EU's AI Act fails to set gold standard for human rights

Wednesday 3 April, 2024

For the last three years, we have worked in coalition as a broad range of digital, human rights and social justice groups to demand that artificial intelligence (AI) works for people, prioritising the protection of fundamental human rights. We have put forward our collective vision for an approach where "human-centric" is not just a buzzword, where people on the move are treated with dignity, and where lawmakers are bold enough to draw red lines against unacceptable uses of AI systems.

Following a gruelling negotiation process, EU institutions are expected to conclusively adopt the final AI Act in April 2024. But while they celebrate, we take a much more critical stance, highlighting the many missed opportunities to make sure that our rights to privacy, equality, non-discrimination, the presumption of innocence and many other rights and freedoms are protected when it comes to AI. Here's our round-up of how the final law fares against our collective demands.

Please note that this analysis is based on the latest available version of the AI Act text, dated 6 March 2024. There may still be small changes made before the law's final adoption.

First, we called on EU lawmakers to empower affected people by upholding a framework of accountability, transparency, accessibility and redress. How did they do?

Some accessibility barriers have been broken down, but more needs to be done:

- Article 16 (ja) of the AI Act fulfills our call for accessibility by stating that high-risk AI systems must comply with accessibility requirements. However, we still believe that this should be extended to apply to low and medium risk AI systems as well, in order to ensure that the needs of people with disabilities are central in the development of all AI systems which could impact them.

More transparency about certain AI deployments, but big loopholes for the private sector and security agencies:

- The AI Act establishes a publicly-accessible EU database to provide transparency about AI systems that pose higher risks to people's rights or safety. While originally only providers of high-risk AI systems were subject to transparency requirements, we successfully persuaded decision-makers that deployers of AI system – those who actually use the system – shall also be subject to transparency obligations.
- Those providers and deployers will be subject to transparency obligations who put on the market or use AI systems in high-risk areas – such as in the areas of employment and education – as designated by Annex III. Providers will be required to register their high-risk system in the database and to enter information about it such as the description of its intended purpose, concise description of the information used by the system and its operating logic. Deployers of high risk AI systems who are public authorities – or those acting on their behalf – will be obliged to register the use of the system. They will be required to enter information in the database such as a summary of the findings of a fundamental rights impact assessment (FRIA) and a summary of the data protection impact assessment. However, deployers of high-risk AI systems in the private sector area will not be required to register the use of high-risk systems – another critical issue;
- The major shortcoming of the EU database is that negotiators agreed on a carve-out for law enforcement, migration, asylum and border control authorities. Providers and deployers of high-risk systems in these areas will be requested to register only a limited amount of information, and only in a non-publicly accessible section of the database. Certain important pieces of information, such as the training data used, will not be disclosed at all. This will prevent affected people, civil society, journalists, watchdog organisations and academics to exercise public scrutiny in these high-stake areas which are prone to fundamental rights violation and hold them accountable.

Fundamental rights impact assessments are included, but concerns remain about how meaningful they will be:

- We successfully convinced EU institutions of the need for fundamental rights impact assessments (FRIAs). However, based on the final AI Act text, we have doubts whether it will actually prevent human rights violations and serve as a meaningful tool of accountability. We see three primary shortcomings:
    - Lack of meaningful assessment and the obligation to prevent negative impacts: while the new rules require deployers of high-risk AI systems to list risks of harm to people, there is no *explicit* obligation to assess whether these risks are acceptable in light of fundamental rights law, nor to prevent them wherever possible. Regrettably, deployers only have to specify which measures will be taken once risks materialise, likely once the harm has already been done;
    - No mandatory stakeholder engagement: the requirement to engage external stakeholders, including civil society and people affected by AI, in the assessment process was also removed from the article at the last stages of negotiations. This means that civil society organisations will not have a direct, legally-binding way to contribute to impact assessments;
    - Transparency exceptions for law enforcement and migration authorities: while in principle, deployers of high-risk AI systems will have to publish the summary of the results of FRIAs, this will not be the case for law enforcement and migration authorities. The public will not even have access to mere information that an authority uses a high-risk AI system in the first place. Instead, all information related to the use of AI in law enforcement and migration will only be included in a non-public database, severely limiting constructive public oversight and scrutiny. This is a very concerning development as, arguably, the risks to human rights, civic space and rule of law are the most severe in these two areas. Moreover, while deployers are obliged to notify the relevant market surveillance authority of the outcome of their FRIA, there is an exemption to comply with this obligation to notify for 'exceptional reasons of public security'. This excuse is often misused as a justification to carry on disproportionate policing and border management activities.

When it comes to complaints and redress, there some remedies, but no clear recognition of "affected person":

- Civil society has advocated for robust rights and redress mechanisms for individuals and groups affected by high-risk AI systems. We have demanded the creation of a new section titled 'Rights of Affected Persons', which would delineate specific rights and remedies for individuals impacted by AI systems. However, the section has not been created but instead, we have a "remedies" chapter that includes only some of our demands;
- This chapter of remedies includes the right to lodge complaints with a market surveillance authority, but lacks teeth, as it remains unclear how effectively these authorities will be able to enforce compliance and hold violators accountable. Similarly, the right to an explanation of individual decision-making processes,

particularly for AI systems listed as high-risk, raises questions about the practicality and accessibility of obtaining meaningful explanations from deployers. Furthermore, the effectiveness of these mechanisms in practice remains uncertain, given the absence of provisions such as the right to representation of natural persons, or the ability for public interest organisations to lodge complaints with national supervisory authorities.

The Act allows a double standard when it comes to the human rights of people outside the EU:

- The AI Act falls short of civil society's demand to ensure that EU-based AI providers whose systems impact people outside of the EU are subject to the same requirements as those inside the EU. The Act does not stop EU-based companies from exporting AI systems which are banned in the EU, therefore creating a huge risk of violating rights of people in non-EU countries by EU-made technologies that are essentially incompatible with human rights. Additionally, the Act does not require exported high-risk systems to follow the technical, transparency or other safeguards otherwise required when AI systems are intended for use within the EU, again risking the violation of rights of people outside of the EU by EU-made technologies.

Second, we urged EU lawmakers to limit harmful and discriminatory surveillance by national security, law enforcement and migration authorities. How did they do?

The blanket exemption for national security risks undermining other rules:

- The AI Act and its safeguards will not apply to AI systems if they are developed or used solely for the purpose of national security, and regardless of whether this is done by a public authority or a private company. This exemption introduces a significant loophole that will automatically exempt certain AI systems from scrutiny and limit the applicability of human rights safeguards envisioned in the AI Act;
- In practical terms, it would mean that governments could invoke national security to introduce biometric mass surveillance systems, without having to apply any safeguards envisioned in the AI Act, without conducting a fundamental rights impact assessment and without ensuring that the AI system meets high technical standards and does not discriminate against certain groups;
- Such a broad exemption is not justified under EU treaties and goes against established jurisprudence of the European Court of Justice. While national security can be a justified ground for *exceptions* from the AI Act, this has to be assessed case-by-case, in line with the EU Charter of Fundamental Rights. The adopted text, however, makes national security a largely digital rights-free zone. We are concerned about the lack of clear national-level procedures to verify if the national security threat invoked

by the government is indeed legitimate and serious enough to justify the use of the system and if the system is developed and used with respect for fundamental rights. The EU has also set a worrying precedent regionally and globally; broad national security exemptions have now been introduced in the newly-adopted Council of Europe Convention on AI.

Predictive policing, live public facial recognition, biometric categorisation and emotion recognition are only partially banned, legitimising these dangerous practices:

- We called for comprehensive bans against any use of AI that isn't compatible with rights and freedoms – such as proclaimed AI 'mind reading', biometric surveillance systems that treat us as walking bar-codes, or algorithms used to decide whether we are innocent or guilty. All of these examples are now partially banned in the AI Act, which is an important signal that the EU is prepared to draw red lines against unacceptably harmful uses of AI;
- At the same time, all of these bans contain significant and disappointing loopholes, which means that they will not achieve their full potential. In some cases, these loopholes risk having the opposite effect from what a ban should: they give the signal that some forms of biometric mass surveillance and AI-fuelled discrimination are legitimate in the EU, which risks setting a dangerous global precedent;
- For example, the fact that emotion recognition and biometric categorisation systems are prohibited in the workplace and in education settings, but are still allowed when used by law enforcement and migration authorities, signal that the EU's will to test the most abusive and intrusive surveillance systems against the most marginalised in society;
- Moreover, when it comes to live public facial recognition, the Act paves the way to legalise some specific uses of these systems for the first time ever in the EU – despite our analysis showing that all public-space uses of these systems constitute an unacceptable violation of everyone's rights and freedoms.

The serious harms of retrospective facial recognition are largely ignored:

- When it comes to retrospective facial recognition, this practice is not banned *at all* by the AI Act. As we have explained, the use of retrospective (post) facial recognition and other biometric surveillance systems (called 'remote biometric identification', or 'RBI' in the text) are just as invasive and rights-violating as live (real-time) systems. Yet the AI Act makes a big error in claiming that the extra time for retrospective uses will mitigate possible harms.;
- While several lawmakers have argued that they managed to insert safeguards, our analysis is that the safeguards are not meaningful enough and could be easily circumvented by police. In one place, the purported safeguard even suggests that simple the suspicion of any crime having taken place would be enough to justify the use of a post RBI system – a lower threshold than we currently benefit from now under EU data protection law.

People on the move are not afforded the same rights as everyone else, with only weak – and at times absent - rules on the use of AI at borders and in migration contexts:

- In its final version, the EU AI Act sets a dangerous precedent for the use of surveillance technology against migrants, people on the move and marginalised groups. The legislation develops a separate legal framework for the use of AI by migration control authorities, in order to enable the testing and the use of dangerous surveillance technologies at the EU borders and disproportionately against racialised people;
- None of the bans meaningfully apply to the migration context, and the transparency obligations present ad-hoc exemptions for migration authorities, allowing them to act with impunity and far away from public scrutiny;
- The list of high-risk systems fails to capture the many AI systems used in the migration context, as it excludes dangerous systems such as non-remote biometric identification systems, fingerprint scanners, or forecasting tools used to predict, interdict, and curtail migration;
- Finally, AI systems used as part of EU large-scale migration databases (e.g. Eurodac, the Schengen Information System, and ETIAS) will not have to be compliant with the Regulation until 2030, which gives plenty of time to normalise the use of surveillance technology.

Third, we urged EU lawmakers to push back on Big Tech lobbying and address environmental impacts. How did they do?

The risk classification framework has become a self-regulatory exercise:

- Initially, all use cases included in the list of high-risk applications would have had to follow specific obligations. However, as a result of heavy industry lobbying, providers of high-risk systems will be now able to decide if their systems is high-risk or not, as an additional "filter" was added into that classification system;
- Providers will still have to register sufficient documentation in the public database to explain why they don't consider their system to be high-risk. However, this obligation will not apply when they are providing systems to law enforcement and migration authorities. This will paving the way for the free and deregulated procurement of surveillance systems in the policing and border contexts.

The Act takes only a tentative first step to address environmental impacts of AI:

- We have serious concerns about how the exponential use of AI systems can have severe impacts on the environment, including through resource consumption, extractive mining and energy-intensive processing. Today, information on the

environmental impacts of AI is a closely-guarded corporate secret. This makes it difficult to assess the environmental harms of AI and to develop political solutions to reduce carbon emissions and other negative impacts;

- The first draft of the AI Act completely neglected these risks, despite civil society and researchers repeatedly calling for the energy consumption of AI systems to be made transparent. To address this problem, the AI Act now requires that providers of GPAI models that are trained with large amounts of data and consume a lot of electricity must document their energy consumption. The Commission now has the task of developing a suitable methodology for measuring the energy consumption in a comparable and verifiable way;
- The AI Act also requires that standardised reporting and documentation procedures must be created to ensure the efficient use of resources by some AI systems. These procedures should help to reduce the energy and other resource consumption of high-risk AI systems during their life cycle. These standards are also intended to promote the energy-efficient development of general-purpose AI models;
- These reporting standards are a crucial first step to provide basic transparency about some ecological impacts of AI, first and foremost the energy use. But they can only serve as a starting point for more comprehensive policy approaches that address all environmental harms along the AI production process, such as water and minerals. We cannot rely on self-regulation, given how fast the climate crisis is evolving.

What's next for the AI Act?

The coming year will be decisive for the EU's AI Act, with different EU institutions, national lawmakers and even company representatives setting standards, publishing interpretive guidelines and driving the Act's implementation across the EU's member countries. Some parts of the law - the prohibitions - could become operational as soon as November. It is therefore vital that civil society groups are given a seat at the table, and that this work is not done in opaque settings and behind closed doors.

We urge lawmakers around the world who are also considering also bringing in horizontal rules on AI to learn from the EU's many mistakes outlined above. A meaningful set of protections must ensure that AI rules truly work for individuals, communities, society, rule of law, and the planet.

While this long chapter of lawmaking is now coming to a close, the next chapter of implementation – and trying to get as many wins out of this Regulation as possible - is just beginning. As a group, we are drafting an implementation guide for civil society, coming later this year. We want to express our thanks to the entire AI core group, who have worked tirelessly for over three years to analyse, advocate and mobilise around the EU AI Act. In particular, we thank the work, dedication and vision of Sarah Chander, of the Equinox Racial Justice Institute, for her leadership of this group in the last three years.

To learn more about our coalition's views and analysis of the final AI Act, check out the following resources:

- Access Now, 'The EU AI Act: a failure for human rights, a victory for industry and law enforcement': https://www.accessnow.org/press-release/ai-act-failure-for-human-rights-victory-for-industry-and-law-enforcement/
- Amnesty International: 'EU: Artificial Intelligence rulebook fails to stop proliferation of abusive technologies': https://www.amnesty.org/en/latest/news/2024/03/eu-artificial-intelligence-rulebook-fails-to-stop-proliferation-of-abusive-technologies/
- European Disability Forum: 'AI Act agreement: partial win on accessibility': https://www.edf-feph.org/ai-act-agreement-partial-win-on-accessibility/
- #ProtectNotSurveil: A dangerous precedent: how the EU AI Act fails migrants and people on the move: https://www.accessnow.org/press-release/joint-statement-ai-act-fails-migrants-and-people-on-the-move/
- Bits of Freedom, 'De AI-verordening is er, maar wij zijn sceptisch': https://www.bitsoffreedom.nl/2023/12/09/de-ai-verordening-is-er-maar-wij-zijn-sceptisch/
- AlgorithmWatch, 'EU Parliament votes on AI Act; member states will have to plug surveillance loopholes': https://algorithmwatch.org/en/eu-parliament-votes-on-ai-act/
- ARTICLE 19, 'EU: AI Act passed in Parliament fails to ban harmful biometric technologies': https://www.article19.org/resources/eu-ai-act-passed-in-parliament-fails-to-ban-harmful-biometric-technologies/

Contributors to this joint analysis:

- Ella Jakubowska, European Digital Rights (EDRi);
- Kave Noori, European Disability Forum (EDF);
- Mher Hakobyan, Amnesty International;
- Karolina Iwańska, European Center for Not-for-profit Law (ECNL);
- Kilian Vieth-Ditlmann, AlgorithmWatch;
- Nikolett Aszodi, AlgorithmWatch;
- Judith Membrives Llorens, Lafede.cat / Algorights;
- Caterina Rodelli, Access Now;
- Daniel Leufer, Access Now;
- Nadia Benaissa, Bits of Freedom;
- Ilaria Fevola, ARTICLE 19;
- With the work, support and collaboration of all the organisations in our coalition.