INT/940

AI / Regulation

WORKING DOCUMENT

Section for the Single Market, Production and Consumption

Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts
[COM(2021) 206 final - 2021/106 (COD)]

For the attention of the study group members

Study group meeting	17/06/2021, 14:30 (video conferencing test starting at 14:00)
Contact	int@eesc.europa.eu
Administrator	Alice Tétu
Document date	../05/2021

Rapporteur: **Catelijne MULLER**

EN

Study Group	AI / Regulation
President	Maurizio MENSI (IT-III)
Rapporteur	Catelijne MULLER (NL-II)
Members	Pietro Francesco DE LOTTO (IT-I)
	Marie-Francoise GONDARD-ARGENTI (FR-I)
	Ileana IZVERNICEANU DE LA IGLESIA (ES-III)
	Veselin MITOV (BG-II)
	Justyna Kalina OCHĘDZAN (PL-III)
	Krzysztof OSTROWSKI (PL-I) (art. 66 - Lech PILAWSKI)
	Claudine OTTO (LU-I)
	Angelo PAGLIARA (IT-II)
	Franca SALIS-MADINIER (FR-II)
	Stoyan TCHOUKANOV (BG-III)
Expert	…, for the rapporteur
Referral	European Parliament, ../../..
	Council, ../../..
Legal basis	Article 114 of the Treaty on the Functioning of the European Union
Section responsible	Section for the Single Market, Production and Consumption
Adopted in section	DD/MM/YYYY
Adopted at plenary	DD/MM/YYYY
Plenary session No	…
Outcome of vote (for/against/abstentions)	…/…/…

1. Conclusions and recommendations

1.1 The **EESC welcomes** the fact that the Commission proposal for the Artificial Intelligence Act (the "AIA") not only addresses the risks associated with AI, but also **raises the bar substantially** as regards to the quality, performance, and trustworthiness of AI that the EU is willing to allow. The EESC is particularly pleased that the AIA puts **health, safety and fundamental rights at the center** of the AIA as well as with the global scope of the AIA.

1.2 The EESC sees **areas for improvement** regarding the scope and definition, the clarity of the prohibited AI practices, the implications of the categorisation choices made in relation to the 'risk-pyramid', the risk-mitigating effect of the requirements for high-risk AI, the enforceability of the AIA and the relation to existing regulation and other recent regulatory proposals.

1.3 The EESC stresses that **AI never operated in a lawless world**. Because of its wide scope as well as its primacy as an EU Regulation, the AIA could create tension with existing national and Community laws and related regulatory proposals at EU level. The EESC recommends the AIA to clarify this.

1.3 The EESC recommends to clarify the definition of AI by removing **ANNEX I** and slightly adjusting art. 3 and to to widen the scope of the AIA so as to **include 'legacy AI-systems'**.

1.4 The EESC recommends to **clarify the prohibitions** regarding 'subliminal techniques' and 'exploitation of vulnerabilities' and to also add 'harm to fundamental rights, democracy and the rule of law' as conditions for these prohibitions.

1.5 The EESC sees **no place in the EU for the scoring** of the trustworthiness of EU citizens based on their social behaviour or personality characteristics, **irrespective of the actor performing the scoring**. The EESC recommends to broaden the scope of this ban so as to include social scoring by private organisations and semi-public authorities. The EESC sees no added value in the sub-conditions attached to this prohibition.

1.6 The EESC **calls for a ban on any use of AI for automated recognition of human (biometric)** features, identities, emotions, behaviour, traits or characteristics in publicly and privately accessible spaces, such as recognition of faces, gait, fingerprints, DNA, voice, keystrokes, heart rate, temperature and other biometric or behavioural signals, in any context, as well as for the categorisation of individuals based thereon, except for very specified cases, such as some health purposes, where the patient emotion recognition is important.

1.7 The 'list-based' approach for high-risk AI runs the risk of **normalising and mainstreaming** a number of AI-systems and uses that are still heavily criticised. The EESC warns that compliance with the requirements set for medium and high-risk AI not necessarily mitigates the risks of harm to health, safety and fundamental rights for all high-risk AI. The EESC recommends the AIA to arrange for this situation. At the very least, the requirements of (i) human agency, (ii) privacy, (iii) diversity, non-discrimination and fairness, (iv) explainability (v) environmental and social well being of the **Ethics Guidelines for Trustworthy AI** should be added.

1.8 In line with its long advocated **'human-in-command'** approach to AI, the EESC strongly recommends the AIA to arrange for certain decisions to remain the **primacy of humans**, particularly in domains where these decisions have legal implications or a societal impact such as in the judiciary, law enforcement, social services, healthcare, housing, financial services, labour relations and education.

1.9 The EESC recommends to make **third party conformity assessments obligatory for all high-risk AI**.

1.10 The EESC recommends to include a **complaints and redress mechanism for organisations and citizens** that have suffered harm from any AI-system, practice or use that falls within the scope of the AIA.

2. Regulatory proposal on Artificial Intelligence - AI Act

2.1 The EESC welcomes the fact that the Commission proposal for the Artificial Intelligence Act (the "AIA") not only addresses the risks associated with AI, but also raises the bar substantially as regards to the quality, performance, and trustworthiness of AI that the EU is willing to allow.

3. Overarching comments AIA

Objective and scope

3.1 The EESC welcomes both the objective and the scope of the AIA. The EESC particularly welcomes the fact the Commission puts health, safety and fundamental rights at the center of the AIA. The EESC also welcomes the external effect of the AIA, making sure that AI that is developed outside of the EU, has to meet the same legal standards if deployed or having an impact within the EU.

Definition of AI

3.2 The definition of AI (art. 3 (1) jo. ANNEX I AIA) has raised the discussion among AI scientist that a number of the examples given in ANNEX I are not considered AI by AI scientists, and a number of important AI techniques are missing. The EESC sees no added value in ANNEX I and recommends to remove it from the AIA entirely. Moreover, the EESC recommends to adjust the definition of art. 3 (I) as follows:

> "'Artificial intelligence system' (AI system) means software that can, in an automated manner, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations or decisions, influencing the environment it interacts with."

Health, safety and fundamental rights - the Risk Pyramid

3.3 The escalating 'risk pyramid' (from low/medium risk, to high risk, to unacceptable risk) used to categorise a number of general AI practices and domain specific AI use cases acknowledges that not all AI poses risks and not all risks are equal or require the same mitigating measures

3.4 The chosen approach puts two important questions in front of us. First, do the mitigating measures (for high risk and low/medium risk AI) indeed sufficiently mitigate the risks of harm to health, safety and fundamental rights? Second, are we ready to allow AI to largely replace human decision making, even in critical processes such as law enforcement and the judiciary?

3.4 As to the first question, the EESC warns that compliance with the requirements set for medium and high risk AI not necessarily mitigates the risks of harm to health, safety and fundamental rights in all instances. This will be further elaborated on in Chapter 4.

3.5 As to the second question, what is missing form the AIA is the notion that the promise of AI lies in augmenting human decision making and human intelligence, rather than replacing it. The AIA works from the premise that, once the requirements for medium and high-risk AI are met, AI can largely replace human decision making.

3.6 The EESC strongly recommends the AIA to arrange for certain decisions to remain the primacy of humans, particularly in domains where these decisions have legal implications or a societal impact such as in the judiciary, law enforcement, social services, healthcare, housing, financial services, labour relations and education.

Relation to existing legislation and recent legislative proposals

3.7 The EESC stresses that AI never operated in a lawless world. Many national and Community regulation already applied (and applies) to the use of AI. Because of its wide scope and depth as well as its primacy as an EU Regulation, the AIA could create tension with these existing (and upcoming) national and Community laws.

4. Specific comments and recommendations AIA

Prohibited AI practices

4.1 The EESC agrees that the AI practices of art. 5 indeed have no social benefit and should be prohibited. The EESC however finds some wordings unclear, which could make some prohibitions difficult to interpret and easy to circumvent.

4.2 Evidence exists that **subliminal techniques** can not only lead to physical or psychological harm (the current conditions for this particular prohibition to kick in), but can, given the environment they are deployed in, lead to other adverse personal, societal or democratic effects, such as altered voting behaviour. Moreover, it is often not the subliminal technique itself, but rather the decision who to target with a subliminal technique that is AI driven.

4.3 In order to capture what the AIA aims to prohibit in art. 5.1 (a), which is nudging people into harmful behaviour without their knowing, the EESC recommends to alter the paragraph as follows: "*(…) an AI system deployed, aimed at or used for, subconsciously (e.g. with subliminal techniques) materially distorting a person's behaviour in a manner that causes or is likely to cause harm to that person's, another person's or group of persons' fundamental rights, including their physical or psychological health and*

safety, or to democracy and the rule of law." In the case of 'nudging' the right to human dignity (art. 1 EU Charter) and the freedom of expression and information (art. 11 EU Charter) are particularly vulnerable.

4.4 The EESC recommends to amend the prohibited practice of **exploitation of vulnerabilities** of article 5.1 (b) in the same manner so as to include harm to fundamental rights, including physical or psychological harm.

4.5 The EESC welcomes the prohibition of '**social scoring**' of article 5.1 (c), however sees no added value in the conditions under subparagraphs (i) and (ii). These conditions complicate the interpretation of the prohibition and create a number of loopholes.

4.6 The EESC recommends the prohibition of social scoring to also apply to private organisations and semi-public authorities, rather than just to public authorities. There that there is no place in the EU for scoring of the trustworthiness of EU citizens based on their social behaviour or personality characteristics, irrespective of the actor performing the scoring. If not, the EU would open the door to multiple area's where social scoring would be allowed, such as at the workplace.

4.7 The AIA aims to ban **real time remote biometric identification** (with for example facial recognition) for law enforcement and categories it as 'high risk' when used for other purposes. The limitation to 'identification' causes concern, whereas many biometrics recognition technologies are not aimed at identifying a person, but rather at assessing a person's behaviour or characteristics (e.g. by evaluating a person's facial features, expressions, eye movement, gate, temperature, hearth rate etc.).

4.8 These AI practices are commonly known under multiple definitions such as 'emotion recognition' or 'affect recognition' and more specifically as 'deception recognition' or 'behaviour recognition'. The AIA categorises 'emotion recognition' generally as low risk, with the exception of a few user domains where they are categorised as high risk. All types of biometrics recognition by AI are extremely invasive, lack any sound scientific basis and pose substantial risks of harm to a number of fundamental rights of the EU Charter, such as the right to human dignity, the right of integrity of the person (which includes mental integrity) and the right to a private life.

4.9 Broadly in line with the call of the EDPS and EDPR, the EESC calls for:
- A ban on any use of AI for automated recognition of human features in publicly **and privately** accessible spaces, such as recognition of faces, gait, fingerprints, DNA, voice, keystrokes and other biometric or behavioural signals, in any context.
- A ban on AI systems using biometrics to categorise individuals into clusters based on ethnicity, gender, political or sexual orientation, or other grounds on which discrimination is prohibited under Article 21 of the Charter.
- A ban on the use of AI to infer emotions, of a natural person, except for very specified cases, such as some health purposes, where the patient emotion recognition is important.

High Risk AI

4.10 In deciding whether an AI practice or use that poses a risk to health, safety or fundamental rights should nevertheless be allowed under strict conditions, the Commission looked at two elements: (i) whether the AI practice or use can have social benefits and (ii) whether the risks to health, safety and whether the risk of harm to health, safety and fundamental rights this use nevertheless poses, can be mitigated by meeting a number of requirements.

4.11 The EESC welcomes the alignement of these requirements with elements of the Ethics Guidelines for Trustworthy AI ("EGTAI"), however, five important EGTAI requirements are not specifically dealt with in the requirements for high-risk AI in the AIA namely: (i) human agency (ii) privacy (iii) diversity, non-discrimination and fairness, (iv) explainability and (v) environmental and social well being. The EESC feels that this is a missed opportunity, because many of the risks AI poses, are those of privacy, bias, exclusion, inexplicability of the outcomes of AI decisions, the undermining of human agency and the environment, and are all reflected in our fundamental rights.

4.12 The EESC recommends to add these requirements to the requirements of Chapter 2 of Title III of the AIA, to improve the ability of the AIA to effectively protect our health, safety and fundamental rights from adverse impact of AI.

4.13 The EESC welcomes the **'interwoven system' between the AIA and Union Harmonisation Legislation.** The EESC recommends to extend the scope of the AIA and the requirements for high risk AI beyond 'AI safety components' or the situation where the AI system is itself a product covered by Union harmonisation legislation listed in ANNEX II. This because AI not only can pose risks when used as safety components of these products, nor is the AI system itself always a product. When used for example as part of a diagnostic or prognostic tool in the medical field or an AI driven thermostat that regulates a boiler.

4.14 The EESC however warns that the chosen 'list-based' approach for high-risk AI of **ANNEX III** can lead to the **normalisation and mainstreaming** of quite a number of AI practices that are still heavily criticised and for which the societal benefits are questionable or lacking.

4.15 Moreover, the risks of harm to health, safety and fundamental rights cannot necessarily always be mitigated by compliance with the 6 requirements for high-risk AI, in particular when it comes to less mentioned fundamental rights that could be impacted by AI, such as the right to human dignity, the presumption of innocence, the right to fair and just working conditions, the freedom of association and assembly, to name a few.

4.16 The EESC strongly recommends to add the management and operation of the t**elecom and internet infrastructure** to point 2 of ANNEX III. The EESC also recommends to extend the scope of this point beyond AI safety components.

4.17 AI systems to **determine acces to education and evaluate students**, pose a number of risks of harm to student health, safety and fundamental rights. Online proctoring tools for example to supposedly flag 'suspicious behaviour' and 'indications of cheating' during online exams by using all kinds of biometrics and behaviour tracking, are truly invasive and lack scientific evidence.

4.18 The use of **AI systems for monitoring, tracking and the evaluation of workers** causes serious concerns as regards workers' fundamental rights to fair and just working conditions, to information and consultation and to justified dismissal. The addition of these AI systems to the high-risk list is likely to cause conflicts with national labour laws for (un)fair dismissal, healthy and safe working conditions and worker information.

4.19 The requirement of '**human oversight**' is particularly relevant in labour relations, because the oversight will be done by a worker or a group of workers. The EESC stresses that these workers should receive training on how to perform this task. Moreover, the fact that these workers are expected to be allowed to disregard the output of the AI-system or even decide not to use it, there should be measures in place to avoid the fear of negative consequences (such as demotion or dismissal) if such a decision is taken.

4.20 The **use of AI systems in relation to access and enjoyment of public services** is more broad than the use of AI systems in relation to access and enjoyment of essential private services, where for the latter, only credit(worthiness) scoring by AI is considered high risk. The EESC recommends to broaden the scope of point 5 (b) of ANNEX III to AI systems intended to evaluate the *eligibility* for essential private services.

4.21 AI used by **law enforcement authorities and in migration, asylum and border control management for making individual (criminal or security) risk assessments**, pose a risk of harm to the presumption of innocence, the right of defence and the right to asylum of the EU Charter. AI systems in general and merely seek correlations that are based on characteristics found in other 'cases'. Suspicion in these instances is not based on actual suspicion of a crime or misdemeanour by the particular person, but merely on characteristics that that person happens to share with other convicted criminals (such as address, income, nationality, debts, employment, behaviour, behaviour of friends and family members and so on).

4.22 **The use of AI in the administration of justice and democratic processes** is particularly sensitive and should be approached with more nuance and scrutiny than it is done now. Merely putting systems to assist a judicial authority in researching and interpreting facts and the law *and* in applying the law to a concrete set of facts overlooks the fact that judging is so much more than finding patterns in historical data (which is in essence what current AI systems do). The text also assumes that these types of AI will only assist the judiciary, while leaving fully automated judicial decision making out of scope. The EESC also misses the mention of AI systems or uses in the realm of democratic processes, such as elections.

4.23 The EESC recommends to add a provision that arranges for the situation where it is either obvious or became clear during the prior conformity assessment that the 6 requirements will not sufficiently mitigate the risk of harm to health, safety and human rights (for example by amending art. 16 (g) AIA).

Governance and Enforceability

4.24 The EESC welcomes the governance structure set up by the AIA. The EESC recommends the AI Board to hold regular **obligatory exchanges with wider society**, including the social partners and NGO's.

4.25 The EESC strongly recommends to widen the scope of the AIA so as to include **'legacy AI-systems'** i.e. systems that are already in use or are deployed prior to the coming into force of the AIA, in order to avoid deployers to fast track any prohibited, high and medium risk AI to avoid compliance requirements.

4.26 The complexity of the requirements and accountability activities plus the self-assessment runs the risk of simplifying this process into check lists where a simple 'yes' or 'no' could suffice to meet the requirements. The EESC recommends to make **third party assessments obligatory for all high-risk AI**.

4.27 The EESC recommends to have appropriate (financial) support measures in place for micro and small organisations as well as civil society organisations, to be able to understand the purpose and meaning of the AIA, as well as to be able to meet the requirements of the AIA. These measures should go beyond supporting Digital Innovation Hubs, and consist of the facilitating of access to high level expertise regarding the AIA, its requirements, its obligations, and particularly the reasoning behind these.

4.28 The EESC recommends to include a **complaints and redress mechanism for organisations and citizens** that have suffered harm from any AI-system, practice or use that falls within the scope of the AIA.

_____