

Ensuring strong equality and fundamental rights safeguards in the political agreement on the AI Act

Joint Equinet and ENNHRI Statement on EU Artificial Intelligence Act Trilogue

Summary

- **Article 6:** Objective high-risk classification - **revert to European Commission approach**
- **Article 68 a and d:** Create a **redress mechanism for the AI Office** and ensure an **effective right to lodge a complaint to the national supervisory authorities and the AI Office** for **individuals, associations representing individual and collective interests of those affected by AI systems**, as well as **Equality Bodies (NEBs) and National Human Rights Institutions (NHRIs)**, including in their own name and without an identifiable victim.
- **Ensure effective cooperation between AIA European enforcement mechanisms (AI Office) with existing public independent equality and fundamental rights enforcement bodies (NEBs and NHRIs): Create obligations for the AI office to notify and cooperate with Article 64(3) bodies** similarly to the way these obligations are created for national supervisory authorities through Article 64 and Article 65.
- **New governance framework for foundation models (FMs) and high-impact foundation models (HIFMs)** – include fundamental rights expertise and place stronger oversight on FMs and generative AI to prevent gaps in fundamental rights protection through a tiered approach.
- **Fundamental rights impact assessments (FRIAs) obligations for deployers** – keep **private sector** in the scope of FRIAs and provide clear, comprehensive and prescriptive guidance on **information to be collected**, as per the original European Parliament proposal.
- **Ban biometric and surveillance practices** risking discrimination.
- Place a clear and effective **ban on predictive policing**.

What are National Equality Bodies (NEBs) and National Human Rights Institutions (NHRIs)?

National Human Rights Institutions (NHRIs) and National Equality Bodies (NEBs) are independent public equality and fundamental rights authorities established by constitution or law across Europe. Equinet (the European Network of National Equality Bodies) and ENNHRI (the European Network of National Human Rights Institutions) represent the collective voices of over 60 independent national authorities in over 40 European states.

What do NHRIs and NEBs do? See end of the document for a brief explainer.

Recommendations for effective equality and fundamental rights protection through the AIA

1. “Ensure robust legal protection for high-risk systems” (Article 6)

Problem: Weakened protection of fundamental rights through the introduction of a subjective qualifier (filter) for high-risk classification which places self-assessment in the hands of AI providers without the required knowledge and experience of equality and fundamental rights risk assessment.

Recommendation: Revert to the original approach by the European Commission for objective high-risk classification based on Article 6.

Rationale:

- Significantly compromises legal certainty and foreseeability.
- Decisions with significant equality and fundamental rights implications are taken by AI providers without relevant knowledge and experience.
- Incentivizes AI providers to downplay the risks posed by their AI systems in order to reduce regulatory scrutiny.
- Negative opinion of the European Parliament’s legal service on this qualifier (filter).

2. Ensure cooperation between AI Office and existing national independent equality and fundamental rights enforcement bodies (NEBs and NHRIs)

Problem: The proposed AI Office lacks specific expertise on equality and fundamental rights enforcement yet will enforce against powerful foundation models with proven potential for systemic harm in relation to equality and fundamental rights.

Recommendation: Create obligations for the AI office to notify and cooperate with Article 64(3) bodies (dedicated independent public non-discrimination and fundamental rights authorities) similar to the way these obligations are created for national supervisory authorities through Article 64 and Article 65.

Rationale:

Ensure effective enforcement against the equality and fundamental rights risks of foundation models and high-impact foundation models by enabling the AI Office to benefit from the specialised expertise and extensive experience of NEBs and NHRIs, including with handling complaints and litigating cases.

3. Guarantee rights of individual and collective redress (Article 68 a and d) through the AI Office and the supervisory authorities: rights to lodge a complaint, representation and public interest action before the supervisory authorities and the AI Office

Problem: Diminished protection for rightsholders through lack of remedies in relation to the AI Office, limited scope of the right to lodge a complaint (Article 68 a) and the limitation of collective redress to the consumer protection framework (Article 68 d).

Recommendation: include the right to lodge a complaint with national supervisory authorities and the AI Office, by:

- Individuals
- Empowered associations representing individual and collective interests
- National authorities for equality and fundamental rights protection (NEBs and NHRIs), also in their own name and without identifiable victim as per Proposed Directives on Equality Bodies.

Rationale:

- Ensure coherence with remedies to individuals or groups under EU fundamental rights and equality law such as the ability of NEBs and NHRIs, and other public interest entities to bring cases to administrative bodies, such as the future AIA supervisory authorities.
- Due to the difficulties for victims of AI-enabled fundamental rights violations to know that they are subject to such violations and claim their rights, it is critical that NEBs, NHRIs and other relevant public interest organizations are explicitly given the power to submit complaints to supervisory authorities and the AI Office in their own name and without any identifiable victims.
- Consumer protection collective redress is not adequate to address AI-enabled breaches of fundamental rights.

4. New governance framework for foundation models (FMs) and high-impact foundation models (HIFMs)

Problem:

- Lack of fundamental rights and equality considerations in the proposed functions and activities of the AI Office.
- Lack of a requirement for expertise on equality and fundamental rights by independent vetted testers for HIFMs.
- Failure to address the fundamental rights risks of foundation models and generative AI – no independent risk assessment and no mitigation measures.

Recommendation:

- Ensure independent risk assessment and mitigation measures for all foundation models and generative AI and not only for HIFMs.
- Include explicit requirement for equality and fundamental rights expertise for oversight mechanisms for FM, HIFM and Generative AI.
- Independent risk assessment should include an obligation to seek inputs from a range of fundamental rights stakeholders, including NEBs, NHRIs, affected persons, and civil society.

Rationale: Ensure effective protection against and prevention of fundamental rights infringements by FMs, HIFMs and Generative AI.

5. Fundamental rights impact assessments (FRIAs) obligations for deployers

Problem: Weakening of the original European Parliament proposal on FRIAs will undermine possibilities for effective equality and fundamental rights enforcement. Under national and EU law, the right to non-discrimination and other fundamental rights are enforced in relation to the users (deployers) of AI systems. Independent public equality and fundamental rights authorities (NEBs and NHRIs) need evidence of impact, which in relation to AI systems is especially difficult to obtain due to the specific nature of AI. FRIAs are indispensable in identifying such impacts in a systemic, transparent and consistent manner.

Weakened effectiveness of equality and fundamental rights protection through excluding the expertise of NHRIs, NEBs, CSOs and other fundamental rights stakeholders.

Weakened enforcement through compromises on scope and content of FRIAs that limit this obligation only to the public sector.

Recommendation:

- Keep private sector in the scope of FRIAs.
- Provide clear and specific instructions on content and scope of collected information (follow European Parliament proposal on FRIAs).
- Keep European Parliament proposal on deployer's obligation by collecting inputs from NEBs, NHRIs, civil society and other stakeholders as part of the FRIA.

Rationale:

- Prevent fundamental rights infringements at the stage of deployment – risk assessment at development might be different from assessment at the deployment stage.
- The private sector is a major source of violations – might result in different protection levels for public versus private sector use. The Corporate Sustainability Due Diligence Directive (CSDDD) poses more limited requirements for the private sector leading to a two-track system of impact assessments on AI systems
- AI opacity and complexity – difficult identification of fundamental rights risks without clear and comprehensive information, need for clear, prescriptive and comprehensive guidance on information gathering requirements
- Obligation for inputs by NHRIs, NEBs and other stakeholders will ensure improved accuracy and reliability of the assessed impact and alignment with existing fundamental rights oversight

6. Ban biometric and surveillance practices risking discrimination

Problem: Concern about exemptions to bans on the use of biometrics to (in)discriminatively categorize persons into groups, "read" their emotions and subject them to live surveillance in public places.

Recommendation: All uses of AI systems that pose unacceptable risks to individuals and their fundamental rights should be prohibited.

Rationale: These exemptions are not proportionate and not backed by compelling evidence as regards the effectiveness of these uses for supporting public order and national security. Also does not meet the requirements for necessity and proportionality according to Article 52.1 of the Fundamental Rights Charter.

7. Place a clear and effective ban on predictive policing

Problem: A prohibition as part of the social scoring ban, that would entail a third prong to current Article 5(1)c, would not capture the harms of predictive policing, which do not stem from an intervention based on a solely automated prediction. These systems embed structural biases, leading to the disproportionate, over-policing of certain groups of people.

Recommendation: A standalone prohibition for predictive policing for criminal and administrative offences.

Rationale: A standalone prohibition is required to address that predictive policing undermines the right to an effective remedy and to a fair trial, the presumption of innocence, and the right to non-discrimination.

What do National Human Rights Institutions (NHRIs) and National Equality Bodies (NEBs) do?

NHRIs and NEBs already play a prominent role in national and international frameworks for human rights oversight, accountability and governance in relation to Artificial Intelligence (AI) systems, for example, through legal work on complaints, raising awareness, providing guidance to businesses developing AI systems, or engaging with governments to map and review the uses of AI systems by public authorities. Crucially, NHRIs and NEBs play a central role in the monitoring and implementation of international and regional human rights legal instruments. They are actively involved in contributing to ensuring oversight and accountability in the context of the EU legal framework, as reflected in the role of NEBs and NHRIs in the implementation of European Union Charter of Fundamental Rights and non-discrimination law, as well as their involvement in a range of fundamental rights and equality critical areas of EU law and policy such as rule of law monitoring mechanism, hate speech, employment and social rights (work life balance, pay transparency and freedom of movement of EU workers) and the protection of whistleblowers.