16 January 2025

Human rights and justice must be at the heart of the upcoming Commission guidelines on the AI Act implementation

The following statement has been written collectively by the AI Act civil society coalition and the #ProtectNotSurveil coalition following the European Commission consultation on the AI Act prohibitions and AI system definition. Check out Access Now's answer to the consultation.

On 11 December 2024, the European Commission's consultation on its Artificial Intelligence (AI) Act guidelines closed. These guidelines will determine how those creating and using AI systems can interpret rules on the types of systems in scope, and which systems should be explicitly prohibited.

Since the final AI Act presents various grave loopholes when it comes to the protection of fundamental rights, particularly in the areas of policing and migration, it is important the guidelines clarify that fundamental rights are the central guiding basis to enable meaningful AI Act enforcement.

More specifically, **we urge the AI Office to ensure the upcoming guidelines on AI Act prohibitions and AI system definition include the following as a necessary basis for fundamental rights-based enforcement:**

- **Clarify that comparatively 'simple' systems are explicitly within scope of the AI system definition: such systems should not be considered out of scope of the AI Act just because they use less complicated algorithms.** We are concerned that developers might leverage the definition of AI and the classification of high-risk AI systems to bypass the obligations of the AI Act. For instance, transforming an AI system into a rule-based system could circumvent the regulations of the AI Act, while maintaining the same functionality and carrying the same risks. Hence, regulation must focus on potential harm, not just technical methods. The Dutch SyRI scandal is a clear example of a system that appeared simple and explainable, but which had devastating consequences for people's rights and lives, especially for racialised people and those with a migrant background.
- Prohibitions of systems posing an 'unacceptable' risk to fundamental rights are clarified to **prevent the weaponisation of technology against marginalised groups and the unlawful use of mass biometric surveillance.** Specifically, the guidelines should reflect:

• The ban on **social scoring must be clarified to include within its scope existing social scoring practices in Europe, especially in welfare and migration procedures.** To this end, the guidelines must: specify that "social behaviour" is interpreted broadly, as several elements can be used as risk indicators, such as 'unusual living arrangements' in the Danish automated welfare case; clarify that "personal or personality" characteristics also include proxy data (e.g. postcode in the Dutch child welfare scandal case) related to race, ethnicity, disability, socio-economic status, and other grounds; limit what can be considered data related to a social context to counter excessive and unlawful data collection, as well as sharing, and merging of datasets, which often exceeds the information necessary for assessing eligibility for benefits or individual risk. Finally, the guidelines must acknowledge the wide-ranging contexts — including in employment, education, welfare, policing, and migration — where social scoring practices are prevalent and should be prohibited.

• Notwithstanding the limited reach of the **ban on predictive policing** — which excludes **event and location-based predictions** **— the guidelines must clarify that predicting 'risk of committing a criminal offence' includes all systems that purport to predict a wide range of behaviours that are criminalised and have criminal law and administrative consequences.** As such, the guidelines should specify that systems making predictions about the likelihood of being registered in a police system (as was the case in the Dutch Prokid system) are within the scope of the prohibition, **as well as predictive systems used in the migration control**, if being irregular or being classified as presenting a risk to public security qualifies as criminal activity. In these cases, such systems must be covered by the ban as they amount to criminal risk assessments, and systems such as risk assessments included in ETIAS shall also be banned..

• The current ban on **non targeted scraping of facial images** leaves room for problematic loopholes. The guidelines must clarify that any derogation from the ban must be in line with the case law of the Court of Justice of the EU, and that any face scraped from the internet or CCTV footage must have a link to the commission of a crime. Otherwise, the facial images of innocent people could be scraped because they appear in the same CCTV footage as the commission of a crime. We further urge the Commission to prevent loopholes by deleting the proposed definition of a facial image database. This could create a loophole where systems like Clearview AI or PimEyes, which claim to store only biographical information or URLs and not the actual facial images, would fall outside of the prohibition.

- Civil society has highlighted the problematic legitimisation of so-called **"emotion recognition"** systems through the AI Act. These tools suffer from fundamental flaws in their scientific underpinnings, are highly intrusive, and could lead to serious life-threatening consequences for persons subjected to these tools, especially in contexts such as policing, migration, and to health and safety applications. The guidelines must therefore clearly establish the difference between legitimate medical equipment (e.g. heart monitors) and systems that aim to infer or identify people's emotions (e.g. "aggression detection"). The latter cannot be claimed as health and safety tools and be exempt from the prohibition.
- For the prohibition on **biometric categorisation** it is important that the guidelines clarify that the ban applies also when deductions are made on "ethnicity" and "gender identity" as they could lead to inferences respectively on "race" and "sex life or sexual orientation", which are in the scope of the ban. With regards to the exception, the guidelines should correct the wrong suggestion made in the consultation text, that labeling or filtering is permissible in the context of law enforcement among others, whereas the AI Act text is clear that this exception applies only in the law enforcement context.
- The guidelines must strengthen the language on **remote biometric identification (RBI)** to prevent forms of biometric mass surveillance. Specifically, it must specify that the development of real-time RBI systems for export should be considered within the scope of the ban; that the "without their active involvement" clause does not mean that law enforcement actors can place posters or flyers in the surveilled space and claim that people are actively involved and therefore the definition does not apply. Finally, while we continue to call for a full ban on retrospective RBI by private and public actors, we urge that the "significant delay" clause should be at a minimum of 24 hours after capture.

- **Concerning the interplay with other Union law, the guidelines must ensure that human rights law, in particular the EU Charter of Fundamental Rights, are the central guiding basis for the implementation** and that all AI systems must be viewed within the wider context of discrimination, racism, and prejudice. For this reason, the guidelines must emphasise that the objective of the prohibitions is to serve a preventative purpose, and therefore must be interpreted broadly in the context of harm prevention.

Lastly, we note the shortcomings of the Commission's consultation process: notably, the lack of advanced notice and a short time frame for submissions, no publication of the draft guidelines to enable more targeted and useful feedback, lack of accessible formats for feedback, strict character limits on complicated, and at times leading, questions which required elaborate answers — for example, Question 2 on the definition of AI systems only asked for examples of AI systems that

should be excluded from the definition of AI, hence allowing to narrow and not widen the definition of AI.

We urge the AI Office and the European Commission to ensure that all future consultations related to the *AI Act* implementation, both formal and informal, give a meaningful voice to civil society and impacted communities and that our views are reflected in policy developments and implementation.

As civil society organisations actively following the AI Act, we expect that the AI Office will ensure a rights-based enforcement of this legislation, and will prioritise human rights over the interests of the AI industry.

Signatories

Organisations

- Access Now
- AlgorithmWatch
- Amnesty International
- ARTICLE 19
- Border Violence Monitoring Network (BVMN)
- Danes je nov dan
- Electronic Frontier Norway (EFN)
- Electronic Privacy Information Center (EPIC)
- Equinox Initiative for Racial Justice
- EuroMed Rights
- European Center for Not-for-Profit Law (ECNL)
- European Digital Rights (EDRi)
- European Disability Forum (EDF)
- European Network Against Racism (ENAR)
- Federación de Consumidores y Usuarios CECU
- Homo Digitalis
- Lafede.cat – Organitzacions per la Justícia Global
- Panoptykon Foundation
- Privacy International
- SHARE Foundation
- Statewatch

Individuals

- Associate prof. Øyvind Hanssen (UiT Arctic University of Norway)
- Douwe Korff, Emeritus Professor of International Law
- Dr. Derya Ozkul (University of Warwick)
- Prof. Jan Tobias Muehlberg (Universite Libre de Bruxelles).