Dear XXX,

[XX June 2021]

In April 2021, the European Commission published a proposal for a new set of rules governing the development, marketing and use of artificial intelligence (AI).[1] As the world's first binding legal framework on AI, the proposal represents a significant turning point and marks a growing recognition by States of the urgent need to address the negative impacts of these emerging technologies on human rights and society. However, as it stands the proposed regulation falls far short of the measures that will be required to meaningfully protect people from harmful AI systems in the EU and globally.

This letter sets out Amnesty International's initial concerns around the primary gaps and shortcomings in the current proposal, and recommendations for strengthening human rights protections in the final regulation. The organisation is conducting a comprehensive analysis of the regulation and will issue a detailed public position in future.

The Commission and EU member states must seize this opportunity to lead the way and set a high bar for AI regulation that truly protects people's rights in the digital age.

1. PROHIBITED ARTIFICIAL INTELLIGENCE PRACTICES

Title II of the proposal sets out a list of prohibited AI practices. While it is welcome that the EU recognises that certain use cases of AI pose an unacceptable risk to fundamental rights, the proposal does not go far enough in prohibiting the most harmful AI systems.

Facial recognition and remote biometric technologies

Facial recognition and remote biometric technologies used for identification purposes are incompatible with human rights and should be subject to an outright prohibition in the EU's AI regulation.[2] These technologies are a mode of mass surveillance and as such represent a disproportionate interference with the right to privacy, as well as posing unacceptable risks of fuelling discrimination and hampering the right to peaceful assembly and the right to freedom of expression.

---

[1] EU Commission, *Proposal for a Regulation laying down harmonised rules on artificial intelligence*, https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence

[2] On 7 June 2021, Amnesty International and more than 170 organisations worldwide issued an open letter calling for a ban on biometric surveillance. See: https://www.amnesty.org/en/latest/news/2021/06/amnesty-international-and-more-than-170-organisations-call-for-a-ban-on-biometric-surveillance/

The EU's proposed prohibition of only 'real-time' remote biometric identification systems by law enforcement in public spaces, with even this use being permitted in certain circumstances, is too narrow and falls far short of an outright ban. Prohibiting only 'real-time' uses enables in principle the use by law enforcement of harmful 'post' remote biometric identification systems such as Clearview AI's biometric photo database, which the Hamburg Data Protection Authority has ruled is illegal under EU data protection law.[3] It would also mean significant gaps remain in banning private uses of 'real-time' and 'post' facial recognition, and many other forms of remote biometric surveillance.

Furthermore, enabling law enforcement to use 'real-time' remote biometric technologies – even only in limited circumstances – means that the technology and associated infrastructure will still have to be adopted and put in place. This creates a serious risk of law enforcement misuse of technologies which are inherently discriminatory and incompatible with human rights. To make the prohibition meaningful in practice, the exemptions set out in Article 5(2) for law enforcement authorities should be removed.

In response to the EU's proposal, the EU's European Data Protection Supervisor has also expressed regret that the regulation does not go further, and calls for a ban on remote biometric identification in public space "whether these are used in a commercial or administrative context, or for law enforcement purposes".[4]

Emotion recognition and biometric categorisation

Emotion recognition systems and biometric categorisation systems are not included in the list of prohibited AI practices under title II but are instead – partially – included in the category of 'high-risk' AI systems under Article 6(2). This is disappointing given the unacceptably high risks such systems pose to human rights in many contexts. Moreover, under Annex III, emotion recognition systems only qualify as 'high-risk' when used by law enforcement and immigration authorities. For all other uses, the only explicit obligation the AI Regulation includes is to inform people that they are exposed to an emotion recognition system. This is not enough.

Emotion recognition and biometric categorisation systems are highly intrusive practices that purport to infer sensitive characteristics about people with high risks of discriminatory outcomes. The potential use cases of such systems are very broad, ranging from commercial use to selection procedures to law enforcement. For example, Spotify has already patented technology that would listen to people's private conversations and

[3] Hamburg Commissioner for Data Protection and Freedom of Information, proceedings against Clearview AI, 27 January 2021, https://noyb.eu/sites/default/files/2021-01/545_2020_Anh%C3%B6rung_CVAI_ENG_Redacted.PDF

[4] European Data Protection Supervisor, *Artificial Intelligence Act: a welcomed initiative, but ban on remote biometric identification in public space is necessary*, 23 April 2021 https://edps.europa.eu/press-publications/press-news/press-releases/2021/artificial-intelligence-act-welcomed-initiative_en

recognize their emotions in order to help recommend its users content such as podcasts and songs.[5] The Chinese government is testing emotion recognition systems on Uyghurs in police stations in Xinjiang purportedly to assess their state of mind, including negative or anxious ones, reportedly intended to reach conclusions "without any credible evidence".[6]

AI systems that claim to be able to determine people's emotional states are also based on fundamentally flawed assumptions that have a highly questionable scientific basis. For example, in 2019 a major scientific metastudy found "insufficient evidence" for the view that emotions can be inferred from facial movements.[7] These systems replicate the logic of discredited eugenicist theories of phrenology and physiognomy, thereby perpetuating discrimination.

In Amnesty International's view, the only way to protect human rights and to prevent harmful practices like ethnic profiling or discrimination against minorities is to prohibit emotion recognition and biometric categorisation systems in contexts where they cannot be used in line with human rights.

Other prohibited AI practices

Under Article 5(1)(a) and (b), the regulation prohibits AI systems that manipulate people or exploit the vulnerabilities of specific groups in a manner that causes or is likely to cause physical or psychological harm. The current wording in the proposal is very broad and unclear, leaving room for (mis)interpretation. The proposal must articulate a much clearer definition if this provision is to be effectively implemented. Furthermore, while Article 5(1)(b) prohibits the exploitation of specific groups of people that are vulnerable due to their age, physical or mental disability, the proposal fails to capture vulnerabilities that go beyond these categories.

It is welcome that Article 5(1)(b) prohibits AI systems used for social scoring, but is too narrow in only prohibiting such practices when used by public authorities or on their behalf. The High-Level Expert Group on AI set up by the European Commission found that *any* form of citizen scoring – by public authorities or private actors - can endanger

[5] Access Now, *Spotify, don't spy: global coalition of 180+ musicians and human rights groups take a stand against speech-recognition technology*, 19 May 2021, https://www.accessnow.org/spotify-spy-tech-coalition/

[6] Jane Wakefield, BBC News, *AI emotion-detection software tested on Uyghurs*, 26 May 2021, https://www.bbc.com/news/technology-57101248

[7] Lisa Feldman Barrett et al*, Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements,* in Psychological Science in the Public Interest, Volume: 20 issue: 1, 1 July 2019, https://journals.sagepub.com/doi/10.1177/1529100619832930

autonomy and the principle of non-discrimination and called for an outright ban.[8] The Regulation should go further to also prohibit social scoring when done by private entities in contexts that impact human rights.

Furthermore, social scoring practices are banned if they lead to detrimental treatment that is "unjustified or disproportionate" to people's "social behaviour or its gravity". However, there is no definition of what is considered "good" or "bad" social behaviour or what is "justified" or "proportionate". This leaves the door open to abuses and has the potential to aggravate discriminatory practices on the basis of race, gender, religion, sexual orientation, and other protected characteristics, for instance in countries where people belonging to LGBTI communities are not socially accepted. It also leaves a loophole for states to interpret "bad" social behaviour to include the participation in protests or legitimate exercise of free expression, in a context where EU authorities have cracked down on protests and stifled dissent.[9]

2. CONFORMITY ASSESSMENT PROCEDURES

A fundamental concern with the draft AI regulation is the over-reliance on the providers of AI systems to self-assess their compliance with the regulation. Under Chapter 5 of Title III, the vast majority of AI systems designated to be 'high-risk' are only subject to a limited conformity assessment procedure based on internal control, without any third-party conformity assessment. This amounts to a very weak safeguard against human rights abuses, especially given the recognised high-risk nature of such systems. This ranges from AI used for evaluating creditworthiness, work performance or the eligibility for public benefits, to crime prediction AI and AI used to examine visa applications. Such uses of AI pose real dangers to human rights yet are left to self-certification.

The only high-risk systems that do require a more stringent conformity assessment including the involvement of an independent third party are AI systems intended to be used as safety components, and AI systems for remote biometric identification. In the latter case, given that as set out above such use should properly be subject to an outright prohibition, the more stringent conformity assessment measures are insufficient.

The EU Commission and EU member states must ensure that competent administrative and judicial authorities have the mandate to enforce compliance with the regulation, and put in place third-party verification for all high-risk AI systems.

---

[8] High-Level Expert Group on AI, *Ethics Guidelines for Trustworthy Artificial Intelligence*, p 34, and *Policy and investment recommendations for trustworthy Artificial Intelligence,* p 20*,* 8 April 2019, https://digital-strategy.ec.europa.eu/en/library/policy-and-investment-recommendations-trustworthy-artificial-intelligence

[9] Marco Perolini, Amnesty International, *Will our right to protest ever be fully returned?,* 29 September 2020 https://www.amnesty.org/en/latest/news/2020/09/will-our-right-to-protest-ever-be-fully-returned/

3. MEASURES IN SUPPORT OF INNOVATION: AI REGULATORY SANDBOXES

Amnesty International is furthermore troubled by the provisions set out in Title V of the AI regulation that permit personal data to be used for developing and testing "innovative AI systems" within controlled AI regulatory "sandboxes".

Human rights apply everywhere. Regardless of what projects are called –experiments, innovation try-outs, or sandboxes – the development of all AI systems and operations must respect human rights. However, by allowing personal data "lawfully collected for other purposes" to be processed in the AI regulatory sandbox, Article 54(1) of the draft proposal is in violation of the fundamental right to protection of personal data as laid down in the Charter of Fundamental rights of the European Union (CFREU), in particular Art. 8(2) of CFREU, which requires that personal data may only be processed for specified purposes. This requirement means in practice that the processing purposes should be identified precisely and fully in order to assist an average data subject, without expert legal or technical knowledge, in the assessment of what processing of data is and is not included in the processing operation.[10] Purpose limitation is also a requirement under Article 5(1)(b) of the General Data Protection Regulation (GDPR). The very nature of AI regulatory sandboxes as laid down in Title V of the draft regulation is conflicting with the purpose specification requirement laid out in the CFREU and the GDPR. Including a provision such as art. 54(1) in the proposal will lead to innovation at the cost of human rights, an outcome that can never be permissible.

5. OVERSIGHT

Oversight of AI systems must be binding and include at all times – for all types of AI systems, regardless of the risk that is assessed at first inspection – an oversight body with the power to review the risks and adverse impacts of the system on all human rights, including economic, social and cultural rights, the right to good governance, and equality and non-discrimination. Oversight must be conducted prior, during and after the deployment of the AI system. A mandatory human rights impact assessment for AI systems must be carried out by governments and private entities as part of their human rights due diligence to ensure transparency and accountability. Governments must be required to properly and effectively resource oversight bodies to the scale on which AI is deployed in their jurisdiction. The oversight body should also have access to the training, development and validation data and models that are deployed.

[10] Article 29 Working Party, Opinion on Purpose Limitation, 2013, WP 203, p. 39; and Article 29 Working Party, Opinion 02/2013 on apps on smart devices, 2013, WP 202, p. 17