Center for AI and
Digital Policy

Center for AI and Digital Policy
(CAIDP.ORG)
Statement on the EU Council's General Approach
On the EU Artificial Intelligence Act (2021/0106(COD)[1]

13 February 2023

To: Swedish Presidency eva.sjogren@gov.se (Minister for EU Affair)
statsradsberedningen.registrator@gov.se (Prime Minister Office)
European Parliament president@ep.europa.eu
European Commission margrethe-vestager-contact@ec.europa.eu
ec-president-vdl@ec.europa.eu
Cc: Co-Rapporteurs in LIBE & IMCO Committees
brando.benifei@europarl.europa.eu
ioan-dragos.tudorache@europarl.europa.eu
Working Party on Telecommunications and Information Society
nina.bjoresten@gov.se ; david.kallstrom@gov.se

The Center for AI and Digital Policy (CAIDP) welcomes the Council's General Approach (The Approach) on the proposed Artificial Intelligence Act (The Proposal) and applauds the Czech Presidency for its important work in this matter. As we stated in our previous comments on the Proposal, "this initiative may be the single most important legal framework for the digital economy to ensure the protection of fundamental rights." [2] Our assessment of the Approach is favourable as the report brings further clarity to the EU AI Act. We appreciate the changes introduced by the earlier Presidency compromise texts, which expanded the transparency and accountability provisions.

We urge the Swedish Presidency to commit to completing the trilogue in the first half of 2023. In the absence of a legal framework, AI systems will be deployed without necessary safeguards, putting at risk public safety and health, and fundamental rights.

CAIDP is a global independent, research and education organisation. We train AI policy advocates and practitioners. We conduct comparative and longitudinal analysis of AI policies and practices in 50 countries in our *Artificial Intelligence and Democratic Values Index*.[3] Our aim is to promote a world where technology promotes broad social inclusion based on fundamental rights, democratic institutions, and the rule of law. We created a global resource page for the EU AI Act, and previously provided in-depth recommendations to the Council of the European Union, the European Parliament, and the European Commission.[4] In Washington DC, we have met with members of both the AIDA Committee and the LIBE Committee.

Below is our assessment of the Approach, and further recommendations for the Council of the European Union, European Parliament, and European Commission to adopt ahead of / during the trilogue process.

[1] Council of the European Union, General approach - Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (25 November 2022)

[2] CAIDP Statements to European Commission, European Parliament, European Council and AIDA Committee regarding the Draft EU AI Regulation of **April 20, 2021, July 28, 2021, 16 December, 2021 and May 10, 2022**. https://www.caidp.org/resources/eu-ai-act/

[3] CAIDP, AI and Democratic Values Index. https://www.caidp.org/reports/aidv-2021/

[4] CAIDP, EU AI Act, https://www.caidp.org/resources/eu-ai-act/

CAIDP commends the Council on the following:

- The Council's acknowledgement that "AI systems providing social scoring of natural persons by public authorities or by private actors **may lead to discriminatory outcomes and the exclusion of certain groups**. They may **violate the right to dignity and non-discrimination** and the values of **equality and justice**. Such AI systems evaluate or classify natural persons **based on their social behaviour in multiple contexts** or **known or predicted personal or personality characteristics**. The social score obtained from such AI systems may lead to the detrimental or unfavourable treatment of natural persons or whole groups thereof in social contexts, which are unrelated to the context in which the data was originally generated or collected or to a detrimental treatment that is disproportionate or unjustified to the gravity of their social behaviour." (Elements of Proposal (EP) 17)
- Addition of "vulnerabilities of a specific group of persons due to their specific social or economic situation" (Art 5b)
- Extension of the prohibition of using AI for social scoring also to private actors (Art 5c)
- Requirement from public authorities, agencies or bodies, or entities acting on their behalf, to register themselves in the EU database before using high-risk AI systems (Art 51)
- Extension of complaint mechanism to include natural persons (Art 63)
- Addition of life and health insurance to high-risk AI systems (Annex III)
- Addition of General-Purpose AI systems to the Approach (Art 4)

CAIDP recommends that the Council of the European Union, the European Parliament, and the European Commission adopt the following changes ahead of / during the trilogue process. These changes would ensure the Proposal is in alignment with the Council's stated goals, further protect fundamental rights and Union values, and bring consistency to the Approach.

PROPOSED PROHIBITIONS TO PSEUDOSCIENTIFIC & DISCRIMINATORY AI SYSTEMS & PRACTICES

As per the Approach, *AI systems should not evaluate or classify natural persons based on their social behaviour (EP 17).* AI systems using biometric data and categorising individuals into groups according to race, ethnicity, gender, political or sexual orientation, constitute discrimination under Article 21 of the Charter. Such categorization assumes genetic traits, diminishing universal rights and transforming social constructs such as race, ethnicity, gender, political or sexual orientation into 'objective' truths. Biometric categorization denies the most fundamental of human rights – the right to freely choose one's identity – and as such is contrary to the values enshrined in Article 2 of the TEU.

These systems also undermine the principle of 'presumption of innocence'.

Therefore, CAIDP recommends:

- **Require scientific validity:** If an AI system is not scientifically valid, it should be prohibited. This requirement should be added to existing requirements for accuracy, representativeness, robustness and cybersecurity.
- **Ban predictive policing:** AI claiming to predict the occurrence or reoccurrence of an actual or potential criminal offence - Annex III

- **Ban emotion recognition system:** AI systems claiming to identify or infer psychological states, emotions or intentions of natural persons on the basis of their biometric data - Annex III
- **Ban biometric categorisation system:** AI systems assigning natural persons to specific categories on the basis of their biometric data - Art 52

Predictive policing, Emotion recognition and Biometric categorization systems do not have scientific validity. Not explicitly banning pseudoscientific AI systems would legitimise these systems. **These systems provide the fundamentals / components of 'social scoring systems'** which the EU acknowledges as unacceptable.

CAIDP's review of country AI practices found that the clearest distinction between AI systems in authoritarian countries and AI systems in democratic countries is the use of facial recognition for mass surveillance.[5] Such indiscriminate ongoing surveillance is intended precisely to coerce social behaviour and to control populations.[6] Biometric recognition techniques are used against political protesters and religious minorities and will almost certainly be more widely deployed unless a clear prohibition is adopted.[7] Therefore, CAIDP recommends:

- **Ban biometric recognition** systems used for **mass surveillance**.[8] [9]
- **Ban predictive risk scoring** system: used in migration, asylum, and border control management to assess security or health risk. Such systems can be discriminatory and can violate the principle of presumption of innocence[10], human dignity and freedom to seek asylum - Annex III
- All of the above-mentioned bans should **apply to both public and private** entities.

REMOVAL OF EXCEPTIONS & EXCLUSIONS IMPACTING FUNDAMENTAL RIGHTS

The current narrow definition and exclusions for the prohibition on "the use of 'real-time' remote biometric identification systems in publicly accessible spaces by law enforcement authorities" will not effectively prohibit the use of these surveillance systems. These exclusions will create loopholes. Such exclusions are also in conflict with the EU's own 'better regulation' resolution.[11]

[5] CAIDP, AI and Democratic Values Index. https://www.caidp.org/reports/aidv-2021/

[6] United Nations Office of the High Commissioner for Human Rights. 21 July 2016. U.N. Doc. A/HRC/33/29, 21 July 2016 Report on best practices and lessons learned on how protecting and promoting rights contribute to preventing and countering violent extremist. https://undocs.org/Home/Mobile?FinalSymbol=A%2FHRC%2F33%2F29&Language=E&DeviceType=Desktop&LangRequested=False,

[7] CAIDP. December 9, 2022. Statement to the UN on AI and the Protection of Fundamental Rights. https://www.caidp.org/statements/

[8] CAIDP Campaign. October 2022. Ban Facial Surveillance Technology and Other Forms of Mass Biometric Identification. 44th Global Privacy Assembly, Istanbul - Turkey

[9] Open Letter from Civil Society. June 7, 2021. Amnesty International and more than 170 organisations call for a ban on biometric surveillance. https://www.amnesty.org/en/latest/press-release/2021/06/amnesty-international-and-more-than-170-organisations-call-for-a-ban-on-biometric-surveillance/

[10] OHCHR. 13 September–1 October 2021. Annual report of the United Nations High Commissioner for Human Rights and reports of the Office of the High Commissioner and the Secretary-General, Human Rights Council, Forty-eighth session. https://www.ohchr.org/sites/default/files/HRBodies/HRC/RegularSessions/Session48/Documents/A_HRC_48_31_AdvanceEditedVersion.docx

[11] European Parliament Resolution of 7 July 2022 on Better regulation: Joining forces to make better laws (2021/2166(INI))

CAIDP therefore recommends that the Council of the European Union, the European Parliament, and the European Commission expand the scope of application of law and narrow the exclusions in both Prohibited practices and High-Risk AI systems.

Remove the Broad Exclusions for Law Enforcement:

- Art 3(40, 41), which provides the basis for an exclusion, defines "law enforcement authority" far more broadly than the term is typically understood. The definition includes any public authority competent for the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties; *or any other body or entity entrusted by Member State law to exercise such authority; or those acting on behalf of the authorities*. If the Law Enforcement Directive (Directive (EU) 2016/680[12] is a reference here, **then exclusions are extremely wide**, and a variety of bodies might fall under the scope.[13]
- Art 5 allows these systems to be used for detection, localisation, identification or prosecution of perpetrators or suspects of 32 criminal offences. The **exclusions are not "exhaustively listed and narrowly defined situations"** as the Elements of Proposal 19 and 20 suggest. In practice, such extensive exclusion means law enforcement can use these systems without limit.
- Art 29 then exempts law enforcement authority from registration obligations (Art 51) when following high-risk AI systems in Annex III: Remote biometric identification systems, Predictive policing, Emotional detection, Profiling or systems assessing Recidivism. The Approach makes reference to Article 3(4) of Law Enforcement Directive (Directive (EU) 2016/680[14] for justification. However, the **Directive does not cover predicting occurrence of a crime or becoming a potential victim of criminal offences**.
- Similarly, "AI systems intended to be used by law enforcement authorities or on their behalf to **assess the risk** for a natural person **to become a potential victim** of criminal offences [emphasis added]" are excluded from transparency obligations. Effectively, this wording will allow authorities to surveil and profile any natural person, causing harm and breach of rights.[15] As mentioned above, **pseudoscientific predictive policing systems claiming to predict risk of becoming a perpetrator or victim should be banned**.

Remove the Ex-ante Excluded Systems:

No system should be excluded ex-ante from the scope of regulation. Such exclusions undermine the protection of fundamental rights and breach the right to an effective judicial remedy (EU Charter of Fundamental Rights - Article 47). Specifically,

- **Existing large-scale EU IT systems:** should have a timeline to conform to obligations for high-risk AI systems, instead of being excluded altogether. These systems, such as Frontex, have significant impact on fundamental rights, and are mainly used on determinations for immigration, migration and asylum, and should be included within relevant use cases under high-risk AI systems (Art 83).

[12] Law Enforcement Directive (Directive (EU) 2016/680 https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016L0680

[13] The Data Protection Commission (DPC) Ireland. Law Enforcement Directive Guidance on Competent Authorities and Scope. https://www.dataprotection.ie/en/organisations/resources-organisations/law-enforcement-directive

[14] DIRECTIVE (EU) 2016/680 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. 27 April 2016. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016L0680&from=EN

[15] Matt Stroud. May 24, 2021. Heat Listed. The Verge. https://www.theverge.com/c/22444020/chicago-pd-predictive-policing-heat-list

- **Existing AI systems:** AI systems on the market or in use by the time of the regulation coming into force should be equally subject to the regulation. Exclusion of such systems gives them an unintended advantage, creating possible monopolies or advantages for larger companies. The exclusion may also lead to rushed development or adoption of systems without due diligence (Art 83).
- **Exceptional reasons:** The Approach allows for Member States to authorise the placing on the market or putting into service of AI systems which have not undergone a conformity assessment for exceptional reasons (Art 47). Without a timeline requirement to conform, or an independent body monitoring these authorizations, this exception can reduce accountability and transparency requirements of public authorities.

Remove the National Security Exclusions:

Any system related to national security is currently excluded from scope of AI Act, regardless of if the entity carrying out those activities is a public or private entity (Art 2). National security is defined by CJEU in La Quadrature du Net (LQDN) judgement as *"primary interest in protecting the essential functions of the State and the fundamental interests of society and encompasses the prevention and punishment of activities capable of seriously destabilising the fundamental constitutional, political, economic or social structures of a country and, in particular, of directly threatening society, the population or the State itself, such as terrorist activities."*[16] The **exclusion should be limited**. The public entity should be subject to the transparency and accountability obligations if an AI system is used to monitor and predict social or political behaviour, track communications and location for 'national security' purposes**.**

Correct the Unequal Protection of Asylum Seekers and Refugee Rights:

The rights should be protected on an equal basis.[17] AI systems used for border control and asylum, refugee populations should not be subject to a different set of transparency and accountability obligations.[18]

- In EP 38, the Council acknowledges "Actions by law enforcement authorities involving certain uses of AI systems are characterised by a significant degree of **power imbalance**" and AI systems not meeting the technical requirements "may single out people in a **discriminatory or otherwise incorrect or unjust manner**", and that "exercise of important procedural fundamental rights, such as the right to an effective remedy and to a fair trial as well as the right of defence and the presumption of innocence, could be hampered, in particular, **where such AI systems are not sufficiently transparent, explainable and documented**."
- In EP 39, the Council notes "AI systems used in migration, asylum and border control management affect people who are often in a particularly vulnerable position and who are dependent on the outcome of the actions of the competent public authorities." **Yet the Approach EXEMPTS** high risk AI systems used for the purpose of law enforcement,

[16] CJEU, Grand Chamber Judgement. 6 October 2020. Joined Cases C-511/18, C-512/18, La Quadrature du Net v. France, and C-520-18, Ordre des barreaux francophones et germanophone, Académie Fiscale ASBL, UA, Liga voor Mensenrechten ASBL, Ligue des Droits de l'Homme ASBL and others v. Belgium. ECLI:EU:C:2020:791

[17] Wojciech Wiewiórowski. January 28, 2023. Privacy and data protection too often suspended at EU borders. Euractiv. https://www.euractiv.com/section/all/opinion/it-is-time-to-tear-down-this-wall/

[18] Niovi Vavoula. August 15, 2021. Artificial Intelligence (AI) at Schengen Borders: Automated Processing, Algorithmic Profiling and Facial Recognition in the Era of Techno-Solutionism. European Journal of Migration and Law.

migration, border control or asylum from certain obligations. These systems **should be included in the registration obligations** (Art 51) as per Art 29 and **have requirement for a separate verification** by at least two natural persons (Art 14).

PROPOSED CHANGES TO THE TRANSPARENCY OBLIGATIONS & ACCOUNTABILITY

CAIDP reiterates the following recommendations in relation to obligations of transparency and accountability:

- **Mandate ex-ante human rights impact assessments:** Providers of high-risk systems in Annex III should be obligated to conduct ex-ante 'human rights impact assessment'. The results of the assessment should be included in the documentation submitted to the EU database and should be publicly accessible.
- **Record Serious Incidents:** Serious incidents reported under Article 62 should be listed under the CE identification number of the Provider.
- **Require Private Users to Register**: Providers of high-risk AI systems and public users of these systems are required to register in the EU database. Similarly, private users of these high-risk systems should also register their use under the CE identification number of the Provider. Without such transparency, it is impossible for individuals, disadvantaged groups and Market Surveillance authorities to understand the impact, prevalence and current status of the particular system in use.
- **Mandate Independent, third-party auditing:** The Proposal should establish a timeline to mandate such audits for high-risk AI systems.
- **Regulate General Purpose AI (GPAI) systems:** GPAI systems have structural issues and can create significant harms. The *"ability of [language models] to pick up on both subtle biases and overtly abusive language patterns in training data, leads to risks of harms, including encountering derogatory language and experiencing discrimination at the hands of others who reproduce racist, sexist, ableist, extremist or other harmful ideologies reinforced through interactions with synthetic language."*[19] Providers of GPAI systems have the most effective means and power to test and modify these systems. While the Approach rightfully includes these systems in the Act, the Approach also removes the liability and responsibility for the main providers of the GPAI systems.[20] By pushing conformity obligations on smaller enterprises, instead of big technology companies developing GPAI systems, the Approach will harm SMEs and benefit larger developers.
- **Remove Additional Horizontal Layer for Annex III systems:** Article 6(3) in the Approach introduces a new qualifier for Providers to assess if their systems should be considered high-risk. This qualifier can result in loopholes and subjective judgements and allow Providers to avoid liability. This additional horizontal layer should be removed. Instead, ex-ante human rights impact assessments with clear methodology should be mandated to identify risks.
- **Establish obligation to terminate AI system no longer under human control:** Where high-risk AI systems generate unacceptable risks to fundamental rights, or if human control of the system is no longer possible, Providers and Users should have an affirmative

[19] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21). Association for Computing Machinery, New York, NY, USA, 610–623. https://doi.org/10.1145/3442188.3445922

[20] Stuart Russell. 22 March 2022. Statement for the IMCO-LIBE Hearing on the AI Act

obligation to terminate the system.[21] As such, the kill-switches should be a key design requirement for high-risk AI systems.

SOCIETAL INTERESTS THAT SHOULD BE SAFEGUARGED

Previously, CAIDP had warned that the Proposal mentions larger risks to society but then leaves out the impact of AI systems on collectives. Unfortunately, the Approach still leaves certain EU commitments to the voluntary acts of Providers. CAIDP insists on the following safeguards:

Protect the Environment: Require AI system providers to document impact of large AI systems (especially training systems) on the environment, emission, and waste (in line with goals of Declaration on A Green and Digital Transformation of the EU).[22]

SafeguardDisability Rights & Accessibility: Require accessibility and inclusiveness design and deployment of AI systems. Accessibility and inclusiveness cannot be voluntary. Obligations must be in line with Directive (EU) 2019/882 of the European Parliament and of the Council of 17 April 2019 on the accessibility requirements for products and services. Additionally, disability rights are protected by the UN Convention on the Rights of Persons with Disabilities. As a Party to the Convention, EU and member states are obliged to ensure accessibility of digital technologies.

Ensure Civil Society Engagement: Ensure meaningful civil society participation in standards-settings work and investigation of AI systems. The role of standardisation organisations should be limited to technical aspects and not fundamental rights or legal obligations. Article 40 should reflect a ‘requirement’ to have multi-stakeholder governance, not a ‘best effort’ from standardisation organisations.

CAIDP reminds again of the following provisions from **UNESCO Recommendations on Ethics of AI[23], which all EU Member States have endorsed.**

- AI systems **should not segregate, objectify, or undermine freedom and autonomous decision-making as well as the safety of human beings and communities**, divide and turn individuals and groups against each other, or threaten the coexistence between humans, other living beings and the natural environment. (Rec #24)

- AI system use must not violate or abuse human rights; and the AI method should be **appropriate to the context and should be based on rigorous scientific foundations.** In scenarios where decisions are understood to have an impact that is irreversible or difficult to reverse or may involve life and death decisions, **final human determination** should apply. In particular, AI systems should not be used for social scoring or mass surveillance purposes (Rec #26)

- All actors involved in the life cycle of AI systems must **comply** with applicable international law and domestic legislation, standards and practices. They should reduce the **environmental impact of AI systems**. (Rec #18)

[21] Universal Guidelines for AI. 2018. https://www.caidp.org/resources/ai-policy-frameworks/
[22] European Commission, Declaration on A Green and Digital Transformation of the EU. https://digital-strategy.ec.europa.eu/en/news/eu-countries-commit-leading-green-digital-transformation
[23] UNESCO Recommendation on the Ethics of AI. 2021. https://unesdoc.unesco.org/ark:/48223/pf0000377897

- Appropriate **oversight, impact assessment, audit and due diligence mechanisms, including whistle-blowers' protection**, should be developed to ensure accountability for AI systems and their impact throughout their lifecycle. (Rec #43)

- Governments should adopt a regulatory framework that sets out a procedure, particularly for public authorities, to carry out **ethical impact assessments on AI systems to predict consequences, mitigate risks, avoid harmful consequences, facilitate citizen participation and address societal challenges**. The assessment should also establish appropriate oversight mechanisms, including auditability, traceability and explainability, which enable the assessment of algorithms, data and design processes, as well as include external review of AI systems. (Rec #53)

Thank you for your consideration of our views. We would welcome the opportunity to discuss these recommendations with you.

Sincerely,

Marc Rotenberg
CAIDP President

Merve Hickok
CAIDP Research Director

Lyantoniette Chua
CAIDP Fellow

Dr. Grace Thomson
CAIDP Fellow

Giuliano Borter
CAIDP Fellow