



Artificial Intelligence Act Amendments

Risk designation and impact assessments: flexibility mechanisms that support innovation and fundamental rights

The proposed AI Act (AIA) has adopted a strictly risk-based and purpose-based approach, with risk categories identified *ex ante* on the basis of specific criteria listed in the Act and related Annexes.

The AIA currently empowers the EU Commission to add new “high-risk” AI systems to the list in Annex III in the future (Article 7), but only as long as such systems are intended to be used in one of the areas already identified by the Act, thus precluding the possibility to include new areas in the future, based on emerging challenges. The AIA also glosses over the review mechanism that the Commission would carry out, referring only to a list of evidence to consider (Article 7, para 2), but omitting to clarify how such evidence would be compiled, collected, and made available to provide a solid impact assessment. On top of that, the AIA calls on the Commission to “assess the need” to amend the high-risk list only once year (Article 84, which the Slovenian Presidency of the Council of the EU’s compromise text would like to change to every two years). Furthermore, the AIA is **silent on the need to consult external stakeholders in this process**, who should have a clear role in **flagging potential adverse impacts or risks**.

The crucial flaw of the current AIA approach is that it **overlooks the complexity of AI systems** and the concrete possibility that when such AI systems are researched, developed and deployed:

- They may also have an intended general or multi-purpose, not just a specific one (e.g.: Multi-Agent AI – MAAI – predictive modelling systems that simulate societies in specific situations such as humanitarian/climate disasters, traffic accidents, crowding, etc. in order to formulate predictions in different scenarios);
- They may have significantly different consequences depending on the specific context and/or environment in which they are used and the individuals or groups affected, regardless of the specific purpose for which they are intended.

Moreover, the **current AIA provisions fail to provide enough flexibility to respond to emerging challenges**. Because of this difficulty of assessing the full scope of risk *ex ante*, as well as the fast-moving nature of AI development, **the risk-based approach taken by the AIA needs to be flexible and easily updatable** so that it can adapt to new applications. The AIA should provide a **more flexible, future proof manner and assess impact of AI systems on rights, people and society on a rolling basis**.



How the AIA should be reviewed and amended

We call on the AIA to facilitate a **shared responsibility on both providers and users to assess impact of AI on a rolling basis to help with flexible, proportionate and agile check-and-balance process**. While some of the risks posed by the AI systems come from how they are designed, significant risks can also stem from how they are used and in which context they are deployed.

- **On the provider's side:** all providers of other AI systems not currently identified or classified as high-risk should carry out a *basic assessment* of whether the system's development or use is likely to result in a high risk to fundamental rights and freedoms. **Providers of AI high-risk systems should conduct a specific impact assessment** within the conformity assessment obligations of the AIA and based on the assessment criteria already envisaged in the AIA.
- **On the user's side:** a mirroring obligation would **require users of AI systems that are not previously identified or classified as high-risk** to carry out a *basic assessment* of whether the system's development or use is likely to result in a high risk to fundamental rights and freedoms. Where that is found to be the case, then these providers should also conduct a *specific impact assessment* as indicated by the AIA. On the other hand, **users of high-risk AI systems as identified and classified in the AIA should conduct a specific impact assessment**.

In the amendments below, **we propose overall criteria for conducting an impact assessment** (in different articles, due to the current structure of the AIA, to cover different types of AI system providers and users). **Based on these overall criteria, we propose the Commission to develop and adopt a concrete and standardised methodology for an impact assessment process** (following a proposal by the European Artificial Intelligence Board). Such methodology should consist of a) *basic module* (for basic, checklist assessment for all providers and users except high-risk AI) and *detailed module*, for high-risk AI providers (as part of the conformity assessment) and users. This methodology should also include rules of engagement for consultation with external stakeholders and publication of key assessment findings.

Why this is doable

Industry and private sector are already working on standardization of impact assessment, risk assessment and certification standards and processes in the design phase (namely, [ISO](#), [IEEE](#)). Facebook and Open Loop conducted a pilot process with 10 EU-based AI start-ups to test the feasibility of an impact assessment framework and developed policy recommendations towards requiring clear legal background for impact assessment along with robust guidance for conducting it (see results and recommendations [here](#)). **A basic impact assessment exercise should not incur onerous costs for providers** – unless they are designing AI that is in fact high risk. This approach is supportive of innovation and trust-building at the same time, as it gives providers an early orientation about potential risks and impacts that could be unacceptable in further deployment process. For the users, **conducting a fundamental rights impact assessment in advance of deployment helps with accountability and potential liability for the use of AI system**. In addition, a number of users, especially the public sector, already have an obligation to conduct Data Protection Impact Assessment under GDPR. AI impact assessment would be an add on to such process, covering important broader scope. Member States have already demonstrated willingness to implement such obligations on users of AI systems. For example, in the Netherlands the Ministry of

Interior and Kingdom Relations has developed the [Impact Assessment Mensenrechten en Algoritmes](#) (IAMA) for users.

Proposed amendments to the AIA

The following amendments are designed to include **obligations on all providers and all users of AI systems, proportionate to the designated risk level, to ensure achieving trustworthy, flexible, future proof and innovation friendly regulation that provides for transparency and accountability to those affected by the use of AI systems.** The obligations outlined in the amendments would require providers and users to delineate the impacts of high risk AI systems, publish the findings and therefore create a crucial tool of flexibility and accountability over how AI systems are developed and deployed. This would also assist in striking a **balance between the aspiration to embrace new AI uses and the minimization of the cost of review.**

<p>Article 7 Amendments to Annex III</p> <p>1. The Commission is empowered to adopt delegated acts in accordance with Article 73 to update the list in Annex III by adding high-risk AI systems where both of the following conditions are fulfilled:</p> <ul style="list-style-type: none">(a) the AI systems are intended to be used in any of the areas listed in points 1 to 8 of Annex III;(b) the AI systems pose a risk of harm to the health and safety, or a risk of adverse impact on fundamental rights, that is, in respect of its severity and probability of occurrence, equivalent to or greater than the risk of harm or of adverse impact posed by the high-risk AI systems already referred to in Annex III. <p>2. When assessing for the purposes of paragraph 1 whether an AI system poses a risk of harm to health and safety or a risk of adverse impact on fundamental rights that is equivalent to or greater than the risk of harm posed by the high-risk AI systems already referred to in Annex III, the Commission shall take into account the following criteria:</p> <ul style="list-style-type: none">(a) the intended purpose of the AI system;(b) the extent to which an AI system has been used or is likely to be used;(c) the extent to which the use of an AI system has already caused harm to the health and safety or adverse impact on the fundamental rights or has given rise to significant concerns in relation to the materialisation of such harm or adverse impact, as demonstrated by reports or	<p>Article 7 Amendments to Annex III - <i>added text</i></p> <p>1. The Commission is empowered to adopt delegated acts in accordance with Article 73 to update the list in Annex III by adding high-risk AI systems where both of the following conditions are fulfilled:</p> <ul style="list-style-type: none">(a) the AI systems are intended to be used in any of the areas listed in points 1 to 8 of Annex III;(b) the AI systems pose a risk of harm to the health and safety, or a risk of adverse impact on fundamental rights, that is, in respect of its severity and probability of occurrence, equivalent to or greater than the risk of harm or of adverse impact posed by the high-risk AI systems already referred to in Annex III. Where an AI system is not intended to be used in any of the areas listed in points 1 to 8 of Annex III, the Commission is empowered to update the list of areas in Annex III by including new areas or extending the scope of existing areas. <p>2. When assessing for the purposes of paragraph 1 whether an AI system poses a risk of harm to health and safety or a risk of adverse impact on fundamental rights that is equivalent to or greater than the risk of harm posed by the high-risk AI systems already referred to in Annex III, the Commission shall take into account the following non-cumulative criteria:</p> <ul style="list-style-type: none">(a) the intended purpose of the AI system;(b) the extent to which an AI system has been used or is likely to be used;
--	---

<p>documented allegations submitted to national competent authorities;</p> <p>(d) the potential extent of such harm or such adverse impact, in particular in terms of its intensity and its ability to affect a plurality of persons;</p> <p>(e) the extent to which potentially harmed or adversely impacted persons are dependent on the outcome produced an AI system, in particular where for practical or legal reasons it is not reasonably possible to opt-out from that outcome;</p> <p>(f) the extent to which potentially harmed or adversely impacted persons are in a vulnerable position in relation to the user of an AI system, in particular due to an imbalance of power, knowledge, economic or social circumstances, or age;</p> <p>(g) the extent to which the outcome produced with an AI system is easily reversible, whereby outcomes having an impact on the health or safety of persons shall not be considered as easily reversible;</p> <p>(h) the extent to which existing Union legislation provides for:</p> <p>(i) effective measures of redress in relation to the risks posed by an AI system, with the exclusion of claims for damages;</p> <p>(ii) effective measures to prevent or substantially minimise those risks.</p>	<p>(e) the extent to which the use of an AI system has already caused harm to the health and safety or adverse impact on the fundamental rights or has given rise to significant concerns in relation to the materialisation of such harm or adverse impact, as demonstrated by reports or documented allegations made available-submitted to national competent authorities;</p> <p>(d) the potential extent of such harm or such adverse impact, in particular in terms of its intensity and its ability to affect a plurality of persons or to affect a particular group of persons disproportionately;</p> <p>(e) the extent to which potentially harmed or adversely impacted persons are dependent on the outcome produced with by a process involving an AI system, in particular where for practical or legal reasons it is not reasonably possible to opt-out of from that outcome;</p> <p>(f) the extent to which potentially harmed or adversely impacted persons are in a vulnerable position in relation to the user of an AI system, in particular due to an imbalance of power, knowledge, economic or social circumstances, or age;</p> <p>(g) the extent to which the outcome produced with an AI system is not easily reversible, whereby outcomes having an impact on the health or safety of persons or on their fundamental rights shall not be considered as easily reversible;</p> <p>a. the extent to which existing Union legislation lacksprovides for:</p> <p>i. effective measures of redress in relation to the risks posed by an AI system, with the exclusion of claims for damages;</p> <p>ii. effective measures to prevent or substantially minimise those risks.</p> <p>Para 3 (new) The Commission, following a proposal by the European Artificial Intelligence Board , will develop and adopt a methodology for an assessment process from the Article 7 (2) within months from the entry into force of this Regulation. The methodology will include rules of engagement for consultation with external</p>
--	--

	<p>stakeholders and publication of key assessment findings by the Commission.</p>
<p>Article 28 Obligations of distributors, importers, users or any other third-party</p>	<p>Article 28a (new) Obligation on all users to define affected persons</p> <p>Before putting into use an AI system the user shall define categories of natural persons and groups likely to be affected by the use of the system.</p> <hr/> <p>Article 28b (new) Fundamental rights impact assessments for AI systems</p> <ol style="list-style-type: none"> 1. All users of AI systems shall conduct an assessment of the systems' impact in the context of use before putting the system into use. This assessment shall include, but is not limited to, the following criteria: <ol style="list-style-type: none"> (a) a clear outline of the intended purpose for which the system will be used; (b) a clear outline of the intended geographic and temporal scope of the system's use; (c) verification of the legality of the system in accordance with Union and national law, fundamental rights law, Union accessibility legislation, and the extent to which the system is in compliance with this Regulation; (d) an assessment of the likely impact on fundamental rights of the AI system, including any indirect impacts or consequences of the systems use; (e) any specific risk of harm likely to impact marginalized, vulnerable persons or those groups at risk of discrimination, or increase existing societal inequalities; (f) the foreseeable impact of the use of the system on the environment, including but not limited to energy consumption; (g) any other negative impact on the public interest; and (h) clear steps as to how the harms identified will be mitigated, and how effective this mitigation is likely to be. 2. If adequate steps to mitigate the risks outlined in the course of the assessment in paragraph 1 cannot be identified, the system shall not be put into use. Market surveillance authorities, pursuant to their capacity under Articles 65 and 67, make take this information into account when investigating systems which present a risk at national level.

	<ol style="list-style-type: none"> 3. The obligation outlined under paragraph 1 applies for each new deployment of the AI system. 4. In the course of the impact assessment, the user should notify relevant national authorities and relevant stakeholders, including but not limited to: equality bodies, consumer protection agencies, social partners and data protection agencies, with a view to receiving input into the impact assessment. 5. Where, following the impact assessment process, the user decides to put the AI system into use, the user shall be required to publish the results of the impact assessment as part of the registration of use pursuant to their obligation under Article 51(2).¹ 6. Where the user is already required to carry out a data protection impact assessment under Article 35 of Regulation (EU) 2016/679 or Article 27 of Directive (EU) 2016/680, the impact assessment outlined in paragraph 1 shall be conducted in conjunction to the data protection impact assessment and be published as an addendum. 7. Users of AI systems shall use the information provided under Article 13 to comply with their obligation under paragraph 1. 8. Where the user, pursuant to their obligation to define affected categories of persons under Article 28a, finds that use of an AI system poses a particular risk to a specific group of natural persons, the user has the obligation to notify established representatives or interest groups acting on behalf of those persons before putting the system into use, with a view to receiving input into the impact assessment. 9. The obligations on users in paragraph 1 is without prejudice to the obligations on users of all high risk AI systems as outlined in Article 29.
<p>Article 43 Conformity Assessment</p>	<p>Article 43, para 7 (<i>new</i>)</p> <ol style="list-style-type: none"> 1. As part of the conformity assessment, high-risk AI system providers shall conduct an assessment of the systems' impact. This shall include, but is not limited to, the following:

¹ This obligation is the subject of an amendment outlined further in the amendment paper 'Ensure consistent and meaningful public transparency' prepared by AlgorithmWatch. This amendment, quoted below, obliges users to register the use of high risk AI systems.

	<ul style="list-style-type: none"> (a) a clear outline of the intended purpose for which the system will be used; (b) verification of the legality of the system in accordance with Union and national law, fundamental rights law, Union accessibility legislation, and the extent to which the system is in compliance with this Regulation; (c) an assessment of the likely impact on fundamental rights of the AI system, including any indirect impacts or consequences of the systems intended purpose; (d) any specific risk of harm likely to impact marginalized, vulnerable persons or those groups at risk of discrimination, or increase existing societal inequalities; (e) the foreseeable impact of the use of the system on the environment, including but not limited to energy consumption; (f) any other negative impact on the public interest; and (g) clear steps as to how the harms identified will be mitigated, and how effective this mitigation is likely to be. <ol style="list-style-type: none"> 2. If adequate steps to mitigate the risks outlined in the course of the assessment of impact cannot be identified, the notified body will not issue a certificate under Article 44. Market surveillance authorities, pursuant to their capacity under Articles 65 and 67, make take this information into account when investigating systems which present a risk at national level. 3. In the course of the impact assessment, the provider should notify relevant national authorities and relevant stakeholders, including but not limited to: equality bodies, consumer protection agencies, social partners and data protection agencies, with a view to receiving input into the impact assessment. 4. The findings of the impact assessment from this Article shall be made public by the provider.
<p>TITLE IV – TRANSPARENCY OBLIGATIONS FOR CERTAIN AI SYSTEMS</p> <p>Article 52 Transparency obligation for certain AI systems</p>	<p>TITLE IV B – TRANSPARENCY OBLIGATIONS FOR CERTAIN AI SYSTEMS</p> <p>Article 52 – added para 5 Transparency obligation for certain AI systems</p>

	<p>5. AI system providers and users shall make public the findings of an assessment of the systems' impact. This assessment shall include, but is not limited to, the following:</p> <ul style="list-style-type: none"> (a) a clear outline of the intended purpose for which the system will be used; (b) verification of the legality of the system in accordance with Union and national law, fundamental rights law, Union accessibility legislation, and the extent to which the system is in compliance with this Regulation; (c) an assessment of the likely impact on fundamental rights of the AI system, including any indirect impacts or consequences of the systems intended purpose; (d) any specific risk of harm likely to impact marginalized, vulnerable persons or those groups at risk of discrimination, or increase existing societal inequalities; (e) the foreseeable impact of the use of the system on the environment, including but not limited to energy consumption; (f) any other negative impact on the public interest; and (g) clear steps as to how the harms identified will be mitigated, and how effective this mitigation is likely to be. <p>2. The obligations on providers in paragraph 1 are without prejudice to the obligations on providers of all high-risk AI systems as outlined in Article 43.7.</p>
<p>Article 58d</p> <p>When providing advice and assistance to the Commission in the context of Article 56(2), the Board shall in particular: [...] (d) issue an advisory opinion on the need for amendment of Annex I and Annex III, including in light of available evidence.</p>	<p>Article 58(d) - <i>added text</i> Tasks of the Board</p> <p>When providing advice and assistance to the Commission in the context of Article 56(2), the Board shall in particular: [...] (d) issue an advisory opinion on the need for amendment of Annex I and Annex III, including in light of available evidence including findings of fundamental rights impact assessment conducted by the provider.</p>
<p>Article 84</p>	<p>Article 84 para (1) c (<i>new</i>)</p> <p>The Commission shall assess the need for amendment of the list in Annex III based on the request of the Board on a rolling basis, when</p>

<p>Article 84 (para 6) In carrying out the evaluations and reviews referred to in paragraphs 1 to 4 the Commission shall take into account the positions and findings of the Board, of the European Parliament, of the Council, and of other relevant bodies or sources.</p> <p>Article 84 (para 7) The Commission shall, if necessary, submit appropriate proposals to amend this Regulation, in particular taking into account developments in technology and in the light of the state of progress in the information society.</p>	<p>supported with evidence of potential adverse impact, within 8 weeks of the request submission.</p> <p>Article 84 para (6) - added text In carrying out the evaluations and reviews referred to in paragraphs 1 to 4 the Commission shall take into account the positions and findings of the Board, of the European Parliament, of the Council, and of equality bodies and other relevant bodies or sources, and shall consult relevant external stakeholders, in particular those potentially affected by the AI system, as well as stakeholders from academia and civil society.</p> <p>Article 84 para (7) - added text The Commission shall, if necessary, submit appropriate proposals to amend this Regulation, in particular taking into account developments in technology, the effect of AI systems on fundamental rights, equality, and accessibility for persons with disabilities, and in the light of the state of progress in the information society.</p>
---	---