

# AIA topics

---

## Fundamental Rights Impact Assessment

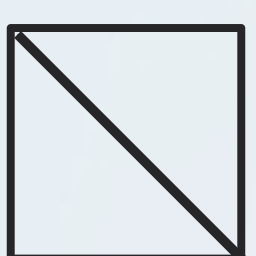
Trilogue paper #1

This topical paper is part of a series of short analyses of important, decisive or divisive topics in the legislative process regarding the European Commission proposal for a Regulation for Artificial Intelligence (AIA).

---

**Author:**

Catelijne Muller, LL.M



# Fundamental Rights and the AIA

---

## What is the issue?

The legislative process around the AIA has recently entered its final phase of the "Trilogue". In this phase the two co-legislators, the European Parliament and the Council of the EU, negotiate the final text of the AIA, with honest brokerage by the European Commission.

In June and December respectively, the EP and the Council set their negotiating positions on the European Commission's proposal of the AIA and fundamental rights have taken up an even more important place in these positions.

While the AIA proposal of the European Commission is already grounded in EU values and aimed at protecting health, safety and fundamental rights, the European Parliament wants to strengthen this objective by **adding the protection of democracy, the rule of law and the environment** to it.

Both the EP and the Council also propose to add the obligation for providers of high risk AI systems (including general purpose AI systems that can be used as such) to conduct a **'Fundamental Rights Impact Assessment'** (FRIA) as part of the obligation to have a risk management system in place (art. 9 AIA). The EP wants to extend this obligation to **deployers** of AI systems as well, whereas, according to Euractiv, the Spanish Presidency of the Council wants to limit this obligation to public actors. As regards foundation models, the EP made proposals that include (ongoing) identification and assessments of fundamental rights impacts of these systems.

These positions show that the protection of fundamental rights has become even more front and center in the run up to the final text of the AIA.

What is also important to note is that both co-legislators suggest adding an extra layer to high risk AI systems. The EP proposes that high risk AI systems listed on ANNEX III shall be considered high-risk **if they pose a significant risk of harm to the health, safety or fundamental rights of natural persons**. Providers can submit a reasoned notification to the national supervisory authority when they conclude that they are not subject to the requirements of Title III Chapter 2 of this Regulation, because their system does not pose such significant risk. A misclassification could result in hefty fines.

The Council takes the opposite approach by proposing that AI systems referred to in Annex III shall be considered high-risk **unless** the output of the system is purely accessory in respect of the relevant action or decision to be taken and **is not therefore likely to lead to a significant risk to the health, safety or fundamental rights**.

We strongly advise against both of these extra layers as they would create continued legal uncertainty, as it leaves too much room for interpretation, which in turn could stifle innovation. However, should an extra layer however find its way into the final text of the AIA, providers and deployers of high risk AI systems, both public and private, could only benefit from a mandatory FRIA, as it will support them in making a correct and well founded assessment.

## AI and Fundamental Rights, Democracy, the Rule of Law and the Environment

While many think of the right to non-discrimination and privacy as the fundamental rights that can be impacted by AI, it should be noted that AI can have an adverse impact on virtually all fundamental rights. It is also important to note that many AI-applications already in use could further exacerbate and amplify this impact at scale, affecting larger parts of society and more people at the same time. Many AI-systems or uses can impact various fundamental rights at the same time, but also can adversely affect one person's or group's rights while positively affecting another's.

One can identify four "Families of Fundamental Rights" under the European Charter of Fundamental Rights (ECFR) that are most likely to be adversely impacted by AI:

1. Human Dignity
2. Freedom of the Individual
3. Equality and Solidarity
4. Justice

Moreover, AI has ample impact on:

5. Democracy
6. The Rule of Law
7. The Environment

Below are some examples of how AI can adversely impact fundamental rights within these 'families'.

### *1. AI & Human Dignity*

The right to human dignity is the foundation of all fundamental rights. It encompasses both the absence of humiliation but also the presence of recognition and is a comprehensive set of 'needs' (Al-Rodhan, 2021), that can all be impacted by AI in complex ways. The right to human dignity is directly reflected in the ECFR, but also in other fundamental rights such as the **right to life** and the **right to integrity of the person**.

The combination of data and computing power has enabled the capture of an unprecedented amount of information about us. The use of AI has made these troves of data a true playground for the categorisation, sifting, sorting, and profiling of our entire lives, behaviour, thoughts, ideas and beliefs. These insights in turn have created ample opportunities to target, nudge, manipulate and deceive us and have led to altered, even harmful beliefs, ideas, and behaviour. As such our human dignity and integrity is jeopardised.

Self-driving cars, AI-driven diagnostics and prognostics, AI to manage critical infrastructure such as electricity grids or emergency dispatch services, can have an adverse impact on the right to life, if they malfunction, are inaccurate, lead to wrong diagnoses and so on.

### *2. AI & Freedom of the individual*

Freedom of the individual is reflected in the ECFR in several principles, freedoms and rights that are relevant for AI, such as **respect for private and family life, freedom of thought, freedom of expression and information, freedom of assembly and association, the right to education** and the **right to asylum**.

Many AI-systems and uses have a broad and deep impact on the right to privacy that goes well beyond our data. Privacy encompasses the protection of a wide range of elements of our private lives, that can be grouped into three broad categories namely: (i) a person's (general) privacy, (ii) a person's physical, psychological or moral integrity and (iii) a person's identity and autonomy.

**AI-driven (mass) surveillance**, for example with facial recognition, affects our 'general' privacy, identity and autonomy in such a way that it creates a situation where we are (constantly) being watched, followed and identified. As a psychological 'chilling' effect, people might feel inclined to adapt their behaviour to a certain norm, for example becoming hesitant to take part in **peaceful demonstrations**. In fact, recently the European Court of Human Rights issued a judgement in *Glukhin v. Russia* that "the use of highly intrusive facial recognition technology in the context of the applicant exercising his Convention right to freedom of expression is incompatible with the ideals and values of a democratic society governed by the rule of law".

**Recognition of micro-expressions**, gate, (tone of) voice, heart rate, temperature, etc. are being used to assess or even predict our behaviour, mental state and emotions, while no sound scientific evidence exists corroborating that a person's inner emotions or mental state can be accurately 'read' from a person's face, gate, heart rate, tone of voice or temperature, or that future behaviour could be predicted by it. These kinds of AI techniques, for example in recruitment, law enforcement, and schools, impact a person's physical, psychological or moral integrity and thus elements of that person's private life.

Moreover, massive amounts of '**data points**' on how we go about our daily lives are used not only to send us targeted advertising, but also to push/influence/induce/nudge us towards certain information and thus influencing our opinions and affecting our moral integrity.

Lastly, recent developments in AI (supposedly) reading our minds, or understanding our thoughts, has (re-)opened the discussion on the need for **neuro rights** or the **right to mental privacy**, as an addition or part of the right to privacy.

The ECFR provides for the **freedom of expression and information**, including the **freedom to hold opinions, and to receive information and ideas**. AI being used to profile, survey, track and identify people and screen, define, sort and influence or nudge behaviour can have a chilling effect on these particular freedoms. AI used in (social) media and news curation, bringing ever more 'personalised' online content and news to individuals, raises concerns. Search engines, social media **recommender systems** and news aggregators often are opaque, both where it comes to the data they use to select or prioritise the content, but also where it comes to the purpose of the specific selection or prioritisation.

Beyond commercial motives, political or other motives might lead to AI-systems being optimised to select or prioritise particular content in an effort to coerce and influence individuals towards certain points of view, for example during election processes. Moreover, AI is becoming very capable of producing media footage (video, audio, images) resembling real people's appearance and/or voice (also known as '**deep fakes**'), enabling the deceptive practices for various purposes. All this can give rise to filter bubbles and proliferation of fake news, disinformation and propaganda, and affects the capacity of individuals to form and develop opinions, receive and impart information and ideas and thus impact our **freedom of expression**.

The internet and social media have shown to be helpful tools for people to exercise their right to peaceful assembly and association. At the same time however, the use of AI could also jeopardise these rights, when people or groups of people are automatically tracked and identified and perhaps even 'excluded' from demonstrations or protests.

Using facial recognition in public areas may not only interfere with a person's freedom of opinion and expression, simply because the **protection of 'group anonymity' no longer exists**, if everyone in the group could potentially be recognised, but could also lead to individuals by no longer partaking in peaceful demonstrations.

### *3. AI and Equality and Solidarity*

One of the most reported impacts of AI on fundamental rights is the impact on the **prohibition of discrimination**. As noted earlier, in many cases, AI has shown to perpetuate and amplify and possibly enshrine discriminatory or otherwise unacceptable biases. Also, AI can enlarge the group of impacted people, when it groups them based on shared characteristics. Moreover, these data-driven systems obscure the existence of biases, marginalizing the social control mechanisms that govern human behaviour.

The principle of solidarity includes a.o. **the right to fair and just working conditions**. AI can have major benefits when used for hazardous, heavy, exhausting, dirty, unpleasant, repetitive, or boring work. AI systems are however also increasingly being used to monitor and track workers, distribute work without human intervention and assess and predict worker potential and performance for hiring, firing and worker evaluations. These applications of AI could jeopardize the right to just conditions of work, safe and healthy working conditions, dignity at work. AI-systems that assess and select job candidates could jeopardise the right to engage in work and the principle of equality between women and men, if the system produces biased outcomes.

### *4. AI & Justice*

The principle of justice is enshrined in the ECHR in several rights and principles that are relevant for AI, such as the **right to a fair trial** and **the presumption of innocence**. The fact that AI can perpetuate or amplify existing biases, is particularly pertinent when used in the judiciary. In situations where physical freedom or personal security is at stake, such as with predictive policing, recidivism risk determination and sentencing, the **right to liberty and security** combined with the right to a fair trial are vulnerable.

When an AI-system is used for recidivism prediction or sentencing it can have biased outcomes, jeopardising the **right to non-discrimination**. When it is a **black box**, it becomes impossible for legal professionals, such as judges, lawyers and district attorneys to understand the reasoning behind the outcomes of the system and thus complicate the motivation and appeal of the judgement, jeopardizing the right to a fair trial.

Less obvious is the impact of AI on the right to the **presumption of innocence**. AI-applications used for predictive policing merely seek correlations based on shared characteristics with other 'cases'. Suspicion in these instances is often not based on actual suspicion of a crime by the particular suspect, but merely on characteristics someone happens to share with others (such as address, income, nationality, debts, (un)employment, behaviour, behaviour of friends and family members and so on). Moreover, the actual characteristics used in the AI-system and the 'weights' given to those characteristics are often unknown.

## 5. Democracy

AI can have (and likely already has) an adverse impact on democracy, in particular where it comes to: (i) social and political discourse, access to information and voter influence, (ii) inequality and segregation and (iii) systemic failure or disruption.

Well-functioning democracies require a well-informed citizenry, an open social and political discourse and absence of opaque voter influence. In information societies citizens can only select to consume a small amount of all the available information. Search engines, social media feeds, recommender systems and many news sites employ AI to determine which content is created and shown to users (information personalization). If done well, this could help citizens to better inform themselves. However, if AI determines which information (including political) is presented to whom, which is suppressed and which is amplified, this could **undermine a shared understanding of issues, mutual respect and trust in democratic institutions.**

The recent developments in LLM's and other generative AI systems have shown an uptick in mis- and disinformation and deep fakes, in the form of text, pictures, audio or video. Since these are hard to identify by citizens, journalists or public institutions, misleading and manipulating the public becomes easier and the level of truthfulness and **credibility of the media and democratic discourse** may deteriorate.

AI could lead to inequality and segregation and thus threaten the necessary level of economic and social **equality** required for a thriving democracy.

AI decisions that previously only humans were able to make, create new challenges for the **security and resilience of societal systems.** In particular, if decisions that previously were made by many decentralised actors are replaced by few centralised AI-driven systems, **systemic risks** increase, where the failure of merely a few centralised systems is enough to potentially create catastrophic results. When critical energy infrastructures, transport systems and hospitals increasingly depend on automated decisions of AI this introduces new vulnerabilities in the form of single points of failure with widespread effects.

Another democratic challenge is private **power concentration.** If too much AI power is concentrated in a few private hands which prioritise shareholder-value over the common good, this can threaten the authority of democratic states.

## 6. Rule of law

Public institutions are held to a higher standard when it comes to their behaviour towards individuals and society, which is reflected in principles such as **justification, proportionality and equality.** AI can increase the efficiency of institutions, yet on the other hand it can also erode the procedural legitimacy of and trust in democratic institutions and the authority of the law, if used inconsiderately.

## 7. Environment

Ever powerful and larger AI systems will consume ever more **energy** for training and operation. The impact of AI on the environment is of serious concern and should be mitigated.

## Why the AIA should include a mandatory Fundamental Rights, Democracy, Rule of Law and Environmental Impact Assessment both for Providers and Deployers?

With AI having a potential impact on virtually all fundamental rights and democracy, the rule of law and the environment, mandating a Fundamental Rights, Democracy, Rule of Law and Environmental Impact Assessment becomes pertinent.

As can be concluded from the above, adverse fundamental rights' impacts by AI can occur when AI is developed or used by **public as well as by private actors**. Moreover, while direct fundamental rights protection exists against violations by state actors, no such direct protection exists against violations by private actors, making a mandatory FRIA even more necessary for the latter, if the EU truly wants to live up to the objective of the AIA.

We recommend to oblige a FRIA **both for providers as well as for deployers** of AI systems. The impact on fundamental rights, democracy, the rule of law or the environment can arise both at the development stage as well as at during the use of a system. The argument that a provider cannot, or should not be held to, do a FRIA because the use of the system cannot be reasonably known, should thus not be followed.

We do realise that mandating a FRIA, including also the impact on democracy, the rule of law and the environment, involves an extra effort for organisations, many of which might not have the expertise or capacity to perform the assessment properly. Hence we also strongly advise the European Commission to provide for practical and financial measures of support.



# ABOUT ALLAI

ALLAI is an independent organisation that aims to foster, promote and achieve the responsible development, deployment and use of AI.

ALLAI's mission is to take a holistic approach to AI, taking into account all impact domains such as economics, ethics, privacy, laws, safety, labour, education, etc. ALLAI aims to involve all stakeholders in its mission: policy-makers, industry, social partners, consumers, NGOs, educational and care institutions, academics from various disciplines.

ALLAI was founded by the three Dutch members of the High Level Expert Group on AI, Catelijne Muller, LLM, Prof. Virginia Dignum and Prof. Aimee van Wynsberghe. Collectively, the founders have a broad expertise in AI: AI sciences, social impact, national and international policy, legal implications, and ethical impact.

## CONTACT



ALLAI  
Amsterdam Science Park  
1089 XH Amsterdam  
The Netherlands



[www.allai.nl](http://www.allai.nl)  
[@ALLAI\\_EU](https://twitter.com/ALLAI_EU)  
[welkom@allai.nl](mailto:welkom@allai.nl)

