus on rights directly challenged by GenAI implications on FoE as per interpretation in line with Article 10 ECHR and case law

Timeline



Scope



Functional definition of Generative AI

Focusing of potential of a system to generate human -like expression or outputs following instructions (prompts) inserted by users, regardless of underlying technology.

Due regard to technological stack

Identify and consider existing technological choices with a clear FoE impact.

Example: Importance of statistical word and vector distribution prediction in today's models

Consider multimodal nature of Generative AI

Include expressions/outputs in various modes such as audio, image, video, text, etc.

Tech-Stack Analysis

Understanding the layers of GenAI supply chain and life cycle

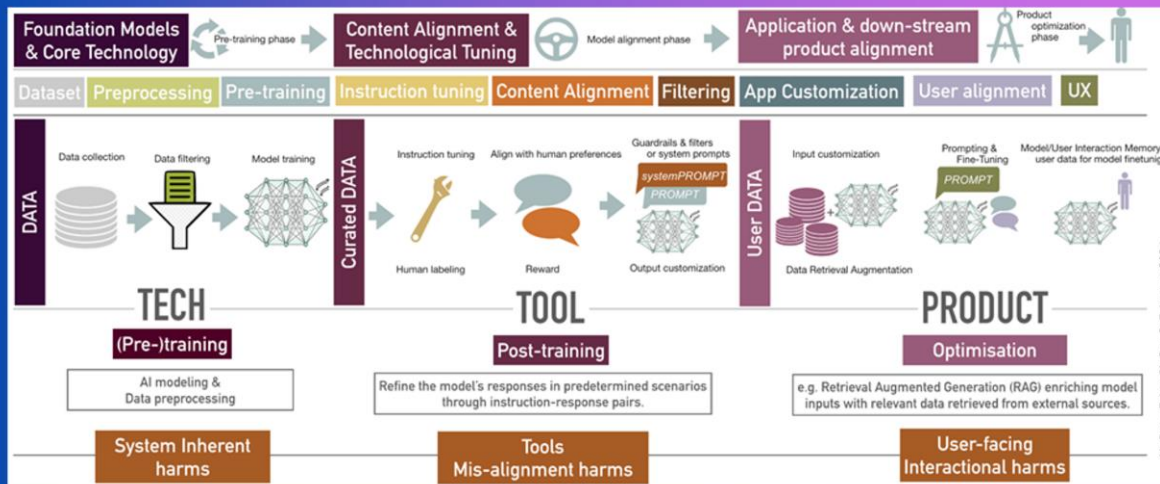
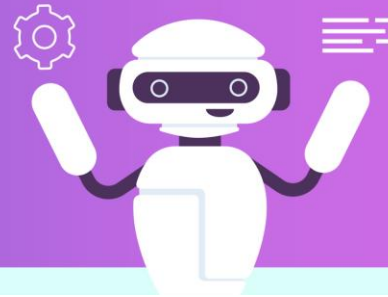
Mapping different implications requires breaking down complexity and unpack the peculiarities that distinguish GenAI

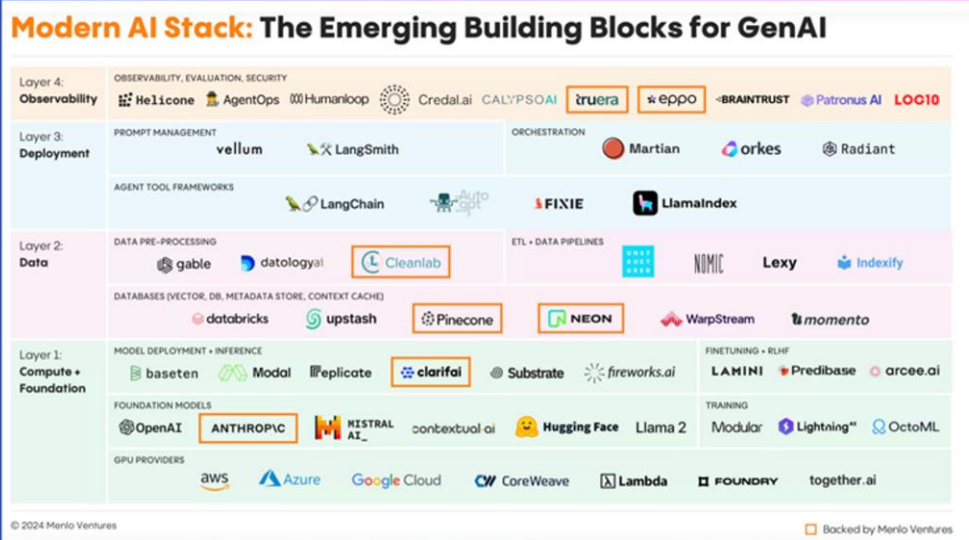
Frame of Analysis: The Tech-Stack-Layers

Layer 1: Foundation Model and core infrastructure

Layer 2: Model tuning and content alignment

Layer 3: Application and deployment product alignment





The positive side of implications

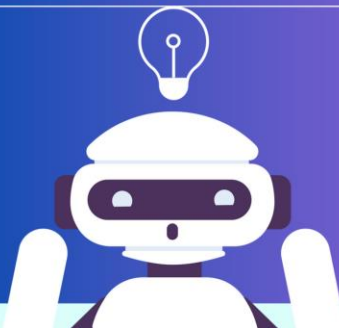
Not only a focus on harms and risks, but also where there are benefits for FoE from GenAI

Dual use cases

Often risks and benefits are different sides of same coins, especially when applications are exploited for malicious purposes

Individual -level and society -level effects

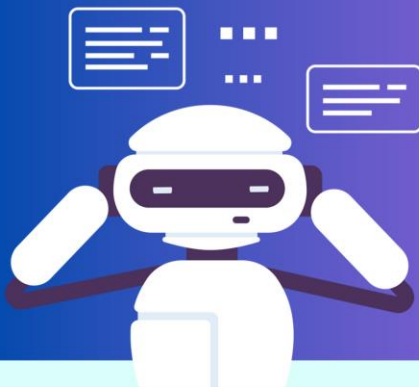
Understanding how GenAI has different impacts at different levels allows us to also devise and fine -tune the necessary reaction



Risks and benefits

FoE - Rights	Benefits	Risks	Use case
Right to be heard Right to impart info	Enhancing personal expression	Biasing personal expression	Co-writing/ co-creating
Right to be heard Right to impart info	Inclusion of vulnerable groups	Exclusion groups in vulnerable situation	Multimodality & automation
Right to impart information	At scale content moderation	FoE interference & biased moderation	GenAI-powered filtering
Right to hold/form opinions	Informational personalization	Persuasiveness & opinion manipulation	Information education social networks /chat
Right to hold/form opinions	Richer content & larger audience involvement	Persuasion influence inaccuracy	Multimodality
Right to be heard Right to impart/receive info	Efficiency and task enhancement	Cognitive-capital loss implying FoE erosion	Cognitive tasks automation
Right to be heard Right to impart/receive info	Multilingualism	Linguistic-capital loss & Cultural bias	Translation
Right to receive information	Automatic content explainers & informationalGrounding in RAG	Hallucination & content approximation	Q&A bot + Augmented search
Right to receive information	Larger scale access to therapy & support	Manipulation	Behavioral therapy with avatars (P&G ex.)
Right to receive information	Contextualizing information	Informational inaccuracy & approximation	Augmented writing content creation
Right to receive information	At scale information dissemination	Informational noise & dis-information	Automatic content generation
Right to be heard Right hold opinion Right to impart/receive info	Counter dis-information (+)	Democratic risk	Political content automation: Biden Voice cloning
Right to be heard Right hold opinion Right to impart/receive info	Knowledge transfer /access knowledge/info	Lack of human/accurate feedback & understanding in automatic content generation	Knowledge content automation: Floridi-bot or Avatars

Summary



Goal

Provide guidance on:

- how to capitalize on Gen AI potential for FoE
- how to mitigate associated risks and harms for FoE

Provide clearly identified and actionable tools for public and private sector actors:

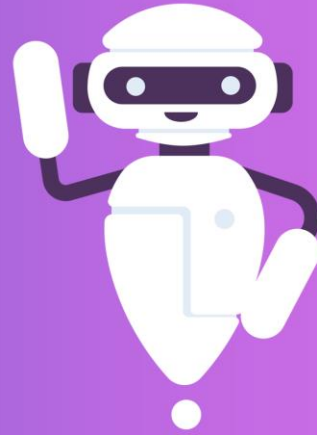
- to identify roles and responsibilities
- to ascertain governance and oversight mechanisms
- to develop and provide access to remedies, standards and capacity building

Provide actionable guidance for users:

- to raise awareness regarding their potential exposure to FoE harms/benefits when using GenAI
- to increase their capability to react accordingly and mitigate potential negative FoE implications

Thank you

andrin.eichin@bakom.admin.ch



Council of Europe instruments and policies concerning the right to privacy and to personal data protection | Peter KIMPIAN



Convention 108

The Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (ETS No. 108)

- is the only legally binding multilateral instrument in the area of the protection of privacy and personal data
- has been open for signature in 1981
- has currently 55 parties and 36 observers, the Conventional committee's work expands on all continents
- provides for a principle-based framework for the protection of individuals' privacy and personal data and a viable forum for cooperation to supervisory authorities.
- its Conventional committee with its 70-80 members and observers offers a unique forum for discussions and deliberations concerning the rights to privacy and data protection at multilateral level and in the meantime represents a viable legal instrument for the convergence towards a high set of standards globally.
- it has recently been modernised by the Protocol CETS No. 223 (Convention 108+)
- it is set to ensure, upon the entry into force of the amending Protocol, hopefully in the coming years that
 - the flow of data is facilitated, including to and from the European Union
 - the respect for human dignity in the digital age is secured



THE UNIVERSAL DECLARATION OF Human Rights

Preamble Recognition of the inherent dignity and of the equal and inalienable rights of all members of the human family is the foundation of freedom, justice and peace in the world.

Article 1 All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood.

Article 2 No one shall be subjected to discrimination on the basis of race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status. Furthermore, no distinction shall be made on the basis of the political, racial or religious opinions of the parents, family or relatives of any individual.

Article 3 Everyone has the right to life, liberty and security of person.

Article 4 No one shall be held in slavery or servitude; slavery or the trade in slaves shall be prohibited in all its forms.

Article 5 No one shall be subjected to torture or to cruel, inhuman or degrading treatment or punishment.

Article 6 Everyone has the right to recognition as a person before the law.

Article 7 All are equal before the law and are entitled without any discrimination to equal protection of the law. All are entitled to a fair and public hearing by an independent and impartial tribunal in the determination of any criminal charge against them, of their rights and obligations, or of any criminal charge against them.

Article 8 Everyone has the right to an effective remedy by the competent national authorities for violations of his or her fundamental rights and freedoms recognized by the constitution and laws of the country.

Article 9 No one shall be subjected to arbitrary arrest, detention or exile.

Article 10 Everyone is entitled in full equality to a fair and public hearing by an independent and impartial tribunal, in the determination of his or her rights and obligations, of any criminal charge against him or her, or of any civil or administrative charge.

Article 11 Everyone charged with a criminal offence shall have the right to be presumed innocent until proved guilty according to law. In the course of a criminal trial, he or she shall have the right to defend himself or herself and, if he or she has no means to do so, to be assisted by a lawyer of his or her own choosing or, if he or she cannot afford one, by one appointed for him or her by the law.

Article 12 No one shall be subjected to arbitrary interference with his or her privacy, family, home or correspondence, nor to attacks upon his or her honour and reputation. Everyone has the right to the protection of the law against such interference or attacks.

Article 13 Everyone has the right to free movement and residence within the frontiers of each State.

Article 14 Everyone has the right to leave any country, including his or her own, and to return to his or her country.

Article 15 1. Everyone has the right to a nationality. 2. No one shall be arbitrarily deprived of his or her nationality. 3. No one shall be denied the right to change his or her nationality.

Article 16 1. Everyone has the right to marry and to found a family, which is the basis of the society. 2. Men and women of full age, without any limitation of race, nationality or religion, shall have the right to marry and to found a family. 3. Marriage shall be entered into only with the free and full consent of the intending spouses. 4. The family is the natural and fundamental group unit of society and is entitled to protection by society and the State.

Article 17 1. Everyone has the right to own property alone as well as in association with others. 2. No one shall be arbitrarily deprived of his or her property.

Article 18 Everyone has the right to freedom of thought, conscience and religion; this right includes freedom to change his or her religion or belief, and freedom, either alone or in community with others and in public or private, of worship, teaching, practice and observance.

Article 19 Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without restriction, to receive and impart information and ideas without any interference by public authorities, by law or by any other person, in writing or in print, orally or in the form of print, press, photography, cinematography, television or by any other means of mass communication.

Article 20 1. Everyone has the right to peaceful assembly and to associate with others in the form of unions, associations or societies for the promotion of their interests. 2. No one shall be compelled to belong to an association.

Article 21 1. Everyone has the right to take part in the government of his or her country, directly or through freely chosen representatives. 2. Everyone has the right of equal access to public service in his or her country. 3. The will of the people shall be the basis of the authority of government; this will shall be expressed through periodic and genuine elections which shall be by secret ballot and shall ensure the free expression of the will of the voters.

Article 22 Everyone has the right to a just and favourable social order.

Article 23 1. Everyone has the right to work, to free choice of employment, to just and favourable conditions of work and to protection of employment. 2. Everyone has the right to equal pay for equal work. 3. Everyone has the right to a just and favourable remuneration which shall include compensation for overtime, and to a just and favourable social security.

Article 24 Everyone has the right to rest and leisure, including reasonable limitation of working hours and periodic holidays with pay.

Article 25 Everyone has the right to a standard of living adequate for the health and well-being of himself or herself and of his or her family, including food, clothing, housing, medical care and necessary social services, and the right to education.

Article 26 1. Education shall be free, at least in the elementary and technical stages. 2. Education shall be directed to the full development of the human personality and to the strengthening of respect for human rights and fundamental freedoms. 3. It shall promote understanding, tolerance and friendship between nations, racial or ethnic groups and religious communities. 4. It shall be directed to the maintenance of peace and to the promotion of world brotherhood.

Article 27 Everyone has the right to the enjoyment of the arts and to the benefit of scientific progress.

Article 28 Everyone has the right to a just and favourable social order.

Article 29 1. Everyone has duties to the community in which alone the free and full development of his or her personality is possible. 2. The exercise of the rights and freedoms set forth in this Declaration entails such duties.

Article 30 Nothing in this Declaration shall be interpreted as authorizing any State, group or person to engage in any activity aimed at the destruction of any of the rights and freedoms set forth in this Declaration.

UNITED NATIONS



ECHR Right to private life Data protection



Article 8 – Right to private life

- 1. Everyone has the **right** to respect for his private and family life, his home and his correspondence.
- 2. There shall be no **interference** by a public authority with the exercise of this right except such as is in accordance with the law and is necessary 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



Data protection : an enabling right

ECHR

Article 1 – Obligation to respect human rights

Article 8 – Right to private life

Article 9 – Freedom of thought, conscience
and religion

Article 10 – Freedom of expression

Article 11 – Freedom of assembly and
association



Protecting data ?

Protecting individuals ?

Convention for the protection of individuals
with regard to the processing of personal
data

CONVENTION 108



Convention 108+

Preamble

Chapter I – General provisions

Article 1 – Object and purpose

Article 2 – Definitions

Article 3 – Scope

Chapter II – Basic principles for the protection of personal data

Article 4 – Duties of the Parties

Article 5 – Legitimacy of data processing and quality of data

Article 6 – Special categories of data



Convention 108+

Article 7 – Data security

Article 8 – Transparency of processing

Article 9 – Rights of the data subject

Article 10 – Additional obligations

Article 11 – Exceptions and restrictions

Article 12 – Sanctions and remedies

Article 13 – Extended protection

Chapter III – Transborder flows of personal data

Article 14 – Transborder flows of personal data

Chapter IV – Supervisory authorities

Article 15 – Supervisory authorities

Chapter V – Cooperation and mutual assistance

Article 16 – Designation of supervisory authorities

Article 17 – Forms of cooperation

Chapter VI – Convention Committee

Chapter VII – Amendments



Convention 108+

Transparency (article 8)

obligation for the controller to provide a detailed list of information, as well as any necessary additional information in order to ensure fair and transparent processing



Convention 108+

Rights of the data subject (article 9)

"... not to be subject to a decision significantly affecting him or her based solely on an automated processing of data without having his or her views taken into consideration"

"...to obtain, on request, knowledge of the reasoning underlying data processing where the results of such processing are applied to him or her"



Convention 108+

Additional obligations (article 10)

- "... take all appropriate measures to comply with the obligations of this Convention and be able to demonstrate (subject to exceptions)...compliance "
- "examine the likely impact ... prior to the commencement (PIA)... and design the processing to prevent or minimise the risk" (**PbD**).
- "implement **technical and organisational measures** at all stages of the processing. (DPO, etc..)
- Adapted according to the context of the data processing (based on potential risksto adapt those measures according to the nature and volume of the data, the nature, scope and purpose of the processing and, where appropriate, the size of the controller or processor.)



Convention 108+

Exception: provided for by law, respects the essence of the fundamental rights and freedoms and constitutes a necessary and proportionate measure in a democratic society for:

the protection of national security, defence, public safety, important economic and financial interests of the State, the impartiality and independence of the judiciary or the prevention, investigation and prosecution of criminal offences and the execution of criminal penalties, and other essential objectives of general public interest; the protection of the data subject or the rights and fundamental freedoms of others, notably freedom of expression.



Convention 108+

Exceptions and restrictions (Article 11.3)

Processing activities for national security and defence purposes

...each Party may provide, by law and only to the extent that it constitutes a necessary and proportionate measure in a democratic society to fulfil such aim, exceptions to:

- the review mechanism (Article 4.3)
- DPAs' powers with regard to transborder data flows (Article 14.5 and 14.6)
- DPAs' powers (Article 15.2 a,b,c,d)



Convention 108+

Exceptions and restrictions (Article 11.3)

processing activities for national security and defence purposes

" ...

This is without prejudice to the requirement that processing activities for national security and defence purposes are subject to independent and effective review and supervision."



Convention 108 modernised

Provisions possibly subject to exception	Purposes for which the exception can be enacted
- General principles (fairness, transparency, purpose limitation, data quality requirements: adequate, relevant, not excessive, accurate, up to date, permits only identification no longer that necessary) - Data breach notification to the supervisory authority - Data controller obligation to inform the data subject - Data subject rights (Article 5 paragraph 4, Article 7 paragraph 2, Article 8 paragraph 1 and Article 9)	a. the protection of national security, defence, public safety, important economic and financial interests of the State, the impartiality and independence of the judiciary or the prevention, investigation and prosecution of criminal offences and the execution of criminal penalties, and other essential objectives of general public interest; b. the protection of the data subject or the rights and fundamental freedoms of others, notably freedom of expression.
- Data controller obligation to inform the data subject - Data subject rights (Articles 8 and 9)	Archiving purposes in the public interest, scientific or historical research purposes or statistical purposes when there is no recognisable risk of infringement of the rights and fundamental freedoms of data subjects.
In addition to the exceptions above, if provided for by law and only to the extent that it constitutes a necessary and proportionate measure in a democratic society: - Committee's power to evaluate the effectiveness of the measures taken - To provide information to the supervisory authority on international transfer - To require by the supervisory authority to demonstrate the lawful conditions for international transfer and its ability to intervene - Supervisory authority's power to investigate and intervene, functions relating to international transfer, power on taking regulatory decisions and sanctions, to turn to the judiciary (Article 4 paragraph 3, Article 14 paragraphs 5 and 6 and Article 15, paragraph 2, letter a, b, c and d.)	Processing activities for national security and defence purposes



Convention 108+

Transborder dataflows (Article 14.1)

Limitation to free flow between Parties where (to be applied in specific cases and narrowly)

- real and serious risk that the transfer would lead to circumventing the provisions of the Convention
- Party bound by harmonised rules of protection shared by States belonging to a regional international organisation



Convention 108+

Transborder dataflows (article 14.3 and 14.4)

- Means to secure an appropriate level of protection (*ad hoc* or approved standardised safeguards provided by legally binding instruments)
- Possibilities to transfer in specific cases where consent, specific interests of the data subject, prevailing legitimate interests provided for by law and are necessary and proportionate in a democratic society



Convention 108+

Supervisory authorities (article 15)

"2. ... such authorities:

- a. shall have powers of investigation and intervention;
- c. shall have powers to issue decisions with respect to violations of the provisions of this Convention and may, in particular, impose administrative sanctions;
- d. shall have the power to engage in legal proceedings or to bring to the attention of the competent judicial authorities violations of the provisions of this Convention;"



Convention 108+

Supervisory authorities (article 15)

(144 privacy legislation in the word)

"5. The Supervisory authorities shall act with complete independence and impartiality in performing their duties and exercising their powers and in doing so shall neither seek not accept instructions."



Convention 108+

Cooperation (article 17)

Supervisory authorities "shall co-operate with one another to the extent necessary for the performance of their duties and exercise of their powers, in co-ordinating their investigations or interventions, or conducting joint actions;"



Convention 108+

Evaluation and follow-up mechanism

Objective: *ensure the credibility of Convention 108+ and establish a genuine dynamic of harmonised protection, guaranteeing that data flows occur among Parties offering an appropriate level of protection*

Transparent, effective and impartial

(independent experts / questionnaire / visits)



Convention 108 today

55 countries bound by the Convention

= Argentina, Cabo Verde, Mauritius, Mexico, Morocco, Senegal, Tunisia, Uruguay, RF + **46 CoE**

pending: Costa Rica

+ 36 observers (Australia, Brazil, Canada, Chile, Gabon, Ghana, Indonesia, Israel, Japan, South Korea, Philippines, USA, New Zealand)

= NEARLY 70 COUNTRIES



Initiatives on AI

Soft law and hard law instruments (standards)

Transversal

- A transversal legal instrument to regulate the design, development and use of artificial intelligence systems – CAI – Opened for signature on 5 September
- Recommendation of the Committee of Ministers to member States on the human rights impacts of algorithmic systems - CM/Rec(2020)1
- Declaration of the Committee of Ministers on the manipulative capabilities of algorithmic processes - Decl(13/02/2019)1
- Unboxing AI: 10 steps to protect human rights - Recommendation of the Commissioner for Human Rights, May 2019
- Recommendation of the Parliamentary Assembly of the Council of Europe about Technological convergence, artificial intelligence and human rights - Recommendation 2102(2017)



Initiatives on AI

Privacy and Data Protection

- [New] Recommendation of the Committee of Ministers on the protection of individuals with regard to automatic processing of personal data in the context of profiling - CM/Rec(2021)8
- Guidelines on Facial Recognition - T-PD(2020)03
- Guidelines on Artificial Intelligence and Data Protection - T-PD(2019)01
- Guidelines on the protection of individuals with regard to the processing of data in a world of Big Data - T-PD(2017)1

Justice and public administration

- [New] [Ongoing] (Until end of 2022) Review of The Administration and the handbook in the light of the use of artificial intelligence (AI) and non-AI algorithmic systems (based on comparative study in member States) [CDCJ]
- [New] Guidelines on electronic court filing (e-filing) and digitalisation of courts - CEPEJ(2021)15 (requirements regarding transparency and accountability regarding the court-filing and decisions supporting systems, and regarding introduction of machine learning based predictive analytics), 9 December 2021
- Guidelines of the Committee of Ministers of the Council of Europe on online dispute resolution mechanisms in civil and administrative court proceedings CM(2021)36add4-final (transparency and accountability guidance for the use of AI systems)
- European Ethical Charter on the use of artificial intelligence (AI) in judicial systems and their environment CEPEJ(2018)14
- Feasibility study on the establishment of a certification mechanism for artificial intelligence tools and services (in the sphere of justice and judiciary) CEPEJ, 8 December 2020

Action against crime

- Artificial intelligence and criminal law responsibility in Council of Europe member states: the case of automated vehicles. concept note CDPC(2018)14rev, October 2018
- Feasibility Study on a future Council of Europe instrument on artificial intelligence and criminal law CDPC(2020)3Rev, 20 September 2020



Initiatives on AI

Healthcare and Biomedicine

- [Strategic Action Plan on technologies and human rights in the field of biomedicine 2020 -2025 \(with AI-specific parts\)](#) - DH-BIO(2018)22
- [Report on the impact of artificial intelligence on the doctor patient relationship](#)

Non-discrimination, Gender equality

- [Ongoing] Recommendation on migrant women (including provisions on the impact of AI) – GEC
- [Ongoing] Draft Recommendation covering the impact of AI on gender inequality - GEC
- [Guidelines of the Committee of Ministers of the Council of Europe on upholding equality and protecting against discrimination and hate during the Covid-19 pandemic and similar crises in the future](#) - CM(2021)37-add1 rev
- [Recommendation of the Committee of Ministers to member States on preventing and combating sexism](#) - CM/Rec(2019)1

Social Rights

- [Declaration by the Committee of Ministers on the risks of computer-assisted or artificial-intelligence-enabled decision making in the field of the social safety net](#) - Decl(17/03/2021)2

Education

- [Recommendation on developing and promoting digital citizenship education](#) - CM/Rec(2019)10

Children's rights

- [Recommendation of the Committee of Ministers to member States on guidelines to respect, protect and fulfil the rights of the child in the digital environment](#) - CM/Rec(2018)7

Freedom of expression and culture

- [Ongoing] Development of Recommendation and Study on the impacts of digital technologies on freedom of expression - MSI-DIG
- [Ongoing] Development of Recommendation on Combating Hate Speech - ADI/MSI-DIG
- [Ongoing] Draft standard-setting instrument with guiding principles for media and communication governance in the context of the new media and information paradigm based on social media distribution, taking account of related risks (manipulation of public opinion, lack of public trust, information disorder) - MSI-REF
- [Recommendation of the Committee of Ministers to member States on the roles and responsibilities of internet intermediaries](#) - CM/Rec(2018)2



[Recommendation of the Committee of Ministers to member States on the human rights impacts of algorithmic systems](#)

The CoE

- to prevent abuses of algorithmic systems and processes
- to contributing to building a doctrine of use that can guarantee effective protection of all human rights, rule of law and democracy

The Recommendation

- addresses human rights impact of AI technologies with guidance on obligations of states and responsibilities of companies.
- provides advices for the private sector in view of determining each actor's responsibilities with respect to human rights and fundamental freedoms in the context of algorithmic systems.
- At the same time warns of significant challenges to human rights related to the use of algorithmic systems, mostly concerning the right to a fair trial; privacy and data protection; freedom of thought, conscience and religion; the freedoms of expression and assembly; the right to equal treatment; and economic and social rights



Recommendation of the Committee of Ministers on the protection of individuals with regard to automatic processing of personal data in the context of profiling

The CM (3 November 2021)

- updated a previous recommendation on the same topic adopted in 2010 with the aim to align its provisions with the modernised Convention 108.
- it highlighted that profiling techniques can impact individuals by placing them in predetermined categories, very often without their knowledge .
- the lack of transparency can pose significant risks to human rights, particularly for vulnerable persons, including children .
- It called on states to promote and make legally binding the use of a "privacy by design" approach during the whole duration of the processing, notably through the use of privacy-enhancing technologies .
- It underlined that processing for the purpose of detecting or predicting racial or ethnic origin, political opinions, trade union membership, religious or other beliefs, health or sexual life should only be allowed where appropriate safeguards are enshrined in law and the data are strictly necessary for the lawful and specific purposes of processing
- It emphasized the necessity for "respect for fundamental rights and freedoms, notably the rights to human dignity and to privacy but also the freedom of expression and the principle of non-discrimination
- It recommended that the imperatives of social justice, cultural diversity and democracy are guaranteed during the processing of personal data, in both the public and private sectors.



Guidelines on Artificial Intelligence and Data
Protection
(25 January 2019)

- AI may be **a useful tool** for decision making in particular for supporting evidence-based and inclusive policies.
- AI development and use shall respect the rights to privacy and data protection
- AI applications shall 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
- ECHR & Convention 108+
- Structure: General guidance, guidance for developers, manufacturers and service providers, guidance for legislators and policy makers



I. General Guidance

- I. **protection of human dignity** and **safeguarding of human rights and fundamental freedoms**, in particular the right to the protection of personal data, especially important when AI applications are used in **decision-making processes**
- II. 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.
- III. **Risk based approach** - avoiding and mitigating the potential risks
- IV. **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 (similarly to the Guidelines on Big Data)
- V. **rights of data subjects** shall be respected fully
- VI. **meaningful control by data subjects** over the data processing and related effects on individuals and on society



II. Guidance for developers, manufacturers and service providers

- I. values-oriented approach in the design of their products and services
- II. possible adverse consequences of AI applications on human rights and fundamental freedoms should be assessed, and, considering these consequences, **a precautionary approach** based on appropriate risk prevention and mitigation measures should be adopted
- III. **human rights by-design** approach should be adopted and 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 should be avoided
- IV. unnecessary, redundant or marginal data during the development, and training phases should be reduced and then the **model's accuracy** as it is fed with new data should be constantly monitored. The use of synthetic data may be considered.
- V. The **risk of adverse impacts** on individuals and society due to de-contextualised data and de-contextualised algorithmic models should be adequately considered
- VI. **Consultation of external bodies** (committees of experts, academic institutions)
- VII. **Participatory forms of risk assessment**
- VIII. 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 should be ensured
- IX. users' freedom of choice over the use of AI, by **providing feasible alternatives to AI** applications should be guaranteed
- X. forms of **algorithm vigilance** (specific duties on transparency, prior assessment of the impact of data processing on human rights and fundamental freedoms) that promote the accountability should be adopted
- XI. **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.
- XII. **The right to object** should be ensured in relation to processing based on technologies that influence the opinions and personal development of individuals.



III. Guidance for legislators and policy makers

- I. principle of accountability, the adoption of risk assessment procedures and the application of other suitable measures, such as **codes of conduct and certification mechanisms** should be privileged
- II. **algorithm vigilance** should be applied during public procurement procedures
- III. **supervisory authority** should be sufficiently empowered to support and monitor the algorithm vigilance programmes
- IV. **Possibility of human intervention** should be preserved
- V. Supervisory authority should be consulted when AI applications have the potential to significantly impact the human rights and fundamental freedoms of data subjects
- VI. **Cooperation among regulatory authorities** such as consumer protection; competition; anti-discrimination; sector regulators and media regulatory authorities should be encouraged
- VII. **Independence of external bodies** should be guaranteed
- VIII. **Inclusion of stake-holders** in the debate on what role AI should play in shaping social dynamics, and in decision-making processes affecting them
- IX. **Digital literacy, awareness raising, education, training and scientific researches** should be supported



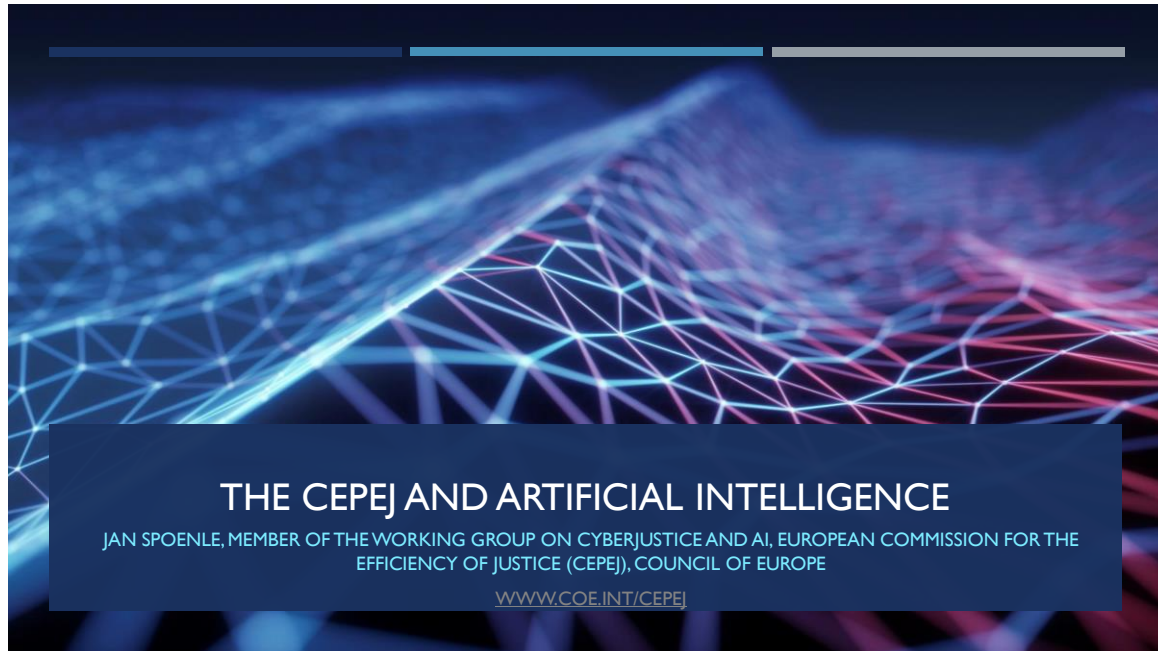
Framework Convention on AI Human Rights, Democracy and the Rule of Law

Main features of the framework convention:

- it covers the use of AI systems in the public and private sector,
- establishes transparency and oversight requirements tailored to specific contexts and risks,
- it ensures accountability and responsibility for adverse impacts and that AI systems with regard to equality, including gender equality, the prohibition of discrimination, and privacy rights, legal remedies for victims of human rights violations.
- also requires Parties to adopt measures to ensure that AI systems are not used to undermine democratic institutions and processes, and that the use of AI for national security purposes ensures that these activities respect international law and democratic institutions and processes.
- it also establishes a follow-up mechanism and an independent oversight mechanism to oversee compliance with the convention,

Several possible cross-cutting issues: methodology to assess risk (art 10 of Convention 108+ and HUDERIA), evaluation of the implementation of new rights of data subjects (Article 9 of the modernised Convention 108+) and of the data protection provisions of the second additional Protocol to the Budapest Convention (article 13,14) → proposing follow-up and evaluation mechanisms which should integrate new emerging trends from all three Conventions .

European Ethical Charter on the use of artificial intelligence in judicial systems and their environment | Jan SPOENLE

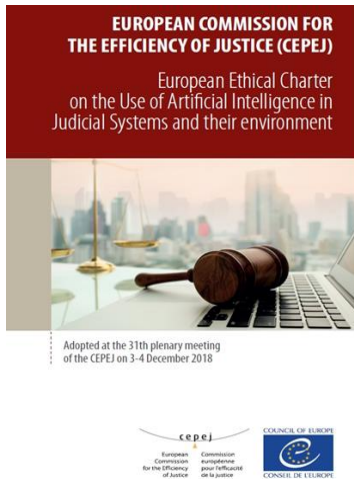


CEPEJ MANDATE

Established 2002 with Resolution Res(2002)12 of the Committee of Ministers of the Council of Europe

- to improve the **efficiency and the functioning of the justice system** of member states, with a view to ensuring that everyone within their jurisdiction can enforce their legal rights effectively
- to enable a **better implementation of the international legal instruments** of the Council of Europe concerning efficiency and fairness of justice.
- The use of **information and communication technologies shall be promoted** in order to strengthen the efficiency of justice, in particular in order to facilitate access to justice, speed up court proceedings, improve the training of legal professionals, as well as the administration of justice and management of courts.

THE EUROPEAN ETHICAL CHARTER ON THE USE OF AI IN JUDICIAL SYSTEMS AND THEIR ENVIRONMENT



1. **Respect of fundamental rights:** ensuring that the design and implementation of artificial intelligence tools and services are compatible with fundamental rights;
2. **Non-discrimination:** specifically preventing the development or intensification of any discrimination between individuals or groups of individuals;
3. **Quality and security:** with regard to the processing of judicial decisions and data, using certified sources and intangible data with models conceived in a multidisciplinary manner, in a secure technological environment;
4. **Transparency, impartiality and fairness:** making data processing methods accessible and understandable, authorising external audits;
5. **“Under user control”:** precluding a prescriptive approach and ensuring that users are informed actors and in control of their choices.

HANDS ON APPROACH

29 main questions designed to raise awareness and identify the potential infringements of an AI system on the principles of the CEPEJ Charter and the AI Framework Convention

5. **Is the AI system likely to enable or facilitate the profiling of judges and/or courts?**
☐ Yes ☐ No
 - 5.1. **If yes, which risk mitigation or management measures are envisaged?**
Multiple answers are possible.
 - ☐ Anonymization of data related to a judge
 - ☐ Anonymization of data related to a court and its members
 - ☐ Exclusion of search criteria facilitating such profiling
 - ☐ Other _____
6. **Is the AI system likely to enable or facilitate the profiling of jurors?**
☐ Yes ☐ No
 - 6.1. **If yes, is the AI system available to all the parties in the lawsuit?**
☐ Yes ☐ No
 - 6.2. **If yes, which risk mitigation or management measures are envisaged?**
Multiple answers are possible.
 - ☐ Anonymization of data related to a juror
 - ☐ Exclusion of search criteria facilitating such profiling
 - ☐ Other _____

MORE DEVELOPMENTS SINCE 2018

- Improved monitoring tools
 - ICT chapter in CEPEJ Evaluation and ICT index (Quantitative)
 - Resource Centre on Cyberjustice and AI (Qualitative)
- Targeted guidance
 - Information note on the Use of Generative AI by judicial professionals in a work-related context



The Draft Methodology for the Risk and Impact Assessment of Artificial Intelligence Systems from the point of view of Human Rights, Democracy and the Rule of Law (HUDERIA) | CAI Secretariat



1st meeting of the CDDH-IA

The Draft Methodology for the Risk and Impact Assessment of Artificial Intelligence Systems from the point of view of Human Rights, Democracy and the Rule of Law (HUDERIA)

Presentation by the CAI Secretariat



COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

Introduction

CAHAI

- Following the adoption of the Feasibility Study in December 2020, the CAHAI, through its Policy Development Group (**CAHAI-PDG**) and specifically Sub-working Group 1, started working on the HUDERIA.
- On the basis of the CAHAI-PDG's work, the **Alan Turing Institute** was asked by the Secretariat to prepare a proposal operationalizing the outline of a Model for a Human Rights, Democracy, and the Rule of Law Impact Assessment.
- The resulting output, *Human Rights, Democracy, and the Rule of Law Assurance Framework for AI Systems: A proposal*, was delivered in September 2021.

Human Rights, Democracy, and the Rule of Law Assurance Framework for AI Systems:

A proposal prepared for the Council of Europe's Ad hoc Committee on Artificial Intelligence



The Alan Turing Institute



The Project Summary (PS) Report

The Project Summary Report (PS Report) is a key document of the HUDERIA used to document preliminary research and each subsequent stage of the HUDERIA process. This document will be updated by the project team throughout the entire AI project lifecycle and serve as a core document for deliberation and accountability. You may consider making a more condensed version of the PS Report publicly available to stakeholders.

The completion of the PS Report and subsequent stages of the HUDERIA process should take place in a multidisciplinary team consisting of experts with a range of matching specialisations and backgrounds.

To lay the foundation for subsequent HUDERIA phases, the project team should begin by answering a set of project scoping and planning questions that form the PS Report (see below for the sample list of questions) aimed at:

- (a) consolidating information about the project-, use-, domain- and data-contexts (including existing bias, structural and historical inequality) of the prospective system;
- (b) identifying relevant stakeholders;
- (c) beginning to scope the potential impacts of the AI system on human rights, democracy and the rule of law, and
- (d) mapping the system's governance processes.

Once the project team has completed the initial scoping exercise and documented it in the PS Report, they may then move onto the Context-Based Risk Analysis (COBRA).

Sample List of Questions for the PS Report

Sample PS Report	
Team members involved in carrying out the HUDERIA:	
Activity 1 Outline Project-, Use-, Domain- and Data-Contexts	
Project	
1) How would you describe your organisation and what sort of services or products do you typically provide (beyond the system under consideration)?	
2) What AI system is being built and what type of product or service will it offer?	
3) What benefits will the system bring to its users and customers, and will these benefits be widely accessible?	
4) Which organisation(s)—yours, other suppliers, or other providers—are responsible for building this AI system?	
5) Which parts or elements of the AI system, if any, will be procured from third-party vendors, suppliers, sub-contractors or external developers?	
6) Which algorithms, techniques and model types will be used in the AI system? (Provide links to technical papers where appropriate)	
7) In a scenario where your project optimally scales, how many people will it impact, for how long, and in what geographic range (local, national, global)? (Describe your rationale)	
8) How diverse is your team (e.g. disciplinary backgrounds, domain expertise, positionality)?	
9) What resources are needed for this project (e.g., costings, power)?	
10) What parties are involved in this project (e.g., project lead, strategic partnerships, donors, providers)?	

5



CAI

- During the 2nd CAI Plenary meeting (September 2022), the Secretariat presented an **outline of the draft Appendix to the "Zero Draft"**, i.e. the Methodology for risk and impact assessment of artificial intelligence systems (HUDERIA).
- A **draft Methodology** was presented by the Secretariat and Dr David Leslie during the 3rd CAI Plenary meeting (January 2023).
- The CAI tasked the Secretariat with preparing a revised draft of the Methodology, on the basis of the comments received, to be examined as a **stand-alone non-legally binding instrument**, separate from (but supporting) the Framework Convention, to be presented after the finalization of the Convention.

3



- The CAI Secretariat, with the Alan Turing Institute, revised the HUDERIA in parallel to the negotiations of the Framework Convention, **consulting with experts in the field and exploring possible synergies with the work of other relevant international organisations** in the area of risk and impact assessment of AI systems.
- **Two rounds of internal consultations** were held within the CoE, through the Secretariat's AI Coordination Group, in October 2022 and May 2024, to gather input and comments on the proposed revised content of the draft Methodology.
- On **23 May 2024**, the Secretariat organized a **virtual workshop on HUDERIA for standards setting organizations** to discuss the relationship between human rights and technical standard-setting processes, take stock of the different initiatives concerning standard setting in the field of AI, and exchange views on the draft Methodology.

6



The HUDERIA combines the **state-of-the art knowledge of the Alan Turing Institute**, the UK's national institute for data science and AI, **and the expertise of the Council of the Europe** in the sphere of human rights, democracy and the rule of law.

Given the intricate nature of AI technologies, **effectively managing risks and impacts of AI systems** necessitates not only **understanding the technical aspects** of these technologies but also **acknowledging the socio-technical context** of their development and application.

7



HUDERIA addresses this complexity by:

- offering an **integrated workflow** that includes multiple processes/steps **to help regulators and project teams** to comprehensively tackle and understand both technical and socio-technical aspects of AI technologies;
- providing **specific guidance on the assessment of risks and impacts** to human rights, democracy and the rule of law, as well as on the mitigation of such risks and impacts in the context of AI;
- implementing a **risk-based approach**, guided by the precautionary principle, **to facilitate the identification of use-cases that warrant pre-emptive measures** to safeguard human rights, democracy and the rule of law;
- employing a **graduated and differentiated approach**, in line with Article 1 of the Framework Convention, to assist AI regulators and relevant stakeholders in tailoring and adjusting the risk and impact assessment process. This includes the description of requirements for iterative reassessment and stakeholder engagement practices.

8



- **support and accompany the application and implementation of the Framework Convention** on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (the Framework Convention);
- it **provides clear and concrete criteria for identifying contexts and intended uses** in which the activities of artificial intelligence (AI) systems have the potential to interfere with human rights, democracy and the rule of law;
- it **ensures and promotes a comprehensive approach towards the identification, assessment, prevention and mitigation of risks** posed by AI systems as well as the consideration of actual and potential impacts to human rights, democracy and the rule of law, but at the same time is **based on the assumption that domestic authorities are better placed to make relevant policy and regulatory choices**, taking into account their country's specific political, economic, social, cultural and technological contexts, and that **they should accordingly enjoy a certain degree of flexibility** when addressing the issues of risk and impact management;
- it **specifies procedures and mechanisms** that could be put in place to ensure the identification, assessment, prevention and mitigation of such risks and impacts;

9



- it **ensures compatibility and interoperability** of the suggested approach **with the existing assessment, governance, and compliance guidelines and frameworks** already followed by the industry and currently developed by the relevant competent bodies (OECD Due Diligence Guidance for Responsible Business Conduct, the NIST AI Risk Management Framework, ISO/IEC, IEEE and other technical standards bodies);
- is **designed to be as “algorithm neutral” and practice-based as possible** so that it can remain maximally future proof and inclusive of various AI applications. The methodology will also need to stay responsive to the development of further AI innovations and use-cases, and should be seen as dynamic and in need of regular revision;
- it may also be used to **identify positive actions to promote human rights**, including non-discrimination;
- the envisaged HUDERIA processes remain **without prejudice to forthcoming sectoral legal standard-setting instruments** that give further guidance on the implementation of human rights principles in the context of AI systems. The relevant legal content of the HUDERIA is sufficiently evolutive and flexible to incorporate ongoing and future development of global and regional human rights standards.

10



- The HUDERIA Pilot Programme was **designed in accordance with an evidence-based approach**.
- **Aim: operational testing of its intended operation in a given, real-world environment.**
- **Form: collaborative test run** of the HUDERIA with a small group of volunteering member and observer States who:
 - chose an existing AI system which is currently in use or has recently been decommissioned (a mock use case created by the Alan Turing Institute was used for the countries without a use case);
 - worked with the Secretariat and the expert team to run through what the design, development, deployment and/or procurement would have looked like had the Methodology been in place from the start.

11



- The HUDERIA Pilot Programme was **designed in accordance with an evidence-based approach**.
- **Aim: operational testing of its intended operation in a given, real-world environment.**
- **Form: collaborative test run** of the HUDERIA with a small group of volunteering member and observer States who:
 - chose an existing AI system which is currently in use or has recently been decommissioned (a mock use case created by the Alan Turing Institute was used for the countries without a use case);
 - worked with the Secretariat and the expert team to run through what the design, development, deployment and/or procurement would have looked like had the Methodology been in place from the start.

11



- For volunteering States, The HUDERIA Pilot Programme consisted in:
 - **5 three-hour workshops** (one for each of the HUDERIA governance steps; with an introductory educational session about the governance step and then a series of team-based exercises that involve carrying out the step in a real-world scenario);
 - **sharing of all relevant information** about the actual process behind the designing, developing, procuring, and using the AI system under consideration;
 - **final meeting** involving all participants to brainstorm the outcomes + **a 45-minute interview** carried out by the Secretariat and the Alan Turing team.

12



- **June 2023: call for volunteers** to participate in the HUDERIA Pilot Programme.
- **7 countries volunteered** to participate (on a confidential basis, 5 with their own use cases and 2 without).
- **June 2024:** launch of the **Pilot workshops**.
- **July 2024:** final meeting and interviews.

13



Following the completion of the HUDERIA Pilot Programme, the CAI Secretariat and the Alan Turing Institute reviewed the very informative input gathered during the Programme and incorporated their findings into the revised draft Methodology which was examined during the 11th Plenary meeting.

14

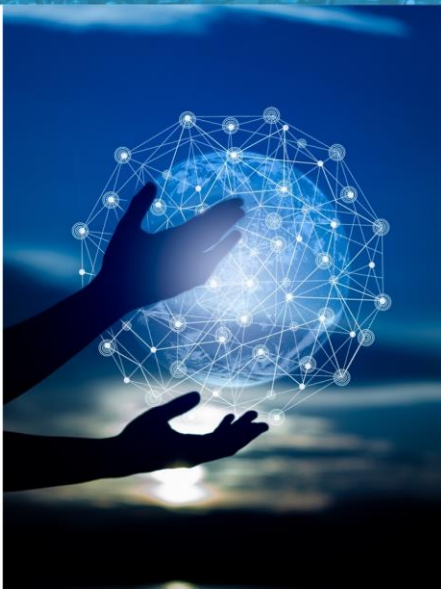


COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

Next steps

11th CAI Plenary

- **General presentation** of the HUDERIA by the Secretariat and ATI
- **Exchange of views and Q&A** with the Committee
- **Step by step** examination of the HUDERIA
- **Stock taking** of general comments





COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

The outcome of the 11th Plenary meeting

- **Overall support for HUDERIA** as a stand alone, non-legally binding instrument which provides detailed guidance on the issues of risk and impact assessment, yet remains flexible and technology neutral.
- **Main points:**
 - readability and useability should be improved, having in mind that the document is intended (also) for technical community;
 - the core description of the HUDERIA process in a shorter document (20 to 25 pages);
 - more detailed guidance on different stages in appendices.



Between 11th and 12th CAI Plenaries

- **Redrafting** by the Secretariat and ATI
- **Distribution** to the Committee
- **Written procedure** for comments and amendments
- **Distribution** to the Committee

12th CAI Plenary

- **Finalisation** of the HUDERIA and **submission** to the Committee of Ministers

