European Commission adoption consultation: Artificial Intelligence Act

Brussels, 3 August 2021

European Digital Rights (EDRi) outlines the following analysis of the Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) 2021/0106(COD).[1] *This input is intended for the European Commission consultation on the adoption of the Artificial Intelligence Act. It builds on previous EDRi positions on the EU's approach to artificial intelligence regulation, including 'Recommendations for a Fundamental rights-based artificial intelligence regulation'*[2] and *'Ban Biometric Mass Surveillance'.*[3]

Section A summarises an initial analysis of the Artificial Intelligence Act, and Section B begins to chart recommendations for improvement and adaption to ensure fundamental rights are duly protected:

A) Analysis: Artificial Intelligence Act
B) Recommendations for a fundamental rights-based Artificial Intelligence Act

EDRi will publish a full response to the Artificial Intelligence Act in autumn 2021, outlining the network's recommendations toward the European Parliament and the Council of the EU.

1 European Commission (2021), Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act): https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206.
2 EDRi (2020), Recommendations for a Fundamental rights-based artificial intelligence regulation: https://edri.org/wp-content/uploads/2020/06/AI_EDRiRecommendations.pdf.
3 EDRi (2020), Ban Biometric Mass Surveillance: https://edri.org/wp-content/uploads/2020/05/Paper-Ban-Biometric-Mass-Surveillance.pdf; for more EDRi resources on artificial intelligence, consult the EDRI AI and fundamental rights document pool: https://edri.org/our-work/artificial-intelligence-and-fundamental-rights-document-pool/.

Summary of recommendations:

EDRi recommends that the European Parliament and the Council of the EU implement the following improvements to the Artificial Intelligence Act (AIA). For the full recommendations, see section B.

1. **Ensure effective protection against prohibited practices and address the full scope of unacceptable risks through AI:**

    a. Strengthen existing prohibitions in article 5 to provide meaningful protection against fundamental rights violations and individual and collective harms;
    b. Comprehensively prohibit the use of remote biometric identification in publicly accessible spaces for any purpose, and implement a general ban on any use of AI for an automated recognition of human features in publicly accessible spaces;
    c. Include new prohibitions on the following practices which are incompatible with fundamental rights and democracy, and thus pose an unacceptable risk:
        i. Uses of AI in the field of law enforcement or criminal justice that purport to predict future behaviour;
        ii. Uses of AI in the field of migration control in ways that undermine the right to claim asylum;
        iii. Uses of AI that implement invasive surveillance, monitoring and algorithmic management in an employment and educational context;
        iv. The use of AI to categorise people on the basis of their human features, which can pose a grave and disproportionate threat to all human rights, in particular equality and non-discrimination;
        v. Placing on the market, putting into service or use of AI to infer, predict, analyse or assess a person's emotions, feelings, emotional state, beliefs, preferences, intentions or otherwise inner thoughts, as well as to use human features, behaviours or expressions to predict future actions or behaviours;
        vi. Uses of AI that constitute mass surveillance.

2. **Adapt the AIA to ensure a holistic, democratic and 'future-proof' framework:**

    a. Introduce a democratic, inclusive and accessible process by for the insertion of new prohibitions. Include criteria for 'unacceptable risk' and the addition of future prohibitions into the AIA;
    b. Ensure the potential to update the high risk use case 'areas' in the future (amending article 7) in addition to updating the use case 'sub-areas';

c. Respond to gaps in regulation with respect to economic and environmental impact, structural forms of inequality and AI and migration control, law enforcement and worker surveillance, mass surveillance, and exports of high-risk or prohibited AI outside the EU;
d. Remove loopholes in articles 2(4) and 83 which currently leave out of scope of the AIA AI systems used as part of international law enforcement agreements;
e. Remove the broad exemption to forgo the duty to conduct a conformity assessment on grounds of public security in article 47;
f. Remove the exemption to the principle of purpose limitation contained in article 54(1)(a) for 'innovative AI' within the regulatory sandbox provisions for uses in the criminal justice context.

3. Ensure responsibility to those subjected to AI systems with enhanced obligations on users of all AI systems:

a. Mandate users to conduct and publish an *ex ante* human rights impact assessment before putting a high risk AI system into use, clearly outlining the stated purpose for which the system will be implemented;
b. Implement on users a duty to cooperate with national competent authorities investigating AI systems for potential threats to fundamental rights or safety under articles 65 and 67 for all AI systems, regardless of risk designation;
c. Implement a duty on users to meaningfully consult with institutions, civil society and social partners representing affected groups before deploying high risk AI systems;
d. When the user of any AI system is a public authority, implement a notification requirement to all those impacted by a decision made by the system.

4. Implement meaningful public transparency for high risk AI systems:

a. Ensure meaningful public transparency by requiring registration in the EU database (article 60) of all high risk AI systems (and potentially also all AI systems to which people are subject) that are put into use. This would enable individuals and civil society to access information about AI systems in operation;
b. Ensure the inclusion of 'instructions for use' for AI systems in law enforcement and migration, asylum and border control management in the public database as per Annex III, points 1, 6 and 7. Remove the exemption contained within Annex VIII, point 11;
c. Require providers to include access to the conformity assessment alongside the instructions for use as per article 13(2)-(3) in the public database under article 60;
d. Require providers to provide more thorough details about the system to the users as part of article 13(3);

e. Remove the exemptions in article 52 relating to the transparency of AI systems used for detection and prevention of criminal offences, (as argued by the EDPB and EDPS) and for the prosecution of people. When AI systems under article 52 are used for investigation, suspects should be notified 'post factum'.

5. **Facilitate accountability: Include oversight and enforcement infrastructures that work for people:**

   a. Ensure a cohesive national enforcement structure;
   b. Include flagging and redress mechanisms allowing individuals and collectives to contest and seek redress for all AI systems that cause harm and threaten fundamental rights;
   c. Implement a more democratic governance infrastructure, with greater independence for the European AI board.

EDRi is the biggest European network defending rights and freedoms online. The EDRi network is a dynamic and resilient collective of 45 NGOs, as well as experts, advocates, and academics working to defend and advance digital rights across Europe and beyond.

Together, the EDRi network builds a movement of organisations and individuals pushing for robust and enforced laws, informing and mobilising people, and promoting a healthy and accountable technology market.

Acknowledgments

With thanks to the EDRi network for their contributions. In particular:

Access Now
ARTICLE 19
Bits of Freedom
Chaos Computer Club (CCC)
Digitale Gesellschaft Schweiz
Free Software Foundation Europe (FSFE)
Panoptykon Foundation
Electronic Frontier Norway (EFN)
Electronic Privacy Information Center (EPIC)
epicenter.works
Homo Digitalis
IT-Political Association of Denmark (IT-POL)
Statewatch

Thanks also to conversation partners: Amos Toh, Fieke Jansen, Jill Toh, Griff Ferris, Jeremias Adams-Prassl, Reuben Binns, Aislinn Kelly-Lyth, Petra Molnar, Alyna Smith, Mher Hakobyan, Marlena Wisniak, Agathe Balayn, Seda Gürses, Mute Schimpf, Jascha Galaski, Mark Brakel and members of the Digital Dignity Coalition for their thoughtful insights.

(A) Analysis: Artificial Intelligence Act

EDRi welcomes the European Commission's globally significant step towards regulating the development and deployment of artificial intelligence (AI) systems. Uses of AI systems have the ability to enable mass surveillance and intrusion into our personal lives, reinforce some of the deepest societal inequalities, fundamentally alter the delivery of public and essential services, shift more power into corporate hands and disrupt the democratic purpose. The proposal thus takes a notable step to acknowledge that some uses of AI are simply unacceptable and must be prohibited. However, we would like to make a number of suggestions to ensure that the AIA is in line with the Charter of Fundamental Rights of the EU (CFEU), 'future proof', and a role model for other rights-protective future AI legislation around the world.

Overall, EDRi raises a number of concerns relating to the AIA as a regulatory framework, specifically in relation to the extent to which it protects fundamental rights and is able to address broader structural, political and economic issues as a result of the widescale promotion and adoption of AI systems in various areas of life. We have argued that any approach which assumes benefits from the widescale uptake of AI will be problematic from a fundamental rights perspective and should not be a policy objective in itself.[4] Further, we highlight throughout that there are structural concerns relating to the extent to which the use of AI systems can systematically target, harm and exclude marginalised communities, exacerbating existing power imbalances in society.

Additionally, there are broad questions as to how far the AIA as a framework is sufficiently comprehensive to address these structural harms, due to its tendency toward de-regulation of all but the most narrowly-defined 'unacceptable' of AI systems, the lack of obligations on users and the lack of provisions for individual or collective redress for those subjected to AI systems.

As such, the following outlines EDRi's main analysis of the proposal for an Artificial Intelligence Act.

1 Inconsistencies with stated objectives of the AIA

The Explanatory Memorandum to the Commission's proposal establishes that the primary purpose of the AIA is "to implement the second objective [of the Commission's White Paper on AI] for the development of an ecosystem of

4 EDRi (2020), Recommendations for a Fundamental rights-based artificial intelligence regulation: https://edri.org/wp-content/uploads/2020/06/AI_EDRiRecommendations.pdf.

trust" by "addressing the risks associated with certain uses of such technology" "based on EU values and fundamental rights" (1.1).
The Memorandum continues that this is separate from the Commission's aim of "promoting the uptake of AI" (1.1) which – while important - is not the main goal of this Act. The four specific objectives which the Memorandum describes are also aligned with the ambition of trustworthy AI: safety and respect for fundamental rights; legal certainty; enhanced governance and enforcement of safety and fundamental rights; and a "lawful, safe and trustworthy" single market (1.1).

Despite this reassurance, the AIA proposal seems, at its core, designed to enable AI uptake rather than to limit or mitigate its harms. The fact that the vast majority of rules apply only to the narrowest sub-set of "high risk AI" - which the Commission explicitly admits is a "minimum necessary" approach (1.1) - is at odds with EU fundamental rights obligations. It also contradicts the "precautionary principle", which civil society has warned is necessary, given the vast contextual harms which may arise from the use of AI systems to which people are subject.

Furthermore, the Memorandum describes that the proposed rules cover "the development, placement on the market *and use* of AI systems [italics for emphasis]" (1.1) when in fact, the use of AI receives insufficient attention in the proposal. Even the "prohibition" of certain forms of remote biometric identification (RBI) under article 5 is recognised in the Memorandum as not being a real ban: the RBI rules, the Commission explains, constitute only "specific restrictions and safeguards" (1.1).

Whilst the AIA seeks to classify the risk level of an AI system "based on the intended purpose of the AI system" (1.2), it has created a set of rules and obligations which are largely unable to achieve this aim. As critics are increasingly pointing out, the AIA proposal attempts to transplant a typical product safety framework into an often novel AI context. By failing to account for the specificities of artificial intelligence, for example the variety of ways in which it can be applied, the importance of context, the fact that it is often sold as a service (not a product), and its self-learning nature, the proposal falls short of what would be needed to anticipate, prevent or at the very least mitigate the myriad ways in which it can cause harm.

There is a broader concern as to the extent to which the AIA's primary objective of harmonising the single market for AI 'products' is compatible with the other objective of safeguarding fundamental rights, and the broader need to mitigate the societal impacts of AI. The promotion of AI's uptake and the push for a harmonised single market, via the act's 'maximum harmonisation' function, may preclude Member States from introducing higher fundamental rights standards than those contained in the AIA.[5] Despite the protection of

5 Veale and Zuiderveen Borgesius (2021), 'Demystifying the Draft EU Artificial Intelligence Act': https://arxiv.org/abs/2107.03721.

personal data being one of the treaty bases of the proposal, it is clear that the proposal goes nowhere near far enough to ensure the protection of fundamental rights, and in doing so, contradicts its own stated aims and objectives to ensure trustworthy AI.

2 Prohibited practices: Incomplete coverage of 'unacceptable' AI and fundamental rights threats

Whilst it is positive that the AIA proposal foresees that some uses of artificial intelligence pose an unacceptable risk to fundamental rights and therefore must be prohibited under article 5, the AIA's approach to unacceptable risks falls short in two main ways.

2.1 Proposed prohibitions are too wide and vague

Firstly, **article 5 leaves a wide scope for interpretation, broad exceptions and in some cases unreasonably high thresholds** for systems to be prohibited. As such, there is a risk that this provision fails to prevent the worst excesses of potential fundamental rights abuses arising from AI systems. Those shortcomings are:

Subliminal and exploitative uses

- **Physical or psychological harm:** Articles 5(1)(a) and (b) are unduly narrow leading to significant concerns that they will fail to prevent against manipulative or exploitative uses of AI. In particular, that both prohibitions only apply when there is or likely to be 'physical or pyschological harm' foresees a burden of proof on invididuals to demonstrate future or actual harm (without creating a legal path to flag or contest such systems, see section 5 below). Whilst the background to the AIA acknowledges the opacity and unpredictability of AI systems, this provision does not incorporate these concerns into the drafting of this provision. The requirement that the use is in order to 'materially distort' behaviour adds an unreasonably high threshold;
- **Individual harm:** Both provisions are drafted in narrow, individualistic terms, not foreseeing that many such systems are unlikely to target specific persons, but rather whole groups of people in society;
- **Limited vulnerabilities:** Article 5(1)(b) attempts to prevent only such uses of AI that may exploit people on the basis of specific vulnerabilities – age and physical or mental disability. It is unclear why the AIA limits only to these vulnerabilities rather than taking a comprehensive approach and prohibiting uses of AI that exploit vulnerabilities based on the full range of protected characteristics under EU law.

Social scoring

- **Limitation to public uses:** The prohibition of social scoring systems contains a number of limitations suggesting an extremely high threshold for its application. Firstly the prohibition is limited to uses in a public context, by public authorities or on their behalf, thus excluding commercial uses, such as scoring of customers on online platforms leading to different service options;
- **Trustworthiness**: The provision is limited to those systems which evaluate or classify trustworthiness, without providing a definition of trustworthiness in the act. This could be potentially limiting for those systems that have a parallel impact but do not purport to map trustworthiness *per se;*
- **General purpose score:** The implicit grounding of the prohibition in the notion of a single score to be used for 'general purposes' (as indicated in Recital 17) suggests that many of the examples of risk scoring used in specific governmental contexts (such as risk-scoring for welfare fraud in the notorious Dutch SyRI case) are to be excluded from the scope of the prohibition, unless they deploy data collected in one context to be used in another or have an 'unjustified or disproportionate impact';
- **Temporal limit:** The provision also includes a temporaral limitation, thus applying to systems which evaluate or classify the trustworthiness of natural persons 'over a certain period of time', another limiting threshold;
- **Added conditions**: The prohibition is limited to uses which lead to detrimental or unfavourable treatment in contexts other than that in which the data was collected, or in a nature that causes 'unjustified or disproportionate' harm. Such conditions suggest that the central principle is not the harm caused, otherwise these conditions would not be relevant. Arguably, there are reasons to contest social scoring systems regardless of the presence of proof of unfavourable outcomes insofar as they reduce the complexity of human experience to a combination of limited, measurable indicators, with potential negative implications for fundamental rights to good administration and human dignity.

Remote biometric identification

The AIA's 'prohibition' of *real-time remote biometric identification (RBI) in publicly accessible spaces for the purpose of law enforcement* addresses only a small range of the many practices that can constitute biometric mass surveillance.[6] As we will recommend further in Section B, **the AI Act must be amended to ensure it does not undermine existing fundamental rights standards;** furthermore all **remote biometric identification and the use of AI for the automated recognition of human features must be prohibited without exceptions**.

6 EDRi (2020), 'Ban Biometric Mass Surveillance': https://edri.org/wp-content/uploads/2020/05/Paper-Ban-Biometric-Mass-Surveillance.pdf.

Despite accepting that real-time RBI can unduly restrict people's fundamental rights (Recital 18), and noting that the majority of respondents to the Commission's Consultation were in favour of new rules, **article 5 of the AIA contradictorily risks creating a blueprint for conducting biometric mass surveillance, instead of a substantive prohibition of these practices**. In its approach to RBI, the Act requires a lot of improvements to bring it in line with existing standards of fundamental rights and data protection:

- **Wide exceptions with low thresholds:** Despite recognising the severe undue fundamental rights risks of real-time RBI, the AI Act allows Member States to adopt three broad and highly discretionary exceptions to the prohibition (5.1.d). In article 5.1.d.i, the exception for "potential victims of crime" suggests that there need only be the *potential* of a crime, creating a dangerously wide and potentially arbitrary scope which may be easily misused to justify perpetual and untargeted use; furthermore, the fallacious reference to "targeted" search fails to recognise that remote biometric identification is by definition always mass / indiscriminate.[7] 5.1.d.iii sets a potentially very low bar to permit RBI to search for perpetrators or suspects of crimes under the European Arrest Warrant, which is a long list of crimes including non-violent ones like counterfeiting currency, forging administrative documents or trafficking endangered plants. This exception is further problematic because it is based on the assumption that facial recognition or other RBI is useful for the "prosecution of a perpetrator or a suspect of a criminal offence". However, due to its inherent probabilistic nature (sometimes referred to as the *base rate fallacy* phenomenon), biometric identification can never and will never provide conclusive identification or inferrence.[8] Thus it cannot be relied upon in a court of law, as shown in a 2019 case in the Netherlands where the defendant was acquited because a facial recognition match could not meet the burden of proof.[9] The Act claims that these exceptions are "narrowly defined" (Recital 19), but these examples show just how wide the exceptions and how low thresholds are to permit the mass infringement of fundamental rights;
- **Safeguards in name only:** The exceptions to the prohibition are furthermore subject to a series of purported safeguards (articles 5.2 and 5.3), including temporal and geographic limitations, and judicial or administrative authorisation. Given that the *ex ante* authorisation safeguard can be waived in the event of each Member State's

7 Garante per la Protezione dei Data Personali (Italian data protection authority) (2021), Riconoscimento facciale: Sari Real Time non è conforme alla normativa sulla privacy: https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/9575842.
8 EDRi (2019), Why EU passenger surveillance fails its purpose: https://edri.org/our-work/why-eu-passenger-surveillance-fails-its-purpose/.
9 EDRi (2021), The Rise and Rise of Biometric Mass Surveillance in the EU: A legal analysis of biometric mass surveillance practices in Germany, the Netherlands, and Poland, [65]: https://edri.org/wp-content/uploads/2021/07/EDRI_RISE_REPORT.pdf.

interpretation of the vague and discretionary threshold of "a duly justified situation of urgency" (5.3), it is possible that in its current form, the AIA may not be able to substantively prevent *any* law enforcement use of real-time RBI. This is especially pertinent in the context of systemic threats to democracy and the rule of law across the EU, evidenced for example in the pending European Court of Human Rights case brought by Panoptykon Foundation against the Polish government for the non-existence of effective supervision over the government's surveillance activities.[10] The authorisation process is thus vulnerable to government pressure and even further weakens the already deficient RBI 'prohibition';

- **The authorisation process**: Dr Nóra Ni Loideain has further noted that the authorisation is a flawed process which does not meet the standards of Charter of Fundamental Rights of the European Union (CFEU): firstly, in the current proposal, prior authorisation is permissible on the basis of a the low evidentiary threshold of "objective evidence *or clear indications presented to it* [the authorising authority]" (italics for emphasis) (5.1.3). This is a low bar in which the decisive factor in whether or not to authorise RBI can be "clear indications" provided by the very entity with a vested interest in using RBI. Furthermore, this authority is compelled to assess "the seriousness, probability and scale of the harm caused *in the absence of the use of the system*" (italics for emphasis) (5.2.a). This coercive and speculative approach seems at odds with fundamental rights principles of necessity and proportionality (CFEU article 52) which require that the burden of proof is on demonstrating that a use case or action does not unduly restrict fundamental rights;[11]
- **Incompatibility with requirements of necessity and proportionality:** In addition to the failure of the proposed safeguards to comply with existing fundamental rights law, the AI Act has further been criticised for its misapplication of the tests of necessity and proportionality for conducting RBI. The EDPS and EDPB Joint Opinion, for example, states that: "[t]he reasoning behind the Proposal seems to omit that when monitoring open areas, the obligations under EU data protection law need to be met for not just suspects, but for all those that in practice are monitored" (paragraph 31).[12] As demonstrated in EDRi's 'Ban Biometric Mass Surveillance' position paper, "real time" and "post" RBI (both of which constitute indiscriminate biometric surveillance) are inherently unnecessary and disproportionate under the CFEU and should be fully prohibited. Conversely, in its current form, the AIA creates the conditions for law enforcement agencies to unduly restrict the

10 Panoptykon Foundation, 'No control over surveillance by Polish intelligence agencies. ECHR demands explanations from the government ', December 2019: https://en.panoptykon.org/government-surveillance-echr-complaint.
11 Ni Loideain, N., University of London, article forthcoming, August 2020.
12 EDPS and EDPB, Joint Opinion on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act): https://edps.europa.eu/data-protection/our-work/publications/opinions/joint-opinion-edps-edps-proposal-regulation-european_en

fundamental rights of whole populations through biometric mass surveillance practices, contravening the CFEU;

- **Threat to existing data protection laws**: Recital 23 clarifies that "this Regulation is not intended to provide the legal basis for the processing of personal data". Therefore, it needs to be clarified that the *lex specialis* status of the prohibition on real-time RBI does not provide a legal basis for law enforcement under one of the exceptions (5.1.d), nor does it weaken existing protections of biometric data under the Data Protection Law Enforcement Directive (LED) or national implementations of the LED;
- **The "post" RBI loophole:** The AIA draws a fundamentally arbitrary distinction between "real-time" and "post" uses of remote biometric identification by virtue of a "significant [temporal] delay" between collection and processing which is, in fundamental rights terms, irrelevant. By doing so, the Act creates a loophole which permits law enforcement agencies to retrospectively apply biometric identification to CCTV footage or photographs. This form of biometric mass surveillance can unduly restrict people's rights equally as profoundly as real-time methods – and sometimes even more invasively so, due to the potential to pool data from many different sources across place and time. This erroneous distinction also leaves the deployment of equally harmful "post" RBI systems free from the restrictions in time, place and authorisation that apply for the exceptional uses of real-time RBI deployments (5.2), meaning that the potential for mass surveillance from "post" RBI is even further strengthened. Similarly, the definition of "remote" (3.36) is overly narrow in ways that may also create loopholes, for example arbitrarily and illogically linking the system definition to the (lack of) prior knowledge of the user. This must be corrected, as we will discuss further in the recommendations laid out in Section B;
- **No ban on other purposes (i.e. other governmental or private purposes):** The AIA limits the RBI prohibition to law enforcement purposes, on the basis that other purposes are already sufficiently prohibited under the General Data Protection Regulation (GDPR). By doing so, the AIA fails to acknowledge the existing wide exemptions under the GDPR, which EDRi has demonstrated have already led to the systematic and sustained violations of people's rights and freedoms across the EU.[13] In this manner, the proposal misses the opportunity to bring in complementary rules which will reinforce and strengthen the provisions on the processing of biometric data in the GDPR and align with the fundamental rights enshrined in the CFEU. Furthermore, by addressing only the *use* of these systems, EU providers may still be able to develop and sell rights-violating RBI systems outside of the EU;
- **No ban on other types of processing:** By limiting the prohibition to real-time RBI, the AIA fails to address other biometric mass surveillance practices such as singling out individuals based on their biometric characteristics. The protection of fundamental rights need a broader

13 EDRi (2021) The Rise and Rise of Biometric Mass Surveillance in the EU: A legal analysis of biometric mass surveillance practices in Germany, the Netherlands, and Poland, https://edri.org/wp-content/uploads/2021/07/EDRI_RISE_REPORT.pdf

prohibition of biometric mass surveillance practices than just remote biometric identification;

- **Infrastructural enablement**: The AIA fails to address the underlying issue of biometric mass surveillance infrastructures and enabling practices. Such infrastructures and practices can proliferate under the Act because "post" uses of RBI, "real-time" exceptional uses (under 5.1.d.i - iii), and uses for non-law enforcement purposes will all ensure that the required databases, software and hardware remain readily accessible. In essence, because only the use is prohibited, and not the development, sale, purchase or deployment, the implication is that biometric devices, software and databases can be bought, installed and maintained, and may be turned on with a simple authorisation (which we have already demonstrated is highly flawed). This doesn't just fail to stop biometric mass surveillance: it may even enable and encourage it by incentivising governments to make greater use of the costly, convenient infrastructures that are already in place;
- **Online spaces out of scope:** The exclusion of online spaces from the definition of publicly-accessible spaces which are subject to the prohibition - contrary to recommendations from EDRi[14] and more recently, the EDPS and EDPB[15] - suggests that the AIA may not prevent the scraping of online sources to develop commercial databases and software such as those offered by Clearview AI to many European law enforcement agencies. This is despite a number of EU data protection authorities (DPAs), including the Hamburg DPA, confirming the inherent rights-violating nature of such practices. The COVID-19 pandemic has made this even more urgent, as large parts of people's everyday lives have necessarily moved online. Online spaces must be included in the definition of publicly-accessible spaces, and the data scraped from online spaces (such as from social media) included in the prohibition;[16]
- **Vague and complicated wording**: The wording relating to RBI is unnecessarily vague and complicated, creating risky grey areas and making it overly onerous for civil society as well as AI providers and users to apply the rules in a consistent and rights-respecting manner.

2.2 Lack of mechanism to add unacceptable uses

Secondly, the **proposal does not introduce a mechanism by which unacceptable uses of AI may be added in the future**, unlike the process outlined in article 7 for updating the list of 'stand alone high risk' use cases.

14 EDRi (2020), Ban Biometric Mass Surveillance: https://edri.org/wp-content/uploads/2020/05/Paper-Ban-Biometric-Mass-Surveillance.pdf.
15 EDPS and EDPB, Joint Opinion on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act): https://edps.europa.eu/data-protection/our-work/publications/opinions/joint-opinion-edps-edps-proposal-regulation-european_en.
16 noyb (2021) Clearview AI's biometric photo database deemed illegal in the EU, but only partial deletion ordered: https://noyb.eu/en/clearview-ai-deemed-illegal-eu.

The lack of more general criteria to establish 'unacceptable risk' is an inconsistency in the framework, leaving a lack of clarity as to why the article 5 prohibitions were included (to the exclusion of others below), whilst not providing a framework for future unacceptable uses cases to be added as the AI market evolves. As outlined in the recommendations below, such criteria might include the impact on fundamental rights, structural power imbalances around the context of deployment (including potential for enhanced discrimination, marginalisation, inequality), lack of capacity for individuals, groups or civil society to contest the usage, etc.

2.3 Unacceptable use cases missing from list of prohibited AI

Thirdly, **the proposal does not put forward a holistic set of prohibitions covering the full range of unacceptable uses of AI**. As highlighted by EDRi alongside 62 human rights organisations,[17] 116 MEPs[18] and the European Data Protection Supervisor and European Data Protection Board,[19] **there are further use cases of AI that pose unacceptable risks to fundamental rights and democracy, and therefore must be prohibited under the AIA.** These are the following:

- **Predictive policing and uses of AI to risk assess for future criminality, offending or re-offending.** The use of predictive modelling to forecast where and by whom certain crimes are likely to be committed, alongside uses of AI to detect risk in the context of criminality, unduly and unnecessarily inpinge on a number of fundamental rights. In particular, the rights to dignity,[20] to an effective remedy and a fair trial,[21] to good administration[22] as well as the presumption of innocence[23] are compromised by practices that attempt to automate the prediction and characterisation of the future behaviour of individuals and groups, with potentially harmful consequences for their liberty and privacy.

  In addition, systems designed to assess risk or predict crimes have been demonstrated to repeatedly score poor, working class, racialised and

17 EDRi (2021) Open letter: Civil society call for the introduction of red lines in the upcoming Commission proposal on artificial intelligence: https://edri.org/our-work/civil-society-call-for-ai-red-lines-in-the-european-unions-artificial-intelligence-proposal/.
18 MEP letter to President von der Leyen, 8th March 2021: https://edri.org/wp-content/uploads/2021/03/MEP-Letter-on-AI-and-fundamental-rights-1.pdf.
19 EDPS and EDPB, Joint Opinion on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act): https://edps.europa.eu/data-protection/our-work/publications/opinions/joint-opinion-edps-edps-proposal-regulation-european_en.
20 Charter of Fundamental Rights of the European Union, article 1.
21 Charter of Fundamental Rights of the European Union, article 47.
22 Charter of Fundamental Rights of the European Union, article 41.
23 Charter of Fundamental Rights of the European Union, article 48; See also EDPS and EDPB, Joint Opinion on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), [34].

migrant communities with a higher likelihood of presumed future criminality,[24] therefore unreservedly compromising the rights to equality before the law and non-discrimination.[25] The use of historical data in practice serves as a proxy for race and other protected characteristics, as well as socio-economic status, reflecting patterns of over-policing of certain communities, exacerbating racial biases, the criminalisation of poverty and affording false objectivity to patterns of racial and other types of profiling. Insofar as such practices reflect ongoing surveillance priorities, it is highly likely that such practices will amplify existing patterns of institutionalised discrimination insofar as they reify presumptions of criminality on the basis of individual or group characteristics, behaviour, or location.

- **The use of AI systems at borders and in migration control.** The proliferation of tests, trials and deployments in the context of migration control is a particular fundamental rights concern, which is not systematically addressed in the AIA. Specific conditions relating to the migration context warrant a higher level of scrutiny and limitations on the use of AI, in particular: the heightened conditions of vulnerability placed on people on the move, including refugees, migrants, non-status individuals, and other categories; the lower procedural safeguards and protection of rights afforded to migrants;[26] and that the migration context has been used as an opportunity to experimenton an already highly marginalised category of persons.[27] It is vital that particular limitations are drawn and higher safeguards applied to ensure that the rule of law and fundamental rights cannot be overridden by national security or other vaguely-defined policy priorities, and that the principles of necessity and proportionality are upheld.

  In particular, myriad uses of AI in the migration control context pose severe risks to fundamental rights of people on the move, as well as comprising potential violations of international refugee and human rights law. The increasing datafication of the migration management process, use of AI systems and big data to predict migration controls in combination with an expansive surveillance infrastructure[28] to detect, intercept and prevent entry into Europe, is an impermissable use of AI systems. It also amounts to mass surveillance and is in contravention of

24 European Network Against Racism (2019), Data Driven Policing: The Hardwiring of Discriminatory Policing Practices Across Europe: https://www.enar-eu.org/IMG/pdf/data-driven-profiling-web-final.pdf,
25 Charter of Fundamental Rights of the European Union, articles 20 & 21.
26 UN Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance (2020), 'Racial discrimination and emerging digital technologies: a human rights analysis' A/HRC/4457.
27 EDRi, Petra Molnar (2020) Technological Testing Grounds: Migration Management Experiments and Reflections from the Ground Up: https://edri.org/wp-content/uploads/2020/11/Technological-Testing-Grounds.pdf.
28 European Parliament Research Service (2021). Artificial Intelligence at EU borders: Overview of applications and key issues. PE 690.706.

obligations under the Geneva Convention as well as a fundamental violation of the right to asylum enshrined in the Refugee Convention and domestic legislation. Further, the growing resort to biometric identification, verification and analysis of migrants' sensitive data in the the context of migration management is deployed in the context of significant power imabalance, particularly given that immigration and border administrative decision-making is already an area rife with opacity and discretion without adequate oversight and accountability.

The use of individual risk assessments and predictive systems to classify security or health risks also pose particular consequences for human dignity, equality and non-discrimination, privacy and data protection risks, as well as due process and good administration rights. The extent to which these systems are used to facilitate processes such as detention and deportation present particular risks to fundamental rights and with vast potential for abuse. In addition, emotion recognition systems (explored further in general below) are particularly harmful in the migration context due to the power imbalance and deep reliance on generally flawed and un-scientific premises,[29] leading to potentially innacurate and discriminatory decision-making processes.

- **Invasive monitoring, surveillance (including of biometric and other human features) and algorithmic management in employment and educational contexts.** As highlighted by unions, there are particular concerns with the deployment of AI to monitor, measure and manage employees, tasks and resources in employment and educational contexts. Firstly, we see a growing resort to invasive monitoring practices, predicated on a vast scale of data collection in extreme power imbalance and subordination,[30] fundamentally undermining notions of consent in data processing, given the contractual subordination of employees to their employers. Many such systems are combined with algorithmic assessments of performance and other forms of algorithmic task management, which are not only very likely to infringe on data protection and privacy rights of workers, but also likely diminish well-being, pose serious physical and psychological harm,[31] limit work autonomy and maintain greater distance and opacity between managers and workers. Further, as demonstrated by a number of cases contested by app-workers, algorithmic management and ranking systems used by large platforms have had severe consequences on the economic situation

29 ARTICLE19, (2021). 'Emotional Entanglement: China's emotion recognition market and its implications for human rights': https://www.article19.org/wp-content/uploads/2021/01/ER-Tech-China-Report.pdf.
30 ETUI (2021) The AI Regulation: entering an AI regulatory winter? Why an ad hoc directive on AI in employment is required: https://www.etui.org/publications/ai-regulation-entering-ai-regulatory-winter.
31 Wood, A. J., Algorithmic Management: Consequences for Work Organisation and Working Conditions, Seville: European Commission, 2021,JRC124874.

of workers as a result of specific decisions, including discriminatory treatment and violation of statutory rights,[32] and major decisions, such as termination, taking through substantively automated means.[33]

The classifcation of such systems as only 'high risk', subject to primarily technical requirements to be fulfilled by providers of AI systems, is wholly insufficient to mitigate these threats to fundamental rights and severe harms to individuals and groups. Rather than restricting these systems, the AIA in its current form rather deems such practices permissable, exacerbating the burden on civil society and affected individuals to seek redress in the event of harm. As outlined below in the recommendations, such practices must instead be prohibited.

2.4 Insufficiency of the limited risk approach

The **following practices are generally considered to have only a limited risk profile under the AIA, despite their vast capacity for harm and violations of fundamental rights**. In practice, at least for uses of AI which process personal data, it is hard to see how such rules go further than existing requirements under the General Data Protection Regulation (GDPR).

Instead, **biometric categorisation must be prohibited wherever it may unduly restrict fundamental rights, most notably equality and non-discrimination**, as set out more extensively in our recommendations in Section B. **Emotion recognition must be fully prohibited due to the fundamentally and unmitigably flawed assumptions on which emotion recognition rests, and its incompatibility with human dignity and many fundamental freedoms**.

Biometric categorisation

- **Ignoring the evidence of harms:** The AIA proposal puts biometric categorisation systems in the category of limited risk, entailing only a small number of mandatory transparency requirements (article 52). It further sets up the possibility that some biometric categorisation use cases could in future be considered high risk under Annex III heading 1, but does not at this point include any such use cases, on the grounds that there is not sufficient evidence. However, the EDRi network and other civil society groups have repeatedly demonstrated that in fact,

32 Tech crunch, 4th January 2021, Italian court rules against 'discriminatory' Deliveroo rider-ranking algorithm: https://techcrunch.com/2021/01/04/italian-court-rules-against-discriminatory-deliveroo-rider-ranking-algorithm/.
33 Personnel Today, 27th October 2020, Uber sued for 'automated' dismissals: https://www.personneltoday.com/hr/uber-sued-for-automated-dismissals/.

such categorisations can create severe and undue fundamental rights restrictions;[34]

- **Threats to equality and non-discrimination:** Biometric categorisation can gravely threaten rights to equality and non-discrimination, in particular when they relate to special categories of data as enshrined in the GDPR and protected under the CFEU and the broader EU equality and non-discrimination acquis. By definition, biometric categorisation is a process that seeks to put people into (often arbitrary, discretionary and stereotyped) boxes, and then to make predictions or decisions about them on that basis. Biometric categorisation has historical roots in systems of oppression and injustice, including the control of enslaved people in the US through the so-called 'lantern laws', the suppression of Indian people under British colonial rule, and even Nazi eugenics.[35] For these reasons, its use in a rule-of-law-respecting society is exceptionally hard to justify;
- **Links with mass surveillance:** Biometric categorisation often forms the technical foundation of other forms of biometric data processing which can lead to mass surveillance, such as in remote biometric identification. Recital 18 of the AIA acknowledges the particularly intrusive and chilling nature of law enforcement performing such RBI practices (albeit only in real-time);
- **Law enforcement exemption:** However, when it comes to the practice of biometric categorisation – which is often inextricable from RBI – article 52.2 contradictorily exempts its use in criminal investigations, detection and prevention from the already very limited transparency requirements that are established in the AIA. Given that law enforcement uses of biometric categorisation can be associated to severe and extensive harms (loss of liberty, denial of access to procedural rights, denial of the presumption of innocence etc) this risks creating a get-out clause for some of the most harmful biometric categorisation practices;
- **Things that cannot be inferred:** Additionally, the biometric categories proposed in article 3.35 treat as equivalent categories that may be predicted with a relatively high degree of accuracy based upon visible human features (such as predicting hair or eye colour, although even these are never absolute) with those that simply cannot be determined on the basis of human features (such as sexual or political orientation) – and instead, risk perpetuating scientifically-discredited and discriminatory theories like phrenology and physiognomy.[36] The self-learning nature of some AI systems could make it even harder to identify

34 EDRi (2020), Ban Biometric Mass Surveillance: https://edri.org/wp-content/uploads/2020/05/Paper-Ban-Biometric-Mass-Surveillance.pdf; All Out (2021) Ban automated recognition of gender and sexuality, https://campaigns.allout.org/ban-AGSR.
35 Najibi, A (2000), Racial Discrimination in Face Recognition Technology :https://sitn.hms.harvard.edu/flash/2020/racial-discrimination-in-face-recognition-technology/; Sengoopta, C (2003) Imprint of the Raj: How Fingerprinting was Born in Colonial India. London: Macmillan.
36 Access Now (2021), Ban Biometric Surveillance: https://www.accessnow.org/cms/assets/uploads/2021/06/BanBS-Statement-English.pdf.

when *prima facie* non-sensitive biometric features are in fact being used as proxies for sensitive characteristics like sexual orientation.

Emotion recognition

- **A scientifically invalid process:** except in the cases of law enforcement or migration uses of polygraphs and "similar tools" (which are designated high risk under Annex III, 6.b and 7.a), the AIA proposal classifies emotion recognition as only limited risk, despite vast evidence of its harms as well as its complete scientific invalidity – such as via EU-funded projects like the much-criticised iBorderCTRL.[37] The Civil Liberties Committee in the European Parliament has already called for the use of emtoion recognition in law enforcement to be "discontinued";[38]
- **Fundamentally incompatible with fundamental rights**: EDRi member ARTICLE19 has demonstrated that, as a practice, emotion recognition is incompatible with international human rights principles and rules.[39] As an incredibly intrusive practice, it can infringe on people's dignity, is often used in discriminatory contexts, and intrudes into people's cognitive liberty by coercing not just how people express themselves, but even how they think. Despite this, emotion recognition is becoming increasingly common in employment contexts, education, border and migration experiments and advertising. It risks infringing on people's rights and freedoms, and has particularly grave impacts on human dignity when important decisions relating to people's free movement, employment and other rights are made upon the basis of an inherently probabilistic and flawed system, which no amount of improvements to accuracy or performance can ever fix.

The problem of definitions within the limited risk category: The definitions in the AIA of an emotion recognition system (3.34) and a biometric categorisation system (3.35) both limit the application of rules for these processes to when it is performed on the basis of biometric data. However, given that the AIA's definition of biometric data (3.33) applies only if it allows or confirms the unique identification of a natural person, there is a risk that certain emotion recognition and biometric categorisation practices could be performed using data sets which avoid or even evade the threshold for being considered biometric data, for example through 'anonymisation' (despite growing scepticism about the credibility of supposedly anonymised biometric

37 Wired (2018), The science behind the EU's creepy new border tech is totally flawed:, https://www.wired.co.uk/article/border-control-technology-biometrics; Jakubowska, E., (2021) Mass facial recognition is the apparatus of police states and must be regulated: https://www.euronews.com/2021/02/17/mass-facial-recognition-is-the-apparatus-of-police-states-and-must-be-regulated.

38 LIBE Committee (2021), Artificial Intelligence in policing: safeguards needed against mass surveillance: https://www.europarl.europa.eu/news/en/press-room/20210624IPR06917/artificial-intelligence-in-policing-safeguards-needed-against-mass-surveillance.

39 ARTICLE19 (2021), Emotional Entanglement: China's emotion recognition market and its implications for human rights: https://www.article19.org/wp-content/uploads/2021/01/ER-Tech-China-Report.pdf.

data). This could allow providers and users to circumvent the already unacceptably low requirements on emotion recognition or biometric categorisation.

The Act should therefore widen the scope for emotion recognition and categorisation to include biometric, physiological and behavioural signals (in a new definition under article 3 of 'human features', as elabourated in section B) in order to ensure that equally harmful uses of data about human features are in scope of the prohibitions even when unique identification may not occur.

Evasive processing practices (e.g. edge or transient processes) may also be employed in attempts to avoid the technical processing threshold for data to be considered biometric (3.33) – although it is important to note that the Italian Data Protection Authority (DPA) has confirmed that even if discarded almost immediately, the practice of scanning the biometric features of everyone in view of a camera is still considered unlawful mass surveillance.[40]

The inadequacy of notification rules to prevent harms arising from biometric categorisation or emotion recognition: For both biometric categorisation and emotion recognition, the AIA fundamentally falters in its presumption that the disclosure of their use to those who are subject to them is a solution to the harms and violations of rights that these practices can entail; and that such notification constitutes genuine transparency and accountability.

Rather, harms such as a trans person being mis-gendered in public, a racialised person being shown (or not shown) adverts on the basis of their predicted ethnicity, or an employee facing disciplicary procedures due to not showing the 'right' emotions at work, are just three of many examples of how the negative impacts of biometric categorisation and emotion recognition will remain just as real, regardless of whether or not the use of AI is disclosed to the subject.

3 The AIA's scope overlooks broader structural harms and impact of AI

It is positive that the AIA proposes a broad definition of artificial intelligence to include in scope a wide range of potentially harmful AI systems. However, despite this broad definition, the AIA has an extremely narrow list-based approach to regulation, which narrowly specifies use cases to be classified as high risk, alongside a process (article 7) for the future inclusion of high risk use cases that fit within existing areas outlined in Annex III.

40 Garante per la Protezione dei Data Personali (Italian data protection authority) (2021), Riconoscimento facciale: Sari Real Time non è conforme alla normativa sulla privacy: https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/9575842.

Yet the following limitations in the AIA's scope are particularly challenging due to the act's primary objective, which is to promote a harmonised single market for AI within the EU. As highlighted by Veale and Zuiderveen Borgesius,[41] this 'maximum harmonisation' function requires that Member States must disapply any conflicting rules with those in the act. Thus, Member States are to be potentially precluded from introducing higher fundamental rights standards than those contained in the AIA.

3.1 A lack of future-proofing for high-risk requirements

The specific challenges as to the scope of the AIA are as follows. Firstly, the core requirements of the AIA apply to a very narrowly defined list of 'high risk' AI systems, as outlined in article 6. As such, with respect to 'stand alone' use cases based on fundamental rights risks, the AIA limits from the start the range of 'high risk' areas, determined solely by the European Commission, which cannot be updated in the future. The limited and caveated nature of the pre-defined areas (for example the processing of biometric data is high risk only if it leads to identification or categorisation, despite the fact that some authentication uses may entail significant risks) casts doubt over how comprehensive and future-proof this Annex can possibly be. Furthermore, new sub-areas can be included only insofar as they are compatible with the criteria outlined in article 7. This falls far from the precautionary principle or rights-based approach called for by civil society.[42] In addition, currently the European Commission has centralised power to update the list of high risk sub-areas in Annex III.

3.2 Exclusion of many harmful use cases from the requirements

Secondly, the risk-based approach, with requirements primarily limited to the narrow list of high risk AI systems, necessarily means that a number of systems with potentially harmful impacts remain unregulated under this act. There is particular concern as to the extent to which the following types of harms are (not) addressed:

- **AI systems which exacerbate structural inequalities and power imbalance:** As outlined above and specifically with respect to deployments of AI in the contexts of law enforcement, migration control and workplace surveillance, uses of AI in certain contexts will necessarily perpetuate structural power imbalances and fundamental rights risks.

41 Veale and Zuiderveen Borgesius (2021) 'Demystifying the Draft EU Artificial Intelligence Act', https://arxiv.org/abs/2107.03721.
42 Access Now (2021), The EU should regulate AI on the basis of rights, not risks: https://www.accessnow.org/eu-regulation-ai-risk-based-approach/; EDRi (2020), Recommendations for a fundamental rights-based Artificial Intelligence Regulation: addressing collective harms, democratic oversight and impermissible use: https://edri.org/wp-content/uploads/2020/06/AI_EDRiRecommendations.pdf.

Further, as outlined by the European Disability Forum (EDF), there are no provisions in the Act to ensure that all AI systems (regardless of risk level) meet international legal obligations relating to the accessibility of persons with disabilities.[43] By promoting the notion that AI systems can be primarily regulated through a series of technical measures, (documentation, human oversight in design and data quality standards), the AIA provides no response to the structural harms outlined in the previous section.[44] This techno-centric framing does not adequately deal with how AI as socio-technical systems become embedded in broader processes of structural discrimination, which the 'examination of possible biases' (article 10), purported improvements in accuracy (Article 15) and more documentation (article 11) will simply not address.[45]

- **Environmental impact:** The AIA wholly underestimates the vast impact of a policy agenda designed to promote the widescale uptake of AI, underpinned by the exponential collection of data and focusing on the presumed benefits, without sufficient regards to the broader implications of the greater resort to AI systems on the environment. In particular, consequences on the environment relating to the vast environmental resources (including the exploitation of natural resources for the hardware required to underpin AI systems) as well as the energy consumption[46] required for many of such systems to be trained and functional, as well as for data to be stored, find no place in the proposed regulatory framework.

- **Economic and infrastructural consequences of AI systems:** By defining 'high risk' primarily with respect to fundamental rights and product safety, the AIA leaves little room for broader political and economic impacts of AI that fall outside of these frameworks. Also overlooked are the labour implications of the AI production pipeline, which often rely on labour exploitation of people in the Global South, but also how the resort to algorthmic management is reshaping the labour market toward 'crowd work' and other more precarious, flexibilised forms of work.[47] In addition, broad scale economic and political impacts following from the increased uptake of AI, including the increased

43 European Disability Forum (2021). Disability Perspective of AI of excellence and of trust (forthcoming).
44 EDRi (2021), EU's AI law needs major changes to prevent discrimination and mass surveillance: https://edri.org/our-work/eus-ai-law-needs-major-changes-to-prevent-discrimination-and-mass-surveillance/.
45 EDRi (2021), Beyond De-biasing: Automated decision making and structural discrimination, authored by Agathe Balayn and Seda Gürses, Delft University of Technology, the Netherlands (forthcoming, September 2021).
46 Emma Strubell, Ananya Ganesh and Andrew McCallum (2019). 'Energy and Policy Considerations for Deep Learning in NLP', accessed via: https://arxiv.org/pdf/1906.02243.pdf.
47 Valerio De Stefano (2016), The rise of the "just-in-time workforce": On-demand work, crowd work and labour protection in the "gig-economy": https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2682602.

dependencies on centralised computational infrastructures,[48] as well as the re-structuring of organisations and democractically accountable institutions[49] are entirely overlooked in the AIA framework, with no foresight of the need for users of AI systems to take such factors into account.

- **Enabling mass surveillance**: The AIA proposal addresses neither the processes nor the infrastructures that may contribute to and normalise mass surveillance in its many forms (including, but not limited to, biometric mass surveillance). By promoting structures for gathering, inferring or predicting ever-more information about people, and connecting it across entities and services, the capacity for states and companies to unjustifiably surveil part or whole populations becomes ever- present and inescapable.

- **Exporting rights-violating AI to the rest of the world**: Concerns are also relevant in the context of export, as the scope of the AIA in article 2 establishes that the Act covers only AI that is put on the market/into service or used *within the EU*. This means that companies based in the EU may nevertheless be able to develop high-risk AI in an unrestricted manner, and even prohibited AI. It is a contradiction of EU rights and values that companies or entities based in the EU should be allowed to develop and then sell systems to states or companies outside of the EU, despite such systems being deemed to pose an unacceptable risk to fundamental rights and safety within the EU.

3.3 Loopholes enabling high risk uses

Military and international law enforcement: Further, a number of loopholes limit the scope of the AIA in areas with crucial fundamental rights implications. article 2(3) leaves out of scope uses of AI developed or used for military contexts, and 2(4) leaves out of scope of the legislative proposal international organisations using AI systems in the framework of international agreements for law enforcement. This poses an unwarranted loophole for uses of AI for organisations such as EUROPOL, yet which still operates with significant fundamental rights implications for individuals in the European Union.[50]

Large scale migration databases: Further, article 83 leaves out of scope AI systems which are components of large scale IT systems, including the Schengen Information system, Visa Information System, Eurodac, the Entry/Exit

48 For an explanation of computational infrastructures, see https://www.tudelft.nl/tbm/program-mable-infrastructures.
49 EDRi (2021), Beyond De-biasing: Automated decision making and structural discrimination, authored by Agathe Balayn and Seda Gürses, Delft University of Technology, the Netherlands (forthcoming, September 2021).
50 EDRi (2021). Recommendations on the revision of Europol's mandate: https://edri.org/wp-content/uploads/2021/06/Recommendations-on-the-revision-of-Europols-mandate.pdf.

system, ETIAS, the European Criminal Records Information System on third-country nationals and stateless persons, and the Interoperability framework (Annex IX). This is major loophole for AI uses within the EU's migration control framework, with significant and severe consequences on the fundamental rights of people on the move should the AI systems that form part of these controls be excluded from the AIA's scope.

Public security exemption: In addition, article 47 provides for a concerning ability for market surveillance authorities to authorise and provide exemptions to the conformity assessment procedure for reasons of 'public security or the protection of life and health of persons, environmental protection and the protection of key industrial and infrastructural assets.' This provides an overly broad basis for market surveillance authorities to eradicate the already limited safeguards provided for in the Act, potentially compromising the principles of necessity and proportionality if not duly respected.

Further processing of data exemptions: Further, article 54 of the proposal sets out a dangerous exemption to the principle of purpose limitation for 'innovative' uses of AI. Specifcially, it allows for further processing of personal data for uses of 'substantial public interest', such as particular uses for law enforcement, public health, and environmental reasons. As argued by the EDPB and EDPS, this provision, alongside others in the proposal, presents a potential disconnect with the underlying principles contained in the GDPR regarding the grounds for further processing.

These loopholes – along with exclusions from rules for certain law enforcement purposes throughout the proposal – also pose an additional risk that certain AI-based processes currently dealt with by administrative authorities may be pushed to law enforcement agencies in order to avoid regulatory scrutiny. The impact of such moves is that certain people – most likely from marginalised groups, for example people on the move – could be criminalised as a result.

4 Focus on providers; limited obligations on users

The core assumption of the AIA is that providers of AI systems are best placed to forecast, identify and mitigate the main harms that may emanate from AI systems. Following this, the AIA centralises the provider in the regulatory framework, with the bulk of requirements in the Act falling on those developing AI systems (articles 8-15).

The regulation allows a very wide scope for self-regulation by companies developing "high risk" AI. For the majority of high-risk AI uses contained in Annex III, the rules in article 43(2) mean that compliance with the Regulation's requirements is primarily ensured through self-assessment by the providers themselves. It is concerning that AI providers (those with a financial interest in

securing compliance and without the expertise to assess the implications on people's rights) themselves to judge if they have sufficiently met the requirements set out on data governance, transparency, accuracy, and more. Further, as higlighted by Veale and Zuiderveen Borgesius[51], the conformity assessment process is likley to be highly influenced by European standardisation organisations such as CEN (European Committee for Standardisation) and CENELEC (European Committee for Electrotechnical Standardisation).

The involvement of such entities in setting broad standards for the fulfilment of the essential requirements on providers further abstracts and weakens the process, creating a 'presumption of conformity'. This process is particularly inappropriate for AI systems falling under article 6(2) of the act relating to fundamental rights impact, which may be incredibly complex, intangible and thus difficult to standardise for, but also have incredibly high potential impact on peoples' rights. Furthermore, article 43(1) establishes a potential loophole, enabling providers to evade the requirement for external conformity assessments of biometric identification or categorisation systems if they comply with the standards. This means that a key purported safeguard may in practice have no effect.

Another crucial flaw of the AIA's approach is that it **overlooks the complexity of AI systems and the importance of context to be able to asess impact on fundamental rights, people and society**. Whilst the provider-led conformity assessment process may identify the core technical shortcomings of the system, the mechanism is fundamentally ill-suited to identify the risks in the context of deployment. For example, a facial recognition system may meet the technical requirements specified in the Act, yet still pose significant fundamental rights violations, compromise data protection and non-discrimination law, and enable mass surveillance in the context of deployment (i.e. in a shopping centre). Because AI systems also by definition learn over time, there are intrinsic limitations to any snapshot-in-time conformity assessment. As such, ensuring that there are greater (and ongoing) obligations on users, in addition to providers, is crucial in order to address the fundamental rights issues that will arise in the use of AI.

Further, we see that the requirements contained in the AIA placed on providers are highly technical in nature, thus largely inappropriate as a mechanism to prevent or mitigate potential risks to fundamental rights or other structural harms, or economic or environmental shifts engendered by the introduction of AI systems in context. Such considerations are inherently better assesed by the users in light of the context of deployment.

4.1 Embedding dominance of AI providers

51 Veale and Zuiderveen Borgesius (2021) 'Demystifying the Draft EU Artificial Intelligence Act': https://arxiv.org/abs/2107.03721.

This focus on providers potentially sets up a situation in which users of an AI system, such as a government agency, will be legally bound to follow the guidelines set out by a private company relating to the use of the AI system which they have procured (article 29). Whilst this article is intended to avoid misuse of a system, an unintended consequence of it may be that technology companies developing AI, whereby many already command disproportionate power over people and markets as established in EDRi's work on platform power, will thus be able to dictate the application of rules to which users – including governments - are bound.

4.2 Insufficient transparency

With respect to the transparency framework, the Regulation (article 13) largely imposes limited transparency obligations on providers toward users, as opposed to transparency requirements directly to people affected by or subject to AI systems (the exception to this is for limited risk uses cases under article 52, however we have explained on page 18-20 the serious shortcomings in the limited risk approach in the AIA). As such, the proposal will have a severely limited effect on people's ability to understand and challenge harmful and opaque AI systems deployed against them.

Whilst the inclusion of an EU database of high risk AI systems as outlined in article 60 is welcomed, currently the provision focuses on registration of high risk applications being put on the EU market. Full public transparency necessitates that this database registers high risk systems being put into use, including details on which actors are deploying them and for which purpose. In addition, Annex VIII, s.11 contains an exception for public transparency for uses of AI in law enforcement and migration control, limiting the efficacy of the tool for public transparency in these sensitive contexts and the extent to which this database provides the necessary democratic checks and balances.[52]

Further, whilst the information currently included in article 13 to users provides a good basis of transparency, there is little obligation on providers to disclose to users information relating to the political assumptions and specific decisions related to the fundamental goals and assumptions of the system, weightings, parameters and standards resulting from these decisions. Articles 13 and 15 refer to accuracy of AI systems, however the definition and standards for accuracy are left to the discretion of the providers. This provides little by way of guarantee to the user as to the validity or efficacy of the AI system for the purpose of use. This self-regulatory approach gives very little certainty with respect to the potential impact on fundamental rights. There is a concern that, if performance metrics conveyed to the user under article 13 are not sufficiently

52 EDPS and EDPB, Joint Opinion on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act): [69]-[70].

detailed and substantive, the broader human rights implications of deploying such a system may not be evident or discoverable by the user.

4.3 Structural discrimination through AI

With respect to discrimination exacerbated by AI systems, the AIA makes the assumption that 'data quality' can solve the harms emenating from high risk AI systems. However, for many of the applications listed in Annex III, whilst AI developers may be able to predict and prevent some negative biases, for the most part such systems will inevitably exacerbate structural inequalities. This is because AI systems are deployed in a wider context of structural discrimination.[53] By relying on technical checks for bias as a response to discrimination, the proposal risks reinforcing a harmful suggestion that removing bias from such systems is even possible, potentially obfuscating the need for structural solutions, such as limitations on certain uses, but also the need for intensive governance related responses.

4.4 Impact on marginalised communities

There are no specific requirements on users intending to put into use an AI system to measure the potential impact of such system on marginalised communities, nor take steps to mitigate those impacts (including ensuring inclusive access, or halting deployments should they have an harmful impact on certain groups). The AIA fails to impose specific requirements on users to ensure the accessibility of AI systems or services that are operational through AI systems.[54] This is a systemic oversight in this Regulation, particularly in light of the broad promise that AI systems are likely to bring benefits to all in society.

4.5 No consultation with affected groups

Futher, the AIA foresees no specific duty on users to consult with affected groups or social partners before deploying AI systems, or modes of democratic oversight of AI systems deployed in contexts vital to the public interest. As already highlighted by Unions, this may serve to dilute existing consultation requirements with social partners in the employment context.[55]

53 EDRi (2021), Beyond De-biasing: Automated decision making and structural discrimination, authored by Agathe Balayn and Seda Gürses, Delft University of Technology, the Netherlands (forthcoming, September 2021).
54 European Disability Forum (2021), Disability Perspective of AI of excellence and of trust (forthcoming).
55 ETUI (2021) The AI Regulation: entering an AI regulatory winter? Why an ad hoc directive on AI in employment is required: https://www.etui.org/publications/ai-regulation-entering-ai-regulatory-winter; UNI Global European Commission consultation response, June 2021.

5 Limited enforcement and governance framework without actionable redress for subjects of AI

The AIA pays insufficient attention to the fundamental interaction between the user and the subject of AI. This relationship is key to any fundamental rights based analysis and regulation of AI systems, and to the crucial question of how harms can be prevented and mitigated.

Aside from article 52 outlining notification requiremented for a few narrowly defined 'limited risk' AI systems (which are in themselves insufficient) the Act does not foresee notification requirements for high risk systems; duties to explain the reasoning behind automated decision making processes; nor, crucically, mechanisms for flagging or contestation of violations or harms as a result of interaction with AI systems. Whilst the AIA foresees the need for coordination between relevant national authorities supervising the regulation (articles 63(7); 64; 65 and 67) it provides no mechanism by which affected individuals or groups may flag to authorities potential harms, breaches of the Act or fundamental rights issues with an AI system.

In addition, there is no mechanism for individual or collective redress for harms in scope of the AIA. This is a particularly fundamental ommission considering the limitations of other legal frameworks to provide effective redress with respect to AI systems, including the limits of article 22 of the GDPR, which is likely to be insufficient as a means to provide a right to explanation for many AI systems.[56] Further, the limits of non-discrimination law, namely the focus on a limited set of protected characterstics, its requirement of a comparator, the focus on individual instances of discrimination as well as the high burden of proof on individuals in practice (exacerbated by the opacity of AI systems) reiterates the need for the inclusion of a redress mechanism in the AIA.[57]

Whilst the Explanatory Memorandum to the proposal claims that "all major stakeholders" were consulted in the course of developing the proposal (3.1), many civil society organisations and communities have challenged the accuracy of this claim, citing their exclusion from this process. It is crucial that going forward in the legislative process, engagement with more affected groups (especially marginalised groups that are most likely to be subjected to AI decisions) is prioritised, in particular to consider the need for meaningful redress measures.

56 Wachter, Mittelstadt, and Floridi, (2017), 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' *International Data Privacy Law*, Volume 7, Issue 2, May 2017, Pages 76–99.
57 Tetyana Krupiy (2021), 'Why the proposed Artificial Intelligence Regulation does not deliver on the promise to protect individuals from harm' European Law Blog: https://europeanlaw-blog.eu/2021/07/23/why-the-proposed-artificial-intelligence-regulation-does-not-deliver-on-the-promise-to-protect-individuals-from-harm/.

Lastly, there is a significant concern with centralised nature of the proposed governance framework for the AIA. Article 56 establishes the European AI Board, however removes the competence of this Board (which appeared in previous verisons) to present additions to the list of high risk AI systems. In the AIA in its current form, this function is centralised with the European Commission only, presenting a significant concern as to the democratic nature of this process. Further, due to the power of the European Commission in the European AI Board, it raises significant questions relating to the independence of national supervisory authorities, which report to the AI board. However, as it is likely that these authorites should be Data Protection Authorities (and following the recommendation of the EDPS and EDPS it is clear that they should), that they should report to the European Commission potentially compromises the independence of these entities.

(B) Recommendations for a fundamental rights-based Artificial Intelligence Act

In light of this analysis, EDRi recommends that the European Parliament and the Council of the EU implement the following improvements to the Artificial Intelligence Act (AIA):

1 Ensure effective protection against prohibited practices and address the full scope of unacceptable risks through AI

Imperative to the goal of a fundamental-rights respecting artificial intelligence regulation is the need to implement meaningful mechanisms geared toward the prevention of harm on individuals, groups and wider society. Civil society has been clear on the need to prevent, rather than to mitigate after the fact, 'impermissible' or 'unacceptable' risks to fundamental rights.[58]

a) Strengthen existing prohibitions in article 5 to provide meaningful protection against fundamental rights violations and individual and collective harms:

i. Ensure that the prohibition on subliminal manipulative techniques in article 5(1)(a) extends to harms which target groups of people as well as individuals;

ii. Remove the caveat that AI systems that deploy 'subliminal techniques in order to materially distort a person's behaviour' (article 5(1)(a)) or 'exploits... vulnerabilities of a specific group due to their age, physical or mental disability' (article 5(1)(b)) must 'cause or be likely to cause psychological and physical harm' in order to be prohibited. Extend the list of vulnerabilities in article 5(1)(b) to at least the protected characteristics outlined in the Charter of Fundamental Rights in the EU, with the explicit inclusion of gender identity;

iii. Ensure wide application of the prohibition on social scoring[59] (article 5(1)(c)). Remove narrow framings, such as the temporal limitation 'over a certain period of time', the limitation to public authorities,

58 EDRi (2021), Open letter: Civil society call for the introduction of red lines in the upcoming Commission proposal on artificial intelligence: https://edri.org/our-work/civil-society-call-for-ai-red-lines-in-the-european-unions-artificial-intelligence-proposal/.
59 EDPS and EDPB, Joint Opinion on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), [29].

the narrow framing of a singular 'score' and replace the reference to trustworthiness to one of 'risk'. Remove references to 'general purpose' in recital 17.

b) **Comprehensively prohibit the use of remote biometric identification in publicly-accessible spaces for any purpose**, and implement "a general ban on any use of AI for an automated recognition of human features in publicly accessible spaces - such as of faces but also of gait, fingerprints, DNA, voice, keystrokes and other biometric or behavioral signals - in any context", as per the EDPS-EDPB Joint Opinion.[60] These prohibitions must apply for all purposes and in any context, including online spaces, and without exception:

i. Furthermore, the putting on the market or placing into service of remote biometric identification software and hardware should be restricted in order to prevent biometric mass surveillance infrastructures being rolled out, and to ensure that EU companies cannot sell products and services which are designed for biometric mass surveillance outside the EU. The purpose limitation principle in the GDPR (article 5.1.b) should be reiterated here, as it already stipulates that CCTV footage, for example, should not be used for other purposes, for example training AI software or for performing re-identification;

ii. The definition of "remote" in RBI (3.36) should add that RBI occurs not just with reference to watchlists but also to general databases. The provision that it applies only if there is no prior knowledge of the user about whether the person of interest will be present and identifiable should be fully removed, in order to avoid creating loopholes;

iii. Human features should be defined under article 3 to include – but not be limited to - biometric, physiological, behavioural and neurological signals;

iv. As called for by the Civil Liberties Committee in the European Parliament, there must be "a ban on the use of private facial recognition databases in law enforcement" such as Clearview AI due to the likely incompatibility of such uses with EU data protection law.[61] EDRi's analysis has shown that many of the same incompatibilities will apply also for databases developed by law enforcement agencies themselves.

60 Ibid [11].
61 LIBE Committee, Artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters 2020/2016(INI), awaiting Plenary vote.

c) Include new prohibitions on the following practices which are incompatible with fundamental rights and democracy, and pose an unacceptable risk:

i. Uses of AI in the field of law enforcement or criminal justice that purport to predict future behaviour, including analysing the risk that individuals will offend or re-offend, and predicting the likelihood that criminal or unfavourable conduct will occur on the basis of personality traits, individual or group characteristics or location;

ii. Uses of AI in the field of migration control in ways that undermine the right to claim asylum, including but not limited to those:

- to risk assess inviduals for factors that do not relate to the substance of their immigration claim, such as risk of terrorism, public health threats, etc;
- to collect data and / or predict patterns in migratory movements for the purpose of preventing the exercise of the right to claim asylum;
- AI systems to assess eligibility for asylum, refugee or visa claims;

iii. Uses of AI that implement invasive surveillance, monitoring and algorithmic management in an employment and educational context;

iv. (Biometric) categorisation, which can pose a grave and disproportionate threat to all human rights, in particular equality and non-discrimination, by comprehensively prohibiting the use of AI to categorise people, on the basis of their human features, to the special categories of data as defined in article 9 of the GDPR;[62] or to categories based on the grounds for unlawful discrimination in article 21 of the Charter of Fundamental Rights of the European Union;[63] or on the basis of mental health status, migration status or gender identity:

- This prohibition must also include the use of potentially non-special or non-personal data, as well as data that does not meet the threshold to be considered biometric, captured from human features when used to categorise people according to *proxies* of

62 Racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation (GDPR, article 9).
63 Sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or other opinion, membership of a national minority, property, disability, age or sexual orientation (CFEU, article 21).

special or protected categories (e.g. by combining eye and hair colour to predict ethnicity or using the wearing of a headscarf to predict religion). In the event that the self-learning nature of some AI systems makes it difficult to know whether people are being assigned to categories that could lead to discrimination, the precautionary principle dictates that such uses should also be prohibited;

- As already explained above, the definition of biometric data (article 3.33) must be complemented with a definition of human features to ensure that it includes all data relevant to biometric or other human feature categorisation (without loopholes for types of data or methods of processing that don't meet the current threshold);[64]

v. Emotion recognition, which is scientifically invalid and can unduly infringe on all human rights, in particular human dignity and free expression, by comprehensively prohibiting the placing on the market, putting into service or use of AI to infer, predict, analyse or assess a person's emotions, feelings, emotional state, beliefs, preferences, intentions or otherwise inner thoughts, as well as to use human features, behaviours or expressions to predict future actions or behaviours;

vi. Uses of AI that constitute mass surveillance should be prohibited. Mass surveillance means the surveillance of, or potential for surveillance of, whole or part populations (including specific groups), and is thus inherently unnecessary and disproportionate.

Note that outside of the proposed **prohibitions of biometric mass surveillance**, **biometric categorisation on the basis of special or protected categories** and **emotion recognition** in general, there are additional AI applications which use human features and which can pose a high risk to fundamental rights. Therefore we additionally recommend that:

1. Heading 1 of Annex III is changed to "Physiological, behavioural, biometric and neurological authentication, identification and categorisation". If human features are defined in article 3 according to our recommendation, then an alternative heading could be "Authentication, identification and categorisation of human features";

2. Under heading 1, the following use cases are added in addition to the existing use case (1.a), but are not necessarily exhaustive at this point:

   i. Physiological, behavioural or biometric authentication, identification or categorisation (i.e. of human features) for law enforcement purposes;

64 See Access Now's Submission to the Consultation on the AI White Paper for a deeper engagement with issues of the definition of (biometric) categorisation.

ii. Physiological, behavioural or biometric authentication, identification or categorisation (i.e. of human features) by private actors for surveillance or security purposes (such as security companies);
iii. Physiological, behavioural or biometric authentication, identification or categorisation (i.e. of human features) for any purpose, where it can determine, solely or in part, people's access to:

- Public services (e.g. getting benefits payments);
- Private or privatised services which are necessary for people to exercise or enjoy their fundamental rights and freedoms (e.g. using e-border gates, entering supermarkets, going to work).

Furthermore, the Act must guarantee that **providers of high-risk uses of AI under Annex III paragraph 1 should not be able to circumvent the obligation for an ex ante third party conformity assessment simply by meeting harmonised EU standards** (as proposed in article 43(1).

2 Adapt the AIA to ensure holistic, democratic and 'future-proof' framework

Noting that the impact of AI systems extend far beyond impact on individual rights and product safety, but also are highly transient and susceptible to rapid change, the following proposals are designed to democratise the framework set out in AIA as well as better respond to structural and infrastructural harms.

a) Introduce a democratic, inclusive and acccessible process by for the insertion of new prohibitions. Include criteria for 'unacceptable risk' and the addition of future prohibitions into the AIA to ensure the enduring relevance of this regulatory instrument:

i. Such criteria might criteria might include the impact on fundamental rights, structural power imbalances around the context of deployment (including potential for enhanced discrimination, marginalisation, inequality), lack of capacity for individuals, groups or civil society to contest the usage, etc.

b) Ensure the potential to update the high risk use case 'areas' in the future (amending article 7) in addition to updating the use cases 'sub-areas'. Further, this process must not be centralised with the European Commission only, but include a range of actors including civil society.

c) **Respond to gaps in regulation with respect to economic and environmental impact, structural forms of inequality and AI and migration control, law enforcement and worker surveillance, mass surveillance, and exports of high-risk or prohibited AI outside the EU.**

d) **Remove loopholes in articles 2(4), and 83 leaving out of scope of the AIA AI systems used as part of international agreements on law enforcement and large scale IT systems in the migration control context.**

e) **Remove the broad exemption to forgo the duty to conduct a conformity assessment on grounds of public security in article 47.**

f) **Remove the exemption to the principle of purpose limitation contained in article 54(1)(a) for 'innovative AI' within the regulatory sandbox provisions for uses in the criminal justice context.**

3 Ensure responsibility to those subjected to AI systems with enhanced obligations on users of all AI systems

In the current AIA framework, the majority of the requirements fall on providers to implement a series of technical measures designed to mitigate harm in the deployment of systems. However, many of these harms are likely to be contextual and are best evaluated and addressed by the user, who has ultimate responsibilty to those subjected to the AI system. To ensure the use of AI systems is accountable to and compliant with fundamental rights, we recommend that the requirements on providers are complemented with obligations on users geared toward greater responsibility to those subjected to AI systems.

a) **Mandate users to conduct and publish an *ex ante* human rights impact assessment before putting a high risk AI system into use, clearly outlining the stated purpose for which the system will be implemented:**

   i. The impact assessment must be published on registration of use of the system in the public database under article 60;

   ii. This impact assessment must involve prior consultation with relevant national authorities, including equality bodies, consumer protection agencies, and data protection agencies. If other impact assessments are also required, these impact assessments must be published together;

iii. The impact assessment must also carry out meaningful consultation with social partners, civil society groups and individuals and groups affected by the use case;

iv. The impact assessment must include full assessment of the fundamental rights that are likely to be impacted by the AI system, in addition to broader, social, political and economic consequences of deploying the AI system in the particualr context for the particular use. This should include 'indirect' consequences of deploying the system, beyond impacts on those directly impacted by a decision generated;

v. The impact assessment must include clear steps as to how the harms identitified will be mititgated, and how effective this mitigation is likely to be. If adequate steps for mitigation cannot be cannot be outlined, the system ought not to be deployed.

b) Implement on users a duty to cooperate with national competent authorities investigating AI systems for potential threats to fundamental rights or safety under articles 65 and 67 for all AI systems, regardless of risk designation.

c) Implement a duty on users to meaningfully consult with institutions, civil society and social partners representing affected groups before deploying high risk AI systems:

i. Documentation of the results of this consultation should be included in the publicly acessible impact assessment.

d) When the user of any AI system is a public authority, implement a notifcation requirement to all those impacted by a decision made by the system:

i. This should include communicating how and why the decision was made, and how other available information or alternative outcomes were considered in reaching a decision.[65]

4 Implement meaningful public transparency for high risk AI systems

65 This proposal was originally made by Melanie Fink (2021), 'The EU Artificial Intelligence Act and Access to Justice': https://eulawlive.com/op-ed-the-eu-artificial-intelligence-act-and-access-to-justice-by-melanie-fink/.

To ensure greater public oversight of AI systems, the existing framework must be complemented by substantive mechanisms for transparency, such that AI systems in use are discoverable by oversight bodies, civil society and individuals.

a) **Ensure meaningful public transparency by requiring registration in the EU database (article 60) of all high risk AI systems**, and potentially **all AI systems to which people are subject that are put into use. This would enable individuals and civil society to access information about AI systems in operation:**

   i. The responsible authority or entity for deploying the high risk AI system should be listed with a contact point;

   ii. This should include information as to the stated purpose of the AI system in clear terms for individuals to understand.

b) **Ensure the inclusion of 'instructions for use' for AI systems in law enforcement and migration, asylum and border control management in the public database as per Annex III, points 1, 6 and 7. Remove the exemption contained within Annex VIII, point 11.**

c) **Require providers to include access to the conformity assessment alongside the instructions for use as per article 13(2)-(3) in the public database under article 60.**

d) **Require providers to provide more thorough details about the system to the users as part of article 13(3). This must include:**

   i. Information relating to the weightings and criteria relevant to choices in automated decision making systems;

   ii. An explanation of the fundamental assumptions and decisions informing the design of the AI system;

   iii. Ensure that information regarding the accuracy of the system under article 13(3)(ii) is precise, allowing the user to objectively assess whether the AI sytem is fit for purpose.

e) **Remove the exemptions in article 52 relating to the transparency of AI systems used for detection and prevention of criminal offences (as argued by the EDPB and EDPS) and for the prosecution of people. When AI systems under article 52**

are used for investigation, suspects should be notified 'post factum'.

5 Facilitate accountability: Include oversight and enforcement infrastructures that work for people

Lastly, the following proposals are designed to ensure that those harmed by the systems regulated under the AIA are able to contest and seek remedies. Further, there must be more independence for the European AI board and more distributed scope of governance funtions.

a) Ensure a cohesive national enforcement structure:

i. Following the recomemndation of the EDPS and EDPB, national Data Protection Authorities (DPAs) should be the designated national supervisory authorities under the act, with a stated duty to work with other relevant enforcement authorities in evaulation and monitoring;

ii. Ensure sufficient resources for national supervisory authorities in order to evaluate AI systems but also to respond and administer complaints.

b) Include flagging and redress mechanisms allowing individuals and collectives to contest and seek redress for all AI systems that cause harm and threaten fundamental rights:

i. This could include a flagging mechanism for those potentially impacted by an AI system to trigger national supervisory authorities' evaluative action under article 67;

ii. This duty to evaluate for fundamental rights risks should not be limited only to high risk systems, but any AI system once the national supervisory authority has received a complaint;

iii. An explicit individual and collective redress mechanism must be introduced specifically to apply to those subjected by all AI systems, in particular noting that many such stand-alone systems are not covered by consumer mechanisms for collective redress.

c) Implement a more democratic governance infrastructure, with greater independence for the European AI board:

i. Ensure that the mandate to make substantive updates to the legal framework (updates to high risk use cases, prohibitions) is held by a representative and democractically accountable European AI Board, not solely with the European Commission;

ii. Include within the structure of the AI Board representatives of social partners and civil society, in particular those representing marginalised groups.