



An Avaaz brief on the role of fundamental human rights impact assessments in the EU AI Act

4 myths busted about Fundamental Rights Impact Assessments (FRIAs) for AI

Introduction

Artificial Intelligence (AI) is already causing [real harm](#) to EU citizens and their rights - some issues, like biometric surveillance, have rightly received a lot of attention, but the harms go well beyond oppressive surveillance. The cases we describe in this brief document detail AI's impacts in Germany and across Europe on our rights to non-discrimination in policing, equal rights to employment, and even our right to life, with worse cancer detection rates for people of colour and domestic violence reports ignored by the police because flawed AI told them they could.

The reason AI is harming human rights is simple: right now AI can be used without any assessment of whether it threatens our rights. That's where Fundamental Rights Impact Assessments (FRIA) come in. These can be conducted against criteria set by the law, to check if the way in which AI is deployed could affect the rights of those subject to its decisions. A **FRIA requires an AI user to consider, and if necessary consult with, people** likely to be the subjects of the decisions their AI makes - to ask the right questions, either of their own development team, or the person selling them the AI, as to whether or not it's been tested and audited to avoid infringing on the rights of those people.

This is the missing link in the chain of protection and the benefit is that we catch the harms **BEFORE** they have affected citizens, instead of conducting a giant experiment on them.

The need to act is urgent or we risk the AI Act failing in its determination to protect human rights. The EU AI Act text does half the job - companies that design and market AI (AI providers) are now under an obligation to respect human rights for the "intended use" of their AI. But the companies that purchase and use these technologies (AI deployers) - are not. Only the proposals by the European Parliament have any measures to close the huge loophole. They propose a complementary FRIA obligation for users of the **highest risk AI**, such as medical use, policing use, and security and border use - in other words, **uses where human rights are simply not negotiable** - [see here](#).

We urge all parties in the trilogue to discuss these ideas and find common ground to achieve comprehensive human rights protection in the AIA.

4 myths busted about Fundamental Rights Impact Assessments (FRIAs) for AI

There are many myths circulating about whether or not we need to protect our rights from AI - including the complacent argument that current legislation, such as GDPR, is sufficient to cover AI's potential harmful effects.

Myth 1: Our rights are already protected by existing EU law

No, they are not. Avaaz commissioned a [legal opinion from the authoritative European law agency, AWO](#). This legal opinion confirms that, although GDPR has some elements that support consideration of human rights, it would not cover the deployment of AI beyond a strict interpretation of data, privacy, and security. GDPR does not address broad risks to fundamental rights and completely fails where there is no processing of personal data. Widespread uses - such as dynamically determined pricing models - use fully anonymised, aggregated data, but these can result in discriminatory price rises, especially for those living in poorer areas, interfering with rights to non-discrimination and to access to services of general economic interest.

Other sector-specific legislation is outdated and piecemeal - for example, the current regulations for medical AI tools in the EU are the 2017/745 Medical Devices Regulation (MDR) and the 2017/746 In Vitro Diagnostic Medical Devices Regulation (IVDR), which were passed in 2017, written at a time when many aspects of AI, such as the identification of algorithmic biases, were not considered¹.

Myth 2: The costs will be a burden on the EU AI industry without significant benefit

This myth rests on the belief that the costs, or administrative burden of FRIAs, will hold back innovation. We don't agree on several counts:

- a) A FRIA by the user only needs to complement and build on the assessment previously done by the provider; and to deal with the actual context of use and likely affected individuals. Any FRIA obligation can be restricted to users of the **highest risk AI**, such as medical use, policing use, and security and border use.
- b) There are real, practical steps that can be taken after a risk is identified by a user. The user can either require the provider to address it, or change the nature of their use, for example, avoid inputting data likely to result in a specific bias. It is also possible for the user to conduct a pre-deployment audit of the AI to determine what forms of mitigation can be applied.²
- c) The savings in terms of brand damage and litigation will be significant. **In the current dominant commercial AI licensing model, current international jurisprudence indicates it would be impossible for EU providers to try to pass liability back up the chain if anything did go wrong, if the issue was reasonably not foreseeable by the provider, even if the user had conformed with "intended use" parameters set by the**

¹ Source: [Artificial intelligence in Healthcare, European Parliamentary Research Service](#).

² See The Danish Institute for Human Rights' [Human rights impact assessment guidance and toolbox](#): EU High-Level Expert Group on AI, Assessment List.

provider.³ This is a recipe for lengthy legal battles that would not only favour global tech giants with deep pockets, but leave the victims stuck for perhaps years not knowing who was actually responsible for what they suffered.

- d) So, for private companies, this is the basic due diligence that any investor would expect, potentially saving companies from the costs of legal liabilities, reputational damage, and operational disruptions caused by human rights violations. Investors agree. One investor group, representing €1.55 trillion in assets, told the Commission, “*Incorporating human rights impact assessment in product and service life cycle will minimize exposure to potential liability, resulting in more stable and sustainable financial returns.*”
- e) SMEs also see real value in the FRIA process. Avaaz conducted a poll in September 2023 with 800 Spanish citizens working in businesses defined as SMEs, of which one in three were currently using or developing AI in their work and three in four were planning to use AI within the next three years.

85.2% of respondees agreed that legal requirements to conduct a fundamental human rights assessment should be introduced for all companies before deploying high-risk AI systems.

The top three benefits they saw in the use of a fundamental human rights assessment from a list we gave them were to:

- Help with long-term planning on the costs and benefits of using AI - 44.5%
- Help companies understand the legal risks of using AI - 40.6%
- Help understand and address the risks caused to their customers by AI use - 38.6%

Only 12.7% of those polled saw no benefit.

Myth 3: Self regulation is the only way to avoid heavy-handed regulation

Self regulation alone is not enough; we can see what that looks like from the evidence of harms emerging across the EU. But we do not need to reinvent the wheel: many examples of FRIAs already exist, and data protection laws provide a model to ensure transparency and regulatory compliance. Impact assessments can be conducted by AI users and made available for public inspection - with fines to follow if their AI harms people.

The Commission, or an AI Office will have to help AI users, especially SMEs, understand how to do an assessment, that's only fair. 47.8% of the group we polled who agreed FRIAs were a good idea did say they would want help and guidance from the Commission or a regulator when they started to

³ <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW16RG2> In this independent assessment of what went wrong when Microsoft licenced an AI product to the NYPD - conducted under the UNGP business model, the assessors stated that - “*From the Assessors’ perspective, there is a **substantial difference between creating generic products that can be licensed to third parties and aiding in the creation of specific products.** Under the scenario in which a domestic law enforcement customer is merely licensing products from Microsoft, without more involvement from Microsoft in the development of the products or services, **the Assessors do not believe that Microsoft would or could be either causing or contributing to any adverse human rights impacts, as those terms are understood under the UNGPs.**”*

conduct FRIAs. But if we are committed to human rights in the age of AI, then everyone will need to learn these skills, just like we did when privacy laws on data were introduced. That's what SMEs have told us they want - support to meet the rights, not avoidance of that duty of care.

Myth 4 - Our rights are only impacted by a narrow set of AI uses - like surveillance

AI harms are already here now and are not just a future existential threat. They affect our rights to liberty, jobs, and even our right to life:

1) Right to non discrimination

a) Policing

[The use of AI in policing in Europe](#) is already affecting our vital right to liberty. In the Netherlands, the **TOP 600 AI automated risk profiling system wrongly labelled citizens as potential criminals. Just take a moment to imagine this.** You can be repeatedly arrested, your home searched, and be constantly followed by police as part of what the police call “ Very Irritating Policing” and you have no idea why. Imagine the injustice you'd feel. **This would have been prevented if the police had to do a Fundamental Rights Impact Assessment (FRIA) before implementing their AI system.** The young people targeted would have never had to suffer this injustice.

2) Right to participate equally in the economy

a) Access to credit

A low credit score from Germany's leading credit bureau, SCHUFA, can result in landlords refusing to rent you an apartment, banks rejecting your credit card application, and Internet service providers saying 'no' to a new contract. SCHUFA collects data on peoples' financial history – unpaid bills, credit cards, loans, fines and court judgments – and uses this information to calculate a SCHUFA score. An investigation by Der Spiegel into SCHUFA's scoring procedure has found a number of flaws in the scoring system, including that it may reinforce discrimination and that it has violated data access provisions. SCHUFA holds data on around 70 million people in Germany, nearly everyone in the country aged 18 or older. Critics allege that SCHUFA's predictive scoring procedure is highly intransparent; it is classified as a trade secret, so no one knows how it works and whether there are errors or biases built into the model or the software. A FRIA by a deployer in this case would need to build on work by a provider to make sure their use had sufficient safeguards in terms of the training of human moderation, and ability to challenge the decision making of the SCHFA programme. If after a FRIA they felt that the lack of transparency was a bar to effective and rights respecting implementation, they would choose another AI product to licence.

b) Employment

The 'My employability' tool in Croatia used the age of users' children as a parameter for employability⁴. When the input was '0-2' years of age, this was reported to have a negative impact on the predictions AI gave on the employability of women, whereas for men the parameter was not even displayed or taken into account. This not only has the potential effect of discouraging women with small children from (re)entering the labour market, but unless considered as part of a FRIA by an employment agency would have the potential of reducing offers of work to these women on the basis of the AI's ranking of their employability.

3) Right to life

Health care - The International Skin Imaging Collaboration (ISIC) archive is used to develop machine learning algorithms for skin cancer diagnosis. Bias and lack of transparency on the source of data sets as well as gaps in accountability have been identified, leading to warnings that AI now used in cancer treatments will lead to [worse outcomes if you have non-white skin](#). A FRIA would force the healthcare provider to look at the balancing rights of those who benefit from overall efficiencies in cancer diagnosis or speeding up access to life-saving care with the risks of missing skin cancer in patients of colour.

⁴ <https://www.equalitylaw.eu/downloads/5361-algorithmic-discrimination-in-europe-pdf-1-975>